

Les implicatures conversationnelles : de l'approche pragmatique à l'analyse grammaticale

Anouch Bourmayan

► **To cite this version:**

Anouch Bourmayan. Les implicatures conversationnelles : de l'approche pragmatique à l'analyse grammaticale. Claire Beyssade. Les implicatures : au delà du sens littéral, 8, ISTE Editions, pp.79-118, 2020, Série Les concepts fondateurs de la philosophie du langage, sous la direction de Beatrice Godard-Wendling, 9781784057107. hal-03087575

HAL Id: hal-03087575

<https://hal.archives-ouvertes.fr/hal-03087575>

Submitted on 12 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapitre 3

Les implicatures conversationnelles : de l'approche pragmatique à l'analyse grammaticale

Anouch BOURMAYAN

3.1. Les implicatures conversationnelles chez Grice

3.1.1. *La distinction gricéenne entre signification naturelle et signification non-naturelle*

Le concept d'implicature conversationnelle a été forgé par le philosophe britannique H.P. Grice dans [GRI 61] et [GRI 75]. Les travaux de Grice s'inscrivent dans ce qu'on appelle « la philosophie du langage ordinaire », courant inspiré des *Recherches philosophiques* de Wittgenstein [WIT 53] mais qui s'est plus largement développé à Oxford avec des philosophes comme Austin ou Strawson. À l'époque, le cadre théorique dominant pour étudier le langage était celui de la philosophie analytique, qui tentait d'expliquer et de comprendre les langues naturelles à l'aide d'outils issus de la logique mathématique. Ainsi, pour nombre de philosophes analytiques, les langues naturelles n'étaient que des réalisations imparfaites d'un langage logique, truffées d'imprécisions et d'ambiguïtés. Cette idée d'une infériorité et d'une imperfection des langues naturelles face aux langages logiques constitue précisément ce que refusent les philosophes du langage ordinaire : pour eux, si les langues naturelles ne sont pas entièrement formalisables, ce n'est pas du fait de leur

imperfection mais parce qu'elles répondent à des principes au moins partiellement autres, non réductibles à la logique mathématique. L'ambition des philosophes du langage ordinaire est de mettre au jour ces lois spécifiques qui régissent la communication et les langues naturelles.

Ainsi, la théorisation du concept d'implicature conversationnelle par Grice s'inscrit dans une réflexion approfondie du philosophe sur la nature spécifique du langage et de la communication. Dans [GRI 57], Grice s'interroge sur les différents sens du verbe anglais *to mean* et en vient à distinguer deux types de signification, correspondant à deux usages sémantiques différents de *mean* : la signification naturelle et la signification non naturelle. La signification non naturelle est celle qu'ont nos énoncés lorsque l'on communique en langage ordinaire. Elle correspond au sens du verbe *mean* dans des énoncés comme (1a) ou (1b), où *mean* peut se traduire en français par la locution verbale *vouloir dire* (*A* est ici un agent humain et *x* une expression linguistique) :

(1) a. *A* veut dire (voulait dire) quelque chose par *x*.

b. *A* veut dire (voulait dire) par *x* que...¹Notons que la traduction de cet usage de *mean* en français par la locution verbale *vouloir dire* est tout à fait édifiante, puisqu'elle vient précisément souligner la dimension d'intentionnalité, centrale dans le concept de signification non naturelle de Grice. En effet, pour Grice, la signification non naturelle est avant tout l'expression d'un sujet psychologique qui conçoit l'intention de communiquer, c'est-à-dire d'induire une certaine croyance chez son interlocuteur, à l'aide d'un outil qui peut être linguistique ou autre. À l'opposé, Grice illustre la notion de signification naturelle grâce aux énoncés (2) et (3), où *mean* peut se traduire en français par la locution *être le signe de* :

(2) Ces boutons sont (étaient) le signe que [le patient] a (avait) la rougeole².

(3) Le récent budget est le signe que nous allons avoir une année difficile³.

Cette fois, la traduction française ne laisse plus apparaître la notion d'intention. Et de fait, pour Grice, la signification naturelle ne met pas en jeu l'intentionnalité d'un sujet psychologique : les boutons n'ont évidemment aucune intention communicative et ne sont pas non plus utilisés par un quelconque sujet psychologique pour exprimer un sens particulier.

1 Les énoncés originaux en anglais sont les suivants : (1) a. *A* means (meant) something by *x*.
b. *A* means (meant) by *x* that... Nous traduisons. Les traductions suivantes dans ce chapitre seront également de notre fait.

2 L'énoncé original en anglais est le suivant : (2) Those spots mean (meant) measles.

3 L'énoncé original en anglais est le suivant : (3) The recent budget means that we shall have a hard year.

Cependant, la notion d'intention n'est pas encore suffisante pour définir précisément la signification non naturelle. Selon Grice, pour qu'un locuteur communique une signification non naturelle grâce à un énoncé, il ne suffit pas qu'il ait l'intention d'induire une certaine croyance chez son auditeur par cet énoncé ; il faut encore que son intention communicative soit reconnue par l'auditeur en tant que telle, et que le locuteur ait l'intention que celle-ci le soit. De fait, la reconnaissance de l'intention communicative par l'auditeur est absolument nécessaire à la bonne transmission du message. Pour illustrer ce fait, Grice compare deux situations. Dans chacun des deux cas, quelqu'un veut communiquer à M. X le fait que Mme X s'est montrée très familière avec M. Y. Mais dans le premier scénario, la personne montre à M. X une photo de sa femme en situation compromettante, tandis que dans le second, la personne dessine Mme X dans cette situation. Selon Grice, il n'y a signification non naturelle que dans le second cas. En effet, dans le premier scénario, il n'est pas nécessaire que M. X reconnaisse l'intention communicative de la personne lui montrant la photo pour qu'il adopte la croyance que sa femme le trompe ; M. X aurait tout aussi bien pu trouver par hasard la photo et penser que sa femme le trompe. Au contraire, dans le second cas, la reconnaissance de l'intention communicative du dessinateur est absolument nécessaire pour que M. X forme la croyance que sa femme le trompe : s'il attribue au dessinateur une ambition purement esthétique et ne saisit pas que celui-ci se sert du dessin pour l'informer de l'adultère, le message ne sera pas transmis. De même, lorsque l'on cherche à communiquer une information à l'aide d'un énoncé, notre message ne peut être transmis avec succès à l'auditeur que si ce dernier reconnaît et comprend notre intention communicative. C'est sur ce fondement théorique, sur cette conception particulière de la signification dans le langage ordinaire, que Grice va ensuite élaborer le concept d'implicature conversationnelle.

3.1.2. La typologie gricéenne des implicatures

Grice évoque pour la première fois la notion d'implicature dans [GRI 61], un article qui porte en fait sur la théorie causale de la perception, mais dans lequel le philosophe établit un ensemble de distinctions qui deviendront centrales dans sa théorie de la communication linguistique. Son point de départ est l'examen de la phrase (4) :

(4) Cela me paraît rouge.

D'après Grice, lorsqu'elle est énoncée, cette phrase véhicule une inférence de forme disjonctive selon laquelle « le locuteur sait ou croit que l'objet auquel il réfère n'est pas rouge, *ou bien* quelqu'un d'autre a nié qu'il était rouge, *ou bien* le locuteur n'est pas sûr qu'il soit rouge, *ou bien* quelqu'un d'autre a exprimé un doute par

rapport au fait qu'il soit rouge, *ou bien* la situation est telle que bien que dans les faits aucun doute n'ait été exprimé et aucune dénégation n'ait été faite, une personne ou une autre pourrait être encline au doute ou à la dénégation si elle était dans la situation de s'interroger sur le fait que l'objet soit réellement rouge » [GRI 61, p. 124]. Quelle est au juste la nature de cette proposition véhiculée par l'énoncé en plus de la proposition correspondant *stricto sensu* à « ce qui est dit », à « ce qui est asserté »⁴? Grice entreprend de répondre à cette question en proposant une typologie des inférences⁵ auxquelles un énoncé peut donner lieu en dehors des inférences logiques. Outre la présupposition, catégorie déjà connue et étudiée, Grice définit ainsi trois catégories qui correspondront à ce qu'il appellera dans [GRI 75] les implicatures conventionnelles, les implicatures conversationnelles généralisées et les implicatures conversationnelles particularisées. En outre, l'essentiel des critères opératoires qui seront développés par la suite sont déjà posés ici : impact sur la valeur de vérité de l'énoncé, dépendance par rapport à un terme linguistique spécifique dans la phrase, persistance lorsque l'énoncé est reformulé (critère que Grice appelle « la détachabilité »), « annulabilité » et nature du contexte.

Ainsi, Grice observe d'emblée que l'inférence en (5b) véhiculée par (5a) affecte la valeur de vérité de (5a) : si (5b) est fausse, (5a) ne peut plus recevoir de valeur de vérité. Au contraire, les inférences en (6b), (7b) et (8b) véhiculées respectivement par (6a), (7a) et (8a) ne doivent pas nécessairement être vraies pour que les énoncés dont elles émanent reçoivent eux-mêmes une valeur de vérité. (5b) est donc une inférence d'un type particulier : il s'agit d'une présupposition.

- (5) a. Smith a arrêté de battre sa femme.
b. Smith battait sa femme.

- (6) a. Elle est pauvre mais elle est honnête.
b. Il y a un certain contraste entre pauvreté et honnêteté, ou entre sa pauvreté et son honnêteté.

- (7) a. [Je suis professeur de philosophie. J'écris une appréciation sur le carnet d'un élève. Tout ce que je dis est :]
Jones a une magnifique écriture et son anglais est correct.

4 L'utilisation des notions de « ce qui est dit » et « ce qui est asserté » reste floue dans le travail de Grice. Mais dans [GRI 75], même si Grice avoue une utilisation très intuitive de ces notions, il indiquera quand même que « ce que quelqu'un a dit est lié de près à la signification conventionnelle des mots (la phrase) qu'il a énoncés ». Dans tous les cas, « ce qui est dit », « ce qui est asserté » grâce à un énoncé définit pour Grice les conditions de vérité de l'énoncé.

5 Le terme original utilisé par Grice dans l'article est *implication*, mais nous traduisons celui-ci par *inférence* pour éviter toute ambiguïté avec la notion d'*implication logique*.

b. Jones n'est pas bon en philosophie.

- (8) a. Ma femme est soit dans la cuisine soit dans la chambre.
b. Je ne sais pas dans laquelle de ces deux pièces elle se trouve.

Une autre spécificité de l'inférence (5b) par rapport à (6b), (7b) et (8b) est qu'elle n'est pas annulable. On ne peut se rétracter en aucune manière par rapport à (5b) lorsque l'on a énoncé (5a). Au contraire, il est facile d'annuler les inférences en (7b) et (8b) ; comme le suggère Grice, il suffit par exemple d'ajouter (9) à (7a) ou (10) à (8a) :

(9) Je ne veux bien sûr pas sous-entendre qu'il n'est pas bon en philosophie.

(10) Cela dit, je ne suis pas en train de dire que je ne sais pas dans laquelle de ces deux pièces elle se trouve.

Concernant (6b), Grice a plus de difficulté à prouver qu'elle est également annulable, mais il affirme pourtant qu'un énoncé comme (11), où la deuxième partie est censée annuler l'inférence véhiculée par la première partie de l'énoncé, ne serait pas complètement inintelligible :

(11) Elle est pauvre mais honnête, quoique bien sûr je ne veuille pas sous-entendre qu'il y ait un quelconque contraste entre pauvreté et honnêteté⁶.

Pour Grice, les inférences comme (6b), (7b) et (8b), qu'il appellera dans [GRI 75] des « implicatures », se distinguent donc des présuppositions en ce qu'elles ne constituent pas une condition préalable pour que l'énoncé reçoive une valeur de vérité et sont de plus annulables.

Cependant, pour Grice, l'inférence en (6b) se distingue de celles en (7b) et (8b) en ce qu'elle est détachable : en d'autres termes, on peut trouver une reformulation de ce qui est dit en (6a) qui ne véhicule pas (6b). Par exemple, l'énoncé en (12) est équivalent à (6a) du point de vue de ce qui est dit, mais ne véhicule pas l'inférence en (6b) :

(12) Elle est pauvre et elle est honnête.

Au contraire, il semble impossible de trouver une reformulation de ce qui est dit respectivement en (7a) et (8a) qui n'induit pas en même temps les inférences en (7b) et (8b). En outre, Grice distingue l'inférence en (6b) par le fait qu'elle dépend

⁶ Notons cependant que dans [GRI 75], le caractère annulable des inférences comme (6b), qui seront alors appelées « implicatures conventionnelles », sera laissé dans l'ombre.

d'un terme linguistique spécifique, *mais*, ce qui selon lui n'est pas le cas des deux autres inférences. (6b) illustre donc une catégorie d'inférences non logiques, détachables et dépendantes d'un item linguistique donné, que Grice appellera dans [GRI 75] « implicatures conventionnelles ». De leur côté, (7b) et (8b) représentent une catégorie d'inférences que Grice nommera dans [GRI 75] « implicatures conversationnelles ». (7b) et (8b) se distinguent essentiellement par la nature du contexte dans lequel elles interviennent, comme le note Grice dès [GRI 61] : tandis qu'une implicature comme (7b) dépend d'un contexte très particulier, en l'occurrence l'évaluation d'un élève, l'implicature en (8b) apparaît dans la plupart des contextes où (8a) est prononcé. Dans [GRI 75], Grice appellera ainsi la première « implicature conversationnelle particularisée » et la seconde « implicature conversationnelle généralisée ». Mais au-delà du caractère spécifique ou non du contexte, comment les implicatures conversationnelles sont-elles identifiées par le destinataire ? Si les implicatures conventionnelles sont déclenchées par la présence d'un item linguistique, comment les implicatures conversationnelles émergent-elles quant à elles ?

3.1.3. Le principe de coopération et les maximes conversationnelles

Dans [GRI 61], tout l'appareil d'étude théorique de Grice est déjà en place quant à la nature des implicatures : seule manque encore la terminologie. Cependant, c'est seulement dans [GRI 75], article issu de deux de ses leçons données lors des « William James lectures » à Harvard en 1967, que Grice formule explicitement une théorie venant non seulement décrire mais aussi expliquer l'émergence des implicatures conversationnelles dans notre pratique du langage ordinaire. Face à l'écart constaté entre la signification des opérateurs logiques \neg , \wedge , \vee , \supset , \forall , \exists , ι et la signification des termes *non*, *et*, *ou*, *si*, *tous*, *des* et *le* que l'on considère habituellement comme leurs équivalents en langue naturelle, Grice renvoie dos à dos ceux qu'il appelle respectivement « les formalistes » et « les informalistes » : tandis que les premiers crient à la supériorité du langage logique et conspuent le flou et l'imprécision des langues naturelles, les seconds défendent la spécificité de ces dernières et leur irréductibilité aux outils formels. Or pour Grice, si les implicatures conversationnelles ne sont en rien une défaillance du langage ordinaire, comme pourraient le prétendre les formalistes, elles ne contreviennent pas non plus à l'analyse logique des énoncés et n'illustrent donc en rien l'irréductibilité des langues naturelles aux outils logiques clamée par les informalistes. La proposition correspondant à « ce qui est dit » reste analysable en termes de conditions de vérité ; les implicatures conversationnelles viennent simplement se surajouter à la première proposition, mais selon des principes qui sont propres à la communication telle qu'elle se manifeste dans la conversation humaine.

En effet, selon Grice, la conversation est « un cas particulier ou une variété de comportement intentionnel, c'est-à-dire rationnel » [GRI 75, p. 47]. Les interlocuteurs n'interviennent pas au hasard, mais selon une direction implicite commune qui peut être plus ou moins définie et peut évoluer au cours de la conversation. Ainsi, chaque participant est tenu de suivre – et présuppose que son interlocuteur suit – ce que Grice nomme « le principe de coopération » et les quatre maximes qui en découlent ; maximes concernant respectivement la qualité, la quantité, la pertinence de l'information transmise et la manière dont celle-ci est communiquée⁷.

Or d'après Grice, c'est précisément sur le principe de coopération et les maximes conversationnelles que l'auditeur s'appuie pour dériver les implicatures conversationnelles qui lui sont communiquées. De fait, face à un locuteur qui semble ne pas – ou ne pas pleinement – respecter les maximes conversationnelles dans un contexte où l'on pourrait pourtant attendre de lui qu'il soit coopératif, l'auditeur est amené à se demander si le locuteur n'est pas en train de lui communiquer un autre type de contenu que simplement « ce qui est dit » par l'énoncé, un niveau de contenu qui serait implicite⁸, et par lequel le locuteur respecterait bel et bien le principe de coopération et les maximes conversationnelles⁹. Pour s'en convaincre, l'auditeur doit cependant penser que le locuteur le croit capable d'identifier l'implicature et le croit également conscient que lui locuteur lui attribue cette capacité. Ainsi, Grice reconstitue le raisonnement de l'auditeur à propos de l'intention communicative du locuteur de la manière suivante :

« Il a dit que p ; il n'y a pas de raison de supposer qu'il ne respecte pas les maximes, ou au moins le PC [principe de coopération] ; il ne pourrait pas le faire à moins qu'il ne pense que q ; il sait (et sait que je sais qu'il sait) que je suis capable de voir que la supposition qu'il pense que q est requise ; il n'a rien fait pour me faire

7 Voir le chapitre 1, section 1.3, pour une présentation détaillée du principe de coopération et des maximes conversationnelles.

8 Ce terme traduit le verbe anglais *to implicate*, forgé par Grice pour désigner le fait de communiquer une implicature.

9 Il ne s'agit évidemment pas de nier que dans certains contextes, les locuteurs transgressent les maximes conversationnelles sans chercher à communiquer d'implicature particulière par ce biais. Le cas le plus évident est celui du mensonge, dans lequel le locuteur viole consciemment la maxime de qualité tout en espérant que son manque de coopérativité ne soit pas identifié par son interlocuteur. Mais il est à noter qu'un tel exemple n'invalide en rien le principe de coopération et les maximes conversationnelles qui lui sont attachées. De fait, si un locuteur ment, c'est bien parce qu'il sait qu'on attend de lui qu'il respecte la maxime de qualité, et qu'il peut ainsi induire de fausses croyances chez son auditoire.

cesser de penser que q ; il a l'intention que je pense, ou veut au moins me permettre de penser, que q ; et ainsi, il a implicite que q . » [GRI 75, p. 50].

Grice illustre ce processus communicatif par un exemple dans lequel A et B discutent d'un ami commun, C, qui travaille dans une banque. A demande à B comment C s'en sort dans son nouveau travail, et B lui fait la réponse suivante :

(13) Oh plutôt bien, je crois ; il apprécie ses collègues et n'est pas encore en prison.

D'après Grice, A, réalisant que B a violé la maxime de pertinence mais n'ayant pas de raison de croire qu'il refuse d'être coopératif, sera porté à penser que B, en prononçant cette phrase dans ce contexte, lui communique l'implicature que C est potentiellement malhonnête. De fait, si cette hypothèse sur l'intention communicative de B est juste, la transgression de la maxime de pertinence n'est plus qu'apparente ; A peut ainsi restaurer sa confiance en la volonté de B d'être coopératif. Mais pour dériver cette implicature, A devra également se convaincre que B le croit capable de faire cette inférence.

On retrouve donc concernant les implicatures une dimension centrale de la signification non naturelle telle que la décrivait Grice dans [GRI 57] : le message du locuteur ne peut être transmis avec succès que si l'interlocuteur reconnaît l'intention communicative du locuteur. Mais cette interprétation des intentions d'autrui intervient à plusieurs niveaux. Tout d'abord, l'auditeur sait que le locuteur veut lui communiquer p . Mais le locuteur lui-même sait que l'auditeur sait qu'il veut lui communiquer p . En outre, l'auditeur sait que le locuteur sait que l'auditeur sait qu'il veut lui communiquer p . Le locuteur sait également que l'auditeur sait qu'il sait que l'auditeur sait qu'il veut lui communiquer p . Enfin, l'auditeur sait que le locuteur sait qu'il sait que le locuteur sait qu'il sait que le locuteur veut lui communiquer p . Cet enchaînement de représentations peut être résumé de la façon suivante :

- (r1) Le locuteur veut communiquer (p) à l'auditeur – c'est-à-dire veut que l'auditeur pense (p).
- (r2) L'auditeur se représente (r1).
- (r3) Le locuteur se représente (r2).
- (r4) L'auditeur se représente (r3).
- (r5) Le locuteur se représente (r4)¹⁰.

¹⁰ Si Grice ne mentionne que ces premiers niveaux, la récursivité impliquée par le mécanisme de reconnaissance de l'intention communicative peut être envisagée comme se prolongeant à l'infini.

Ainsi, même si l'analyse de l'émergence des implicatures conversationnelles proposée par Grice est compatible avec le maintien d'un premier niveau propositionnel correspondant à « ce qui est dit », interprétable selon les outils formels de la logique, la nature même de l'analyse de Grice fait basculer l'étude des langues naturelles dans un champ tout à fait étranger aux analyses proprement logiques : celui de la théorie de l'esprit. De fait, l'analyse gricéenne présuppose la capacité des interlocuteurs à se représenter de manière potentiellement très complexe les états mentaux des autres participants, inscrivant ainsi proprement l'étude des implicatures conversationnelles dans le domaine de la psychologie cognitive.

À partir de cette reconstitution du mécanisme interprétatif des implicatures conversationnelles, Grice distingue trois cas différents. Selon lui, dans le premier cas, il n'est pas clair qu'une maxime ait été réellement violée. Ainsi, en (14), la réponse de B suggère que Smith a ou pourrait avoir une compagne à New York :

- (14) A : Smith n'a pas l'air d'avoir de copine en ce moment.
B : Il est beaucoup allé à New York récemment.

Or pour Grice, dans cet exemple, il serait excessif de dire qu'une maxime a été réellement transgressée. Cependant, pour s'assurer que la maxime de pertinence soit complètement respectée, il faut considérer non pas seulement ce qui est dit par le locuteur mais aussi ce qu'il communique implicitement ; c'est donc en identifiant l'implicature selon laquelle Smith aurait potentiellement une compagne à New York que la réponse de B apparaîtra pleinement coopérative.

Dans le second cas, une maxime conversationnelle est cette fois nettement violée, mais on peut expliquer ce fait par la nécessité de respecter une autre maxime. Grice en donne l'exemple suivant : A est en train d'établir avec B un itinéraire pour des vacances en France. Tous deux savent que A veut voir son ami C, si cela ne rallonge pas trop leur parcours :

- (15) A : Où habite C ?
B : Quelque part dans le sud de la France.

D'après Grice, il n'y a pas de raison de supposer que B ne veut pas être coopératif, mais force est de constater que sa réponse est moins informative que ne le requerrait la question de A : B transgresse ainsi la première sous-maxime de quantité selon laquelle il faut donner autant d'informations que nécessaire. Pour expliquer cela, on peut alors supposer que B a préféré être moins informatif pour ne pas donner d'information fautive ; en d'autres termes, B ne voulait pas transgresser la première sous-maxime de qualité selon laquelle il ne faut dire que ce que l'on sait

être vrai. On en infère ainsi l'implicature suivante : B ne sait pas où C habite précisément.

Enfin, dans le troisième cas répertorié par Grice, une maxime est très nettement mise à mal au niveau de ce qui est dit ; or le respect de celle-ci n'entre pas en conflit avec une autre maxime. La maxime est en fait exploitée par le locuteur pour communiquer une implicature conversationnelle spécifique, et c'est donc au niveau du contenu implicite que le principe de coopération et les maximes conversationnelles se trouvent réellement respectés. Grice illustre ce cas pour chaque maxime et sous-maxime, citant par exemple des figures de rhétorique comme l'ironie, la métaphore, la litote ou l'hyperbole pour la première sous-maxime de qualité. En effet, dans un cas où l'on sait pertinemment – et le locuteur sait que nous savons – que X vient de trahir un secret du locuteur à l'un de ses rivaux dans le domaine professionnel, si le locuteur affirme (16), il est évident qu'il ne croit pas ce qu'il dit ou affirme :

(16) X est un bon ami.

Il viole donc manifestement la première sous-maxime de qualité qui exige que l'on ne dise que ce que l'on croit vrai. Pour maintenir l'hypothèse que le locuteur est coopératif et respecte les maximes, il faut alors supposer qu'il est en fait en train de communiquer à un niveau implicite une autre proposition, relativement proche de la première ; et celle qui apparaît en l'occurrence la plus pertinente est la proposition contradictoire, selon laquelle X n'est pas du tout un bon ami.

Dans [GRI 75], Grice distingue ainsi différents mécanismes interprétatifs permettant d'identifier les implicatures conversationnelles. Mais tous ces processus se ressemblent par un point central : l'idée que toute personne prenant part à une conversation est censée respecter certains principes communicationnels subsumés sous ce que Grice appelle le principe de coopération. Les successeurs de Grice gardent cette idée, mais proposent de remanier la nature et l'importance respectives des maximes conversationnelles.

3.2. Les maximes conversationnelles dans les approches néo-gricéennes

3.2.1. Horn

S'inscrivant pleinement dans l'héritage gricéen, Horn [HOR 84] reprend à son compte l'idée que les implicatures conversationnelles sont identifiées par l'auditeur

sur la base de maximes conversationnelles orchestrant la communication. Cependant, Horn propose de remanier la nature et le nombre de ces maximes.

S'il ne touche pas à la maxime de qualité et lui accorde même un statut particulier, la présentant comme un véritable présupposé à toute communication linguistique, Horn propose de réduire les trois autres maximes gricéennes – celles de quantité, de pertinence et de manière – à deux uniques maximes reflétant l'une l'économie du locuteur et l'autre l'économie de l'auditeur. En effet, Horn reprend à Zipf [ZIP 49] l'idée que la communication voit s'affronter deux dynamiques opposées dessinant deux économies antagonistes, celle du locuteur et celle de l'auditeur. D'un côté, l'auditeur attend que le message lui soit communiqué de la manière la plus précise et la plus détaillée possible, de façon à ce que ses efforts d'interprétation s'en trouvent réduits ; de l'autre, le locuteur cherche au contraire à économiser les moyens par lesquels il transmet son message. Horn définit ainsi deux principes conversationnels opposés : le principe Q – pour *quantity* (« quantité ») – fondé sur l'économie de l'auditeur, et le principe R – pour *relevance* (« pertinence ») – reflétant l'économie du locuteur¹¹ :

(17) Principe Q

Rends ta contribution suffisante. Dis-en autant que tu peux (étant donné R).

(18) Principe R

Rends ta contribution nécessaire. N'en dis pas plus que ce que tu dois (étant donné Q).

Or d'après Horn, ces deux principes antagonistes donnent lieu à des implicatures conversationnelles de types également opposés. Le principe Q induit des implicatures définissant une limite supérieure. Par exemple, si quelqu'un affirme (19a), on est en droit d'inférer (19b), puisque selon le principe Q, si le locuteur savait que tous ses amis sont bouddhistes, cette information étant pertinente pour l'interlocuteur, le locuteur l'aurait donnée :

(19) a. Certains de mes amis sont bouddhistes.

¹¹ Selon [HOR 84], le principe Q correspond à la première maxime de quantité de Grice, tandis que le principe R traduit les autres maximes et sous-maximes conversationnelles gricéennes, à savoir la seconde sous-maxime de quantité, la maxime de pertinence et la maxime de manière. Cependant, par la suite, Horn reviendra sur cette équivalence proposée entre ces deux principes et les maximes de Grice. Ainsi, dans [HOR 89], il associera le principe Q non seulement à la première maxime de quantité de Grice mais aussi aux deux premières sous-maximes de manière, présentant à l'inverse le principe R comme uniquement équivalent à la deuxième maxime de quantité, à la maxime de relation et aux deux dernières sous-maximes de manière.

b. Tous mes amis ne sont pas bouddhistes.

De même, le principe Q peut donner lieu aux implicatures suivantes :

- (20) a. Il a mangé trois carottes.
> Il n'a pas mangé plus de trois carottes.
- b. Il est possible qu'elle gagne.
> Il n'est pas certain qu'elle gagne.
- c. Marie est patriote ou idéaliste.
> Marie est soit patriote soit idéaliste, mais pas les deux.
- d. Je suis heureux.
> Je suis heureux mais pas fou de joie.

Au contraire, le principe R induit des implicatures définissant une limite inférieure ; en d'autres termes, l'auditeur est en droit d'inférer plus que le contenu du message communiqué s'il est mutuellement manifeste que l'auditeur est capable de dériver par lui-même une telle implicature conversationnelle dans le contexte en jeu. Le meilleur exemple est selon Horn celui des actes de langage indirects : si quelqu'un vous demande si vous pouvez lui passer du sel dans un contexte où votre capacité à le faire est évidente pour tous, vous êtes en droit d'inférer sur la base du principe R que le locuteur vous demande en fait de vous exécuter.

Cependant, ces deux principes strictement opposés l'un à l'autre entrent potentiellement en conflit dans la dérivation des implicatures. Comment résoudre cette tension ? Par quel mécanisme l'auditeur sélectionne-t-il le principe satisfaisant ? D'après Horn, celui-ci est guidé dans son choix par la forme du message. Si le locuteur utilise une expression non marquée, c'est-à-dire une expression simple, nécessitant peu d'efforts de la part de l'auditeur, ce dernier est en droit d'inférer que le locuteur cherche à lui décrire une situation stéréotypique, dont il sera en mesure d'identifier les détails spécifiques ; l'auditeur dérive alors des implicatures selon le principe R. Au contraire, si le locuteur utilise une expression marquée, c'est-à-dire complexe ou prolix, quand une expression synonyme non marquée est disponible, c'est le signe qu'il décrit une situation non stéréotypique n'autorisant pas l'enrichissement du contenu vers une limite supérieure : il faut alors appliquer le principe Q.

Horn illustre ce mécanisme interprétatif par différents exemples, dont celui des actes de langage indirects. De fait, pour exprimer une requête comme (22), on utilisera de manière conventionnelle des phrases comme (21a) et (21b), mais non (21c) et (21d) :

- (21) a. Peux-tu (s'il te plaît) fermer la fenêtre ?
b. Pourrais-tu (s'il te plaît) fermer la fenêtre ?
c. Es-tu capable de fermer (? s'il te plaît) la fenêtre ?
d. As-tu la capacité de fermer (* s'il te plaît) la fenêtre ?
- (22) (S'il te plaît) ferme la fenêtre.

Comment expliquer que seul le modal *pouvoir* soit devenu une forme conventionnalisée pour exprimer indirectement une requête, et non les expressions périphrastiques équivalentes *être capable de* et *avoir la capacité de* ? Horn souligne que *pouvoir* est une forme plus fréquente, phonologiquement plus simple et plus courte que les deux autres expressions. En d'autres termes, le modal constitue une forme linguistique non-marquée, quand *être capable de* et *avoir la capacité de* sont des formes linguistiques marquées. Ainsi, lorsqu'un auditeur entend (21a) ou (21b), suivant l'heuristique explicitée ci-avant, il interprète l'énoncé comme décrivant la situation stéréotypique qui lui est associée et se considère en droit d'enrichir l'énoncé selon le principe R, dérivant l'acte de langage indirect correspondant à une requête. À l'inverse, en entendant (21c) ou (21d), l'auditeur reconnaît une forme marquée : il considère alors nécessaire de s'en tenir au principe Q et interprète l'énoncé littéralement.

Or un autre néo-Gricéen, Levinson, va reprendre l'idée de ces deux principes antagonistes Q et R, mais en érigeant au rang de véritable maxime conversationnelle l'heuristique hornienne censée guider l'auditeur dans sa dérivation d'implicatures.

3.2.2. Levinson

Levinson [LEV 87, 00] reprend explicitement à Horn ses principes Q et R, rebaptisant le second « principe I »¹². En outre, il manifeste la même attention que lui au caractère dual de la communication, prenant à chaque fois en compte le pôle du destinataire aussi bien que celui du locuteur dans sa formulation des maximes.

Le principe Q de Levinson peut être résumé ainsi :

(23) Principe Q :

Maxime du locuteur : Fais une contribution aussi informative que ce que ta connaissance du monde te permet, dans la mesure où celle-ci ne contrevient pas au principe I.

¹² La lettre I renvoie à *informativeness*, qui signifie « valeur informative ».

Corollaire du destinataire : Considère que le locuteur a fait une contribution aussi informative que ce que sa connaissance du monde lui permet.

D'après Levinson, ce principe permet au destinataire d'un message comme (24a) d'inférer l'implicature en (24b), puisque si tous les meilleurs amis du locuteur étaient linguistes, en vertu du principe Q, le locuteur l'aurait indiqué plutôt que de faire une assertion moins informative comme (24a). De même, d'après Q, le destinataire d'un message comme (25a) est en droit d'inférer (25b) car si le locuteur était certain qu'il y a de la vie sur Mars, il aurait dit que Jean le « sait » et non simplement qu'il le « croit »¹³.

- (24) a. Certains de mes meilleurs amis sont linguistes.
b. Le locuteur pense que tous ses meilleurs amis ne sont pas linguistes.
- (25) a. Jean croit qu'il y a de la vie sur Mars.
b. Le locuteur pense qu'il est possible mais pas certain qu'il y ait de la vie sur Mars.

Selon Levinson, les implicatures dérivées grâce au principe Q ont donc deux caractéristiques principales : elles sont contrastives car elles nécessitent de s'interroger sur ce que le locuteur aurait pu dire mais n'a pas dit ; elles sont négatives, puisqu'elles sont des postulats sur ce qui n'est pas le cas.

Le principe I de Levinson est quant à lui le suivant :

(26) Principe I :

Maxime du locuteur : la maxime de minimisation. Dis-en aussi peu que nécessaire ; en d'autres termes, produis l'information linguistique minimale qui soit suffisante pour atteindre tes buts communicatifs (en gardant Q en tête).

Corollaire du destinataire : la règle d'enrichissement. Amplifie le contenu informationnel de l'énoncé du locuteur en trouvant l'interprétation la plus *spécifique*, jusqu'à ce que tu juges être l'intention communicative du locuteur, à moins que le locuteur n'ait enfreint la maxime de minimisation en utilisant une expression marquée ou prolixe.

13 Notons qu'une implicature plus forte de (25a) serait (25c) :

(25c) Le locuteur ne croit pas qu'il y ait de la vie sur Mars.

Cette inférence correspond à ce que Percus [PER 06] nomme « une anti-présupposition » et Sauerland [SAU 08] « une présupposition implicite ».

I constitue pour le locuteur un principe de minimisation de la quantité d'informations données, qui se traduit chez le destinataire par un mouvement d'enrichissement du contenu de l'énoncé. Il permet selon Levinson d'expliquer les implicatures suivantes :

- (27) a. Jean a déballé le pique-nique. La bière était tiède.
b. La bière faisait partie du pique-nique.
- (28) a. J'ai vu les enfants de Wendy.
b. J'ai vu les enfants dont Wendy est la mère.

En effet, si le locuteur de (27a) ne prend pas la peine d'indiquer de quelle bière il s'agit, en vertu du principe I, le destinataire est en droit d'adopter l'hypothèse la plus évidente quant à la nature de cette bière, à savoir qu'elle faisait partie du pique-nique. De même, si le locuteur d'une expression comme (28a) ne spécifie pas le lien précis entre Wendy et les enfants, le destinataire infèrera la réponse la plus évidente, à savoir qu'il s'agit de ses propres enfants. À la différence des implicatures dérivées grâce à Q, les implicatures dérivées par le principe I sont donc des propositions positives, puisqu'elles constituent des postulats sur ce que le locuteur n'a pas pris la peine de préciser mais qui est bien le cas.

Cependant, tandis que Horn décrit le caractère marqué ou non marqué d'un énoncé uniquement comme un indice guidant l'auditeur vers le principe conversationnel adéquat, pour Levinson l'attention portée à la forme du message relève d'un principe conversationnel spécifique, distinct des précédents, le principe M pour « manière »¹⁴ :

(29) Principe M :

Maxime du locuteur : indique une situation anormale ou non stéréotypée en utilisant des expressions marquées qui contrastent avec celles que tu utiliserais pour décrire la situation normale et stéréotypée correspondante.

Corollaire du destinataire : ce qui est dit d'une manière anormale indique une situation anormale, ou les messages marqués indiquent des situations marquées.

Selon Levinson, ce principe permet par exemple d'inférer de (30a) l'implicature en (30b), ou encore de l'énoncé en (31a) l'implicature en (31b) :

- (30) a. Jean avait la capacité de résoudre le problème.

¹⁴ Le terme anglais original est *manner*.

b. Jean n'a pas résolu le problème.

(31) a. Les coins de lèvres de Sue se sont légèrement tournés vers le haut.

b. Sue a fait un rictus ou une grimace.

En effet, l'énoncé en (30a) apparaît comme une forme marquée de l'énoncé « Jean pouvait résoudre le problème ». Selon le principe M, c'est donc que le locuteur décrit une situation elle-même marquée, c'est-à-dire atypique : l'auditeur infère alors que bien que Jean ait eu la capacité de résoudre le problème, paradoxalement il n'en a rien fait. De même, la phrase en (31a) apparaît comme une version marquée de la phrase « Sue a souri ». En choisissant cette forme, le locuteur indique qu'il est en train de décrire une situation non stéréotypée ; ainsi, l'auditeur infère que Sue n'est pas véritablement en train de sourire mais que les mouvements qui composent généralement un sourire sont ici le signe d'autre chose, rictus ou grimace.

3.2.3. Sperber et Wilson

Tandis que Levinson réduit au nombre de trois principes et Horn au nombre de deux principes les maximes conversationnelles de Grice, Sperber et Wilson [SPE & WIL 89, WIL & SPE 04] proposent quant à eux de n'en garder qu'un seul, le « principe de pertinence ».

Le principe de pertinence de Sperber et Wilson semble à première vue renvoyer à la maxime de relation de Grice, dont la définition est « sois pertinent ». Cependant, comme le reconnaît Grice lui-même, le contenu précis de cette maxime reste flou du fait du vague de la notion de pertinence. Or en reprenant cette notion, Sperber et Wilson lui donnent cette fois une valeur très précise. Selon eux, la pertinence d'un input, qu'il soit linguistique ou autre, est une affaire de degré relativement à un individu. Ainsi, plus un input produit des effets cognitifs positifs importants chez un sujet donné, plus il est pertinent pour cet individu ; à l'inverse, plus l'input demande un effort de traitement important, moins il est pertinent pour l'individu. Selon Sperber et Wilson, l'ensemble de la cognition humaine est guidé par la notion de pertinence, car les individus sont constitués de manière à maximiser la pertinence des inputs qu'ils traitent, c'est-à-dire qu'ils cherchent à traiter des inputs produisant des effets positifs aussi grands que possible et requérant des efforts aussi réduits que possible : il s'agit là du principe de pertinence cognitif. Or ce principe présente un versant communicatif plus spécifique : selon Sperber et Wilson, des inputs comme les énoncés véhiculent la présomption de leur propre pertinence optimale. En d'autres termes, lorsqu'un locuteur produit un énoncé à destination d'un public, il s'engage tacitement à formuler son message de la façon la plus pertinente possible pour son auditoire, c'est-à-dire de manière à ce que le traitement de cet input

requière le moins d'efforts possible à son public, compte tenu des propres capacités et préférences du locuteur. En outre, le locuteur s'engage à ce que son message soit suffisamment pertinent pour valoir la peine d'être traité : un énoncé doit produire plus d'effets cognitifs positifs qu'il ne requiert d'efforts à son destinataire. Sur la base de ce principe communicationnel de pertinence, le destinataire interprète un énoncé – et infère notamment les implicatures conversationnelles qui lui sont communiquées par cet énoncé – en suivant un chemin de moindre effort : il teste les hypothèses interprétatives dans leur ordre d'accessibilité, c'est-à-dire de la moins coûteuse à la plus coûteuse, et s'arrête lorsque ses attentes de pertinence sont satisfaites.

Sperber et Wilson érigent donc le principe de pertinence en une super-maxime capable à elle seule de subsumer l'ensemble des maximes conversationnelles proposées par Grice et d'expliquer toutes les implicatures conversationnelles communiquées par un énoncé. En outre, ils donnent à leur principe communicationnel de pertinence une teneur résolument cognitive et l'inscrivent dans une théorie plus vaste de la cognition humaine qu'ils nomment « théorie de la pertinence ». Leur analyse s'inscrit donc dans le champ de la psychologie cognitive, confirmant une tendance qui n'était encore qu'amorcée chez Grice.

3.3. Les implicatures scalaires : nature et problèmes

3.3.1. Les implicatures scalaires, le problème dit « de la symétrie » et la notion d'échelle

Certaines implicatures conversationnelles sont dites *de quantité*, par référence à la première sous-maxime de quantité de Grice selon laquelle le locuteur doit être aussi informatif que nécessaire. Si l'on suit l'approche gricéenne, les implicatures de quantité peuvent en effet s'expliquer par l'existence d'une proposition alternative plus informative, menant l'auditeur à effectuer le raisonnement suivant : (i) le locuteur a énoncé un message, or il aurait pu être plus informatif ; (ii) s'il ne l'a pas été, en vertu de la première maxime de quantité, c'est qu'il ne croit pas que le contenu plus informatif soit vrai, soit parce qu'il ignore si celui-ci est vrai, soit parce qu'il sait qu'il ne l'est pas ; (iii) si l'on considère que le locuteur est suffisamment bien informé, on peut en conclure qu'il pense que le contenu plus informatif est faux.

Cette approche permet par exemple d'expliquer la lecture exclusive de la conjonction de coordination *ou*. Considérons (32a) : le *ou* est littéralement inclusif, c'est-à-dire que la proposition est vraie si seulement l'un de ses deux membres est vrai mais aussi si les deux le sont. Considérons maintenant (32b) : celle-ci est une

proposition logiquement plus forte que (32a). Or, en vertu de la maxime de quantité, le locuteur est tenu d'être aussi informatif que nécessaire. Ainsi, s'il croit le locuteur suffisamment bien informé, le destinataire infère que le locuteur croit (32b) faux. Il dérive l'implicature en (33a), et en conclut que le message complet communiqué par (32a) correspond à (33b), c'est-à-dire que le *ou* a une valeur exclusive : la phrase est vraie si seulement l'un des membres de la disjonction est vrai, mais pas si les deux le sont.

(32) a. A ou B.
b. A et B.

(33) a. Il est faux que (A et B).
b. A ou B mais pas (A et B).

De même, l'exemple de Horn cité en (19) est typiquement un cas d'implicature de quantité : lorsqu'un locuteur affirme que certains de ses amis sont bouddhistes, il donne un message moins informatif que s'il affirme que tous ses amis le sont, car la proposition « tous mes amis sont bouddhistes » implique asymétriquement la proposition « certains de mes amis sont bouddhistes » (elle l'implique logiquement mais l'inverse n'est pas vrai). En vertu de la première maxime de quantité de Grice, l'auditeur est donc en droit d'inférer que si le locuteur n'a pas donné cette information plus forte, c'est qu'il ne la croit pas vraie, soit parce qu'il ne sait pas si tous ses amis sont bouddhistes, soit – si l'on considère qu'il est suffisamment informé sur le sujet – parce qu'il sait que ce n'est pas vrai : l'auditeur infère ainsi l'implicature de quantité « tous les amis du locuteur ne sont pas bouddhistes ». Les exemples cités en (20) sont également tous des cas d'implicatures de quantité.

Un problème émerge cependant, qui ne nous permet pas de nous en tenir à cette analyse. Si l'on reprend l'exemple en (32a), il apparaît que (32b) n'est pas la seule proposition plus informative que (32a) : (33b) l'est tout autant. Pourquoi dès lors ne pas considérer (33b) comme une proposition concurrente, et inférer que le locuteur de (32a) fait l'implicature que (33b) est fausse ? Or, si l'on ajoute à (32a) la négation de (33b), on prédit que le contenu global de l'énoncé correspond à (32b), ce qui de toute évidence est une prédiction fautive. Il faut donc admettre que (33b) n'est pas un compétiteur naturel de (32a). En l'état, l'analyse gricéenne surgénère : c'est le problème dit « de la symétrie ».

Pour résoudre cette question, Horn [HOR 72], suivi de Gazdar [GAZ 77] et Levinson [LEV 83], définit les notions d'*échelle* et d'*alternative scalaire*. Horn appelle *échelle* une série d'items lexicaux de même catégorie grammaticale ordonnés les uns par rapport aux autres en terme de force logique, comme illustré en (34a)-(34d) :

- (34) a. <ou, et>
b. <peut, doit>
c. <1, 2, 3... n>
d. <il est possible que p , il est nécessaire que p >

Les expressions appartenant à une échelle sont appelées des *items scalaires*. Pour une phrase donnée contenant un item scalaire i , il est possible de définir ses *alternatives scalaires* : il s'agit de toutes les phrases obtenues en remplaçant i par un autre item de la même échelle. (32b) est par exemple une alternative scalaire de (32a) parce que (32b) résulte de la substitution dans (32a) du terme scalaire *ou* par *et*, terme qui appartient à la même échelle. À l'inverse, (33b) n'est pas une alternative scalaire de (32a) : la proposition ne résulte pas de la simple substitution d'un item scalaire par un autre de la même échelle. Les implicatures dérivées à partir d'une alternative scalaire sont précisément dites *implicatures scalaires*. Gazdar [GAZ 77], Horn [HOR 72] et Levinson [LEV 83] définissent l'algorithme suivant pour les calculer : pour toute proposition p , et pour toute alternative scalaire p' telle que p' implique asymétriquement p , la proposition p implicite la négation de p' . Or (32b) étant une alternative scalaire de (32a) logiquement plus forte et (33b) n'étant pas une alternative scalaire de (32a), on peut dès lors expliquer pourquoi un énoncé de la forme de (32a) communique l'implicature scalaire correspondant à la négation de (32b) mais pas l'implicature correspondant à la négation de (33b). Le problème de la symétrie paraît résolu.

D'après la description donnée ci-avant, les implicatures apparaissant en (19) et (20) sont des implicatures scalaires. Cependant, il faut noter que toutes les implicatures de quantité ne sont pas nécessairement scalaires. Pour ne prendre qu'un exemple, l'effet d'exhaustivité de la réponse de B en (35a), glosé en (35b), correspond bien à une implicature de quantité :

- (35) a. A : Qui as-tu vu ?
B : Marie.
b. Le locuteur a vu seulement Marie.

De fait, l'implicature en (35b) est dérivée à partir du raisonnement suivant : le locuteur n'a pas énoncé de proposition plus informative indiquant qu'il aurait vu Marie et une ou plusieurs autres personnes ; si l'on suppose qu'il est à la fois coopératif et conscient des personnes qu'il a vues, on peut en conclure que le locuteur considère ces propositions alternatives comme fausses, et qu'il n'a donc vu que Marie. Cependant, il apparaît ici difficile de parler d'*item scalaire* pour un terme comme *Marie*, et l'on voit mal dans quelle échelle au sens défini par Horn [HOR 72] celui-ci s'insérerait pour expliquer la dérivation de l'implicature correspondant à

(35b). Celle-ci est donc une implicature de quantité, mais non une implicature scalaire.

3.3.2. Vers une redéfinition de la notion d'échelle

La notion précise d'échelle a été cependant beaucoup discutée. Plusieurs linguistes ont affirmé que la définition donnée par Horn était insuffisante et que d'autres critères lexicaux devaient être spécifiés pour définir les termes appartenant à une même échelle. Gazdar [GAZ 77] soutient par exemple que les items scalaires d'une échelle donnée doivent présenter des présuppositions identiques. Son raisonnement se fonde sur le couple de verbes <regretter, savoir>. Ces deux verbes satisfont *a priori* aux critères de l'échelle hornienne, puisqu'ils sont de même catégorie grammaticale et que l'un est logiquement plus fort que l'autre. De fait, une proposition simple avec *regretter* implique asymétriquement la même proposition avec *savoir*. (36a) implique par exemple (36b), tandis que l'inverse n'est pas vrai :

- (36) a. Ils regrettent qu'il y ait eu un accident.
b. Ils savent qu'il y a eu un accident.

Or dans la majorité des cas, *savoir* ne semble pas déclencher d'implicature scalaire liée à *regretter*, contrairement aux prédictions fondées sur les échelles horniennes. Ainsi, si l'on considère le dialogue en (37a), (37b) n'est pas une implicature de la réponse de B :

- (37) a. A : Dois-je les appeler pour leur raconter l'accident ?
B : Ce n'est pas la peine. Ils savent déjà qu'il y a eu un accident.
b. Ils ne regrettent pas qu'il y ait eu un accident.

D'après Gazdar [GAZ 77], ce résultat est dû au fait que *regretter* et *savoir* n'ont pas les mêmes présuppositions. De fait, une phrase comme (36b) déclenche la présupposition en (38a), tandis que (36a) déclenche non seulement la présupposition en (38a) mais aussi celle en (38b) :

- (38) a. Il y a eu un accident.
b. Ils savent qu'il y a eu un accident.

Gazdar [GAZ 77] propose donc d'ajouter comme contrainte supplémentaire sur les échelles horniennes le fait que les items scalaires déclenchent les mêmes présuppositions. Atlas et Levinson [ATL 81] affirment en outre que les items scalaires d'une même échelle doivent appartenir au même champ sémantique, présenter la même longueur et être lexicalisés au même degré.

Mais plutôt que d'ajouter ainsi des contraintes lexicales sur les items scalaires, Matsumoto [MAT 95] propose d'adopter une autre approche, selon laquelle les échelles ne seraient pas des entités figées prédéfinies lexicalement mais des éléments en construction qui varient selon les différents contextes. Ainsi, Matsumoto propose d'expliquer le fait que (37a) ne déclenche pas l'implicature en (37b) en termes purement gricéens. Il note tout d'abord que d'autres couples de verbes présentant des présuppositions différentes peuvent déclencher des implicatures scalaires. De fait, si l'on considère les verbes *savoir* et *penser*, seul le premier déclenche la présupposition selon laquelle la proposition qu'il introduit est vraie. Or comme l'observe Matsumoto, ces deux verbes peuvent former une échelle donnant lieu à une implicature scalaire. Par exemple, la réponse de B en (39a) va déclencher soit l'implicature en (39c) soit celle en (39d), selon le degré de compétence que l'on assigne au locuteur sur l'état de santé de Jean :

- (39) a. A : Jean est-il vraiment malade ?
 B : Il pense qu'il l'est.
 b. Il sait qu'il l'est.
 c. B ne sait pas si Jean est vraiment malade.
 d. B croit que Jean n'est pas vraiment malade.

Or les implicatures (39c) et (39d) sont dérivées toutes deux à partir de l'alternative scalaire (39b), où *penser* est remplacé par *savoir*. Dans un contexte où le locuteur est considéré comme peu informé sur l'état de santé de Jean, l'auditeur s'expliquera que B ne prononce pas la proposition concurrente (39b) en faisant l'hypothèse que B ne sait tout simplement pas si la présupposition de *savoir* est vraie ; il dérivera alors l'implicature en (39c). En revanche, dans un contexte où l'auditeur juge le locuteur très informé sur l'état de santé de Jean, il ira plus loin et fera l'hypothèse que B ne prononce pas (39b) parce qu'il pense la présupposition de *savoir* fautive ; il dérivera alors l'implicature en (39d).

Contrairement à ce qu'affirme Gazdar [GAZ 77], l'absence d'implicature scalaire en (37) ne s'expliquerait donc pas par les différentes propriétés présuppositionnelles des deux verbes. Selon Matsumoto, si (39a) déclenche une implicature scalaire et non (37a), c'est parce que parmi les verbes *savoir* et *regretter*, seul *regretter* exprime l'état émotionnel du sujet. Les deux verbes sont donc utilisés pour résoudre des questions très différentes : *savoir* répond à la question du degré de confiance du sujet dans une proposition, tandis que *regretter* décrit l'attitude émotionnelle du sujet. Or dans un contexte comme (37a), seul le premier point est en jeu. On peut donc expliquer l'absence de l'implicature (37b) par la première maxime de quantité de Grice : la question de l'attitude émotionnelle du sujet n'est pas en jeu, donc le fait que le locuteur ne donne pas d'information sur ce

point n'est pas interprété comme le signe qu'il veut nier l'implication émotionnelle du sujet. Au contraire, les verbes *penser* et *savoir* concernent tous deux la question de la vérité de la proposition en jeu et du degré de confiance du sujet ; ainsi, le fait que le locuteur utilise le verbe *croire* et non *savoir* est interprété comme le signe que le locuteur est soit agnostique sur la vérité de la proposition en jeu, soit persuadé de sa fausseté.

De la même manière, Matsumoto [MAT 95] conteste la nécessité de postuler les contraintes lexicales proposées par Atlas et Levinson [ATL 81]. D'après lui, il est notamment inutile et trop contraignant de spécifier une condition selon laquelle les termes d'une même échelle appartiennent nécessairement au même champ sémantique : inutile, parce que les faits que cette condition cherche à capturer sont déjà subsumés sous la maxime de relation de Grice ; trop contraignant, parce qu'elle prédit – à tort – que la réponse de B en (40a) ne peut déclencher une implicature comme (40b), dans la mesure où *taper* et *poster* n'appartiennent pas au même champ lexical (Matsumoto emprunte cet exemple à Hirschberg [HIR 85]) :

- (40) a. A : Où en es-tu avec ce courrier ?
 B : Je l'ai dactylographié.
 b. B ne l'a pas encore posté.

De fait, Hirschberg [HIR 85] montre que les échelles ne sont pas nécessairement lexicales, mais peuvent se composer potentiellement de n'importe quelle série d'items constituant un ensemble ordonné au sein duquel on peut déterminer pour un item particulier s'il est plus ou moins élevé qu'un autre sur l'échelle. Ainsi l'exemple (40) met-il en jeu une échelle d'ordre chronologique correspondant aux différentes étapes d'une situation stéréotypique sur laquelle l'étape de poster le courrier apparaît après celle de le taper. Matsumoto rejoint précisément Hirschberg sur le fait que les échelles responsables des implicatures scalaires sont avant tout contextuelles, et ne consistent pas nécessairement en des échelles lexicales mettant en jeu une relation d'implication logique¹⁵.

3.3.3. Les contextes monotones décroissants

Qu'elles soient de nature lexicale ou contextuelle, les échelles présentent cependant un avantage important, celui de rendre compte de l'inversion de la valeur des implicatures scalaires dans les contextes dits « monotones décroissants ».

¹⁵ Cependant, alors que Hirschberg [HIR 85] redéfinit la notion d'échelle en termes de *saillance contextuelle*, Matsumoto [MAT 95] maintient une analyse strictement gricéenne, selon laquelle les implicatures scalaires sont le fruit d'un raisonnement du destinataire mené à partir des maximes conversationnelles de Grice.

En effet, Horn [HOR 89] et Fauconnier [FAU 75] ont mis en évidence l'existence de contextes qui renversent le schéma habituel des implications logiques. Dans les cas standard, l'inférence logique va par exemple de *chien* vers *animal*. Ainsi, (41a) entraîne asymétriquement (41b).

- (41) a. Jean a un chien.
b. Jean a un animal.

Or dans certains contextes comme sous la négation (voir (42a)), le restricteur et la portée nucléaire des quantificateurs *aucun* ou *peu* (voir (43a) et (44a)), le restricteur du quantificateur universel *tous les* (voir (45a)), ou encore dans l'antécédent des phrases conditionnelles (voir (46a)), ces schémas logiques se trouvent inversés : l'inférence logique ne va plus de *chien* vers *animal* mais de *animal* vers *chien*. Ainsi, de (42) à (46), ce n'est pas la phrase a qui implique la phrase b mais l'inverse.

- (42) a. Jean n'a pas de chien.
b. Jean n'a pas d'animal.
- (43) a. Aucun chien ne vit dans cette maison.
b. Aucun animal ne vit dans cette maison.
- (44) a. Peu de chiens vivent dans ce domaine.
b. Peu d'animaux vivent dans ce domaine.
- (45) a. Tous les chiens sont sortis.
b. Tous les animaux sont sortis.
- (46) a. S'il avait un chien, il serait heureux.
b. S'il avait un animal, il serait heureux.

Ces contextes inversant les relations logiques sont appelés des contextes « monotones décroissants ».

Or on remarque que dans ce type de contextes, le schéma des implicatures scalaires se trouve également inversé. Considérons par exemple l'échelle partielle <beaucoup de, tous les> : dans les contextes les plus simples, une phrase contenant *tous les* entraîne asymétriquement son alternative scalaire avec *beaucoup de*. Suivant l'algorithme néo-gricéen standard, une phrase contenant *beaucoup de* déclenche ainsi une implicature scalaire correspondant à la négation de l'alternative scalaire plus forte avec *tous les*, et *beaucoup de* est finalement interprété comme *beaucoup mais pas tous*. Par exemple, (47b) est une alternative scalaire de (47a), et (47b) est logiquement plus forte que (47a) puisqu'elle l'implique asymétriquement :

(47a) déclenche donc l'implicature scalaire en (47c) correspondant à la négation de (47b), si bien que (47a) est *in fine* interprété comme en (47d).

- (47) a. Jeanne a lu beaucoup de livres de Balzac.
- b. Jeanne a tous les livres de Balzac.
- c. Jeanne n'a pas lu tous les livres de Balzac.
- d. Jeanne a lu beaucoup de livres de Balzac, mais pas tous.

Or dans un contexte monotone décroissant, les relations logiques s'inversent : c'est la phrase avec *beaucoup de* qui entraîne logiquement la phrase avec *tous les*. Ainsi, l'algorithme néo-gricéen prédit qu'en contexte monotone décroissant, une phrase avec *beaucoup* ne déclenche pas l'implicature *pas tous* ; au contraire, c'est la phrase avec *tous les* qui doit déclencher l'implicature *pas beaucoup*. Est-ce juste ? Considérons l'exemple en (48). Comme *il est impossible que* crée un contexte monotone décroissant, la phrase en (48b) avec *beaucoup* entraîne asymétriquement l'alternative scalaire en (48a) avec *tous les*. Suivant l'algorithme néo-gricéen, (48a) devrait donc déclencher l'implicature en (48c) correspondant à la négation de (48b). Inversement, la phrase en (48b) ne devrait pas déclencher d'implicature correspondant à la négation de (48a), à savoir (48d), puisque (48a) n'est pas une alternative scalaire plus forte.

- (48) a. Il est impossible que Jeanne ait lu tous les livres de Balzac.
- b. Il est impossible que Jeanne ait lu beaucoup de livres de Balzac.
- c. Il est possible que Jeanne ait lu beaucoup de livres de Balzac.
- d. Il est possible que Jeanne ait lu tous les livres de Balzac.

Ces prédictions sont vérifiées : (48a) déclenche bien l'implicature en (48c), et (48b) ne déclenche pas l'implicature en (48d). La théorie néo-gricéenne des échelles rend compte adéquatement de la spécificité des implicatures scalaires produites en contexte monotone décroissant.

Mais qu'en est-il plus généralement des implicatures déclenchées par un item scalaire se trouvant dans la portée d'un autre opérateur logique, voire d'un autre terme scalaire ? L'analyse néo-gricéenne en termes d'échelle peut-elle en rendre compte systématiquement ?

3.3.4. La théorie localiste

Chierchia [CHI 04] observe que certaines implicatures scalaires semblent apparaître localement. En effet, le terme scalaire qui les déclenche se trouve dans la portée d'un autre opérateur logique, or l'implicature normalement attachée à la proposition enchâssée semble être préservée dans la portée de cet opérateur. Par

exemple, le quantificateur *certain*s présent sous la portée du verbe d'attitude *penser* en (49a) ou sous la portée du connecteur *soit... soit...* en (50a) peut être interprété dans chaque cas comme signifiant « certains mais pas tous ». Ainsi, les deux énoncés suscitent respectivement les implicatures scalaires en (49b) et (50b) :

- (49) a. Jean pense que certains de ses étudiants l'attendent.
b. Jean pense qu'il est faux que tous ses étudiants l'attendent.
- (50) a. Marie est soit en train de travailler à son article soit en train de voir certains de ses étudiants.
b. Marie n'est pas en train de voir tous ses étudiants.

Tout se passe comme si l'implicature était déclenchée localement au niveau de la proposition enchâssée, à savoir *certain*s étudiants l'attendent en (49) et *Marie est en train de voir certains de ses étudiants* en (50), indépendamment de l'opérateur logique dans la portée duquel elle se trouve. On parle alors d'implicature « enchâssée ».

Or une approche dite « globaliste » comme la théorie néo-gricéenne des alternatives scalaires n'apparaît *a priori* pas capable de rendre compte de ce type d'implicature. En effet, dans la théorie néo-gricéenne, le raisonnement de l'auditeur est le suivant : le locuteur n'a pas employé telle alternative scalaire *S'* qui aurait été plus informative, or en vertu de la maxime de quantité, il doit chercher à être aussi informatif que nécessaire ; c'est donc qu'il ne croit pas que *S'* soit vraie, voire – s'il est bien informé – qu'il croit *S'* fautive. Le raisonnement est effectué à propos d'un acte de langage complet, et les alternatives scalaires sont calculées sur la base de l'ensemble de l'énoncé. Examinons maintenant ce que cette analyse prédit pour les énoncés (49a) et (50a). En (49a), reproduit ci-après en (51a), si l'on remplace le terme scalaire *certain*s par *tous*, terme logiquement plus fort sur l'échelle pertinente, on obtient l'alternative scalaire en (51b) qui, une fois niée, produit l'implicature scalaire en (51c) :

- (51) a. Jean pense que certains étudiants l'attendent.
b. Jean pense que tous ses étudiants l'attendent.
c. Il est faux que Jean pense que tous ses étudiants l'attendent.

Clairement, cette prédiction est plus faible que le résultat réellement obtenu, indiqué en (49b). De même, en (50a), reproduit en (52a), l'alternative scalaire obtenue par le remplacement de *certain*s par *tous* est (52b). Une fois niée, cette alternative produit l'implicature scalaire en (52c), or celle-ci implique logiquement (52d), une prédiction manifestement insatisfaisante :

- (52) a. Marie est soit en train de travailler à son article soit en train de voir certains de ses étudiants.
b. Marie est soit en train de travailler à son article soit en train de voir tous ses étudiants.
c. Il est faux que Marie est soit en train de travailler à son article soit en train de voir tous ses étudiants.
d. Il est faux que Marie est en train de travailler à son article.

Une approche « globaliste » telle que l'analyse néo-gricéenne standard produit donc des résultats erronés concernant ce type d'implicatures.

Cette observation mène Chierchia [CHI 04] à proposer une théorie alternative selon laquelle les implicatures scalaires ne sont pas dérivées par un raisonnement de l'auditeur sur l'acte de langage effectué par le locuteur, mais calculées par un mécanisme compositionnel d'interprétation similaire et parallèle à celui par lequel nous calculons le sens littéral des expressions. Selon lui, chaque expression est associée à sa valeur sémantique littérale, mais également à une valeur pragmatique renforcée calculée à partir des alternatives scalaires de l'expression. Par exemple, *certain* reçoit le sens de « certains et possiblement tous » au sein de la computation prenant en compte le sens littéral des expressions. Cependant, dans la seconde computation, qui s'effectue en parallèle et calcule la valeur scalaire des expressions, *certain* est associé au sens de « certains mais pas tous », qui intègre cette fois l'implicature scalaire établie à partir de l'alternative scalaire <tous>. À chaque étape de ce dernier processus de composition sémantique, le renforcement scalaire se maintient et se propage, à moins que la sélection du sens renforcé de l'expression ne crée une incohérence dans le contexte en jeu, auquel cas l'implicature est défaite et l'on revient au sens littéral de l'expression. Certes, ceci semble créer une situation d'ambiguïté, mais Chierchia précise que dans une situation de communication réelle, les expressions sont interprétées par défaut avec leur valeur renforcée et non leur valeur littérale¹⁶. Ainsi, en (51a), *certain étudiants* est interprété comme « certains étudiants, mais pas tous », puis *certain étudiants l'attendent* comme « certains étudiants l'attendent, mais pas tous », et *in fine*, l'ensemble de la proposition reçoit le sens de « Jean pense que certains étudiants l'attendent, mais pas tous ».

16 Il est intéressant de noter que pour Chierchia, cette préférence pour la valeur renforcée des expressions résulte du fait que l'on attend du locuteur qu'il utilise les valeurs fournies par le composant grammatical « de manière coopérative » [CHI 04, p. 64]. Ainsi, même s'il rejette le caractère globaliste de l'approche gricéenne et réinterprète le calcul des implicatures comme un mécanisme relevant du système computationnel grammatical, Chierchia ne semble pas rejeter fondamentalement la validité de la maxime de quantité de Grice. En outre, Chierchia n'écarte l'analyse globaliste gricéenne qu'en ce qui concerne les implicatures scalaires, et non pour tout type d'implicature.

L'implicature scalaire est déclenchée localement et n'est pas dépendante de l'opérateur linguistique, en l'occurrence le verbe d'attitude propositionnelle *penser*, dans la portée duquel elle se trouve¹⁷.

En outre, Chierchia entend également rendre compte des cas où les implicatures scalaires semblent absentes, comme cela se produit par exemple dans les contextes monotones décroissants. Comme nous l'avons vu, l'implicature scalaire est calculée d'après lui localement selon un mécanisme linguistique compositionnel qui s'applique par défaut. Mais lorsque l'expression associée à une implicature scalaire est composée avec une expression qui crée un contexte monotone décroissant, en vertu de l'inversion logique suscitée par ce type de contexte, le sens renforcé correspondant à l'implicature scalaire devient moins informatif que le sens littéral de l'expression ; l'implicature scalaire est alors automatiquement annulée, et le sens « renforcé » devient équivalent au sens littéral. Ainsi, dans une phrase comme (53a), le sens renforcé « Jean a vu certains et pas tous ses étudiants » associé à la subordonnée devient moins informatif que son sens littéral « Jean a vu au moins certains de ses étudiants » lorsque l'expression est composée avec le verbe *douter*, puisque celui-ci crée un contexte monotone décroissant. L'implicature scalaire calculée au préalable est alors annulée et le sens renforcé de la subordonnée devient similaire à son sens littéral ; le sens renforcé de (53a) correspond lui-même à (53b), conformément à nos intuitions selon lesquelles l'implicature scalaire n'est pas présente :

- (53) a. Je doute que Jean ait vu certains de ses étudiants.
- b. Je doute que Jean ait vu certains et possiblement tous ses étudiants.

Si elle a pour but de résoudre une question technique, celle de l'émergence d'implicatures scalaires enchâssées, la théorie localiste proposée par Chierchia se trouve donc opérer une remise en cause fondamentale de la conception néo-gricéenne traditionnelle des implicatures scalaires. Dans cette perspective, l'identification de ces dernières n'implique plus une réflexion pragmatique sur les intentions communicatives du locuteur mais résulte d'un mécanisme qui fait partie intégrante du système computationnel de la grammaire. Les implicatures scalaires ne caractérisent plus des actes de langage mais des expressions linguistiques.

¹⁷ L'implémentation formelle précise et détaillée de ces idées est exposée dans l'annexe de [CHI 04].

3.3.5. L'interprétation néo-gricéenne des implicatures enchâssées et le pas épistémique

Sauerland [SAU 04] affirme qu'il n'est pas nécessaire d'effectuer une rupture si radicale avec le système gricéen pour rendre compte des implicatures scalaires enchâssées. Celui-ci se concentre tout particulièrement sur les cas où un terme scalaire apparaît dans la portée d'un autre item scalaire, comme en (54), où *une partie de* se trouve dans la portée de *ou* :

(54) Kai a mangé les brocolis ou une partie des petits pois hier soir.

Pour calculer les implicatures scalaires des phrases contenant plusieurs termes scalaires dans un cadre d'analyse gricéen, Sauerland fait trois postulats. Tout d'abord, il définit les alternatives scalaires de ce type de phrases tel qu'en (55), où $\text{ScalAlt}(\varphi)$ est l'ensemble de toutes les alternatives scalaires d'une phrase φ , et $\varphi(X, Y)$ une phrase contenant deux termes scalaires X et Y tels que X est un élément de l'échelle quantitative Q_X et Y un élément de l'échelle quantitative Q_Y :

(55) $\text{ScalAlt}(\varphi(X, Y)) = \{\varphi(X', Y') \mid X' \text{ un élément de } Q_X, Y' \text{ un élément de } Q_Y\}$

Selon cette définition, les alternatives scalaires d'une phrase comme (54) contenant les items scalaires *ou* et *une partie de* correspondent respectivement à (56a), (56b) et (56c) :

- (56) a. Kai a mangé les brocolis ou tous les petits pois hier soir.
- b. Kai a mangé les brocolis et une partie des petits pois hier soir.
- c. Kai a mangé les brocolis et tous les petits pois hier soir.

Cependant, cela n'apparaît pas suffisant pour rendre compte des implicatures scalaires des phrases comprenant un item scalaire dans la portée d'une disjonction. Reprenons l'exemple (54). D'après l'algorithme néo-gricéen, puisque les alternatives scalaires (56a)-(56c) sont logiquement plus fortes que (54), leur négation correspond pour chacune à une implicature scalaire communiquée par (54). Les implicatures correspondant respectivement à la négation de (56b) et à celle de (56c) ne posent pas de problème. En revanche, l'alternative en (56a) est censée donner lieu à l'implicature scalaire (57a). Or (57a) est logiquement équivalente à (57b) :

- (57) a. Il est faux que Kai a mangé les brocolis ou tous les petits pois hier soir.
- b. Il est faux que Kai a mangé les brocolis hier soir et il est faux qu'il a mangé tous les petits pois hier soir.

En d'autres termes, cette analyse prédit que (54) communique l'implicature que Kai n'a pas mangé les brocolis hier soir. Or ce résultat est empiriquement faux. En outre, (54) semble communiquer une autre implicature, donnée en (58), que cette analyse ne prédit absolument pas :

(58) Il est faux que Kai a mangé tous les petits pois hier soir.

Pour résoudre cette question tout en maintenant le cadre d'analyse néo-gricéen, Sauerland propose deux amendements supplémentaires à la théorie néo-gricéenne. D'une part, il pose que les alternatives scalaires d'une phrase disjonctive de la forme *A ou B* sont les suivantes : $\langle A, B, A \text{ et } B \rangle$, ajoutant à la liste d'alternatives scalaires des phrases disjonctives le premier et le second membre de la disjonction. D'après cette analyse, les alternatives scalaires de (54) sont donc non seulement (56a), (56b) et (56c) mais aussi (59a), le premier membre de la disjonction, (59b), le second membre de la disjonction, et (59c), alternative scalaire du second membre de la disjonction (le premier n'en a pas) :

- (59) a. Kai a mangé les brocolis hier soir.
b. Kai a mangé une partie des petits pois hier soir.
c. Kai a mangé tous les petits pois hier soir.

D'autre part, Sauerland met en avant le fait que les implicatures ont différents statuts épistémiques. Il est souvent affirmé par commodité qu'une phrase *S* présentant une alternative scalaire *S'* déclenche une implicature scalaire selon laquelle le locuteur de *S* croit que *S'* est fausse. Cependant, il s'agit là d'un raccourci, du moins dans une approche qui se veut gricéenne, car si l'on suit la maxime de quantité, on peut seulement conclure de *S* que le locuteur ne croit pas que *S'* soit vraie. En termes logiques, si *S'* exprime la proposition *p'*, *S* implicite (60a), où *K* est l'opérateur qui exprime la certitude. Cependant, il se peut que le locuteur n'ait pas d'opinion sur la valeur de vérité de *p'*. Pour inférer de *S* l'implicature – plus forte – en (60b) selon laquelle le locuteur est certain de la fausseté de *p'*, il faut faire l'hypothèse additionnelle que le locuteur est bien informé sur la valeur de vérité de *p'*.

- (60) a. $\neg Kp'$
b. $K\neg p'$

Sauerland défend ainsi l'idée que les implicatures scalaires sont calculées en deux étapes. En premier lieu sont calculées les implicatures du type de (60a), signifiant « le locuteur ne croit pas que *p'* soit vraie ». Sauerland appelle ce type d'implicatures « implicatures primaires ». Puis en second lieu seulement sont calculées les implicatures scalaires du type de (60b), signifiant « le locuteur croit

que p' est fausse ». Sauerland nomme ces dernières « implicatures secondaires ». Pour passer des implicatures primaires aux implicatures secondaires, il faut accomplir ce qu'il appelle « le pas épistémique ». En outre, Sauerland postule que toute implicature secondaire entrant en contradiction avec l'assertion littérale ou l'une des implicatures primaires est annulée¹⁸.

Selon cette analyse, les implicatures scalaires primaires de (54), calculées à partir des alternatives scalaires définies en (56) et (59), sont les propositions suivantes :

- (61) a. $\neg K$ (Kai a mangé les brocolis hier soir.)
- b. $\neg K$ (Kai a mangé une partie des petits pois hier soir.)
- c. $\neg K$ (Kai a mangé les brocolis et une partie des petits pois hier soir.)
- d. $\neg K$ (Kai a mangé tous les petits pois hier soir.)
- e. $\neg K$ (Kai a mangé les brocolis ou tous les petits pois hier soir.)
- f. $\neg K$ (Kai a mangé les brocolis et tous les petits pois hier soir.)

Considérons maintenant les implicatures secondaires correspondantes en (62) :

- (62) a. $K \neg$ (Kai a mangé les brocolis hier soir.)
- b. $K \neg$ (Kai a mangé une partie des petits pois hier soir.)
- c. $K \neg$ (Kai a mangé les brocolis et une partie des petits pois hier soir.)
- d. $K \neg$ (Kai a mangé tous les petits pois hier soir.)
- e. $K \neg$ (Kai a mangé les brocolis ou tous les petits pois hier soir.)
- f. $K \neg$ (Kai a mangé les brocolis et tous les petits pois hier soir.)

(62a) et (62b) sont toutes deux bloquées car elles contredisent l'assertion initiale enrichie des implicatures primaires (61a) et (61b). En effet, si le locuteur croit à la vérité de la disjonction en (54), c'est qu'il croit qu'au moins l'un de ses membres est vrai. Or s'il adopte la croyance en (62a) que le premier membre de la disjonction est faux, c'est qu'il croit que le second est vrai, ce qui contredit l'implicature primaire en (61b). Inversement, s'il croit comme en (62b) que le second membre est faux, en vertu de l'assertion initiale, il croit que le premier membre est vrai, ce qui contredit cette fois l'implicature primaire en (61a). L'implicature secondaire en (62c) est quant à elle satisfaisante : elle ne contredit ni l'assertion initiale ni les implicatures primaires et prédit le caractère exclusif de la disjonction. De même, (62d) est un résultat satisfaisant, car il s'agit d'une implicature empiriquement juste que ne prédisait pas l'analyse néo-gricéenne standard. (62e) doit être bloquée, car elle revient à dire que le locuteur pense qu'il est faux que Kai ait mangé les brocolis et

18 Si Sauerland [SAU 04] ne mentionne le pas épistémique que pour les implicatures scalaires, celui-ci apparaît pertinent pour toutes les implicatures de quantité, c'est-à-dire toutes les implicatures dérivées à partir d'une alternative plus informative.

qu'il est également faux que Kai ait mangé tous les petits pois ; or l'affirmation que Kai n'a pas mangé les brocolis est insatisfaisante, car elle contredit l'implicature primaire en (61a). Enfin, (62f) ne pose pas de difficulté mais n'apporte pas d'information nouvelle, car elle est déjà impliquée par (62c). Ainsi, seules les implicatures secondaires (62c) et (62d) subsistent, ce qui apparaît comme un résultat empiriquement juste.

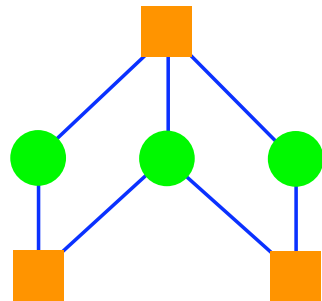
L'analyse de Sauerland [SAU 04], fondée sur une modification de l'échelle de la disjonction et sur la mise en évidence du pas épistémique, fait donc des prédictions satisfaisantes sur les implicatures scalaires des phrases contenant un item scalaire sous une disjonction, et propose en cela une véritable alternative gricéenne à l'analyse grammaticale de Chierchia [CHI 04].

Cependant, si la nature des implicatures enchâssées est sujette à controverse, leur existence même fait débat. Différents travaux expérimentaux entendent ainsi éclairer la controverse théorique entre approche grammaticale et approche pragmatique en précisant la réalité des données sur un plan empirique.

3.3.6. Les implicatures enchâssées : quelle réalité empirique ?

Geurts et Pouscoulous [GEU 09] entendent invalider le postulat général de l'approche grammaticale selon lequel les implicatures scalaires émergent librement et systématiquement dans les positions enchâssées. Ceux-ci proposent ainsi une expérience dans laquelle ils présentent aux participants la phrase (63a) où le terme scalaire *certain*s se trouve sous la portée d'un quantificateur universel, c'est-à-dire dans un contexte monotone croissant. Dans son sens littéral, (63a) est équivalente à (63b). Mais d'après les théories localistes comme celle de Chierchia [CHI 04], *certain*s devrait être interprété en (63a) dans le sens de « certains mais pas tous », comme explicité en (63c). Au contraire, une approche globaliste comme la théorie néo-gricéenne ne peut *a priori* prédire la lecture en (63c), mais seulement l'interprétation plus faible en (63d), obtenue par négation de l'alternative scalaire avec *tous*.

- (63) a. Tous les carrés sont reliés à certains des cercles.
- b. Tous les carrés sont reliés à au moins certains cercles.
- c. Tous les carrés sont reliés à certains des cercles et aucun carré n'est relié à tous les cercles.
- d. Tous les carrés sont reliés à certains des cercles et il est faux que tous les carrés sont reliés à tous les cercles.



All the squares are connected
with some of the circles.

true false

Figure 3.1. Figure tirée de la troisième expérience
de Geurts et Pouscoulous [GEU 09]

Geurts et Pouscoulous associent la phrase (63a) à la figure 3.1 qui rend la phrase vraie selon une lecture littérale et selon sa lecture globale décrite en (63c), mais la rend fausse selon la lecture locale donnée en (63b). Or, lorsque l'on demande aux participants si la situation représentée en figure 3.1 rend (63a) vraie ou fausse, ceux-ci répondent unanimement que la phrase est vraie. Geurts et Pouscoulous en concluent que la lecture locale n'est pas détectée par les participants. Cependant, ils envisagent l'éventualité que celle-ci existe néanmoins : par exemple, (63a) pourrait être ambiguë entre plusieurs lectures, dont la lecture locale, et pour une raison ou pour une autre, les sujets préféreraient choisir une lecture rendant la phrase vraie. Geurts et Pouscoulous répliquent ainsi l'expérience en proposant aux participants l'option supplémentaire de répondre que la phrase est soit vraie soit fausse ; cependant, aucun ne choisit cette option pour (63a). Geurts et Pouscoulous en concluent que les implicatures locales n'existent pas.

Cependant, d'après Chemla et Spector [CHE 11], les expériences menées par Geurts et Pouscoulous ne sont pas convaincantes. Tout d'abord, en vertu du principe qu'un résultat négatif n'est pas informatif, Chemla et Spector affirment que le fait que la lecture locale ne soit pas détectée ne prouve pas que cette lecture n'existe pas. De fait, selon eux, plusieurs raisons peuvent expliquer l'incapacité des participants à déceler la lecture locale dans l'expérience de Geurts et Pouscoulous, telles le caractère difficilement déchiffrable de la figure 3.1 ou le manque de saillance de la lecture locale de (63a). En outre, Chemla et Spector font valoir un principe de charité établi dans la littérature, selon lequel les sujets ont tendance à interpréter une

phrase comme vraie dès lors que celle-ci est vraie dans une de ses lectures. Il se pourrait donc que (63a) soit ambiguë entre plusieurs lectures dont la lecture locale (qui rend la phrase fausse), mais que les sujets jugent néanmoins la phrase vraie parce qu'elle l'est selon une autre de ses lectures. Et pour Chemla et Spector, le fait que les participants jugent (63a) non ambiguë dans la seconde expérience de Geurts et Pouscoulous ne montre en rien que la phrase n'est pas réellement ambiguë : les sujets ne sont peut-être simplement pas conscients de cette ambiguïté.

Chemla et Spector proposent ainsi une version modifiée de l'expérience de Geurts et Pouscoulous. Ils choisissent la phrase en (64a), similaire à (63a) puisque *certain* apparaît également dans la portée nucléaire d'un quantificateur universel. À nouveau, il s'agit d'examiner si la lecture locale, correspondant à (64c), peut être détectée par les participants, ou si seules sont perçues les lectures littérale et globale en (64b) et (64d) :

- (64) a. Chaque lettre est reliée à certains de ses cercles.
- b. Chaque lettre est reliée à au moins certains de ses cercles.
- c. Chaque lettre est reliée à certains de ses cercles et aucune lettre n'est reliée à tous ses cercles.
- d. Chaque lettre est reliée à certains de ses cercles et il est faux que chaque lettre est reliée à tous ses cercles.

(64a) est associée alternativement à quatre images correspondant respectivement à une première condition dans laquelle la phrase est fausse, à une seconde condition dans laquelle la phrase est vraie uniquement dans sa lecture littérale, à une troisième condition « faible » dans laquelle la phrase est vraie dans sa lecture littérale et dans sa lecture globale mais pas dans sa lecture locale, et à une quatrième condition « forte » dans laquelle la phrase est vraie dans ses lectures littérale, globale et locale. La figure 3.2 illustre ainsi la troisième condition, explorée dans Geurts et Pouscoulous [GEU 09] avec la figure 3.1.

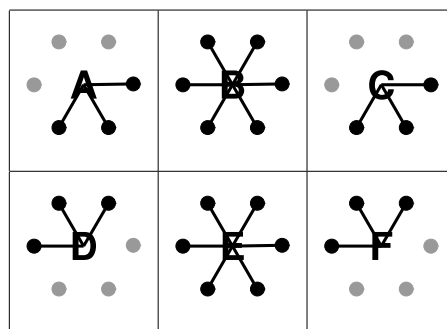


Figure 3.2. Figure proposée par Chemla et Spector [CHE 11] comme contrepartie à la figure de Geurts et Pouscoulous [GEU 09] (figure 3.1)

Chemla et Spector considèrent ce type d'image plus lisible que celui de la figure 3.1. En outre, en remplaçant les carrés – tous identiques – par des lettres différentes les unes des autres, Chemla et Spector entendent rendre la lecture locale plus saillante, en incitant les participants à examiner distinctement chaque lettre. Enfin, au lieu de demander aux participants des jugements binaires en termes de vérité et de fausseté, ils leur proposent d'évaluer le degré de vérité de la phrase sur une échelle continue allant du faux au vrai. Leur hypothèse est que plus le nombre de lectures vraies d'une phrase est grand dans une situation donnée, plus le degré de vérité attribué à la phrase est lui-même élevé. Chemla et Spector entendent ainsi contourner le principe de préférence pour la vérité, en vertu duquel les participants jugent une phrase vraie dès lors qu'elle l'est sous l'une de ses lectures. Or les résultats de leur expérience sont frappants : le degré de vérité attribué à (64a) est de 12 % quand la phrase est fautive, de 44 % quand la phrase est vraie uniquement dans sa lecture littérale, de 68 % dans la condition faible, c'est-à-dire lorsque la phrase est vraie dans ses lectures littérale et globale mais pas dans sa lecture locale, et de 99 % dans la condition forte, lorsque la phrase est vraie sous ses trois lectures, littérale, globale et locale.

En d'autres termes, conformément à l'hypothèse initiale de Chemla et Spector, plus le nombre de lectures vraies de la phrase est élevé, plus le degré de vérité attribué par les participants à la phrase est grand. En outre, et c'est là le résultat majeur de ce travail, la lecture locale est visiblement détectée par les participants, puisque le taux attribué à la condition forte est plus élevé que celui donné à la condition faible. Les résultats sont d'ailleurs similaires pour le terme scalaire *ou* testé en parallèle. Chemla et Spector prouvent ainsi qu'en améliorant le dispositif expérimental, la lecture locale, invisible dans l'expérience de Geurts et Pouscoulous, peut être détectée¹⁹.

3.3.7. Les implicatures scalaires : bilan

Si la notion d'implicature scalaire n'apparaît pas en tant que telle chez Grice, elle est initialement pensée comme une notion éminemment gricéenne par les linguistes qui s'en emparent, parmi lesquels Horn, Gazdar ou Levinson. De fait, les implicatures scalaires sont analysées par les néo-gricéens comme un type particulier d'implicature de quantité, au sens où elles sont dérivées à partir d'une proposition alternative plus informative qui n'a pas été dite par le locuteur et dont on peut donc

¹⁹ Dans le même article, Chemla et Spector [CHE 11] proposent en outre une seconde expérience portant sur un contexte non-monotone, sur laquelle ils s'appuient pour invalider l'approche pragmatique et conforter l'analyse grammaticale des implicatures enchâssées.

penser – en vertu de la première sous-maxime de quantité de Grice – que le locuteur ne la croit pas vraie. Les implicatures scalaires sont cependant considérées comme spécifiques en ce qu'elles sont analysées à partir de la notion d'échelle : les alternatives plus informatives dont elles sont la négation apparaissent construites par remplacement dans la phrase initiale d'un terme scalaire par un autre item logiquement plus fort appartenant à la même échelle. Dans les contextes dits monotones décroissants, l'orientation logique des échelles s'inverse, ce qui explique également l'effet d'inversion du schéma des implicatures scalaires. La notion d'implicature scalaire suscite cependant de nombreux débats. La nature des échelles, tout d'abord, est extrêmement débattue, qu'il s'agisse des éventuelles contraintes lexicales qu'elles subissent ou de l'importance du contexte dans leur établissement. En outre, l'émergence du phénomène des implicatures scalaires dites « enchâssées », c'est-à-dire des implicatures qui semblent apparaître localement, indépendamment de l'opérateur logique dans la portée duquel se trouve l'item scalaire, donne lieu, nous l'avons vu, à une proposition de rupture radicale avec l'approche néo-gricéenne standard des implicatures scalaires : un linguiste comme Chierchia propose ainsi de réanalyser celles-ci comme un phénomène essentiellement grammatical ; les implicatures scalaires ne seraient plus dérivées à partir de la proposition complète par un raisonnement portant sur les intentions du locuteur accomplissant un tel acte de langage, mais calculées localement, à chaque étape de la composition sémantique. Cette approche apparaît en outre confortée par certains travaux expérimentaux, dont ceux de Chemla et Spector.

Le débat est cependant loin d'être clos. D'une part, parce que d'autres analyses s'inscrivent dans une approche gricéenne ont été proposées pour rendre compte du phénomène des implicatures scalaires enchâssées. D'autre part, parce qu'admettre que les implicatures scalaires puissent être calculées localement ne revient pas nécessairement à reconnaître leur dimension grammaticale ; de fait, il se pourrait que les implicatures locales soient dérivées de manière pragmatique, quoique par un mécanisme différent de celui par lequel sont obtenues les implicatures globales²⁰.

Il reste à expliquer pourquoi les implicatures scalaires, qui constituent de loin les implicatures les plus étudiées parmi les implicatures conversationnelles, ont suscité et continuent de susciter un tel intérêt. Bien qu'il soit de toute évidence impossible de proposer une réponse univoque, nous aimerions ici faire trois suggestions. Tout d'abord, les implicatures scalaires apparaissent peut-être plus facilement formalisables que d'autres types d'implicatures conversationnelles grâce aux notions d'échelle et d'alternative scalaire, d'où l'intérêt initial des néo-gricéens mais aussi par la suite de chercheurs s'inscrivant plus nettement dans une approche formelle. D'autre part, les implicatures scalaires semblent souligner davantage les limites de

20 C'est une piste qu'explore par exemple Recanati [REC 03].

l'approche gricéenne que d'autres types d'implicatures conversationnelles, ce qui contribue également à expliquer que l'intérêt des chercheurs pour ce phénomène n'ait pas faibli avec la contestation grandissante du modèle gricéen. Enfin, le caractère aisément identifiable des implicatures scalaires se réduit à mesure que leur contexte d'apparition se complexifie ; or seuls des contextes très spécifiques peuvent permettre de – ou du moins aider à – départager les différentes analyses proposées. Ainsi, les approches expérimentales, de plus en plus prégnantes dans le champ de la pragmatique et plus généralement de la linguistique ces dernières décennies, vont également peu à peu s'emparer du phénomène et assurer sa pérennisation comme sujet de recherche de premier plan.

3.4. Conclusion

La notion d'implicature conversationnelle dans le système gricéen est indissociable du constat que la communication en langue naturelle ne se fait pas par simple codage et décodage de l'information, mais repose de manière cruciale sur l'expression et la compréhension d'une intention communicative. Dans l'approche gricéenne, les interlocuteurs disposent en effet d'une théorie de l'esprit sophistiquée, c'est-à-dire d'une capacité à se représenter les états mentaux d'autrui à un niveau de complexité élevé. Pour identifier les éventuelles implicatures conversationnelles communiquées par le locuteur, l'auditeur doit non seulement prendre en compte la forme linguistique de l'énoncé, mais aussi s'appuyer de manière cruciale sur le contexte. En outre, il doit présupposer que le locuteur respecte le principe de coopération et les maximes conversationnelles, et que ceci relève des connaissances partagées. Si les successeurs de Grice, tels Horn ou Levinson, proposent des remaniements plus ou moins importants des maximes conversationnelles, c'est bien Sperber et Wilson qui vont le plus loin dans l'actualisation de cette orientation cognitive latente du modèle gricéen, puisque ceux-ci proposent de réduire les maximes à un unique principe de pertinence qu'ils érigent en fondement d'une théorie générale de la cognition.

Or, de manière relativement ironique, ce sont les implicatures conversationnelles les plus étudiées par les néo-gricéens, à savoir les implicatures scalaires, qui vont mener à une remise en cause du modèle « intentionnaliste » proposé par Grice. Les implicatures scalaires sont initialement interprétées comme des implicatures typiquement gricéennes car relevant de la première sous-maxime de quantité, leur spécificité étant de dépendre de la présence dans la phrase d'un terme scalaire. L'attention des néo-gricéens s'est ainsi d'abord focalisée sur la nature précise des échelles. Cependant, l'identification du phénomène des implicatures scalaires dites enchâssées conduit certains linguistes à remettre en cause la nature

fondamentalement pragmatique des implicatures scalaires : celles-ci ne seraient pas le fruit d'un raisonnement permettant à l'auditeur de mettre au jour l'intention communicative du locuteur à partir du contexte, des maximes conversationnelles et de l'ensemble de la proposition exprimée par le locuteur, mais le résultat d'un mécanisme de calcul grammatical consistant à appliquer un algorithme prédéfini parallèle au processus de composition sémantique du sens littéral des expressions. Contrairement à l'approche gricéenne qui se veut globale en ce qu'elle postule une comparaison de l'ensemble de la proposition exprimée par rapport aux alternatives scalaires, cette dernière approche est dite « localiste » car elle suppose que le calcul des implicatures scalaires se fait « de bas en haut », à partir des expressions simples puis par composition progressive des expressions les unes avec les autres. Pour départager les approches pragmatique et grammaticale, le débat s'est donc resserré sur la question de savoir si les implicatures scalaires pouvaient réellement émerger localement. Dès lors, les recherches expérimentales se sont avérées fondamentales pour circonscrire la réalité des implicatures enchâssées, car les intuitions sont insuffisantes dans des contextes complexes. Cependant, si les travaux récents semblent conforter l'existence d'implicatures scalaires émergeant localement, cela ne préjuge en rien de leur nature grammaticale : même si l'approche gricéenne présupposant un raisonnement de l'auditeur sur l'ensemble de la proposition exprimée par le locuteur ne peut s'appliquer dans des cas spécifiques, rien n'empêche qu'un autre mécanisme pragmatique, local cette fois, puisse rendre compte de l'émergence des implicatures en jeu.

Une dernière question se pose avant de clore ce chapitre : comment expliquer la présence d'implicatures conversationnelles dans la communication humaine ? Plusieurs hypothèses ont été avancées (voir notamment [DAV 19] pour un résumé sur ce point). Soulignons-en trois. Un premier avantage des implicatures scalaires peut être de contribuer à l'efficacité de la communication en permettant au locuteur de communiquer une ou plusieurs propositions en plus de la proposition littéralement exprimée. Un autre de leurs avantages serait de permettre au locuteur de tromper son interlocuteur sans risquer d'être accusé de mensonge, puisque les implicatures conversationnelles sont toujours annulables. Enfin, certains linguistes comme Sperber et Wilson (voir [SPE 89]) mettent en avant le rôle fondamental des implicatures conversationnelles dans l'expression poétique²¹.

21 Voir le chapitre de Louis de Saussure dans ce volume.

3.5. Bibliographie

- [ATL 81] ATLAS J.D., LEVINSON S.C., « It-clefts, informativeness, and logical form: Radical pragmatics (revised standard version) », dans P. COLE (DIR.), *Radical pragmatics*, Academic Press, New York, 1981.
- [CHE 11] CHEMLA E., SPECTOR B., « Experimental evidence for embedded scalar implicatures », *Journal of Semantics*, n° 28(3), p. 359-400, 2011.
- [CHI 04] CHERCHIA G., « Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface », dans A. BELLETTI (DIR.), *Structures and Beyond*, vol. 3, Oxford University Press, 2004.
- [DAV 19] DAVIS W., « Implicature », dans E. D. ZALTA (DIR.), *The stanford Encyclopedia of Philosophy*, disponible à l'adresse : <https://plato.stanford.edu/archives/fall2019/entries/implicature/>, automne 2019.
- [FAU 75] FAUCONNIER G., « Pragmatic scales and logical structure », *Linguistic Inquiry*, p. 353-375, 1975.
- [GAZ 77] GAZDAR G., *Implicature, Presupposition and Logical Form*, Mimeograph, Indiana University Linguistics Club, 1977.
- [GEU 09] GEURTS B., POUSCOULOUS N., « Embedded implicatures ?? », *Semantics and Pragmatics*, n° 2, p. 1-34, 2009.
- [GRI 57] GRICE H.P., « Meaning », *Philosophical Review*, n° 66, p. 377-88, 1957.
- [GRI 61] GRICE H.P., « The Causal Theory of Perception », *Proceedings of the Aristotelian Society*, n° 35 (suppl.), p. 121-52, 1961.
- [GRI 75] GRICE H.P., « Logic and conversation », dans P. COLE, J. L. MORGAN (DIR.), *Syntax and Semantics*, vol. 3, p. 41-58, Academic Press, New York, 1975.
- [HIR 85] HIRSCHBERG J., *A Theory of Scalar Implicature*, thèse, Université de Pennsylvanie, 1985.
- [HOR 72] HORN L., *On the Semantic Properties of Logical Operators in English*, thèse, UCLA, 1972.
- [HOR 84] HORN L., « Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature », dans D. SCHIFFRIN (DIR.), *Meaning, Form, and Use in Context : Linguistic Applications*, p. 11-42, Georgetown University Press Washington, 1984.
- [HOR 89] HORN L., *A Natural History of Negation*, University of Chicago Press, Chicago, 1989.
- [LEV 83] LEVINSON S. C., *Pragmatics*, Cambridge University Press, Cambridge, 1983.
- [LEV 87] LEVINSON S. C., « Pragmatics and the grammar of anaphora », *Journal of linguistics*, n° 23, p. 379-434, 1987.
- [LEV 00] LEVINSON S. C., *Presumptive Meanings: The Theory of Generalized Conversational Implicature*, Cambridge, MIT Press, 2000.

- [MAT 95] MATSUMOTO Y., « The conversational condition on Horn scales », *Linguistics and Philosophy*, n° 18(1), p. 21-60, 1995.
- [PER 06] PERCUS O., « Antipresuppositions », dans A. UEMAYA (DIR.), *Theoretical and Empirical Studies of Reference and Anaphora: Toward the establishment of generative grammar as an empirical science*, Report of the Grant-in-Aid for Scientific Research (B), project n° 15320052, Japan Society for the Promotion of Science, p. 52-73, 2006.
- [REC 93] RECANATI F., « Embedded Implicatures ? », Institut Jean-Nicod, 1993.
- [SAU 04] SAUERLAND U. « Scalar Implicatures in Complex Sentences », *Linguistics and Philosophy* n° 27(3), p. 367–391, 2004.
- [SAU 08] SAUERLAND U., « Implicated presuppositions », dans A. STEUBE (DIR.), *The discourse potential of underspecified structures*, p. 581-600, Mouton De Gruyter, Berlin, 2008.
- [SPE 89] SPERBER D., WILSON D., *La Pertinence : communication et cognition*, Éditions de Minuit, Paris, 1989.
- [WIL 04] WILSON D., SPERBER D., « Relevance Theory », dans L. HORN, G. WARD (DIR.), *The Handbook of Pragmatics*, p. 607-632, Blackwell, Oxford, 2004.
- [WIT 53] WITTGENSTEIN L., *Philosophische Untersuchungen. Philosophical Investigations* [I, 1936-1945 ; II, 1947-1949], G. E. M. ANSCOMBE et R. RHEES (DIR.), trad. angl. par G. E. ANSCOMBE, Blackwell, Oxford, 1953.
- [ZIP 49] ZIPF G. K., « Human behavior and the principle of least effort », Addison-Wesley, Cambridge, 1949.

