

## Utilité et conventions : les deux faces d'une même pièce pour l'attribution d'intentions\*

Jean-François Bonnefon<sup>1,2</sup>      Dominique Longin<sup>2,1</sup>      Éric Raufaste<sup>1,2</sup>  
bonnefon@univ-tlse2.fr      Dominique.Longin@irit.fr      raufaste@univ-tlse2.fr

<sup>1</sup>Laboratoire Travail et Cognition (UMR 5551)  
Université Toulouse-le-Mirail, Maison de la Recherche,  
5, allées Antonio-Machado, 31058 Toulouse cedex 9

<sup>2</sup>Institut de Recherche en Informatique de Toulouse (UMR 5505)  
Université Paul Sabatier  
118 route de Narbonne, 31062 Toulouse cedex 4

### Résumé :

Le travail présenté ici est le fruit d'une recherche pluridisciplinaire sur les actes de langage indirects. L'approche adoptée pour la reconnaissance d'intentions indirectes intègre une « heuristique utilitariste » et une « heuristique conventionnaliste » comme les deux faces d'une même pièce. La première est basée sur la distribution d'utilité du locuteur inférée à partir du contexte, et a fait l'objet d'expérimentations sur des sujets humains par les psychologues cognitivistes ; la seconde, basée sur la reconnaissance de propriétés spécifiques de l'énoncé cible, se situe dans le prolongement des travaux de Searle, travaux utilisés et étendus dans une perspective informatique. L'idée générale que nous développons dans cet article est que la reconnaissance d'intention est parfois basée sur une représentation (chez l'auditeur) des préférences du locuteur et que cette représentation des préférences va orienter le processus interprétatif de l'auditeur ; parfois l'auditeur « décode » l'interprétation d'un énoncé en fonction de la forme linguistique de ce dernier, se basant sur les conventions langagières. Cette approche est mise en perspective avec d'autres approches récentes de la littérature en sciences cognitives.

**Mots-clés :** actes de langage indirects, psychologie cognitive, logique modale

### Abstract:

This article is the product of a multidisciplinary approach to indirect speech acts. Our approach to the recognition of indirect intentions includes a « utilitarian heuristic » and a « conventional heuristic » as two sides of the same coin. The former is based on the perceived utility distribution of the speaker in the context, and has been experimentally investigated by cognitive psychologists ; the latter, based on the recognition of some specific properties of the target statement, relates to Searle's work, extended from a computer science perspective. The central claim of this article is that hearers sometimes recognize intentions based on their representation of the preferences of the speaker, which will guide the interpretative process ; and sometimes « decode » the intention from the linguistic form of the statement, drawing on linguistic conventions. We discuss our approach within the context of recent trends in the cognitive science literature.

**Keywords:** indirect speech acts, cognitive psychology, modal logic

\*Ce travail a été soutenu financièrement par le Programme Interdisciplinaire TCAN. Une version antérieure de ce travail est en cours de publication dans la revue « Psychologie de l'Interaction ». Merci aux rapporteurs pour leurs remarques constructives et leur aide à éclaircir certains points.

## 1 Introduction

Cet article concerne les processus de reconnaissance de l'intention sous-jacente à la production d'un énoncé isolé. Nous supposons néanmoins que l'agent effectuant cette reconnaissance dispose de toutes les informations d'arrière-plan nécessaires, en particulier celles issues des échanges ayant précédé l'énoncé en cours de traitement. Nous supposons en particulier que toutes ces informations sont contenues dans la distribution d'utilité de l'agent. Nous supposons que locuteur et auditeur sont deux agents distincts : nous ne traitons pas les cas de monologue, ni les situations où le locuteur découvre un sens à ce qu'il dit au moment où il le dit (en revanche, notre approche est compatible avec le cas où l'intention reconnue par l'auditeur ne correspond pas à celle du locuteur). 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 [30]), la gestion des tours de parole, etc. Nous ne traiterons pas non plus des aspects paraverbaux comme l'utilisation de la prosodie. Afin de présenter notre problématique, nous commençons par introduire brièvement la théorie des actes de langage.

**La théorie des actes de langage.** La notion d'acte de langage (« speech act ») remonte à Austin [4]. Celui-ci dresse 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). En d'autres termes, l'illocution est la production d'acte(s) de langage particulier(s) destiné(s) à transmettre à l'auditeur l'effet perlocutoire que le locuteur a l'intention de réaliser en produisant

cet acte de langage.<sup>1</sup>

Searle [23] va montrer la faiblesse de la taxinomie d'Austin et en proposer une nouvelle (celle que nous allons utiliser) déterminée par le *but illocutoire* des énoncés. Ce but partitionne les énoncés en cinq types d'actes de langage<sup>2</sup>. 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. 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'énonciation<sup>3</sup>. Nous considérerons dans cet article les actes (à but illocutoire) directif(s).

Dans ce qui suit, nous nous focalisons sur les effets de l'acte en terme d'intention (la distinction entre effet illocutoire et perlocutoire ne nous est pas nécessaire ici), et précisément sur l'attribution au locuteur par l'auditeur de cette intention. Nous ne considérons donc pas sa réaction (physique ou linguistique) suite à cette attribution.

Nous appelons « acte principal » l'acte illocutoire 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). Par exemple, dans « Peux-tu me passer le sel ? » l'acte littéral correspond à une question dont la réponse attendue est « oui » ou « non », et si le locuteur parle littéralement, alors l'acte principal est l'acte littéral. Dans certains contextes (par exemple, à table), le locuteur peut avoir l'intention de signifier quelque chose de plus, par exemple : « Passe-moi le sel » ; dans ce dernier cas, l'acte principal, qui correspond par définition à ce que le locuteur cherche principalement à signifier, est donc distinct de l'acte littéral, et on parle alors « d'acte indirect ». Dans cet exemple, quand l'intention du locuteur est d'obtenir le sel, il demande indirectement à l'auditeur de le lui passer (en lui demandant *s'il peut* le lui passer). Le locuteur réalise donc dans ce contexte un *acte directif indirect*<sup>4</sup>.

<sup>1</sup>Il est important de spécifier ici qu'Austin a introduit (à tort) la notion d'acte de langage perlocutoire. Searle [23] a montré que c'était un non sens que de parler d'*acte de langage* concernant le perlocutoire, lui préférant la notion d'*effet* perlocutoire.

<sup>2</sup>Nous assimilons ici, comme il est fréquent, et bien que ce soit un abus de langage, actes illocutoires et actes de langage.

<sup>3</sup>Les « performatifs » initialement identifiés par Austin.

<sup>4</sup>La communication non littérale ne se borne pas à la communication indirecte. Même si nous jugeons notre modèle compatible avec des formes de communication non littérale autres que la forme indi-

**Problématique.** Le problème de l'interprétation d'énoncé en présence d'actes de langage indirects (et plus généralement, non littéraux) est d'identifier quel est l'acte principal. Ce problème est intimement lié à la détermination des éléments à prendre en compte. Et l'une des questions sous-jacentes à cette identification est le rôle joué par l'acte littéral : est-il inféré systématiquement ? ou bien seulement quand aucune interprétation non littérale n'est disponible ? ou encore parallèlement à toutes les interprétations candidates ? Cet article se propose de contribuer à la résolution de ce problème. Nous montrerons comment les deux visions classiques de l'interprétation en philosophie du langage peuvent être vues de façon complémentaire (Sect. 2) *via* deux heuristiques que nous détaillerons (Sect. 3). Nous montrerons en quoi nos travaux antérieurs permettent de formaliser ces deux heuristiques au sein d'une logique modale de la croyance, du choix, du temps et de l'action.

## 2 L'attribution d'intentions

Selon Sadock [22], 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 [4] et Searle [23], mettent l'accent sur le côté conventionnel de la relation entre actes locutoires et illocutoires. Les secondes (e.g. [13, 28]) mettent l'accent sur le rôle de l'intention suggérée par le locuteur. Rejetant l'idée que les conventions puissent traiter seules l'ensemble du problème, Grice [12, 13] a suggéré que les conventions ne sont pas autant déterminantes que les intentions que le locuteur rend manifestes à son auditoire. Par exemple, dans une salle où la fenêtre est fermée, l'énoncé « il fait chaud » peut être interprété comme une requête indirecte afin que quelqu'un ouvre la fenêtre. 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

recte (voir [18] pour plus de détails), nous nous en tiendrons ici à cette dernière.

intentionnaliste met l'accent sur l'apport informationnel des contextes linguistique et extralinguistique. De notre point de vue, l'opposition entre les approches conventionnaliste et intentionnaliste témoigne en fait d'une complémentarité.

Dans le cadre de la théorie de la pertinence, Sperber et Wilson [27, p. 208] avancent 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-ensembles disjoints<sup>5</sup>, et considérer que chacun de ces sous-ensembles est traité dans le contexte de l'autre. ».

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, ou bien celle activée par l'information nouvellement entrante (si l'on reconsidère la représentation préalable). Nous allons nommer « énoncé focal » l'énoncé entrant et réserverons le terme « contexte » à la représentation déjà activée au moment de l'énonciation.

L'analyse de Sperber et Wilson nous conduit à remarquer que 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. Ainsi, plutôt que de voir une opposition fondamentale entre les deux approches, nous suggérons qu'un même auditeur sera capable de passer d'un processus à l'autre selon les occasions, parfois même au cours du traitement d'un même énoncé.

Une fois admise la complémentarité de ces deux approches, nous sommes naturellement amenés à postuler la coexistence de deux mécanismes. Le premier va des inférences préalablement réalisées sur les intentions candidates (dites « activées ») vers les propriétés formelles de l'énoncé focal. Nous appellerons « heuristique utilitariste » ce mécanisme qui utilise un contexte contenant une gamme d'intentions plausibles activées. Le second mécanisme part des propriétés formelles de l'énoncé focal pour activer de façon privilégiée les intentions typiquement associées. Il concerne le décodage de la forme

<sup>5</sup>Ces deux sous-ensembles disjoints servent à décrire d'une part l'information déjà présente dans le système, et d'autre part l'information nouvellement entrante.

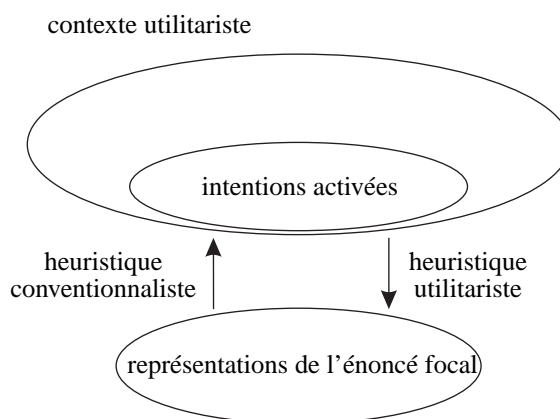


FIG. 1 – Esquisse d'une théorie de l'attribution d'intention

linguistique de l'énoncé, et nous l'appellerons : « heuristique conventionnaliste ». Dans la Figure 1, ces deux mécanismes sont replacés par rapport à leurs espaces d'entrée et de sortie.

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 [8, 16, 29]). Une autre question est celle de la gestion de la compétition entre les deux mécanismes. Ainsi, Strawson [28] 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. [2]).

### 3 Les deux heuristiques liées à notre approche utilitariste

#### 3.1 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, au 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., [17]), 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 sa 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 qu'il serait plus utile pour *S* d'obtenir une cigarette que d'obtenir l'information, et si

*H* 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 contextuels 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 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 ([21]), 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.

Le contexte utilitariste a fortement influencé les évaluations des participants, 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 même direction.

### 3.2 L'heuristique 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. Cette vision est à la base même du travail de Searle [24, 25], travail étendu par Virbel [6, 32] et raffiné et formalisé par Longin [6, 19, 18].

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 conventionnaliste 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 tient à ce qu'elles peuvent fonctionner en l'absence d'un contexte utilitariste clair. Avec un énoncé extrêmement conventionnaliste 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.*

### 3.3 Formalisation

Ces définitions ne seraient pas d'une grande utilité sans une méthode appropriée d'analyse des structures linguistiques et des contextes. Dans la lignée des travaux de Searle et de Vanderveken sur la formalisation des actes de langage [26, 31] les notions de *forme d'indirection* et de *schéma d'indirection* introduites par Longin [6, 19], peuvent s'appliquer pour modéliser les heuristiques conventionnaliste et utilitariste.<sup>6</sup>

**Formes et schémas d'indirection.** Searle [24, 25] 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

<sup>6</sup> Afin de ne pas surcharger inutilement ce qui suit, le langage logique utilisé est introduit au fur et à mesure des besoins. On peut trouver le détail de sa sémantique et de son axiomatique dans [14] par exemple.

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, ainsi qu'une troisième catégorie portant sur des conditions de planification de l'action visée dans l'énoncé [6, 32].<sup>7</sup>

Longin a lui-même systématisé l'approche de Virbel, en prenant en compte la combinatoire de certaines caractéristiques constitutives des catégories de Virbel [18]. Selon ce point de vue, Longin définit une *forme d'indirection* 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* (CMI)<sup>8</sup>, 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).

Ainsi, si  $a$  est un acte de langage constituant une forme d'indirection,  $b$  un acte de langage, et  $\varphi$  une formule du langage objet représentant un CMI, alors :

$$a, \varphi \rightsquigarrow b \quad (1)$$

est un schéma d'indirection qui signifie que l'accomplissement de  $a$  peut, dans un certain contexte qui requiert au minimum que  $\varphi$  soit vrai, produire les effets habituellement associés à une exécution de  $b$ .

Dans [19] Longin montre que les schémas d'indirection peuvent être décrits au sein d'une logique modale de la croyance, du choix, du temps, et de l'action, *via* des lois d'actions conditionnelles. Ces dernières caractérisent formellement ce qu'est une action en termes de

<sup>7</sup>Par exemple, pour pouvoir passer le sel à quelqu'un, il faut tenir le sel dans sa main. On peut donc demander à quelqu'un de nous passer le sel en lui demandant s'il a le sel, ou s'il peut l'attraper par exemple...

<sup>8</sup>Le 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.

préconditions et d'effets. Ainsi, les lois d'effets sont de la forme

$$A_i \rightarrow \text{After}_a B_i \text{ (pour } i \geq 0) \quad (2)$$

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,  $A_0$  représente le contexte dans lequel l'acte doit être interprété de façon littérale, et  $A_i$  (pour  $i > 0$ ) les différents contextes où l'acte doit être compris de façon indirecte.

Le *contexte minimal d'interprétation*  $\varphi_i$  dérive de  $A_i$  (i.e.  $A_i \rightarrow \varphi_i$  pour tout  $i > 0$ ) sans pour autant lui être équivalent.

$B_0$  est l'effet de l'acte littéral  $a$  (il représente formellement l'intention du locuteur associée à son énoncé pris au sens littéral), alors que les  $B_i$  (pour  $i > 0$ ) représentent les différents effets associés normalement aux différents actes indirects  $b$  possibles.

De cette façon, nous pouvons rendre compte du fait qu'un même énoncé peut, dans des contextes différents, être associé à des interprétations indirectes différentes. Selon le lien logique existant entre  $A_0$  et les  $A_i$  ( $i > 0$ ) on peut rendre compte du fait que l'acte littéral est accompli en même temps ou indépendamment de l'acte indirect associé. Cela rend compte en particulier de résultats en neuropsycholinguistique infirmant le *modèle pragmatique standard* selon lequel l'inférence de l'acte indirect nécessite en premier lieu celui de l'acte littéral (cf. [10, 9, 11]).

Par exemple : soient  $a$  l'acte réalisé par l'énoncé « peux-tu me passer le sel ? » et  $b$  celui réalisé par l'énoncé « passe-moi le sel ». Dans certains contextes,  $a$  peut être interprété comme  $b$ . On a alors les lois d'effet suivantes :

$$A_0 \rightarrow \text{After}_a \text{Effets}(a) \quad (3)$$

$$A_1 \rightarrow \text{After}_a \text{Effets}(b) \quad (4)$$

où  $\text{Effets}(a)$  et  $\text{Effets}(b)$  sont des abréviations représentant respectivement les effets associés à l'action  $a$  (et traduisant donc l'interprétation littérale de l'énoncé) et ceux associés à l'action  $b$ . Ainsi, dans le contexte où  $A_1$  est vrai, les effets associés à l'acte de langage  $a$  sont ceux d'une demande de passer le sel.

Chaque individu dispose de ses propres lois d'action, qui ne sont pas nécessairement les mêmes que celles des autres.

**De l’heuristique conventionnaliste.** 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.

Formellement, les schémas d’indirection correspondent exactement à l’heuristique conventionnelle : selon la forme syntaxique de l’énoncé cible, celui-ci est associé à l’acte indirect correspondant.

Dans [6, 19] Longin exhibe les schémas correspondant aux trois catégories de Virbel. Par exemple, la condition préparatoire de la demande « passe-moi le sel » est, selon la théorie des actes de langage, que l’auditeur peut passer le sel au locuteur. Selon la Catégorie 1, poser une question sur cette condition est une manière de réaliser indirectement cette demande. Ainsi : « peux-tu me passer le sel » est une forme d’indirection. Ici, le « peux-tu » est caractéristique d’une question sur la condition préparatoire. On peut donc exhiber le schéma d’indirection suivant ( $s$  représente le locuteur, et  $h$  l’auditeur,  $c$  est une action dont  $h$  est l’auteur) :

$$\langle s, \text{QueryYn}, h, \text{Feasible}_c \top \rangle \rightsquigarrow \langle s, \text{Request}, h, \text{Done}_c \top \rangle \quad (5)$$

Il s’agit d’une instance du schéma d’indirection (1) où le CMI est vide, l’action  $a$  est un acte de langage dont l’auteur est  $s$ , le destinataire  $h$ , la force illocutoire de type question fermée, et le contenu propositionnel la formule  $\text{Feasible}_c \top$  (signifiant « l’agent  $h$  peut exécuter l’action  $c$  ») ; l’action  $b$  (relative à (1)) est également un acte de langage mais de force illocutoire directive ( $c$  est une demande). Cette instance (5) signifie que si  $s$  demande à  $h$  si celui-ci peut faire l’action  $c$  alors dans certains contextes cela peut signifier que  $s$  demande à  $h$  de faire  $c$ . (Dans notre exemple,  $c$  est l’action :  $h$  passe le sel à  $s$ .)

Dans [18], Longin exhibe différents schémas dans le cadre des actes de type directif et de Catégorie 2 (pressentie comme englobant les deux

autres — les catégories 1 et 3 faisant état de raisons particulières : le fait qu’on veuille, qu’on puisse etc. faire une certaine action). La catégorie 2 initiale de Virbel est raffinée en 64 sous-catégories. Les schémas correspondant sont systématiquement donnés.

**De l’heuristique utilitariste.** Dans nos travaux antérieurs, les schémas d’indirection étaient destinés à traiter le problème de l’attribution d’intentions dans une perspective conventionnaliste. Comme nous venons de le montrer dans les sections précédentes, notre hypothèse est que nous utilisons l’heuristique utilitariste et, quand celle-ci fait défaut (du fait de l’absence de connaissance chez l’auditeur à propos de la distribution d’utilité du locuteur), que nous utilisons l’heuristique conventionnaliste.

L’idée est alors de proposer des lois d’action dont le contexte rend compte de cette connaissance sur la distribution d’utilité du locuteur. Dans cette perspective, nous étendons notre langage mulimodal par un ensemble d’opérateurs  $U_i$  tels que que  $U_i a$  signifie « dans la situation présente, l’exécution de l’action  $a$  a pour l’agent  $i$  une utilité espérée maximale » ( $i$  n’étant pas nécessairement l’auteur de  $a$  — l’agent qui l’exécute).

Dans ces conditions, si  $a$  est une forme d’indirection,  $b$  l’acte potentiellement réalisé indirectement suite à l’accomplissement de  $a$ , et  $c$  l’action que  $s$  cherche à faire faire à l’auditeur, les lois d’action peuvent être représentées de la façon suivante :

$$U_s c \rightarrow \text{After}_a \text{Effets}(b) \quad (6)$$

Autrement dit, « si l’on est dans un état où l’utilité pour  $s$  à ce que l’action  $c$  soit accomplie est maximale, alors lorsque  $s$  accomplit l’acte de langage  $a$  il a l’intention de produire l’effet perlocutoire attaché à  $b$  » (où  $b$  est supposé être une demande de faire  $c$ ).

Par exemple, l’heuristique utilitariste correspondant au schéma d’indirection (5) est ( $c$  étant l’action de  $h$  de passer le sel à  $s$ ) :

$$U_s c \rightarrow \text{After}_{\langle s, \text{QueryYn}, h, \text{Feasible}_c \top \rangle} \text{Effets}(\langle s, \text{Request}, h, \text{Done}_c \top \rangle) \quad (7)$$

Autrement dit, dans le contexte où l’utilité pour  $s$  à ce que  $h$  lui passe le sel est maximale, quand  $s$  demande à  $h$  si celui-ci peut lui passer le sel,  $s$  souhaite en fait lui demander de lui passer le sel.

Du point de vue sémantique, l'opérateur  $U$  peut être défini de manière simple et classique<sup>9</sup>. Soit  $AGT$  l'ensemble non vide des agents et  $ACT$  l'ensemble non vide des actions. Pour tout agent  $i$ , soit  $\mathcal{U}_i : W \times ACT \times W \longrightarrow N$ .  $\mathcal{U}_i(w, a, w')$  est l'utilité pour  $i$  du monde  $w'$  atteint par une exécution de l'action  $a$  (dont  $i$  n'est pas forcément l'auteur) à partir du monde  $w$ .

Soit  $R_a$  la relation d'accessibilité associée à l'action  $a$ . Alors  $w' \in R_a(w)$  ssi le monde  $w'$  peut être atteint à partir de  $w$  en exécutant l'action  $a$ . Chaque fonction  $\mathcal{U}_i$  est isomorphe à  $\bigcup_{a \in ACT} R_a$  : l'utilité  $\mathcal{U}_i(w, a, w')$  du monde  $w'$  est définie ssi  $w' \in R_a(w)$ , pour tout  $i \in AGT$ ,  $w \in W$  et  $a \in ACT$ .

De la même manière,  $\mathcal{P}(w, a, w')$  est la probabilité que le monde  $w'$  soit atteint *via* l'exécution de l'action  $a$  à partir du monde  $w$ .

La fonction  $\mathcal{P}$  est isomorphe à  $\bigcup_{a \in ACT} R_a$ . Nous imposons de plus que la somme des probabilités associées à l'ensemble des mondes pouvant être atteints par l'exécution d'une action donnée soit égale à 1. Soit :

$$\left( \sum_{w' \in R_a(w)} \mathcal{P}(w, a, w') \right) = 1$$

pour tout  $w \in W$  et  $a \in ACT$ .

Pour tout  $i \in AGT$ , soit  $\mathcal{EU}_i(a, w)$  la fonction retournant l'utilité espérée par l'agent  $i$  de l'exécution de l'action  $a$  (dont il n'est pas forcément l'auteur) à partir du monde  $w$ . Dans ces conditions :

$$\mathcal{EU}_i(a, w) = \sum_{w' \in R_a(w)} (\mathcal{U}_i(w, a, w') \times \mathcal{P}(w, a, w'))$$

Soit  $a_i$  une notation pour signifier que l'auteur de l'action  $a$  est l'agent  $i$ , avec  $i$  non nécessairement différent de  $j$ . Alors, l'utilité pour  $i$  à ce que  $j$  fasse l'action  $a$  dans le monde  $w_0$  se définit de la manière suivante :

$$\begin{aligned} w_0 \Vdash U_i a_j \text{ ssi } & R_{a_j}(w_0) \neq \emptyset \ \& \\ & \forall b_j (b_j \in ACT \ \& \ b_j \neq a_j \ \& \\ & \quad R_{b_j}(w_0) \neq \emptyset \Rightarrow \\ & \quad \mathcal{EU}_i(a_j, w_0) > \mathcal{EU}_i(b_j, w_0)). \end{aligned}$$

<sup>9</sup>On pourrait également lui donner une sémantique plus qualitative (ordinaire), en terme de rang.

## 4 Discussion

Notre approche amène à reconsidérer en profondeur 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, 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 considérées 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 [15], les auditeurs en inféreraient que l'interprétation la plus plausible serait celle potentiellement la plus menaçante pour la face sociale dans ce contexte. Ceci pourrait être 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 co-existantes.

Fodor (cité dans [33]) 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 : 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 fondée sur les travaux en philosophie d'Austin [4], Grice [12, 13], Searle [23] et Bratman [5]. 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 applications au traitement du langage de cette approche visaient la conformité au modèle pragmatique standard (cf. [1]), 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 [20]. 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. [7]. 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, il s'agit désormais d'intégrer à la logique un mécanisme d'évolution des utilités, et de définir clairement les liens que ces opérateurs entretiennent avec les autres (croyance, choix, temps, et action).

En termes de développements 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 [3] 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 [2].

## Références

- [1] J.F. ALLEN et R. PERRAULT. Analyzing intention in utterances. *Artificial Intelligence*, 15 :143–178, 1980.
- [2] J.R. ANDERSON, D. BOTHELL, M. D. BYRNE, S. DOUGLASS, C. LEBIERE, et Y. QIN. An integrated theory of the mind. *Psychological Review*, 111 :1036–1060, 2004.
- [3] J.R. ANDERSON et C. LEBIERE. *The Atomic Components of Thought*. Erlbaum, Mahwah, NJ, 1998.
- [4] J.L. AUSTIN. *How To Do Things With Words*. Oxford University Press, 1962.
- [5] M.E. BRATMAN. What is intention ? Dans Ph.R. COHEN, J. MORGAN, et M.E. POLLACK, éditeurs, *Intentions in Communication*, chapitre 2, pages 15–31. MIT Press, Cambridge, MA, 1990.
- [6] M. CHAMPAGNE, A. HERZIG, D. LONGIN, J.-L. NESPOULOUS, et J. VIRBEL. Formalisation pluridisciplinaire de l'inférence d'actes de langage non littéraires. *Information, Interaction, Intelligence*, Hors série :197–225, 2002.
- [7] E.P. FRANCIK et H.H. CLARK. How to make requests that overcome obstacles to compliance. *Journal of Memory and Language*, 24 :560–568, 1985.

- [8] M.A. GERNSBACHER et T. GIVÒN. *Coherence in spontaneous text*. John Benjamins, Amsterdam, 1995.
- [9] R.W.Jr GIBBS. Do people always process the literal meaning of indirect requests? *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 9 :524–533, 1983.
- [10] R.W.Jr GIBBS. Figurative thought and figurative language. Dans M. A. GERNSBACHER, éditeur, *Handbook of psycholinguistics*, pages 411–446. Academic Press, San Diego, CA, 1994.
- [11] S. GLUCKSBERG, P. GILDEA, et M.B. BOOKIN. On understanding non-literal speech : Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, 21 :85–98, 1982.
- [12] H.P. GRICE. Meaning. *Philosophical Review*, 66 :377–388, 1957.
- [13] H.P. GRICE. *Studies in the way of words*. Harvard University Press, Cambridge, MA, 3rd édition, 1989.
- [14] A. HERZIG et D. LONGIN. C&L intention revisited. Dans D. DUBOIS, C. WELTY, et M.-A. WILLIAMS, éditeurs, *Proc. 9th Int. Conf. on Principles on Principles of Knowledge Representation and Reasoning (KR2004)*, Whistler, Canada, pages 527–535. AAAI Press, 2–5 juin 2004.
- [15] T. HOLTGRAVES. Comprehending indirect replies : When and how are their conveyed meanings activated? *Journal of Memory and Language*, 41 :519–540, 1999.
- [16] W. KINTSCH. The role of knowledge in discourse comprehension : A construction-integration model. *Psychological Review*, 95 :163–182, 1988.
- [17] A.M. LESLIE, O. FRIEDMAN, et T.P. GERMAN. Core mechanisms in ‘theory of mind’. *Trends in Cognitive Sciences*, 8 :528–533, 2004.
- [18] D. LONGIN. Des raisons qu’ont certains actes à être indirects. *Psychologie de l’interaction*, 2005. à paraître.
- [19] D. LONGIN et É. RAUFASTE. Actes de langage indirects : co-construction d’un modèle logico-psychologique. Dans A. HERZIG, B. CHAIB-DRAA, et Ph. MATHIEU, éditeurs, *Proc. 2ndes Journées Francophones Modèles Formels de l’Interaction (MFI’03)*, pages 169–178. Cepaduès-Éditions, 2003.
- [20] W. MAO et J. GRATCH. A utility-based approach to intention recognition. Paper presented at the AAMAS 2004 Workshop on Agent Tracking : Modeling Other Agents from Observations, 2004.
- [21] É. RAUFASTE, J.F. BONNEFON, et D. LONGIN. The utilitarian heuristic : a bounded rationality framework for intention recognition in indirect directive speech acts. Manuscrit en soumission.
- [22] J. SADOCK. Speech acts. Dans L. R. HORN et G. WARD, éditeurs, *Handbook of pragmatics*, pages 53–73. Blackwell, Oxford, 2004.
- [23] J.R. SEARLE. *Speech acts : An essay in the philosophy of language*. Cambridge University Press, New York, 1969.
- [24] J.R. SEARLE. Indirect speech acts. Dans J. P. COLE et J. L. MORGAN, éditeurs, *Syntax and Semantics*, volume 3, *Speech Acts*. Academic Press, New York, 1975.
- [25] J.R. SEARLE. *Expression and Meaning. Studies on the Theory of Speech Acts*. Cambridge University Press, 1979.
- [26] J.R. SEARLE et Daniel VANDERVEKEN. *Foundation of illocutionary logic*. Cambridge University Press, 1985.
- [27] D. SPERBER et D. WILSON. *La pertinence : communication et cognition*. Proposition. Les éditions de Minuit, Paris, 1989.
- [28] P.F. STRAWSON. Intention and convention in speech acts. Dans J.R. SEARLE, éditeur, *The philosophy of language*, pages 23–38. Oxford University Press, London, 1971.
- [29] P. THAGARD. *Coherence in thought and action*. MIT Press, Cambridge, MA, 2000.
- [30] A. TROGNON et R. GHIGLIONE. *Où va la pragmatique ? De la pragmatique à la psychologie sociale*. Presses Universitaires de Grenoble, Grenoble, 1993.
- [31] D. VANDERVEKEN. *Meaning and Speech Acts*, volume 2, *Formal semantics of success and satisfaction*. Cambridge University Press, 1991.
- [32] J. VIRBEL et J.-L. NESPOULOUS. Des raisons de (ne pas) faire l’action. approches logico-pragmatiques et perspectives (neuro)psycholinguistiques. *Psychologie de l’interaction*, 2005. à paraître.
- [33] H. WIMMER et V. WEICHBOLD. Children’s theory of mind : Fodor’s heuristics examined. *Cognition*, 53 :45–57, 1994.