

Utilitarisme pragmatique et reconnaissance d'intention dans les actes de langage indirects*

Éric Raufaste^{1,2}, Dominique Longin^{2,1}
Jean-François Bonnefon^{1,2}

1-Laboratoire Travail et Cognition (UMR 5551)

Université de Toulouse-Le-Mirail (Toulouse II)

2-Institut de Recherches en Informatique de Toulouse (UMR 5501)

Université Paul Sabatier (Toulouse III)

Résumé : Après un rappel sur les actes de langage indirects et le traitement classique de ce problème, nous proposons une approche utilitariste de la reconnaissance d'intention dans le langage verbal. Cette approche combine une « heuristique utilitariste » basée sur la distribution d'utilité du locuteur, inférée à partir du contexte, et une « heuristique conventionnaliste » basée sur la reconnaissance de propriétés spécifiques de l'énoncé cible. Cette approche est mise en perspective avec d'autres approches récentes de la littérature en sciences cognitives.

Mots-clés : reconnaissance d'intention, actes de langage, pertinence, heuristiques, compréhension.

Abstract: The so-called “standard pragmatic view” of intention recognition in indirect speech acts was largely disconfirmed by psycholinguistic research (Gibbs, 1994). We propose a utilitarianist approach to pragmatics. Basically, the framework proposed here assumes that both utterance production and intention recognition can be viewed as a decision-making problem using subjective expected utility.

We assume that, as part of the standard equipment of a normal adult's theory of mind, is included the idea that individuals act so as to maximize their well-being while minimizing suffering. We then define the idea of *utilitarian context* as a collection of contextual elements that provide information about the speaker's distribution of utility over the possible behavioral responses of the listener. Hence, the utilitarian context enables the listener building a representation of the speaker's distribution of utility. For the listener, using the *utilitarian heuristic* consists of attributing the intent to obtain the state of the world that presumably maximizes the speaker's expected utility. In some neutral contexts, however, the distribution of utility cannot be built and therefore, the utilitarian heuristic cannot be used. In such cases, the speaker may recognize a *form of indirection* in the utterance, form that is conventionally associated with a specific intent. The listener uses the *conventionalist heuristic* when attributing the intent that is conventionally associated with that particular form of indirection. We present the notion of “indirection schemes” defined as logical rules that select (1) the intent conventionally associated with the form of indirection of the focal utterance, or (2) the intent with the maximum expected utility to the meaning of the focal utterance.

* Ce travail a été soutenu financièrement par le Programme Interdisciplinaire TCAN.

Finally, our approach is contrasted with other approaches to intention recognition that have been used in cognitive science.

Keywords: intention recognition, speech acts, relevance, heuristics, understanding.

Introduction

Cet article concerne les processus de reconnaissance de l'intention sous-jacente à la production d'un énoncé isolé. Nous ne traiterons pas de l'ensemble des processus de l'interlocution, ce qui impliquerait de modéliser la négociation des enjeux de la communication (voir Trognon & Ghiglione, 1993), la gestion des tours de parole, etc. Nous ne traiterons pas non plus des aspects paraverbaux comme l'utilisation de la prosodie.

La notion d'acte de langage (« speech act ») remonte à Austin (1962). Rompant avec le courant positiviste du langage qui considérait que les seuls énoncés bien formés étaient les propositions vraies ou fausses (les autres énoncés étant qualifiés de « pseudo-affirmations » ou « non-sens »), Austin constate qu'une bonne part de ces « non-sens » ne sont pas des pseudo-affirmations pour la simple raison... qu'elles n'ont jamais été destinées à être des affirmations ! Austin distingue alors les *constatifs* (les affirmations) des *non-sens*, et de ce qu'il appelle alors les *performatifs*, dont la caractéristique est de viser à faire quelque chose (baptiser, marier, ouvrir une séance...). En ce sens, *dire, c'est faire*, ce qui le pousse à dresser une taxinomie de ce qu'il appelle « actes illocutoires », *i.e.* les actes produits *en* disant quelque chose, et consistant à rendre manifeste *comment* les paroles doivent être comprises au moment de l'énonciation).

Searle (1969) va montrer la faiblesse de la taxinomie d'Austin et en proposer une nouvelle (qui est celle que nous utilisons présentement) déterminée par le *but illocutoire* des énoncés. Ce but partitionne les énoncés en cinq types d'actes de langage¹. Le *but assertif* est d'affirmer la véracité d'une proposition. Le *but commissif* détermine un engagement du locuteur vis-à-vis de l'auditeur à faire quelque chose. Le *but directif* vise à

faire faire quelque chose à l'auditeur. Le *but expressif* est de communiquer des informations sur l'état psychologique du locuteur. Enfin le *but déclaratif* vise à accomplir des transformations du monde social par le seul fait de l'énonciation². Dans le présent article, nous considérerons essentiellement les actes directifs (*i.e.* à but illocutoire directif).

Nous appelons « acte principal » l'acte que le locuteur cherche à accomplir de manière ultime. Parfois, l'acte principal correspond à l'acte littéral (*i.e.* l'acte accompli en interprétant littéralement ce qui est dit), parfois non. Par exemple, dans « Peux-tu me passer le sel ? » l'acte littéral correspond à une question dont la réponse attendue est « oui » ou « non ». Mais en disant cela dans certains contextes (par exemple, à table), le locuteur peut avoir l'intention de signifier « Passe-moi le sel ». Dans ce contexte, l'acte principal est donc cet acte-ci.

Quand l'acte principal ne correspond pas à l'acte littéral, on parle « d'acte indirect ». Dans l'exemple précédent, si l'intention du locuteur est d'obtenir le sel, il a indirectement demandé à l'auditeur de lui passer le sel (en lui demandant s'il pouvait le lui passer). Le locuteur a donc, dans ce contexte, réalisé un *acte directif indirect*.³

1. L'approche classique de l'attribution d'intention

Le « modèle pragmatique standard » (Gibbs, 1994) repose sur l'idée du « principe de coopération » (Grice, 1975, 1978). Selon Grice, les interlocuteurs coopératifs adhèrent, et s'attendent à ce que les autres adhèrent, à

¹ Nous assimilons ici, comme il est fréquent, actes illocutoires et actes de langage.

² Les « performatifs » initialement identifiés par Austin.

³ En fait, la communication non littérale ne se borne pas à la communication indirecte, mais nous n'étudions pour l'instant qu'elle (même si nous pensons que notre modèle est compatible avec les autres formes de communication non littérale). Voir l'article de Longin dans le présent ouvrage pour plus de détails.

quatre « maximes conversationnelles » : il faut transmettre ni trop ni trop peu d'information (maxime de quantité), que cette information soit fiable (maxime de qualité), pertinente (maxime de relation), exprimée clairement et sans ambiguïté (maxime de manière). Une violation de maxime traduit la volonté de transmettre *indirectement* une information. L'auditeur déclenche alors un processus d'inférence (Grice, 1975) afin de déterminer l'intention réelle du locuteur. Ce principe implique (1) que l'interprétation du sens littéral est obligatoire ; (2) que le sens non littéral n'est recherché qu'en cas de découverte d'un sens littéral défectif ; (3) que des inférences additionnelles sont ensuite réalisées pour atteindre le sens non littéral. Dans la pratique, le modèle standard est battu en brèche par de nombreux résultats expérimentaux qui montrent, entre autres, que dans la compréhension des requêtes indirectes on n'accède pas toujours au sens littéral (Gibbs, 1983) ; ou encore qu'on peut accéder au sens non littéral alors même que le sens littéral est parfaitement acceptable (*i.e.* non défectif) dans le contexte (Glucksberg, Gildea, & Bookin, 1982).

Selon Sadock (2004), on peut observer dans le développement de la théorie des actes de langage un mouvement de balancier entre des approches que nous qualifierons ici de « *conventionnalistes* » et d'autres que nous appellerons « *intentionnalistes* ». Les premières, à commencer par Austin (1962) et Searle (1969), mettent l'accent sur le côté conventionnel de la relation entre actes locutoires et illocutoires. Les secondes (e.g., Grice, 1989; Strawson, 1971) mettent l'accent sur le rôle de l'intention suggérée par le locuteur. Ainsi, rejetant l'idée que les conventions puissent traiter seules l'ensemble du problème, Grice (1957, 1989) a introduit l'idée que les conventions ne sont pas tant déterminantes que les intentions que le locuteur rend manifestes à son auditoire. Par exemple, l'énoncé « il fait chaud » dans une salle où la fenêtre est fermée peut être interprété comme une requête indirecte pour que quelqu'un ouvre la fenê-

tre. Mais rien dans la forme linguistique de l'énoncé ne permet d'inférer cette intention. Autrement dit, l'énoncé n'est pas directement porteur du sens mais fait partie d'un ensemble d'indices permettant à l'auditoire d'accéder à l'intention sous-jacente. En résumé, l'approche conventionnaliste met l'accent sur les propriétés syntaxico-sémantiques de l'énoncé tandis que l'approche intentionnaliste met l'accent sur l'apport informationnel des contextes linguistique et extralinguistique.

2. L'approche utilitariste de la reconnaissance d'intention

De notre point de vue, l'opposition entre les approches conventionnaliste et intentionnaliste témoigne en fait d'une complémentarité. Ainsi, dans le cadre de la théorie de la pertinence, Sperber et Wilson (1989) énoncent que : « On peut diviser les hypothèses qui se trouvent, au début d'un processus de déduction, dans la mémoire du dispositif déductif en deux sous-ensemble disjoints, et considérer que chacun de ces sous-ensembles est traité dans le contexte de l'autre. » (p. 208). Ces deux sous-ensembles disjoints servent à décrire l'information déjà présente dans le système d'une part, et l'information nouvellement entrante d'autre part. Cette analyse suggère que, du point de vue des processus d'inférence, le contexte interprétatif peut être tour à tour la représentation préalablement activée dans le système, et celle activée par l'information nouvellement entrante (si l'on reconsidère la représentation préalable). Nous appellerons « *énoncé focal* » l'énoncé entrant et réserverons le terme « *contexte* » à la représentation déjà activée au moment de l'énonciation. Ainsi les processus qui s'appuient sur le sous-ensemble des données contextuelles tombent dans le champ de l'approche intentionnaliste tandis que les processus qui s'appuient sur le sous-ensemble des informations entrantes tombent dans le champ de l'approche conventionnaliste. Loin d'y voir une opposition fondamentale, rien n'empêche de supposer qu'un même

auditeur sera capable de passer de l'un à l'autre selon les occasions, parfois même au cours du traitement d'un même énoncé.

Une fois admise la complémentarité des deux approches, on se trouve à postuler la coexistence de deux mécanismes, l'un partant des inférences préalablement réalisées sur les intentions candidates, et l'autre partant des propriétés formelles de l'énoncé pour activer de façon privilégiée les intentions typiquement associées. Nous appellerons « *heuristique utilitariste* » le mécanisme postulé pour l'utilisation d'un contexte contenant une gamme d'intentions plausibles activées, et « *heuristique conventionnaliste* » celui concernant le décodage de la forme linguistique de l'énoncé. Dans la Figure 1, ces deux mécanismes sont replacés par rapport aux deux espaces qu'ils ont en entrée et en sortie.

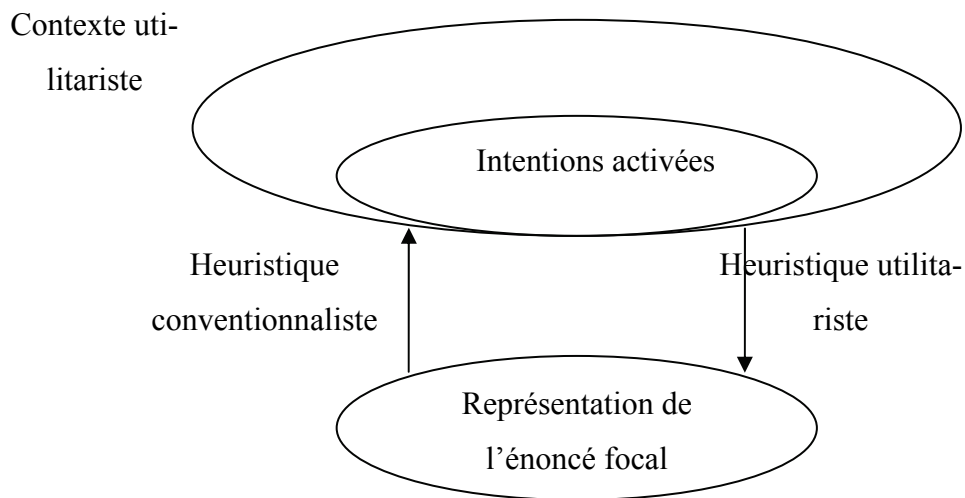


Figure 1: Esquisse d'une théorie de l'attribution d'intention

Un tel schéma n'exclut pas la possibilité d'un aller-retour entre les deux mécanismes, ce qui pose la question du critère d'arrêt du cycle. La cohérence, au sens de l'intégration de l'énoncé focal et du contexte utilitariste avec le minimum de contradiction résiduelle, pourrait être ce critère (voir Gernsbacher & Givón, 1995; Kintsch, 1988; et Thagard, 2000). Une

autre question est celle de la gestion de la compétition entre les deux mécanismes. Ainsi, Strawson (1971) distinguait entre des actes cérémonieux (comme le mariage) où la convention est le mécanisme illocutoire principal, et les actes communicatifs (questionner, affirmer) pour lesquels l'intention est primordiale. Au-delà de cette distinction, on peut envisager que cette compétition soit gérée par les processus généraux de gestion des conflits entre règles (cf., Anderson et al., 2004).

Nous allons maintenant détailler successivement l'heuristique utilitariste et l'heuristique conventionnaliste.

3. L'hypothèse de l'heuristique utilitariste

En réponse à un acte de langage, l'auditeur va produire un comportement dont les conséquences seront plus ou moins favorables au locuteur. Évidemment, rien ne garantit que l'état du monde résultant soit celui attendu par le locuteur. Au moment de choisir son énoncé, celui-ci se trouve donc dans une situation de décision sous incertitude. Il paraît raisonnable d'assumer qu'un agent rationnel produira alors l'énoncé ayant l'espérance d'utilité subjective maximale. Ce calcul, étant de nature profondément subjective, reste bien sûr soumis aux biais classiques liés à l'estimation des probabilités subjectives, du cadrage de la situation de décision, etc. L'heuristique utilitariste est un processus qui repose sur l'existence d'une représentation de la distribution d'utilité du locuteur dans l'esprit de l'auditeur. Cette heuristique suppose (1) la construction d'une représentation de la distribution d'utilité du locuteur ; et (2) la sélection d'une intention candidate selon une fonction de cette distribution d'utilité et des propriétés linguistiques de l'énoncé.

Tout auditeur adulte normalement constitué possède une théorie de l'esprit, nécessaire à la vie sociale (e.g., Leslie, Friedman, & German, in press), et qui lui permet de construire une représentation des intentions

possibles du locuteur. Nous postulons que la théorie de l'esprit d'un individu adulte inclut normalement le principe utilitariste classique de Bentham selon lequel tout individu agit de manière à maximiser son bien-être propre et à minimiser la souffrance. Ainsi équipé, l'auditeur peut s'appuyer sur sa représentation de la distribution d'utilité du locuteur pour inférer que l'intention qui a motivé un énoncé correspond à la recherche d'un effet sur l'auditeur dont l'utilité sera maximale pour le locuteur.

Nous définissons *l'heuristique utilitariste* comme ***l'attribution à un acteur A par un observateur O de l'intention d'obtenir l'effet plausiblement le plus utile pour A.*** Cette heuristique utilitariste est formulée d'une manière qui ne qualifie pas l'acte à interpréter comme spécifiquement langagier. Autrement dit, nous pensons que cette heuristique peut s'appliquer pour l'interprétation de n'importe quel type d'acte. Néanmoins, dans le cadre du présent article, nous donnerons une définition plus spécifique des actes de langage : nous définirons une « *heuristique utilitariste conversationnelle* » comme ***l'attribution à un locuteur S par un auditeur H de l'intention d'obtenir l'effet perlocutoire plausiblement le plus utile pour S dans le contexte de l'élocution.*** Par exemple, la question « Est-ce que vous fumez ? » peut induire plusieurs réponses comportementales. Une information peut être fournie « oui, je fume », ou bien une cigarette peut être offerte « Oui, vous en voulez une ? ». Si H a des raisons de penser que pour le S, il serait plus utile d'obtenir une cigarette que d'obtenir l'information et s'il applique l'heuristique utilitariste alors H attribuera à S l'intention d'obtenir une cigarette.

L'heuristique utilitariste ne peut s'appliquer que si l'on possède au préalable une distribution d'utilité permettant de discriminer entre les différentes intentions possibles. Un second concept doit donc rendre compte de la construction de cette distribution d'utilité. Nous définissons le *contexte utilitariste* d'un énoncé comme *l'ensemble des indices contex-*

tuels permettant à l'auditeur de construire la fonction d'utilité du locuteur sur les réponses comportementales potentielles de l'auditeur.

Les indices composant le contexte utilitariste de la question sont ensuite intégrés aux connaissances préalables de l'auditeur, permettant ainsi à celui-ci de construire une représentation des préférences du locuteur. Par exemple, si le contexte mentionne un goût du locuteur pour la consommation de cigarettes, la représentation de la distribution d'utilité de ce dernier va automatiquement attribuer de la valeur aux états du monde permettant au locuteur de fumer. La seconde réponse (la proposition d'une cigarette) peut seulement être obtenue si l'énoncé est conçu comme une requête indirecte, à travers la reconnaissance de son sens non littéral. Nous prévoyons que le contexte utilitariste sera un fort prédicteur d'une telle reconnaissance, et ce particulièrement face à des requêtes indirectes *non conventionnalisées*, c'est-à-dire dont la structure ne permet pas un traitement purement à base de conventions.

Dans une expérience récente (Raufaste, Bonnefon, & Longin, manuscrit en préparation), nous avons manipulé le contexte utilitariste de quatre énoncés, en fournissant un contexte minimal renseignant les préférences du locuteur. Les énoncés n'étaient pas des requêtes indirectes conventionnelles (du type « Pouvez-vous me passer le sel ? ») mais des phrases ambiguës (telles que « Je n'ai pas eu le temps de nettoyer les vitres », potentiellement, une requête adressée à l'auditeur pour qu'il le fasse). À titre d'exemple de notre manipulation expérimentale, la question « Fumez-vous ? » était posée, selon le scénario, dans deux contextes utilitaristes différents. Dans un premier scénario, le locuteur était un fumeur occasionnel interviewant une candidate à un poste de baby-sitter pour son enfant. Cela suggérait donc aux participants que le locuteur avait une distribution d'utilité telle que le fait d'obtenir l'information demandée était plus important que d'obtenir d'une cigarette. Dans un second scénario, le

contexte présentait l'énoncé focal comme prononcé par un gros fumeur n'ayant plus de cigarette et interviewant une candidate à un poste de femme de ménage. Ce dernier contexte suggérait une utilité du fait d'obtenir une cigarette supérieure à l'utilité d'obtenir simplement l'information. Les participants devaient évaluer leur accord avec le fait que le sens littéral de l'énoncé (ou, alternativement, son sens non littéral) était l'intention principale du locuteur.

Les résultats ont montré une forte influence du contexte utilitariste sur ces évaluations, et ce dans la direction attendue. Toutefois les participants n'ont pas marqué de désaccord sévère avec le sens littéral de l'énoncé, même quand le contexte utilitariste était supposé favoriser le sens non littéral. En revanche, ils étaient fortement en désaccord avec le sens non littéral quand le contexte utilitariste était supposé défavoriser ce dernier. Afin de vérifier si cette asymétrie relevait d'un artefact de la procédure expérimentale, une seconde expérience a été conduite dans laquelle les participants étaient libres de générer leur interprétation de l'énoncé, au lieu de se prononcer sur un sens littéral ou non littéral fourni. L'influence du contexte utilitariste a été encore plus forte que lors de la première expérience, dans la direction prédite par l'heuristique utilitariste.

4. L'hypothèse conventionnaliste

Ce que nous nommons ici « heuristique conventionnaliste » répond à une idée très classique dans la littérature sur l'interprétation d'actes de langage, à savoir que le traitement par l'auditeur de certaines propriétés linguistiques de l'énoncé peut suffire à inférer l'intention du locuteur. L'heuristique conventionnaliste est un processus qui repose sur l'existence de conventions sociales basées sur la structure linguistique des énoncés focaux. Ce processus comprend deux étapes (1) la construction d'une représentation linguistique de l'énoncé focal ; (2) l'application

d'une règle conventionnelle d'inférence générant en sortie une représentation de l'intention du locuteur.

Un exemple typique d'énoncé très conventionnalisé est « Peux-tu me passer le sel ? ». Face à un tel énoncé, et en l'absence d'un contexte particulier susceptible de bloquer l'inférence, un adulte ayant une maîtrise normale du français comprendra immédiatement que le locuteur ne souhaite pas savoir si l'auditeur a la possibilité de passer le sel mais veut que celui-ci accomplisse physiquement l'action suggérée. La force de ces conventions est qu'elles peuvent fonctionner en l'absence d'un contexte utilitariste clair. Avec un énoncé extrêmement conventionnalisé comme le précédent, on peut aller jusqu'à supposer que cette phrase donnée seule au début d'un roman activera immédiatement dans l'esprit du lecteur un schéma comportant une situation de repas avec au moins deux personnes, le fait qu'il est plus facile à l'auditeur qu'au locuteur d'attraper le sel, etc.

Nous définissons donc *l'heuristique conventionnaliste* comme *l'attribution à un acteur A par un observateur O de l'intention d'obtenir l'effet conventionnellement associé à la structure de l'acte.*

Cette heuristique est formulée d'une manière qui ne qualifie pas l'acte à interpréter comme langagier. Dans le cadre plus spécifique des actes de langage, nous définirons donc une *heuristique « conventionnaliste langagière »* comme *l'attribution à un locuteur S par un auditeur H de l'intention d'obtenir un effet conventionnellement associé à la structure linguistique de l'énoncé focal.*

5. Formalisation des deux heuristiques

Ces définitions ne seraient pas d'une grande utilité sans une méthode appropriée d'analyse des structures linguistiques et des contextes. Les notions de *forme d'indirection* et de *schéma d'indirection* introduites par Longin (Champagne, Herzig, Longin, Nespoulous, & Virbel, 2002; Lon-

gin & Raufaste, 2003), peuvent s'appliquer pour modéliser les heuristiques conventionnaliste et utilitariste.

Searle (1979) a montré qu'un acte de langage indirect pouvait être accompli *via* une assertion ou une question fermée (dont la réponse attendue est « oui » ou « non ») sur deux types principaux d'arguments : des conditions de succès de l'acte illocutoire indirect visé (Catégorie 1) ou l'existence de raisons pour faire l'action dont il est question dans l'énoncé considéré (Catégorie 2). Par exemple, « Peux-tu me passer le sel ? » est une question sur la condition préparatoire de l'acte illocutoire « Passe-moi le sel ? ». Cela suffit, d'après Searle, à faire du premier énoncé une manière possible de signifier le second (Catégorie 1). Par ailleurs, « Tu devrais moins manger » évoque l'existence d'une raison de faire une certaine action (moins manger). C'est donc un énoncé susceptible de signifier « Mange moins », et il fait partie de la Catégorie 2. Virbel a étendu les travaux de Searle en ajoutant une sous-catégorie supplémentaire à chacune des catégories évoquées par Searle. Longin a lui-même étendu la classification de Virbel, en prenant en compte d'autres types d'arguments (cf. l'article « Des raisons qu'ont certains actes à être indirects » dans le présent ouvrage). Selon ce point de vue, Longin définit une *forme d'indirection* (Champagne et al., 2002; Longin & Raufaste, 2003) comme un énoncé appartenant à l'une ou l'autre des catégories des classifications susmentionnées. Un *schéma d'indirection* est la représentation du fait que, dans un *contexte minimal d'interprétation*⁴, l'énonciation d'une forme d'indirection est susceptible de produire sur l'auditeur les effets associés à un certain acte de langage (l'acte illocutoire indirect visé).

⁴ Le *contexte minimal d'interprétation* (CMI) représente l'ensemble minimal des informations nécessaires à une interprétation indirecte dans le cadre de l'évocation de partie de raisons (pour les trois autres types d'arguments, il est vide). Le CMI n'est cependant pas suffisant pour inférer que l'acte indirect ait effectivement eu lieu.

Dans l'article précité du présent numéro, Longin se centre sur les actes directifs, et sur la Catégorie 2 des formes d'indirection, les catégories 1 et 3 étant pressenties comme des cas particuliers de la Catégorie 2. L'idée sous-jacente est que s'il y a pour le locuteur une utilité quelconque à ce que son énoncé réalise indirectement une demande de faire une certaine action, c'est que son énoncé évoque l'existence, le contenu, ou une partie de contenu, de raisons pour faire cette action, ou l'inexistence de raison pour ne pas la faire. Par exemple : « Tu devrais me passer le sel » (évocation de l'existence d'une raison pour passer le sel) ; « De par la position que j'occupe, tu es tenu de me passer le sel » (évocation d'une raison de passer le sel) ; « C'est fade » (évocation d'une partie de raison, associée au *CMI* selon lequel il n'est pas souhaitable de manger un plat qui est fade, et que mettre du sel réduit la fadeur) ; « Il n'y a aucune raison à ce que tu ne me passes pas le sel » (évocation de l'inexistence de raisons pour ne pas passer le sel) sont autant d'énoncés pouvant signifier, dans un certain contexte d'énonciation, « Passe-moi le sel ». Ces quatre types d'arguments se dérivent eux-mêmes en plusieurs sous-catégories, selon que (1) le mode d'énonciation est l'assertion ou la question ; (2) la *portée* du *type d'argument* est *une demande de faire* ou *de ne pas faire une action*, *une certaine action*, ou *autre chose qu'une certaine action* ; (3) que le *lien entre le type d'argument et portée* est *une demande de faire* ou *de ne pas faire*. Au total, donc, 64 cas entrent dans la Catégorie 2. Pour plus de détails et d'exemples, voir l'article de Longin (cf. ce numéro).

Dans les deux expériences précédemment évoquées, les énoncés focaux proposés aux sujets étaient présentés dans un contexte utilitariste. Afin de tester l'hypothèse conventionnaliste, des énoncés focaux ont aussi été présentés en contexte neutre, c'est-à-dire ne fournissant aucune indication sur la distribution d'utilité du locuteur. Ces énoncés focaux, toutefois, possédaient une structure linguistique correspondant à différentes formes

d'indirection connues. Conformément à l'hypothèse conventionnaliste, les sujets des deux expériences ont très majoritairement attribué au locuteur une intention correspondant à une interprétation non littérale.

6. Perspectives de modélisation logique

Actuellement, les schémas d'indirection sont intégrés à une logique de la croyance, de l'intention, et de l'action, *via* des lois d'actions conditionnelles. Les lois d'action caractérisent formellement ce qu'est une action. Elles sont de la forme $A_i \rightarrow \text{After}_a B_i$ (pour $i = 1, 2, \dots$) ce qui signifie que « si A_i est vrai, alors toute exécution de l'action a engendrera l'effet B_i ». Dans ce type de lois, A_1 représente le contexte dans lequel l'acte doit être interprété de façon littérale, et A_i (pour $i > 1$), ce sont les différents contextes où l'acte doit être compris de façon indirecte. Le *contexte minimal d'interprétation* φ_i dérive des A_i (*i.e.* $A_i \rightarrow \varphi_i$ pour tout $i = 1, 2, \dots$) sans pour autant leur être équivalent. B_1 est l'effet littéral de l'acte a (il représente formellement l'intention du locuteur associée à son énoncé pris au sens littéral), alors que les B_i (pour $i > 1$) représentent les différents effets indirects (intentions liées au sens non littéral) associés à chaque interprétation indirecte possible.

Chaque individu dispose de ses propres lois d'action, qui ne sont pas nécessairement les mêmes que celles des autres. Jusqu'à présent, nous ne disposons pas des croyances précises d'un individu donné (en l'occurrence, le destinataire d'un acte de langage) permettant à ce dernier de caractériser de façon unique une interprétation donnée, c'est-à-dire que formellement nous ne pouvons pas fournir une formule plus précise que la simple existence des A_i (pour $i = 1, 2, \dots$). L'hypothèse utilitariste vient pallier ce manque : si nous supposons que chaque individu i_1 dispose d'une modalité supplémentaire (par exemple U) indicée par un individu i_2 telle que $U_{i_2} A$ signifie « dans la situation présente, l'état dans lequel A est

vrai a une utilité maximale pour l'agent i_2 par rapport aux états où A est faux ». Dans ces conditions, les lois d'action peuvent être représentées de la façon suivante (S représentant le locuteur) : $U_S B_i \rightarrow After_a B_i$ (pour $i = 1, 2, \dots$). Autrement dit, « si l'état dans lequel l'effet B_i de l'action a conduit à un état dont l'utilité est maximale pour le locuteur, alors c'est que cet effet s'est produit. Techniquement, cette loi est précédée de l'agent ayant cette croyance (en l'occurrence, l'auditeur H). Dans le cas particulier de l'évocation de parties de raisons, le *contexte minimal d'interprétation* φ_i doit également être vérifié pour que l'effet puisse être dérivé. Soit : $(U_S B_i \ \& \ \varphi_i) \rightarrow After_a B_i$ (pour $i = 1, 2, \dots$), où « & » est le symbole de la conjonction logique.

7. Discussion

Notre approche remet fondamentalement en cause le modèle standard. Nous postulons une construction incrémentale de l'intention, sans statut particulier à l'interprétation littérale, mais, plus profondément, nous n'admettons pas le principe de coopération de Grice. Du point de vue utilitariste, la distribution d'utilité des différents acteurs de la communication sert à déterminer s'il est raisonnable ou non d'admettre qu'il y a coopération. Il s'ensuit que si un acteur a un intérêt fort à ne pas être réellement coopératif, il est de la plus haute importance de tenir compte de cet intérêt pour éviter d'éventuelles manipulations. Imaginons par exemple une situation de vente immobilière. Il serait naïf d'assumer que le vendeur est *a priori* sincère et respecte la maxime de qualité. Il paraît plus prudent de supposer une part d'enjolivement de la réalité, voire de mensonge, car l'utilité d'une conclusion positive de la transaction est majeure pour le vendeur. En d'autres termes, nous tenons que le principe de coopération n'est admis et utilisé qu'à condition de ne pas entrer en conflit avec la distribution d'utilité attribuée au locuteur.

D'autres hypothèses importantes en matière d'interprétation des actes indirects peuvent être réinterprétées à la lumière de l'approche utilitariste. Ainsi, selon l'hypothèse de la préservation de la face, les auditeurs s'appuieraient sur le fait qu'une partie des requêtes indirectes sont choisies de manière à ce que les différents interlocuteurs évitent de perdre la face. Selon Holtgraves (1999), les auditeurs en inféreraient que l'interprétation la plus plausible serait celle potentiellement la plus menaçante pour la face sociale dans ce contexte. De notre point de vue, il ne s'agit en fait que d'un cas particulier de l'heuristique utilitariste : celui où l'enjeu majeur est négatif (motivation à éviter les mauvaises conséquences d'une action), ce qui néglige complètement les enjeux positifs (où l'intention du locuteur est d'obtenir un résultat plaisant par exemple). Néanmoins, ces travaux soulèvent l'intéressante question de la compétition entre des intentions positives et négatives coexistantes.

Fodor (1992, cité dans Wimmer & Weichbold, 1994) a envisagé l'hypothèse que les enfants, par économie cognitive, appliqueraient une heuristique du type « Prédire que l'agent se comportera d'une façon qui satisfera ses désirs ». Cette heuristique est très proche de l'heuristique utilitariste proposée ici. Nous notons cependant que l'heuristique utilitariste n'est pas limitée aux désirs mais inclut aussi une composante « incertitude ». Cet élément est essentiel. En effet, l'omettre peut conduire à attribuer au locuteur l'intention d'obtenir un état du monde extrêmement désirable même si l'on pense par ailleurs que le locuteur ne croit pas que cet état du monde soit possible !

L'approche par plans en IA est principalement fondées sur les travaux en philosophie d'Austin (1962), Grice (1957, 1989), Searle (1969) et Bratman (1990). Elle considère que les énoncés sont des actions finalisées au sein d'un plan plus global que l'on va chercher à découvrir (d'où le terme de « reconnaissance de plans »). Bien que les premières applica-

tions au traitement du langage de cette approche visaient la conformité au modèle pragmatique standard (cf. Allen & Perrault, 1980), des développements récents de cette approche, indépendants de nos propres travaux, intègrent maintenant explicitement la notion d'utilité espérée dans le calcul du jugement social (Mao & Gratch, 2004). À notre connaissance, ces développements n'ont pas encore été appliqués au traitement des actes de langage mais notre approche indique qu'elles pourraient l'être avec profit.

Une autre approche intéressante est l'hypothèse du blocage, selon laquelle, dans la production et l'interprétation des actes de langage indirects, les locuteurs feraient mention explicite de l'obstacle le plus important qui pourrait empêcher l'auditeur d'accomplir la requête. (Francik & Clark, 1985). L'énoncé même de cette hypothèse la pose comme une sous-classe de l'heuristique conventionnaliste puisqu'elle est fondée sur une propriété de l'énoncé : la mention d'un point potentiel de blocage. Par contre le caractère « conventionnel » de la mention d'un point de blocage est questionnable. Quoi qu'il en soit, il est clair que notre approche permet des interprétations d'acte directifs indirects correspondant à des désirs et non à des blocages (par exemple « il fait chaud » pour demander d'ouvrir la fenêtre). Certains cas de blocages rentrent clairement dans le cadre de notre heuristique conventionnaliste. « Peux-tu me passer le sel », par exemple, mentionne un blocage potentiel. Est-ce toujours le cas ? Cette étude reste à faire afin de préciser si notre approche inclut complètement ou seulement partiellement l'hypothèse du blocage.

Les perspectives offertes par notre approche sont très riches. Au plan empirique, il reste de nombreux aspects du modèle à tester relativement à la cohérence interne (mécanismes d'arrêt du cycle, choix du niveau d'entrée dans une hiérarchie d'intentions, etc.). Au plan de la généralisabilité, les limites de notre modèle ne sont pas encore claires. Il paraît probable que l'approche utilitariste puisse se généraliser à d'autres cas de

langage figural que les directifs indirects (par exemple, au traitement des implicatures, ou à celui de certaines métaphores), mais un important travail d'analyse reste à réaliser à cet égard.

Du point de vue logique, l'intégration d'un opérateur d'utilité et l'analyse formelle de l'interaction d'un tel opérateur modal avec les autres opérateurs (croyance, intention, action) restent à élaborer rigoureusement.

En termes de développement computationnels, l'approche proposée peut s'implémenter dans une architecture cognitive, à ceci près qu'à notre connaissance, celles existant actuellement ne traitent pas encore le problème de l'utilité de façon satisfaisante. Néanmoins, une architecture comme ACT-R (Anderson & Lebiere, 1998) permet de représenter des connaissances sur les utilités. ACT-R permet aussi de traiter de façon satisfaisante les aspects liés à l'incertitude ainsi que les règles d'inférences conventionnalisées. Cela devrait être suffisant pour réaliser une première implémentation de nos idées dans une simulation par ailleurs compatible avec une théorie intégrée de l'esprit (Anderson et al., 2004).

RÉFÉRENCES

- Allen, J. F., & Perrault, R. (1980). Analyzing intention in utterances. *Artificial Intelligence*, 15, 143-178.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111, 1036-1060.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Austin, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.
- Bratman, M. E. (1990). What is intention? In P. R. Cohen, J. Morgan & M. E. Pollack (Eds.), *Intentions in communication* (pp. 15-31). Cambridge, MA: The MIT Press.

- Champagne, M., Herzig, A., Longin, D., Nespoulous, J.-L., & Virbel, J. (2002). Formalisation pluridisciplinaire de l'inférence de certains types d'actes de langage non littéraires. *Information-Interaction-Intelligence (I3)*, numéro hors-série, 197-225.
- Francik, E. P., & Clark, H. H. (1985). How to make requests that overcome obstacles to compliance. *Journal of Memory and Language*, 24, 560-568.
- Gernsbacher, M. A., & Givón, T. (1995). *Coherence in spontaneous text*. Amsterdam: John Benjamins.
- Gibbs, R. W., Jr. (1994). Figurative thought and figurative language. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 411-446). San Diego, CA: Academic Press.
- Gibbs, R. W. J. (1983). Do people always process the literal meaning of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 524-533.
- Glucksberg, S., Gildea, P., & Bookin, M. B. (1982). On understanding non-literal speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, 21, 85-98.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377-388.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41-58). New York: Academic Press.
- Grice, H. P. (1978). Some further notes on logic and conversation. In P. Cole (Ed.), *Syntax and semantics* (Vol. 9, pp. 113-128). New York: Academic Press.
- Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Holtgraves, T. (1999). Comprehending indirect replies: When and how are their conveyed meanings activated? *Journal of Memory and Language*, 41, 519-540.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163-182.

- Leslie, A. M., Friedman, O., & German, T. P. (in press). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*.
- Longin, D., & Raufaste, E. (2003). Actes de langage indirects: Co-construction d'un modèle logico-psychologique. In A. Herzig, B. Chaib-Draa & P. Mathieu (Eds.), *Actes des secondes journées francophones modèles formels de l'interaction, mfi'03*, (pp. 169-178). (Lille, 20-22 mai 2003), Toulouse: Cépaduès-Éditions.
- Mao, W., & Gratch, J. (2004). *A utility-based approach to intention recognition*. Paper presented at the AAMAS 2004 Workshop on Agent Tracking: Modeling Other Agents from Observations.
- Sadock, J. (2004). Speech acts. In L. R. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 53-73). Oxford: Blackwell.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, J. R. (1979). *Expression and meaning*. Cambridge: Cambridge University Press.
- Sperber, D., & Wilson, D. (1989). *La pertinence: Communication et cognition* (A. Gerschenfeld & D. Sperber, Trans.). Paris: Editions de Minuit.
- Strawson, P. F. (1971). Intention and convention in speech acts. In J. R. Searle (Ed.), *The philosophy of language* (pp. 23-38). London: Oxford University Press.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MA: The MIT Press.
- Trognon, A., & Ghiglione, R. (1993). *Où va la pragmatique? De la pragmatique à la psychologie sociale*. Grenoble: Presses Universitaires de Grenoble.
- Wimmer, H., & Weichbold, V. (1994). Children's theory of mind: Fodor's heuristics examined. *Cognition*, 53, 45-57.