

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LE CONCEPT DE CONSÉQUENCE LOGIQUE CHEZ TARSKI ET SA
CRITIQUE

MÉMOIRE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN PHILOSOPHIE

PAR

JULIEN OUELLETTE-MICHAUD

NOVEMBRE 2016

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

La rédaction de ce mémoire aurait sans doute été impossible sans la rencontre de nombreux modèles et sans l'aide de nombreuses personnes. J'aimerais, maintenant la rédaction achevée, prendre le temps d'adresser quelques remerciements. Je tiens d'abord à exprimer ma gratitude envers mon directeur de recherche, Serge Robert, pour avoir accepté d'influencer l'ensemble de mon parcours académique universitaire. J'aimerais également remercier Mathieu Marion et Alain Voizard pour leurs enseignements et pour leurs précieux commentaires sur mon travail.

Je voudrais ensuite remercier mes ami-e-s et mes collègues de l'université et du Collège de Maisonneuve, dont la présence à un moment ou à un autre a facilité beaucoup de choses. J'aimerais remercier en particulier Stéphanie Schanck, William-J. Beauchemin et Simon Brien, pour avoir enrichi ma vie intellectuelle par de nombreux échanges d'idées et par le partage d'un cocktail d'émotions. J'aimerais aussi remercier Alexandre Romano pour sa relecture et Louis Chartrand pour son aide avec la mise en page.

Mes remerciements vont enfin à ma famille, pour avoir toujours su se montrer compréhensive à mon égard malgré mes écarts et pour m'avoir toujours encouragé dans mon travail.

J'ai eu la chance de bénéficier du soutien financier du *Conseil de recherches en sciences humaines* ainsi que du *Fonds de Recherche du Québec - Société et Culture*. J'en suis extrêmement reconnaissant.

Table des matières

Remerciements	i
Résumé	iv
Introduction	1
0.1 Théorie des modèles et théorie de la preuve	2
0.1.1 Caractéristiques et problèmes	2
0.1.2 Développement historique	4
0.2 Aperçu et découpage du mémoire	7
CHAPITRE I	
La conséquence logique selon Tarski	10
1.1 Le projet de Tarski	10
1.1.1 La notion de définition et les métamathématiques	11
1.1.2 Intuitions, formalisme et <i>intuitionnisme formaliste</i>	16
1.2 Des tentatives jugées infructueuses par Tarski	21
1.2.1 Les approches syntaxiques et leur rejet	21
1.2.2 La tentative « sémantique » insuffisante de Carnap	23
1.3 L'approche sémantique de Tarski	28
1.3.1 Deux conditions (insuffisantes) d'adéquation	28
1.3.2 Modèles et satisfaction	31
1.3.3 La définition est-elle adéquate?	33
1.4 Conclusion	38

CHAPITRE II	
La critique d'Etchemendy	41
2.1 Qu'est-ce que la sémantique ?	41
2.1.1 L'approche représentationnelle et l'approche interprétationnelle	46
2.1.2 Les deux approches coïncident-elles ?	54
2.2 Quantification et nécessité	58
2.2.1 Logique du premier ordre, quantificateurs et restrictions	58
2.2.2 Le « sophisme de Tarski » et la question des constantes logiques	62
2.3 Le problème de la <i>sur-génération</i>	67
2.3.1 La variation du domaine d'interprétation	67
2.3.2 Le principe de réduction	70
2.4 Conclusion	75
CHAPITRE III	
Le dépassement de la critique d'Etchemendy	79
3.1 Les distorsions d'Etchemendy	80
3.1.1 Le faux dilemme d'Etchemendy ?	80
3.1.2 Le <i>problème</i> de la sur-génération	84
3.1.3 Comprendre l'argument de Tarski	87
3.2 (Comment) Peut-on sauver Tarski du « sophisme de Tarski » ?	90
Conclusion	95
Références	101

RÉSUMÉ

Il est commun, dans les manuels de logique, de présenter une définition *sémantique* du concept de conséquence logique. Cette approche est le fruit d'une tradition qu'on peut faire remonter au moins jusqu'aux travaux d'Alfred Tarski dans les années 1920 et 1930, lequel propose une définition du concept de conséquence en termes de modèles et de satisfaction : une conclusion est une conséquence logique d'un ensemble de prémisses si et seulement si tous les modèles des prémisses sont aussi un modèle de la conclusion. Autrement dit, une conclusion est une conséquence logique d'un ensemble de prémisses si la vérité est nécessairement préservée des prémisses à la conclusion. Cette définition a le mérite, selon Tarski, de rapporter le concept de conséquence aux critères de formalité et de nécessité.

John Etchemendy a remis à l'ordre du jour l'analyse du concept de conséquence logique, dans les années 1980 et 1990, par sa critique de la définition de Tarski. Ses arguments visent à identifier des problèmes de nature conceptuelle et de nature extensionnelle dans la définition tarskienne. Selon lui, la définition repose d'abord sur une confusion entre les approches représentationnelle et interprétationnelle de la sémantique. Elle échoue ensuite à caractériser adéquatement la nécessité du concept de conséquence logique. Enfin, la définition tarskienne déclare ou bien trop, ou bien trop peu d'arguments comme étant valides.

Ce mémoire porte sur la définition tarskienne du concept de conséquence logique et sur la littérature critique qu'elle a suscitée, particulièrement depuis les années 1990 et les travaux de John Etchemendy. Après des présentations philosophiques détaillées de la définition de Tarski et de sa critique par Etchemendy, je tente de réhabiliter la définition tarskienne en montrant des limites de cette critique sur chacun des trois axes.

Mots-clés : Conséquence logique, Alfred Tarski, John Etchemendy, définition, sémantique, théorie des modèles.

INTRODUCTION

Une question fondatrice de la logique est de savoir ce qu'est une inférence valide ; une inférence consiste à tirer un énoncé, la conclusion, d'un ensemble d'énoncés, les prémisses¹. La relation de conséquence porte sur les prémisses et la conclusion des inférences, c'est-à-dire que l'on juge qu'une inférence est valide si la conclusion est une conséquence logique des prémisses. Nos inférences valides reposent ainsi sur la reconnaissance de la relation de conséquence. C'est à la relation de conséquence logique que s'intéressera ce mémoire².

1. Aux fins de la présente introduction, l'inférence est ici entendue comme *produit* et non comme *acte*, bien qu'il soit possible de critiquer cette position (Peregrin, 2014). Nous n'insisterons pas non plus sur la distinction qu'il est possible d'établir entre inférence valide et conséquence logique. Voir par exemple les points de vue opposés de Dag Prawitz, Stewart Shapiro et Göran Sundholm dans (Prawitz, 2005), (Shapiro, 2011), (Sundholm, 2002) et (Sundholm, 2012).

2. On pourrait bien sûr mener une réflexion sur la structure même de la relation de conséquence logique. Puisque la conséquence logique est souvent pensée comme une relation binaire dans laquelle une chose s'ensuit d'un ensemble d'autres choses, l'on pourrait s'interroger sur la nature des éléments reliés, c'est-à-dire des choses à lier entre elles par la relation de conséquence. On pourrait ici penser à la conséquence logique comme relation entre ensembles, entre multi-ensembles, entre séquences de formules, et ainsi de suite. Une question corolaire consisterait à se demander quelles sont les composantes de ce que lie la relation de conséquence : s'agit-il d'énoncés ? de propositions ? d'assertions ? Qui plus est, ces composantes prennent-elles la forme sujet-prédicat ? fonction-arguments ? Enfin, une question serait de savoir si la conclusion (ou le conséquent) peut être une combinaison de plusieurs éléments. C'est le problème des logiques à conclusions multiples (Restall, 2005). Ces questions sont toutes dignes d'être étudiées, mais seront secondaires et abordées seulement indirectement dans le cadre de ce mémoire.

0.1 Théorie des modèles et théorie de la preuve

0.1.1 Caractéristiques et problèmes

De nombreuses approches ont été développées pour étudier la notion de conséquence logique, mais on peut dégager des similitudes et regrouper ces approches en deux grandes familles mettant l'accent sur des aspects différents du concept, à savoir la *théorie des modèles* et la *théorie de la preuve* (McKeon, 2010). Selon la première de ces familles, une conclusion est une conséquence logique de prémisses si cette conclusion est vraie dans tous les modèles possibles où les prémisses sont vraies, un modèle étant une interprétation qui assigne une signification aux expressions extra-logiques d'un schéma d'inférence (Caret et Hjortland, 2015, p.4).

Selon la seconde approche, on dit plutôt qu'une conclusion est une conséquence logique des prémisses si celle-ci peut-être dérivée ou produite à partir des prémisses en appliquant de manière adéquate des règles d'inférence (Caret et Hjortland, 2015, p.8). L'application des règles permet de transformer les énoncés initiaux, les prémisses, et de parvenir, par une suite d'étapes, à la conclusion. Ces règles d'inférence sont considérées comme primitives et comprises comme permettant de définir les expressions logiques qu'elles contiennent, dont les connecteurs logiques et les quantificateurs. Pour simplifier grossièrement le portrait, on pourrait dire que les approches issues de la théorie des modèles sont des approches *sémantiques*, en ce qu'elles tiennent d'abord et avant tout compte des valeurs de vérité des opérations, tandis que celles qui suivent la théorie de la preuve sont des approches *syntactiques*, en ce qu'elles s'intéressent principalement à la manipulation des symboles³.

3. Jean-Yves Girard (Girard, 1989) opère d'autres distinctions entre ces deux familles : la théorie des modèles adopte un point de vue plutôt *statique* de la logique, tandis que la théorie de la preuve adopte un point de vue plus *dynamique*. Ou encore : la théorie des modèles s'intéresse surtout à la *dénotation* des connecteurs logiques, tandis que la théorie de la preuve s'intéresse davantage à leur *sens*. Il ne faudrait toutefois pas exagérer la distinction et penser par exemple que la théorie de la preuve exclut les considérations sémantiques ; bien au contraire, on parle de sémantique preuve-théorique depuis au moins les travaux de Peter Schroeder-Heister. Voir par exemple (Wansing, 2000) et (Caret et Hjortland, 2015). On pourrait paraphraser Kant et dire qu'une syntaxe sans sémantique est vide et qu'une sémantique sans syntaxe est aveugle.

Évidemment, aucune des deux familles d'approches ne fait consensus auprès des philosophes ou des logiciens, car chacune a ses mérites respectifs, mais aussi, et surtout, car aucune d'entre elles n'est exempte de problèmes (Caret et Hjortland, 2015). Parmi les problèmes à résoudre, il y a le fait que l'énonciation des règles est souvent associée à des langages particuliers, rendant difficile de capturer la notion de conséquence dans toute sa généralité, et plus généralement le fait que les inférences sanctionnées par un système déductif devront recevoir une justification externe, c'est-à-dire qu'il faudra justifier que tel ensemble de règles soit adopté plutôt que tel autre ensemble, ce qui exige de recourir à un appareillage méta-théorique parfois très sophistiqué. Toutefois, ces problèmes ne sont ni insurmontables ni propres à l'une ou l'autre des orientations.

Concernant les problèmes plus spécifiques, les approches modèle-théorétiques se caractérisent notamment par leur esprit *réductif* : la conséquence logique, une notion à laquelle on reconnaît un aspect modal, est expliquée en termes non modaux et de manière purement extensionnelle par les modèles (Caret et Hjortland, 2015, p.4). On pourrait tout aussi bien dire qu'un schéma d'inférence est valide si on ne peut pas lui trouver de contre-exemple, c'est-à-dire si on ne peut pas trouver d'instanciation qui le falsifie. Cette réduction ne va pas de soi et doit être défendue pour être acceptée. Un autre problème pour ces approches est d'indiquer comment il est possible de savoir qu'une inférence donnée préserve la vérité, des prémisses à la conclusion, sous toutes les interprétations possibles, sans savoir préalablement et d'une manière indépendante qu'elle est valide. C'est le problème de la *quantification sur les domaines d'interprétation* (Gomez-Torrente, 2009). Les approches suivant la théorie des modèles doivent de plus indiquer comment, ou par quel principe, on peut sélectionner le vocabulaire à faire varier d'une interprétation à l'autre. C'est le problème des *constantes logiques*, ou de la *démarcation des termes logiques et extra-logiques* (MacFarlane, 2015). Nous reviendrons sur ces problèmes en détail dans ce qui suit.

Les approches par la théorie de la preuve font, quant à elles, appel à la notion de preuve dans un système déductif, censée rendre la conséquence logique plus facile à aborder parce que plus proche de nos pratiques courantes. Une objection

sérieuse à l'approche par la théorie de la preuve vient d'Arthur Prior (Prior, 1960), qui tente en effet de montrer que certaines caractérisations des règles d'inférence déterminent des connecteurs logiques inadmissibles. C'est le cas, par exemple, du fameux connecteur *tonk*⁴. Bien entendu, de nombreuses réponses ont été formulées donnant suite au problème qu'a soulevé Prior (Belnap, 1962; Dummett, 1991). La plupart d'entre elles visent à imposer des contraintes supplémentaires sur les systèmes déductifs et leurs règles d'inférence, comme l'*harmonie* entre règles d'introduction et d'élimination des connecteurs, ou encore la *conservativité* de l'extension des systèmes logiques, et ainsi de suite. Toutefois, les réponses fournies soulèvent à leur tour des questions supplémentaires, par exemple de savoir quel effet a le critère d'harmonie sur le sens des connecteurs logiques, ou encore de savoir si l'harmonie rend ceux-ci, à la différence du connecteur *tonk* de Prior, cohérents et consistants. Ces questions constituent autant de défis à relever pour les défenseurs des approches preuve-théorétique.

0.1.2 Développement historique

On peut remonter au moins jusqu'aux *Premiers Analytiques* d'Aristote, en particulier dans sa théorie du syllogisme, pour voir à l'oeuvre une réflexion sur la notion de conséquence logique. Cependant, ce n'est qu'au XXe siècle que les systématisations les plus satisfaisantes de la notion de conséquence ont vu le jour. On peut comprendre les approches modèle-théorétiques comme des héritières directes des travaux de Gottlob Frege (Frege, 1999), notamment en vertu de son développement d'une sémantique reposant sur la distinction entre le sens et la dénotation des expressions. Rudolf Carnap a quant à lui mis au point, dans sa *Syntaxe logique du langage* (Carnap, 1937), ce qui est considéré comme l'une des premières caractérisations syntaxiques de la conséquence logique⁵. Par contre,

4. Ce connecteur logique, faut-il le rappeler, se définit de la manière suivante : *tonk* s'introduit à la manière d'une disjonction (extensionnelle), c'est-à-dire qu'on peut affirmer $A \text{ tonk } B$ dès qu'on a A ou B , et s'élimine à la manière d'une conjonction (extensionnelle) c'est-à-dire qu'on peut affirmer A ou affirmer B dès qu'on a $A \text{ tonk } B$.

5. Nous verrons dans le premier chapitre du présent mémoire que Tarski considère un peu différemment la contribution de Carnap.

ce sont Alfred Tarski et Gerhard Gentzen qui sont les véritables piliers des approches sémantique et syntaxique, respectivement. En effet, ce sont les travaux de Tarski qui ont eu la plus grande influence sur le développement de la théorie des modèles, et ceux de Gentzen pour ce qui est du développement de la théorie de la preuve. Dans sa thèse doctorale (Gentzen, 1955), Gentzen met au jour deux systèmes de déduction qui deviendront des outils essentiels du théoricien de la preuve contemporain, à savoir la déduction naturelle et le calcul des séquents. On peut remarquer l'influence des idées de Gentzen sur la logique contemporaine en s'intéressant au domaine des logiques sous-structurales, qui connaît un essor depuis les années 1980 ; le développement des travaux en ce domaine repose presque entièrement sur les techniques formelles développées par Gentzen, qu'il s'agisse de travaux à saveur philosophique, linguistique, ou informatique (Restall, 2000).

La théorie des modèles, en tant qu'étude mathématique des langages formalisés par le biais des structures de la théorie des ensembles⁶, prend source dans les travaux de Tarski des années 1920 et 1930 (Tarski, 1983). Un exemple paradigmatique est sans doute la définition récursive qu'il donne du concept de vérité, entendu maintenant comme propriété mathématique des énoncés d'un langage, et non en un sens philosophique quelconque. Dans la même veine, Tarski présente, dans son article de 1936 (Tarski, 2002), une définition sémantique de la conséquence logique en termes de vérité dans un modèle faisant appel à la notion de satisfaction, relation entre des fonctions propositionnelles et des séquences d'objets. Pour toute interprétation d'un schéma d'inférence valide, il est impossible que les prémisses soient vraies et que la conclusion soit fausse. Dans les mots de Tarski : « the sentence σ follows logically from the sentences of the class Γ if and only if every model of the class Γ is also a model of the sentence σ » (Tarski, 2002, p.186)⁷. Cette définition a le mérite, selon Tarski, de rapporter le concept de conséquence logique aux critères de formalité et de nécessité. Plus précisément, la

6. Au sens technique, la théorie des modèles est aujourd'hui une branche des mathématiques, mais en un sens plus relâché, elle est tout simplement l'étude de l'interprétation des langages.

7. Il existe une traduction française faite par Gilles-Gaston Granger, révisée par Denis Bonnay et Mikaël Cozik, de l'article de 1936 (Tarski, 2009). Toutefois, comme la version anglaise est celle qui fait autorité et qui est citée dans la littérature sur Tarski, c'est à elle que nous ferons référence tout au long du mémoire. Par souci d'uniformité, nous utiliserons les méta-variables Γ et σ plutôt que K et X .

conséquence logique est dite formelle en ce qu'elle s'intéresse aux propriétés structurelles des inférences plutôt qu'au *contenu* particulier de chacune des prémisses ; elle est dite nécessaire en ce qu'on doit inexorablement accepter la vérité de la conclusion d'une inférence valide si l'on accepte la vérité de ses prémisses. En ce sens, la notion de conséquence agit à titre de norme de rationalité (Shapiro, 2005 ; Murzi et Carrara, 2014).

Le travail conceptuel et formel de Tarski est à l'évidence l'une des contributions à la logique les plus importantes du XXe siècle, comme en témoignent le fait que son analyse de plusieurs notions, moyennant quelques ajustements, est passée à la postérité, et le fait que ses travaux ont suscité d'innombrables critiques (Field, 1972; Kripke, 1975; Hodges, 1986). Sa définition de la conséquence logique ne fait pas exception ; elle se retrouve, au moins à l'état implicite, dans de nombreux manuels d'introduction à la logique qui font autorité, mais prête également flanc aux critiques en ce qu'elle repose sur, et doit donc élucider, des notions aussi centrales que la vérité ou la nécessité pour être pleinement satisfaisante. Elle laisse également ouvertes plusieurs questions, comme nous l'avons déjà souligné, dont celle de la distinction entre signes logiques et signes extra-logiques, c'est-à-dire la distinction des termes relativement aux différents niveaux de langage.

C'est sans doute à Etchemendy, dans les années 1980 et 1990, qu'il revient d'avoir le plus vigoureusement remis à l'ordre du jour l'analyse de la conséquence logique par sa véhémence critique de la définition de Tarski (Etchemendy, 1983; 1988a; 1990). La critique d'Etchemendy pourrait être qualifiée de *critique interne*, par opposition à *critique externe*, en ce sens qu'Etchemendy critique la définition de Tarski bien qu'il adopte lui-même une approche modèle-théorique. Bien entendu, il serait possible de formuler une critique externe de la définition tarskienne ; c'est par exemple ce que propose Stephen Read en défendant une logique de la pertinence selon l'approche preuve-théorique (Read, 1988). C'est spécifiquement à une partie du débat interne à la théorie des modèles, et non pas au débat externe, qu'est consacré le présent mémoire, où je procède d'abord à un examen de la définition donnée par Tarski avant de présenter sa critique dans la lignée des travaux d'Etchemendy. Je termine enfin en donnant une évaluation de cette

critique. Le point focal de la recherche est ainsi la définition sémantique du concept de conséquence logique.

0.2 Aperçu et découpage du mémoire

Le premier chapitre est consacré à la définition de la conséquence logique par Tarski dans l'article de 1936 (Tarski, 2002). Il s'agit, dans ce chapitre, de mettre en évidence l'importance conceptuelle de la conception sémantique de la conséquence logique pour l'analyse des langages formalisés et ce qui fait son originalité sur le plan historique. Dans ce chapitre, je tente d'abord d'exposer certaines caractéristiques du projet général de Tarski dans le domaine de la logique et des mathématiques, de manière à situer le contexte de l'élaboration de la définition du concept de conséquence (§1.1). Je tente ensuite d'explicitier les motivations ayant mené Tarski à rejeter les approches syntaxiques (§1.2.1) ainsi que la tentative à saveur plus *sémantique* développée par Carnap (Carnap, 1937)(§1.2.2), à laquelle Tarski fait explicitement référence dans son article. J'examine enfin la définition elle-même du concept de conséquence donnée par Tarski (§1.3), en m'intéressant aux conditions nécessaires (mais non suffisantes) du concept identifiées par Tarski (§1.3.1) et aux notions de modèle et de satisfaction (§1.3.2). Une attention particulière est portée à l'évaluation de la définition, puisque j'examine les liens de celle-ci avec la définition de Carnap (§1.3.3.1) et avec la question de la démarcation entre expressions logiques et extra-logiques (§1.3.3.2). Au fil de la discussion de la définition de Tarski, je m'efforce de faire ressortir les avantages et limites d'une définition sémantique de la conséquence logique.

Le second chapitre est consacré à la présentation de la critique principale — du moins dans la littérature contemporaine — de la définition de Tarski, à savoir celle de John Etchemendy (Etchemendy, 1990). Cette critique est en fait constituée d'arguments ciblant plusieurs aspects distincts mais reliés de la définition de Tarski. Premièrement, la définition de Tarski est erronée sur le plan conceptuel, en ce que, d'une part, elle repose sur une confusion entre deux approches de la sémantique (§2.1), et, d'autre part, en ce qu'elle échoue à caractériser adéquatement

la nécessité du concept de conséquence (§2.2). Plus précisément, Etchemendy juge qu'il faut adopter l'approche représentationnelle de la sémantique (§2.1.1.1) plutôt que l'approche interprétationnelle (§2.1.1.2), comme le fait Tarski, puisque les deux approches ont des implications différentes sur l'étude des langages formalisés (§2.1.2). Le fait que la définition tarskienne repose sur l'approche interprétationnelle rend problématique la caractérisation des quantificateurs (§2.2.1), ce qui entraîne la définition, selon Etchemendy, vers un sophisme modal, à savoir le « sophisme de Tarski » (§2.2.2).

Deuxièmement, la définition est aussi erronée sur le plan extensionnel, en ce qu'elle n'arrive pas à déclarer tous et seulement les arguments valides comme tels (§2.3). Autrement dit, pour Etchemendy, certaines inférences qui sont admises comme valides selon la définition de Tarski ne devraient pas l'être, et d'autres inférences devraient être jugées valides alors qu'elles ne sont pas considérées comme telles. Le problème est que Tarski, contrairement aux théoriciens des modèles contemporains, ne fait pas varier le domaine d'interprétation des modèles (§2.3.1) et fait reposer sa définition sur le principe de réduction, qui veut que la vérité logique d'un énoncé soit réduite à la vérité d'une généralisation universelle d'une forme propositionnelle associée (§2.3.2). Ce chapitre conclut en montrant que les problèmes qu'identifie Etchemendy soulèvent des doutes quant à la question de la démarcation des termes logiques, composante importante de la définition tarskienne, ainsi que, plus généralement, sur notre conception de la logique (§2.4).

Enfin, suite à cet examen minutieux de la définition de Tarski et de critiques qu'il est possible de lui adresser, je tente de montrer, dans le troisième chapitre, des limites de la critique d'Etchemendy. Je présente d'abord des défauts présents sur chacun des trois axes de la critique d'Etchemendy (§3.1). Plus précisément, j'essaie de montrer que la distinction entre approches représentationnelle et interprétationnelle de la sémantique est un faux dilemme (§3.1.1). J'essaie aussi de montrer que Tarski dispose des ressources pour obtenir l'équivalent de la variation du domaine d'interprétation (§3.1.2), et qu'il y a une manière différente de comprendre le raisonnement modal de Tarski (§3.1.3). À partir de cette critique à l'endroit d'Etchemendy, je tente de montrer, sur la base d'un argument de logique

modale, comment le « sophisme de Tarski » peut être évité (§3.2).

CHAPITRE I

LA CONSÉQUENCE LOGIQUE SELON TARSKI

1.1 Le projet de Tarski

L'article de 1936 de Tarski s'ouvre avec la phrase suivante :

« The concept of following logically belongs to the category of those concepts whose introduction into the domain of exact formal investigations was not only an act of arbitrary decision on the side of this or that researcher : in making precise the content of this concept, efforts were made to conform to the everyday 'pre-existing' way it is used. »
(Tarski, 2002, p.176)

Cette remarque peut paraître, à première vue, bien anodine, mais elle témoigne pourtant d'une approche méthodologique sur laquelle beaucoup de commentateurs de Tarski n'ont pas suffisamment insisté. En effet, il n'est pas inhabituel de présenter Tarski comme un logicien ayant pour objectif de fournir une *analyse conceptuelle* ou d'éclaircir la nature métaphysique de notions logiques fondamentales comme la vérité, la conséquence logique, et ainsi de suite. Ce qu'il entend faire lorsqu'il propose une définition est pourtant quelque chose de bien précis ayant bien peu à voir avec ce genre d'analyse. Comme comprendre le projet de Tarski jusqu'à l'article de 1936 est important pour comprendre la définition du concept de conséquence qu'il y fournit, nous commencerons par en faire une brève présentation.

1.1.1 La notion de définition et les métamathématiques

Par analyse conceptuelle, on veut souvent dire qu'un concept gagnerait à être élargi ou réduit à des concepts plus fondamentaux et plus simples. C'est en quelque sorte une manière d'*éliminer* un concept de notre vocabulaire. Un exemple caricatural de ce que pourrait vouloir dire *analyser* un concept est le concept « être humain », qu'on peut décomposer en concepts plus simples comme « animal » et « rationnel ». De manière moins triviale, on pourrait par exemple penser à l'ouvrage *The Concept of Law* (Hart, 1961) comme un exercice d'analyse conceptuelle : dans ce livre, l'idée générale de Herbert Hart est de considérer le droit (concept analysé) comme l'union de règles primaires et de règles secondaires (composantes du concept analysé), comprises comme spécifiant des conditions pour la reconnaissance de la validité juridique, pour le changement des lois et pour la l'arbitrage en cas de violation des lois¹. Ce qu'il est important de noter ici, c'est que ce n'est *pas* à ce genre d'exercice auquel s'adonne Tarski. Bien au contraire, ses travaux *présupposent* que de telles analyses conceptuelles sont disponibles.

Plusieurs autres articles de la même période (du début des années 1930) permettent d'arriver au même constat. Nous en citerons pour preuve deux autres extraits, sur le concept de vérité et, dans la section suivante, sur les ensembles définissables de nombres réels, respectivement. Commençons par le premier extrait :

« A thorough analysis of the meaning current in everyday life of the term 'true' is not intended here. Every reader possesses in greater or less degree an intuitive knowledge of the concept of truth and he can find detailed discussions on it in works on the theory of knowledge. »
(Tarski, 1983, p.153)

Dans cet extrait, on remarque que Tarski s'en remet au lecteur pour déterminer la signification du concept à étudier, concept dont il devrait déjà avoir une compréhension plus ou moins précise². Tarski spécifie également qu'il ne procédera pas à une analyse du terme « vrai » ; ses définitions ne sont donc pas des ana-

1. Pour d'autres exemples non triviaux de cette approche méthodologique, voir notamment les ouvrages de Gilbert Ryle (Ryle, 1949) et Hanna Pitkin (Pitkin, 1967).

2. Il ne faut pas oublier que Tarski publie ses travaux dans des revues scientifiques et que son lectorat est donc familier avec les sciences formelles.

lyses conceptuelles. Ceci soulève alors la question suivante : quel est précisément l'objectif que se fixe Tarski en *définissant* des concepts ?

Avant de poursuivre, une remarque sur la notion de définition s'impose. Bien qu'une définition puisse généralement être considérée comme un bi-conditionnel, c'est-à-dire comme une équivalence entre deux parties, il est coutume de distinguer au moins deux points de vue sur les définitions, à savoir le point de vue *abrégatif* et le point de vue *explicatif* (Belnap, 1993). Dans le cas abrégatif, le but d'une définition est de fixer à la pièce l'usage d'un mot, qu'il s'agisse d'un terme d'usage courant (définition lexicale) ou d'un nouveau terme (définition stipulative). Par exemple, le mathématicien dira à son élève « on obtient le *carré* d'un nombre en multipliant ce nombre par lui-même » ou dira à l'auditoire de sa conférence « soit *E* un espace vectoriel normé complet ». Dans ce cas, la définition sert d'outil pour qu'une communauté de locuteurs, par exemple un professeur et ses élèves, s'entende sur le sens, technique ou non, d'un mot ou d'une expression.

Par opposition, les définitions explicatives sont plutôt une affaire globale que locale et reposent sur une conception des théories prises comme des tous. La définition d'une expression revient dans ce cas à prendre un concept d'usage relativement courant, mais qu'on juge insatisfaisant sur le plan scientifique en regard d'autres concepts théoriques, tout en l'infléchissant de sorte qu'il soit plus exact et moins problématique. Il s'agit ainsi de *réviser* ou de *corriger* l'usage qui est fait du vocabulaire ; une définition revient alors à dire « voici l'usage qui devrait être fait d'un mot ». L'exemple du droit entendu comme union de règles primaires et secondaires pourrait être vu comme une définition de ce type³, ou encore l'analyse traditionnelle de la connaissance comme croyance vraie justifiée, par laquelle Socrate vise à corriger la conception de son interlocuteur⁴.

Pour comprendre la nature des définitions qu'il propose, il faut garder en tête le fait que Tarski s'intéresse d'abord et avant tout aux *langages formels* plutôt qu'aux

3. Rappelons qu'un objectif important de Hart dans *The Concept of Law* est de pallier aux lacunes du droit entendu comme application de menaces et de sanctions dans le positivisme juridique de John Austin.

4. Je laisse de côté les problèmes bien connus de ces définitions, l'idée étant simplement de présenter la teneur de ce *type* de définition.

langues naturelles. Bien qu'il y ait de bonnes raisons de penser que cette distinction en est une de degré et non de nature⁵, une caractéristique importante des langages formels semble être l'utilisation calculatoire ou computationnelle pouvant en être faite. Lorsqu'on ajoute des règles de transformation à un langage formel, on en fait un *système déductif* où il est possible de déduire des formules à partir d'autres formules. Autrement dit, on peut alors « follow 'mechanically' the instructions contained in a calculus, i.e., a well-defined notational system with clear rules of formation and transformation » (Dutilh Novaes, 2012, p.17). Les définitions que Tarski donne des concepts de vérité ou de conséquence logique ne sont donc pas destinées à se retrouver dans un dictionnaire régissant notre utilisation ordinaire ou scientifique des mots, mais concernent plutôt des systèmes déductifs qui permettent de formaliser des branches de la connaissance comme la physique, la géométrie ou l'arithmétique. Les définitions qui occupent Tarski ne peuvent ainsi pas être comprises directement à la lumière des définitions abrégatives et explicatives telles que nous venons de les présenter. Pour mieux comprendre, voyons d'abord ce que comprennent les systèmes déductifs. Les systèmes déductifs comprennent :

1. Des symboles, dont les signes de ponctuation et les termes primitifs. La signification des termes primitifs est en lien avec la branche des connaissances dont il est question (par exemple, pour la géométrie, on aura comme primitifs « point », « droite », etc.) ;
2. Une syntaxe, qui indique comment construire l'ensemble de formules à partir des symboles ;
3. Des axiomes, qui sont, parmi l'ensemble des formules, celles apparaissant évidentes (par exemple, dans sa géométrie, Euclide considère « une droite peut être formée à partir de n'importe quelle paire de points distincts » comme un axiome) ;
4. Des règles d'inférences nous permettant de dériver, à partir des axiomes, des théorèmes, qui forment un sous-ensemble de formules vraies parmi l'ensemble des formules.

5. Un langage pourrait par exemple être considéré comme *plus ou moins* formel (Dutilh Novaes, 2012).

Avec l'idée de système déductif vient la distinction importante entre des *niveaux* auxquels il est possible d'étudier scientifiquement un objet d'étude⁶. Pour illustrer ceci, prenons l'exemple de la géométrie. Au premier niveau se trouvent le *système déductif* et ses théorèmes, c'est-à-dire des énoncés vrais auxquels on voudra arriver à partir de nos axiomes et de règles d'inférences. Par exemple, le théorème de Pythagore est un théorème de la géométrie euclidienne. Une fois qu'on indique des règles pour la formation des formules et une relation de conséquence entre celles-ci, on obtient, à un niveau d'ordre supérieur, une *théorie déductive*. On peut alors arriver à des figures comme les triangles, construites à partir de points puis de droites, puis procéder à des démonstrations complexes sur celles-ci. On peut par exemple montrer que « la somme des angles intérieurs d'un triangle est égale à deux droits ». Au troisième niveau se trouve l'étude des niveaux précédents, à l'aide d'une méthodologie formelle. C'est ce que Tarski appelle une *méta-théorie*. À ce niveau, on vérifiera par exemple que notre théorie déductive est consistante, qu'elle ne permet pas de dériver de contradictions, etc. Par contre, Tarski croit que les langages ordinaires contribuent également à notre méthodologie, de manière informelle. On le voit chaque fois qu'un mathématicien utilise le français, l'anglais, etc., lorsqu'il écrit par exemple ses articles scientifiques. Cette contribution informelle ne doit surtout pas être négligée, notamment puisqu'elle est directement liée au contenu intuitif qu'il s'agira d'étudier dans une théorie déductive.

Les métamathématiques, dans la lignée des travaux de David Hilbert sur les fondements des mathématiques, visent à développer des méthodes pour étudier les propriétés des théories et systèmes déductifs, en prenant les mathématiques elles-mêmes comme objet d'étude. En tant que *méta*-mathématicien, Tarski se propose donc d'étudier la méthodologie des théories déductives :

« Our object in this communication is to define the meaning, and to establish the elementary properties, of some important concepts belonging to the *methodology of deductive sciences* [les métamathématiques] [...] These disciplines are regarded, from the standpoint of metamathematics, as sets of *sentences*. » (Tarski, 1983, p.30)

6. Cette présentation se base en grande partie sur la contribution de Wilfrid Hodges, dans (Patterson, 2008).

Les définitions que propose Tarski ne sont donc pas des stipulations *locales*, concernant des preuves particulières ou la démonstration de formules spécifiques, mais visent plutôt l'usage *global* de termes au niveau méta-théorique par la formalisation de ceux-ci. Le travail de Tarski est ainsi d'indiquer comment des concepts dont on comprend *déjà* la signification de manière intuitive peuvent être introduits ou *importés*, pour ainsi dire, et *exprimés* dans des systèmes formels⁷. Pour le concept de conséquence, Tarski a par exemple proposé cette axiomatisation :

« From the sentences of any set X certain other sentences can be obtained by means of certain operations called *rules of inference*. These sentences are called *consequences of the set X* . The set of all consequences is denoted by the symbol ' $Cn(X)$ '. [...] [Consequence] will here be regarded as primitive and will be characterized by means of a series of axioms. In the customary notation of general set theory these axioms can be formulated in the following way :

Axiom 1. $\bar{S} \leq \aleph_0$

Axiom 2. If $X \subseteq S$, then $X \subseteq Cn(X) \subseteq S$.

Axiom 3. If $X \subseteq S$, then $Cn(Cn(X)) = Cn(X)$.

Axiom 4. If $X \subseteq S$, then $Cn(X) = \sum_{Y \subseteq X \text{ and } \bar{Y} \leq \aleph_0} Cn(Y)$

Axiom 5. There exists a sentence $x \in S$ such that $Cn(x) = S$. »

(Tarski, 1983, p.30-31)

Cependant, Tarski ne vise pas à donner des conditions nécessaires et suffisantes pour l'usage des concepts, ou à réduire des concepts à d'autres concepts plus simples, puisqu'il ne s'intéresse pas tant à leur contenu qu'à leur formalisation. Dire qu'il vise, par ses définitions, à fournir des *analyses conceptuelles* de notions logiques et mathématiques est ainsi une erreur⁸.

7. En cela, Tarski s'inscrit en continuité avec ses professeurs Stanislaw Lesniewski et Tadeusz Kotarbinski; voir (Patterson, 2012, Chapitre 1).

8. D'ailleurs, quiconque a déjà jeté un oeil à l'ensemble des travaux de Tarski est en mesure de voir que ses définitions, contrairement à celles de Ryle, Hart ou Pitkin, ne permettent pas vraiment de *simplifier* ou de nous *conforter* dans l'usage d'un concept. La définition qui vient d'être donnée du concept de conséquence en est un exemple, comme la *définition 22* de l'article de 1933 sur le concept de vérité : « The sequence f satisfies the sentential function x if and only if f is an infinite sequence of classes and x is a sentential function and if f and x are such that either (α) there exist natural numbers k and l such that $x = \iota_{k,l}$ and $f_k \subseteq f_l$; (β) there is a sentential function y such that $x = \bar{y}$ and f does not satisfy the function y ; (γ) there are sentential functions y and z such that $x = y + z$ and f either satisfies y or satisfies z ; or finally (δ) there is a natural number k and a sentential function y such that $x = \cap_y k$ and every

1.1.2 Intuitions, formalisme et *intuitionnisme formaliste*

Dans un article sur les ensembles définissables de nombres réels, Tarski fait une remarque similaire à ce qui vient d'être dit sur les définitions. L'extrait qui suit nous permettra également de voir quelle est l'importance méthodologique du contenu *intuitif* des concepts dans le projet de Tarski :

Notre intérêt se porte sur un terme dont nous nous rendons compte plus au moins [sic] précis quant à son contenu intuitif, mais dont la signification n'a pas été jusqu'à présent (tout au moins dans le domaine des mathématiques) établie de façon rigoureuse. Nous cherchons donc à construire une définition de ce terme qui, tout en satisfaisant aux postulats de la rigueur [sic] méthodologique, saisisse en même temps avec justesse et précision la signification "trouvée" du terme. C'étaient bien les problèmes de cette nature que résolvaient les géomètres qui établissaient pour la première fois le sens des termes "mouvement", "ligne", "surface" ou "dimension". (Tarski, 1931, p.212)

Dans cet extrait, on peut d'abord noter, en lien avec ce qui a été soutenu dans la section précédente, que la signification « trouvée » d'un terme a un contenu intuitif pouvant être rendu précis dans une théorie mathématique, c'est-à-dire que la signification trouvée peut être saisie de manière rigoureuse par un formalisme. Ceci est vrai des notions de vérité et de conséquence comme de diverses notions étudiées par les mathématiques ; on a, par exemple, déjà une idée de ce qu'est un point ou une droite avant de commencer à faire de la géométrie analytique. Tarski poursuit l'article de la manière suivante en indiquant la relation entre intuition et formalisation :

Je commencerai donc par présenter au lecteur le contenu du terme à envisager [ensemble définissable de nombres réels], et notamment tel qu'on l'entendait jusqu'à présent en Métamathématique. Les remarques que je vais faire à ce sujet ne sont point indispensables pour les considérations qui vont suivre — pas plus que la connaissance empirique des lignes et surfaces ne l'est pour une théorie mathématique de ces notions. Elles nous permettront toutefois de nous emparer plus

infinite sequence of classes which differs from f in at most the k -th place satisfies the function y » (Tarski, 1983, p.193).

facilement des constructions exposées aux §§ ultérieurs, et, avant tout, de juger si elles répondent en effet à la signification “trouvée” de la notion. (Tarski, 1931, p.213)

On remarque alors que les formalismes servant à exprimer les intuitions sont en quelque sorte indépendants de celles-ci, c'est-à-dire qu'une personne pourrait très bien comprendre la théorie mathématique visant à capturer les intuitions sans avoir les intuitions en question. Réciproquement, les intuitions sont relativement indépendantes de leur expression rigoureuse. Par contre, malgré l'indépendance du concept intuitif et de la définition rigoureuse qu'on vise à en donner, Tarski défend qu'on peut évaluer le succès de la définition formelle en la mesurant au concept intuitif. Une définition peut alors être comprise comme une certaine interaction entre des niveaux de langage, entre l'usage et la mention des termes à l'intérieur d'un système déductif.

Sans aller dans le détail de la conception du langage, de la logique et des mathématiques derrière les propos de Tarski, quelques remarques sont de mise pour rendre plus clair ce point. Premièrement, une idée importante attribuée à Tarski est la distinction entre des niveaux de langage, introduite notamment pour éviter les paradoxes comme le paradoxe du menteur⁹. Tarski croit que le paradoxe du menteur résulte du fait que les langues naturelles sont sémantiquement fermées, c'est-à-dire qu'elles disposent de prédicats sémantiques comme « vrai », « faux », etc., qu'on peut appliquer aux énoncés des langues naturelles elles-mêmes. Par opposition, les langues sémantiquement ouvertes n'ont pas d'énoncés pouvant être prédiqués d'eux-mêmes, ce qui permet d'éviter la formulation du paradoxe du menteur. Pour pouvoir donner la définition de concepts comme la satisfaction ou la vérité pour un langage L sémantiquement ouvert, appelé *langage objet*, on procédera en donnant la définition en recourant à un autre langage, sémantiquement ouvert lui aussi, à savoir le *métalangage* M, contenant les prédicats référant aux propriétés sémantiques du premier langage. Le concept de vérité, par exemple, est ainsi un concept du métalangage exprimant la propriété d'être-vrai-dans L¹⁰.

9. On arrive au paradoxe du menteur en déclarant une chose comme « Je mens », énoncé ne pouvant pas recevoir de valeur de vérité. Pour une discussion de ce paradoxe et de la distinction entre niveaux de langage, voir (Kirkham, 1992, chapitre 9).

10. Cette distinction entre niveaux de langages, introduite par Tarski pour éviter l'apparition du paradoxe du menteur, vient elle-même avec ses problèmes, dont une version plus complexe

Deuxièmement, Tarski croit qu'un des objectifs de l'utilisation des théories et systèmes déductifs est l'expression de la pensée. Ainsi, les formules d'un langage formel expriment des relations entre elles — en ce qu'on peut dériver les unes à partir des autres conformément à des règles d'inférence, en ce que certaines sont incompatibles entre elles, etc. — mais elles servent également à exprimer des pensées. Ceci est surtout vrai pour ses premiers travaux dans les années 1920, mais on trouve également des traces de cette conception dans ses écrits des années 1930¹¹.

Troisièmement, les théories et systèmes déductifs étudiés par les métamathématiques visent à gouverner l'interprétation des termes primitifs, conformément aux intuitions, par l'obtention de théorèmes qui contraignent la signification des termes de la théorie¹². Autrement dit, l'usage des concepts doit être régi par les théorèmes de notre théorie ; on obtiendra la signification des concepts que l'on souhaite capturer lorsque notre théorie permettra d'arriver aux théorèmes souhaités. On peut alors remarquer l'importance de la capture du contenu intuitif de concepts et de leur expression comme moteur du développement des formalismes.

Enfin, les formules dérivées à l'intérieur d'un système déductif doivent être intuitivement contraignantes (*intuitively binding*). En d'autres mots, on devrait reconnaître la validité intuitive des inférences à l'intérieur du système déductif. En ce sens, il doit y avoir un aspect épistémique associé aux formalismes développés par les mathématiciens.

Cette brève description de quelques caractéristiques de la position de Tarski, loin d'en épuiser la finesse et la richesse, permet toutefois d'en dessiner les contours et du paradoxe, comme l'a remarqué Saul Kripke (Kripke, 1975).

11. Cet aspect, comme nous le verrons, revêt une moins grande importance dans l'article de 1936, mais s'est avéré crucial pour le développement des méthodes de définition tarskiennes et le rejet de la définition carnapienne.

12. La nature exacte de la notion d'intuition en jeu mériterait elle-même un examen minutieux. Je crois cependant qu'on peut faire une lecture *minimaliste* de la notion d'intuition présente chez Tarski, c'est-à-dire qu'on peut comprendre cette notion sans se commettre à des thèses philosophiques trop fortes. Ignacio Jané offre par exemple des raisons convaincantes de penser que le concept *intuitif* de Tarski est simplement celui associé au développement des méthodes axiomatiques (Jané, 2006).

d'en comprendre l'orientation¹³. Tarski appelle cette position le *formalisme intuitionniste* (*intuitionistic formalism*), en référence directe aux idées du directeur de sa thèse de doctorat, Stanislaw Lesniewski (Tarski, 1983, p.62). Pourtant, le formalisme, étant associé aux travaux de l'école de Hilbert¹⁴, est souvent considéré comme une approche relevant davantage de la théorie de la preuve, en ce qu'une de ses caractéristiques principales est une *abstraction* du contenu sémantique des concepts ou de la signification des expressions. Il peut alors paraître surprenant de voir que Tarski lui-même a embrassé une variante de cette position, lui qui est pourtant connu pour ses définitions sémantiques et ses travaux en quelque sorte fondateurs de la tradition de la théorie des modèles.

Il ne faut toutefois pas oublier, comme nous l'avons précisé dans l'introduction, qu'on ne doit pas séparer de manière trop appuyée les notions de syntaxe et de sémantique, qui sont des faces d'une même médaille. Il ne faut pas non plus oublier que la plupart des travaux de Tarski des années 1920 et du début des années 1930 sont menés suivant le modèle axiomatique ; les premières définitions que Tarski a données du concept de conséquence sont développées de manière axiomatique, en mettant davantage l'accent sur l'aspect syntaxique¹⁵. Enfin, notons que les travaux proprement mathématiques de Tarski sont surtout de nature algébrique, ce qui est un aspect central dans l'histoire du formalisme comme courant méthodologique.

Par contre, pour mieux voir comment Tarski s'est toujours démarqué des travaux de l'école de Hilbert, voyons la définition que donne Michael Detlefsen du formalisme de ce dernier, en particulier concernant la relation entre contenu et axiomes dans un système déductif :

« The mathematician, on [Hilbert's] view, is free to stipulate of a concept she's introducing that it have exclusively the properties pro-

13. Pour un exposé détaillé de cette position, on consultera avec grand profit (Patterson, 2012) et (Patterson, 2008), en particulier les contributions d'Arianna Betti, de Solomon Feferman, de Wilfrid Hodges et de Douglas Patterson.

14. Au moins en tout cas pour ce qui est de la crise des fondements des mathématiques au tournant du XXe siècle.

15. Voir par exemple la définition donnée dans la section précédente. Nous verrons plus loin les motivations de Tarski derrière le rejet de cette approche pour la définition de ce concept.

vided for by the axioms she uses to introduce it. There is no content belonging to a concept introduced in this way except that which is provided for by the introducing axioms. [...] Concepts are thus identified with the roles they occupy in mathematical thinking [...] They do not have to have an intuitive content in order to be significant. Hilbert believed that the inferential roles of concepts are determined not by contents given prior to the axioms which introduce them, but by those introducing axioms themselves. » (Detlefsen, 2005, p.294)

Bien que beaucoup d'autres aspects permettent par ailleurs de l'y identifier, on peut voir que Tarski se démarque d'un pan important de la tradition formaliste, puisque les caractéristiques soulevées par cette définition du formalisme entrent en conflit avec la position que nous avons décrite précédemment. L'une des oppositions majeures entre ces variantes de formalisme réside dans l'acceptation — respectivement, le refus — des langages non interprétés, une interprétation étant ce qui permet de déterminer la valeur de vérité des formules d'un langage donné¹⁶. Ainsi pour Tarski, de la période allant des années 1920 jusqu'au milieu des années 1930, les langages formels doivent impérativement être interprétés, puisqu'ils doivent rendre compte des pensées et donc d'un contenu intuitif¹⁷. Ceci veut dire que le métamatématicien pour Tarski travaille avec des langages dont les expressions sont douées de sens (*meaningful*). Au contraire, pour Hilbert ou pour Carnap, le mathématicien est libre de construire ses systèmes formels comme il l'entend, dans la mesure où ceux-ci sont minimalement consistants et féconds, puisqu'aucune considération extra-systémique touchant le contenu intuitif des concepts ne pèse sur lui. C'est d'ailleurs en grande partie pour combler des lacunes présentes dans les travaux de Carnap que Tarski a voulu définir le concept de conséquence en des termes proprement sémantiques dans son article de 1936, comme nous le verrons dans les prochaines sections.

16. On peut donc considérer ceci comme une composante sémantique importante. Nous présenterons plus loin de manière plus détaillée en quoi consiste la notion d'interprétation et en quoi consiste son importance pour les approches modèle-théorétiques.

17. Ceci est par exemple confirmé par cette remarque de Tarski : « It remains perhaps to add that we are not interested here in 'formal' languages and sciences in one special sense of the word 'formal', namely sciences of the signs and expressions to which no meaning is attached. For such sciences the problem here discussed has no relevance, it is not even meaningful. We shall always ascribe quite concrete and, for us, intelligible meanings to the signs which occur in the languages we shall consider » (Tarski, 1983, p.166-167). Voir (Jané, 2006), (Mancosu, 2010) et (Mancosu *et al.*, 2010).

1.2 Des tentatives jugées infructueuses par Tarski

Après avoir soulevé quelques problèmes pouvant surgir dans la définition du concept intuitif de conséquence, tel que le fait qu'il puisse être confondu avec d'autres notions ou encore le fait que son usage par les mathématiciens puisse être vague et conduire à des ambiguïtés, Tarski poursuit l'article de 1936 en se penchant sur des tentatives ayant été faites pour le capturer dans une théorie déductive. Le constat auquel il arrive est que ces tentatives débouchent toutes sur des problèmes, auxquels il souhaitera lui-même remédier avec sa définition.

1.2.1 Les approches syntaxiques et leur rejet

Le recours à des concepts et méthodes de nature sémantique a longtemps rencontré une grande opposition, par exemple de la part de Carnap et des autres membres du Cercle de Vienne. Selon eux, les notions sémantiques ne peuvent pas faire partie d'une conception scientifique du monde digne de ce nom. Ceci est notamment dû au fait de leur adhésion à un empirisme strict et à la crainte de la « métaphysique spéculative ». De plus, les concepts sémantiques, sans traitement formel adéquat, peuvent déboucher sur des paradoxes (sémantiques), comme le paradoxe du menteur et le paradoxe de Russell¹⁸. Enfin, le scepticisme à l'endroit des notions sémantiques peut aussi être attribué au fait que beaucoup de travaux réalisés au courant des années 1910 et 1920 ont débouché sur des résultats syntaxiques importants¹⁹. Parmi ces travaux, on peut par exemple penser à ceux d'Emil Post, de Thoralf Skolem ou encore de David Hilbert et Wilhelm Ackerman, sur la complétude de la logique du premier ordre. Ainsi, il a été suggéré par de nombreux logiciens que l'approche syntaxique, en particulier sous sa forme axiomatique consistant à appliquer successivement des règles d'inférences pour dériver

18. On arrive au paradoxe de Russell, en théorie des ensembles, en demandant si l'ensemble des ensembles ne se contenant pas eux-mêmes se contient lui-même, question à laquelle on ne peut répondre ni par l'affirmative ni par la négative.

19. L'influence de Tarski sur Carnap sera par contre si grande que celui-ci publiera en 1942 un ouvrage d'introduction à la sémantique, même si leur conception de cette notion est légèrement différente.

les théorèmes à partir des axiomes, permettait de rendre compte entièrement du concept de conséquence. Par contre, des problèmes que Tarski ne manque pas de soulever se posent pour ces approches, dont celui concernant les ω -règles et l' ω -incomplétude.

Les ω -règles, connues également comme *règles d'induction infinie*, permettent d'inférer une généralisation universelle à partir d'un ensemble (infini dénombrable) de ses instances. Ces règles sont d'une grande importance, notamment parce qu'elles sont associées à la consistance des théories déductives, mais surtout pour le rôle qu'elles remplissent dans la preuve du théorème d'incomplétude de Gödel, en 1931. Une théorie est dite ω -incomplète s'il faut lui ajouter une ω -règle pour la rendre complète, c'est-à-dire pour pouvoir dériver *toutes* les conséquences intuitivement valides. On peut saisir ce dont il est question en présentant l'exemple de Tarski d'un système ayant certaines propriétés particulières. Dans certains systèmes comprenant les énoncés de l'arithmétique, on peut par exemple prouver que :

1. Le nombre 0 possède la propriété P ;
2. Le nombre 1 possède la propriété P ;
3. Le nombre 2 possède la propriété P ;
4. Le nombre n possède la propriété P , pour tout nombre naturel n arbitraire.

Par contre, certains de ces systèmes ne permettent pas de prouver, uniquement à l'aide des règles finitaires standards, que :

5. *Tous* les nombres naturels n possèdent la propriété P .

Tarski croit pourtant que l'énoncé (5) est une *conséquence intuitive* des énoncés (1)-(4), en ce sens qu'en présence d'énoncés respectant cette forme, on passera toujours d'énoncés vrais à un énoncé vrai (Tarski, 2002, p.178). Une solution à ce problème, à savoir au fait de ne pas reconnaître comme valide le passage de la totalité des énoncés singuliers à l'énoncé universel, est d'introduire une nouvelle catégorie de règles à notre théorie, appelées ω -règles²⁰. Toutefois, Tarski en vient

20. Les ω -règles sont également associées à l'idée de *catégoricité*, c'est-à-dire que leur introduction permet d'assurer que tous les modèles d'une théorie sont isomorphes, ce qui permet en retour d'assurer que la signification des concepts peut être exprimée précisément, conformément à l'intuitionnisme formaliste. Voir Tarski, 1983, Chapitre IX et (Patterson, 2012, §2.3).

à considérer que le théorème d'incomplétude de Kurt Gödel (Gödel, 1967) donne le coup de grâce à ce genre de tentatives de définition syntaxique :

« The supposition suggests itself that on the route sketched above — supplementing the rules of inference used in the construction of deductive theories with further rules of a structural character [les ω -règles] — we would succeed finally in capturing the “essential” content of the concept of following, which has by no means been exhausted by the rules used until now [les règles finitaires]. Relying on the investigations of K. Gödel [sur l'incomplétude], one can demonstrate that this supposition is mistaken : if we abstract from certain theories with a very elementary structure, then always — no matter how we enrich the stock of rules of inference — we shall be able to construct sentences which follow in the everyday sense [le concept intuitif] from the theorems of the deductive theory under consideration, but which cannot be proven in this theory on the basis of the accepted rules. » (Tarski, 2002, p.181)

Pour formuler simplement l'idée générale présente dans le théorème d'incomplétude, on pourrait dire qu'il y aura toujours, dans une théorie déductive suffisamment complexe pour dériver une partie des théorèmes de l'arithmétique élémentaire, des énoncés, et plus particulièrement des énoncés *vrais*, qu'il est néanmoins impossible de prouver. C'est donc qu'il n'est pas possible de capturer *toutes* les conséquences intuitives à l'aide de règles d'inférence, finitaires ou non, en raison du problème de l' ω -incomplétude. Les définitions axiomatiques de la conséquence fournies par Tarski des années 1920 et du début des années 1930 échouent donc à capturer adéquatement le concept de conséquence. Loin de conclure de ceci que tous les travaux de nature syntaxique doivent être rejetés, Tarski prend plutôt acte du fait que c'est à l'aide d'autres méthodes qu'il faut procéder pour fournir une définition satisfaisante du concept de conséquence.

1.2.2 La tentative « sémantique » insuffisante de Carnap

Bien qu'il soit plus commun de voir Carnap comme travaillant, pour ainsi dire, du côté *syntaxique* de la clôture, Tarski, dans l'article de 1936, semble référer à ses travaux comme s'approchant plutôt d'une définition sémantique. En effet,

le fait que Tarski traite de Carnap seulement *après* avoir parlé des approches syntaxiques et de Gödel, pour les écarter, laisse penser qu'il traite celui-ci différemment de celles-là, tout comme le fait qu'il soit l'unique logicien auquel Tarski réfère explicitement comme ayant tenté de définir sémantiquement la notion de conséquence. Un autre indice est que Tarski commence sa présentation de Carnap de la manière suivante : « The *first* attempt at the formulation of a precise definition for the *proper concept of following* comes from R. Carnap » (Tarski, 2002, p.182, je souligne), comme si les travaux de Carnap, dans sa *Syntaxe Logique du Langage*, représentaient un pas important dans la bonne direction ²¹.

La *Syntaxe Logique du Langage* (Carnap, 1937) peut être comprise comme une réponse à Ludwig Wittgenstein, en ce que d'une part Carnap tente de déterminer ce qui peut être *dit* en distinguant les modes de discours « formel » et « matériel » ²², et d'autre part Carnap tente de présenter les problèmes *philosophiques* comme des problèmes *logiques*, plutôt que de les rejeter comme de pseudo-problèmes tel que l'auteur du *Tractatus* le fait. L'ouvrage de Carnap peut aussi être vu comme une réponse aux travaux de Gödel, en ce qu'il conçoit le théorème d'incomplétude comme menaçant la distinction entre les énoncés analytiques et synthétiques, ce qui est problématique pour la construction d'une *logique de la science* comme veut le faire Carnap ²³. Dans les mots de Carnap :

« The book itself makes an attempt to provide, in the form of an exact syntactical method, the necessary tools for working out the problems of the logic of science. This is done in the first place by the formulation of the syntax of two particularly important types of language which we shall call, respectively, 'Language I' and 'Language II'. » (Carnap, 1937, p.xiii-xiv)

21. Dans son livre sur Tarski (Patterson, 2012), c'est l'hypothèse que propose Patterson pour comprendre le concept de conséquence de Tarski. Ceci permet notamment d'expliquer la transition d'une version forte de l'intuitionnisme formaliste, des années 1920 au début des années 1930, vers l'article de 1936.

22. Rappelons que, pour Wittgenstein dans le *Tractatus Logico-Philosophicus*, il est impossible de référer à la structure de la logique ou du langage elle-même ; cela ne peut se *dire* mais se *montrer* (Wittgenstein, 1961).

23. Cette question dépasse les limites de ce mémoire ; pour un traitement détaillé, voir (Wagner, 2009), en particulier l'introduction de Pierre Wagner et la contribution de Philippe de Rouilhan.

Tarski critique Carnap en affirmant que la généralisation de la définition du Langage I est difficilement concevable, c'est-à-dire qu'elle est « essentially tied to the specific properties of the formalized language which was selected as object » (Tarski, 2002, p.182). Mais elle a au moins le mérite d'échapper au problème de l' ω -incomplétude, ce qui est un gain par rapport aux conceptions associant le concept de conséquence à la dérivation de théorèmes par l'application de règles. Par contre, comme le Langage II de Carnap a une plus grande capacité expressive, c'est surtout à la définition qui y est donnée que Tarski s'intéresse. Voici la reconstruction donnée par Tarski de cette définition²⁴ :

« We say that the sentence σ follows logically from the class of sentences Γ if and only if the class consisting of all sentences of the class Γ and of the negation of the sentence σ is contradictory. » (Tarski, 2002, p.182)

On peut remarquer que cette définition n'est pas tellement éloignée de celle donnée par Tarski²⁵, au moins dans la forme. On peut aussi remarquer que c'est sur la notion de contradiction que repose tout le poids de cette définition. C'est d'ailleurs ce qui est jugé problématique par Tarski. Carnap, dans le Langage II, n'utilise pas directement le vocabulaire, et plus précisément le vocabulaire *sémantique*, utilisé par Tarski dans sa définition (référence, satisfaction, vérité, etc.). Ces concepts se trouvent toutefois à l'état implicite dans ses travaux par le biais des notions d'analyticité, de contradiction et d'évaluation, qui tiennent lieu respectivement des notions sémantiques de vérité logique, de fausseté logique et de satisfaction. Ceci est dû, rappelons-le, au fait des préjugés défavorables de Carnap à l'endroit des notions sémantiques, mais aussi au fait qu'il ne reconnaît pas qu'un métalangage *scientifique* puisse contenir à la fois les termes du langage objet et ses propres termes syntaxiques. En d'autres mots, Tarski considère le travail de Carnap comme un pas vers la sémantique, car celui-ci recourt à des notions sémantiques déguisées en notions syntaxiques. Toutefois, à la différence de Carnap (en 1934), Tarski (après 1933) dispose d'une définition sémantique du concept de vérité, à partir de laquelle il peut travailler et introduire, sémantiquement, le

24. Nous laissons ici de côté la question de savoir s'il s'agit d'une reconstruction fidèle ou non des idées de Carnap.

25. Voir l'introduction pour la citation de la définition de Tarski.

concept de conséquence dans une théorie déductive.

On trouve chez Tarski plusieurs notions apparentées mais différentes de la sémantique. De manière générale, il faut par contre reconnaître que la sémantique a à voir avec « the totality of considerations concerning those concepts which, roughly speaking, express certain connexions between the expressions of a language and the objects and states of affairs referred to by these expressions » (Tarski, 1983, p.401). Cependant, comme les prémisses et conclusions d'un argument sont des énoncés d'un langage particulier, on peut se demander comment la conséquence logique, si elle doit être comprise sémantiquement, a à voir avec les objets (*objects*) et états de chose (*states of affairs*) du monde. La réponse de Tarski se trouve dans la distinction que nous avons présentée dans la section précédente entre niveaux de langage, c'est-à-dire entre langages desquels on fait usage et desquels on fait mention²⁶. Parmi les concepts sémantiques, certains, comme les notions de référence ou de dénotation, sont utilisés pour relier directement les objets et les expressions linguistiques. Par contre, d'autres concepts, comme les notions de vérité et de conséquence logique, sont seulement indirectement sémantiques, en ce qu'ils sont définis médiatement par les concepts directement sémantiques. Par exemple, le nom « Robert Brandom » réfère au philosophe américain — c'est-à-dire qu'il lie directement une expression du langage et un objet (une personne) du monde —, tandis que « Robert Brandom est barbu » — un énoncé de la langue française — est vrai parce qu'il existe des relations appropriées entre les objets du monde. En d'autres mots, l'objet (la personne) à laquelle on réfère a la propriété d'être barbu, rendant l'énoncé vrai. Comme nous le verrons dans la section suivante, la conséquence logique peut, de manière similaire, être comprise en termes de satisfaction.

Un problème majeur pour Carnap est qu'il utilise une version substitutionnelle de la quantification, plutôt qu'une version objectuelle comme celle de Tarski. Selon l'interprétation objectuelle de la quantification, des énoncés quantifiés sont vrais en vertu d'*objets* qui satisfont les énoncés quantifiés, c'est-à-dire qui comprennent des variables liées par des quantificateurs. Pour donner un exemple, un énoncé

26. Des remarques similaires peuvent être trouvées dans l'article de 1933 (Tarski, 1983, p.252).

comme « Il existe un x noir » ou, symboliquement, « $\exists x N(x)$ », est vrai s'il existe un objet ayant la propriété d'être noir, comme un corbeau. Selon l'interprétation substitutionnelle de la quantification, les énoncés quantifiés sont plutôt vrais en vertu des *noms* satisfaisant les énoncés ouverts. Dans ce cas, un énoncé n'est pas vrai si, pour reprendre notre exemple, un objet donné a la propriété d'être noir, mais plutôt s'il existe une instance substitutionnelle de l'énoncé ouvert « $N(x)$ » qui est vraie. Autrement dit, la vérité ne dépend pas ici de l'existence d'un objet noir mais plutôt d'un énoncé singulier « Na », où a est le nom permettant d'individuer par exemple un corbeau²⁷.

Comme nous l'avons déjà dit, Carnap recourt, dans ses définitions, à la notion d'évaluation, à la notion de contradiction et à sa contrepartie, la notion d'analyticité. Or ces relations ne peuvent exister qu'entre des expressions, c'est-à-dire au niveau linguistique. La notion de quantification associée aux définitions de Carnap ne peut donc tomber que sous l'interprétation substitutionnelle. En effet, la notion d'évaluation consiste simplement à remplacer des expressions par d'autres expressions lors de la généralisation à des ensembles d'énoncés. Tarski croit que les définitions de Carnap, en ce qu'elles font appel à une interprétation substitutionnelle de la quantification, ont donc le résultat suivant : « they make the denotation of the defined concepts dependent in an essential way on the richness of the language which is the object of consideration » (Tarski, 2002, p.193)²⁸. Ce qu'il faudrait plutôt est une notion de variable dont la portée est constituée d'ensembles d'objets, c'est-à-dire une conception objectuelle de la quantification. Il faut donc suppléer à la conception du concept de conséquence de Carnap une conception sémantique des variables et de la quantification. C'est précisément ce que fera Tarski dans l'article de 1936, en remplaçant la notion (substitutionnelle) d'évaluation de Carnap par la notion (objectuelle) de satisfaction. L'idée, pour la formuler simplement, est que la notion de satisfaction est une relation entre d'un

27. Ces deux interprétations, objectuelle et substitutionnelle, sont motivées notamment par des considérations ontologiques et liées au développement de théories de la signification. Pour une étude plus détaillée de cette distinction, voir par exemple (Hand, 2007).

28. Ceci est dû au fait que la notion d'analyticité ne peut être que relative à un langage particulier. Notons que c'est le même problème qu'identifie Etchemendy lorsqu'il compare les approches de Bolzano et Tarski (Etchemendy, 1990, Chapitre 3).

côté des expressions et de l'autre des entités non-linguistiques²⁹.

Un moteur important pour la définition sémantique donnée par Tarski semble donc être Carnap : en voyant les problèmes dans la conception de la *Syntaxe Logique du Langage*, Tarski a réalisé qu'il manquait une sémantique, ou plutôt qu'il fallait insister spécifiquement sur l'aspect sémantique pour capturer adéquatement le concept de conséquence. C'est alors que l'intérêt porté par Tarski à l'*expression de la pensée* par les systèmes déductifs, tel qu'on peut le trouver dans l'intuitionnisme formaliste, a commencé à s'estomper (Patterson, 2012, p.179). Grâce à Carnap, Tarski a vu tout ce que pouvaient accomplir les concepts sémantiques qu'il avait déjà commencé à élaborer dans ses travaux des années 1920 et 1930, mais qu'il ne leur faisait pas accomplir jusque là. Les concepts sémantiques serviraient l'objectif que Carnap voulait voir la syntaxe remplir, mais de manière beaucoup plus efficace.

1.3 L'approche sémantique de Tarski

1.3.1 Deux conditions (insuffisantes) d'adéquation

À la lumière des problèmes soulevés dans les définitions présentées jusqu'ici, Tarski propose d'offrir une définition alternative du concept de conséquence logique. C'est précisément l'objet de l'article de 1936. Bien qu'il n'insiste plus, à présent, sur l'expression de la pensée comme une condition devant être remplie par sa définition, il n'abandonne pas pour autant tous les aspects de l'intuitionnisme formaliste. Ainsi, il continue par exemple de mettre l'accent sur l'aspect *intuitif* que doit revêtir la définition technique du concept de conséquence. C'est d'ailleurs ceci qui constitue le point de départ de sa définition : « From the point of view of everyday intuitions it is clear that it cannot happen that all the sentences of the class Γ would be true but at the same time the sentence σ would be false » (Tarski, 2002, p.183). Le concept de conséquence visé n'est donc pas un concept quelconque,

29. Nous reviendrons plus en détail sur la notion de satisfaction dans la section suivante.

mais bien celui utilisé par les logiciens et les mathématiciens au sein des systèmes formels.

Ceci n'est toutefois qu'une des conditions à respecter. En effet, comme on s'intéresse à la relation de conséquence logique, qu'il s'agit de capturer à l'intérieur d'une théorie déductive, le fait de reconnaître comme valide ou non une inférence, c'est-à-dire de considérer la conclusion comme une conséquence logique des prémisses, ne devrait pas dépendre de faits empiriques concernant le monde extérieur. Ainsi, Tarski croit que la définition de la conséquence logique doit également pouvoir être considérée comme *formelle*.

Comme Tarski, conformément à ce que nous avons vu de l'intuitionnisme formaliste, veut que les théories deductives arrivent à des théorèmes contraignant les expressions de manière à ce qu'elles expriment des concepts intuitifs, et comme il veut des techniques métamathématiquement acceptables permettant d'introduire par des définitions les concepts en question dans les théories deductives, il doit assurer un traitement « matériellement adéquat » des concepts sémantiques et « formellement correct » des définitions. Ce sont ces conditions qu'il s'agit maintenant de remplir pour le concept de conséquence, comme c'était le cas pour le concept de vérité dans l'article de 1933, pour lequel il a recouru aux désormais célèbres *Conventions T*. Pour le concept de conséquence logique, ces deux conditions, à savoir que la définition soit « matériellement adéquate » et « formellement correcte », sont exprimés par Tarski dans la *Clause-F* :

« If in the sentences of the class Γ and in the sentence σ we replace the constant terms which are not general-logical terms correspondingly by arbitrary other constant terms (where we replace equiform constants everywhere by equiform constants) and in this way we obtain a new class of sentences Γ' and a new sentence σ , then the sentence σ' must be true if only all sentences of the class Γ' are true. » (Tarski, 2002, p.183-184)

Bien que ces deux conditions soient, selon Tarski, essentielles pour définir la conséquence logique, elles ne sont pas à elles seules suffisantes, puisqu'il pourrait arriver que le langage pour lequel on définit la conséquence logique ne possède pas suffisamment de constantes extra-logiques pour exprimer tous les objets possibles. Ce

serait le cas, par exemple, s'il y avait plus de propriétés d'objets dans le monde qu'il n'y a de prédicats dans un langage pour exprimer les propriétés en question (Tarski, 2002, p.185). Il existe par exemple beaucoup plus de longueurs d'ondes associées à des nuances de couleurs différentes pouvant être perçues que de concepts pouvant exprimer chacune de ces nuances. Pour mieux comprendre ce dont il s'agit, illustrons ce problème à l'aide d'un autre exemple.

Imaginons un langage arithmétique se limitant aux nombres naturels, c'est-à-dire ne pouvant exprimer qu'un sous-ensemble des nombres réels. La division d'un nombre naturel par un autre (par exemple la division de 8 par 4) peut avoir pour résultat un autre naturel (dans ce cas-ci 2), mais on peut aussi obtenir un nombre non naturel (par exemple si l'on divise 8 par 3)³⁰. Une solution, pour éviter de parler des nombres rationnels dans un langage n'ayant pas suffisamment de ressources expressives pour exprimer par exemple les fractions (sous-ensemble différent des nombres réels), est de parler du *reste* d'une division ; ainsi, lorsqu'on divise 8 par 3, le reste est égal à 2³¹. Mais il s'agit d'une solution bien locale et limitée au problème ; plutôt que de parler de *reste*, on pourrait également enrichir notre langage (en y ajoutant les entiers relatifs, les nombres rationnels, les nombres complexes, etc.) pour capturer plus adéquatement les propriétés générales des ensembles de nombres. La division de 8 par 3 donne alors $2\frac{2}{3}$, ou encore $2,\bar{6}$.

L'idée de Tarski est similaire : la Clause-F pourrait être satisfaite dans des cas où un énoncé σ n'est pourtant pas une conséquence de l'ensemble d'énoncés Γ , parce que « this condition [Clause-F] can in fact be satisfied only as a result of an insufficient stock of constant terms in the language which our considerations concern » (Tarski, 2002, p.184). En d'autres termes, il se pourrait qu'un contre-exemple existe sans que notre langage, trop pauvre sur le plan des possibilités expressives, puisse l'exprimer. Loin d'être un problème banal, cette question indique que le concept de conséquence logique ne dépend pas de la richesse expressive des langages, puisque cette richesse semble être un fait contingent.

30. On dit que l'ensemble des nombres naturels n'est pas *fermé* ou *clos* sous la division.

31. C'est-à-dire qu'il y a $(2 \times 3) + 2$ dans 8. 2 étant inférieur à 3 et ne pouvant pas être divisé simplement par celui-ci, la division ne peut pas résulter en un entier.

1.3.2 Modèles et satisfaction

Pour remédier à cette lacune du manque de ressources expressives, Tarski recourt à la notion de satisfaction, que nous avons évoquée plus haut. C'est précisément en ceci que consiste l'originalité historique et conceptuelle de la définition tarskienne. La notion de satisfaction est en fait une relation entre des énoncés ouverts, qu'on appelle des fonctions propositionnelles, et des séquences ou ensembles d'objets; elle est donc une notion sémantique au sens contemporain du terme. On comprendra mieux ceci à l'aide d'exemples.

L'énoncé « Jules César a conquis la Bretagne » est un énoncé composé de deux noms (« Jules César » ; « Bretagne ») et d'un verbe conjugué (« conquérir »)³². On peut ainsi le comprendre comme une fonction à deux places d'argument, mettant en relation deux parties distinctes. Dans ce cas-ci, les choses mises en relation par l'énoncé sont une personne et un lieu, mais on pourrait tout aussi bien imaginer des énoncés mettant en relation des objets, des adjectifs, et ainsi de suite. Pour obtenir une *fonction propositionnelle* à partir de l'énoncé initial, il suffit d'y remplacer les noms de personne et de lieu utilisés initialement par des variables. On obtient par exemple « x a conquis la Bretagne » ou encore « x a conquis y », selon qu'on choisisse d'inclure plus ou moins de variables³³.

Lorsque les variables, c'est-à-dire les places d'argument vides de ces fonctions propositionnelles, sont saturées par des objets des catégories appropriées, l'on obtient des énoncés pouvant être vrais ou faux, selon ce que ces énoncés expriment. Ainsi, si l'on sature les variables x et y de la fonction propositionnelle « x a conquis y » respectivement par « Jules César » et « Bretagne », ce qui donne l'énoncé initial, on obtient un énoncé vrai, en vertu du fait que César a bel et bien conquis la Bretagne. Par contre, si l'on effectue la même opération, mais cette fois avec les objets « Jules César » et « Italie », on obtient un énoncé faux, puisque Jules César n'a jamais conquis l'Italie.

32. Nous laissons de côté ici les technicalités concernant par exemple la présence de déterminants, les temps de verbes, et ainsi de suite.

33. Dans le cas limite, on pourrait voir un énoncé comme une fonction propositionnelle n'ayant aucune place d'argument.

On peut généraliser cette idée pour toute fonction propositionnelle, peu importe le nombre d'arguments qu'elle contient, de la manière suivante. D'abord, on désigne par une expression de la forme $f(x_1, x_2, \dots, x_n)$ les fonctions propositionnelles, où f indique la relation entre les variables x_1, x_2, \dots, x_n ³⁴. On dit ensuite de cette fonction qu'elle est *satisfaite* par un n -tuple ordonné $\langle a_1, a_2, \dots, a_n, a_{n+1} \rangle$, c'est-à-dire par une séquence bien précise d'objets, lorsqu'elle est satisfaite par les n premiers membres de la séquence.

Dans son article, Tarski précise que la notion de satisfaction est toujours relativisée à un langage particulier³⁵, mais croit qu'il existe une méthode rendant possible la construction de définitions de ce concept pour une grande catégorie de langages formels. Néanmoins, il n'en fournit pas le détail pour la raison que « it would be impossible to sketch here the method mentioned even in its most general features » (Tarski, 2002, p.185)³⁶. On peut par contre trouver des remarques plus substantielles sur le concept de satisfaction dans les articles « The Concept of Truth in Formalized Languages » et « The establishment of Scientific Semantics » (Tarski, 1983).

Il semble possible d'affirmer que, comme c'est le cas pour tous les autres concepts sémantiques présentés par Tarski, la relativisation n'est pas un obstacle au caractère général de la notion. Pour reprendre l'exemple de Patterson (Patterson, 2012, p.59), bien qu'on parle des lois chinoises et des lois américaines et que l'on reconnaisse d'importantes différences entre celles-ci, c'est-à-dire que ces lois sont en un certain sens *relatives*, on considère les lois de différents États comme spécifiant des conditions particulières d'une *unique* notion de *légalité*³⁷. De même, bien que la notion de satisfaction doive être relativisée aux langages particuliers,

34. Ces fonctions peuvent être aussi complexes que souhaité : x a volé la pomme de y à l'endroit z au moment t pour la raison u puis l'a donnée à v à l'endroit w , et ainsi de suite.

35. Comme c'est le cas pour d'autres notions, une définition générale déboucherait sur des paradoxes sémantiques.

36. Il ne faut pas oublier que l'article de 1936 a été présenté initialement sous forme de conférence, avec toutes les contraintes y étant associées. Tarski travaille ainsi avec une notion de satisfaction plutôt informelle. Les développements de cette notion par les théoriciens des modèles contemporains sont quant à elles très techniques ; voir par exemple Manzano, 1999, p.47-48.

37. Il serait pour le moins étrange de dire qu'il existe autant de conceptions de la légalité qu'il existe de codes de lois différents.

il n'en demeure pas moins qu'elle ne désigne pas plusieurs notions distinctes.

La notion de modèle est l'une des notions pouvant être définies à partir de la notion de satisfaction (Tarski, 2002, p.185). Cette définition procède de la manière suivante, avec les étapes (1) et (2) pouvant être illustrées par les exemples que nous venons de donner :

1. On associe à chaque constante extra-logique de notre langage L une variable distincte ;
2. On remplace dans un ensemble d'énoncés de L les constantes extra-logiques par les variables associées, de manière uniforme. On obtient alors un ensemble de fonctions propositionnelles L' ³⁸ ;
3. Une séquence arbitraire d'objets *satisfaisant* toutes les fonctions propositionnelles de L' sera appelée un *modèle* de L .

Autrement dit, un modèle est une séquence pour laquelle tous les énoncés sont rendus vrais. Comme nous l'avons déjà mentionné, Tarski définit le concept de conséquence logique comme préservation de la vérité des prémisses à la conclusion, c'est-à-dire que les modèles des prémisses sont également des modèles de la conclusion : « the sentence σ follows logically from the sentences of the class Γ if and only if every model of the class Γ is also a model of the sentence σ » (Tarski, 2002, p.186).

1.3.3 La définition est-elle adéquate ?

1.3.3.1 Adéquation et lien avec la définition de Carnap

Après avoir donné cette définition, Tarski s'empresse de montrer que celle-ci est adéquate, en ce qu'elle capture les intuitions se manifestant dans son usage courant, mais aussi en ce qu'elle débouche sur les implications recherchées par son

38. Notons que c'est en raison de cette distinction entre signes logiques et extra-logiques qu'on peut dire que la relation ainsi définie est *formelle*. Nous reviendrons sur cette distinction plus loin lorsque nous tenterons de donner une appréciation de la définition tarskienne et lorsque nous présenterons les critiques d'Etchemendy à l'endroit de celle-ci.

introduction dans les théories déductives :

« [O]ne can prove on the basis of the definition accepted above that a sentence which follows logically from true sentences must itself be true ; further, that the relation of following logically is completely independent of the sense of the extra-logical constants occurring in the sentences among which this relation obtains. » (Tarski, 2002, p.186-187)

On peut donc penser que Tarski juge sa définition, d'une part, « matériellement adéquate », puisqu'elle assure la préservation de la vérité, et d'autre part, « formellement correcte », en ce qu'elle est menée dans le métalangage sans référer aux constantes extra-logiques. Ces affirmations sont toutefois problématiques, car Tarski n'en donne pas vraiment de justification. Autrement dit, Tarski ne fournit pas d'argument à proprement parler nous enjoignant à accepter ces implications de la définition supposées en garantir le caractère adéquat³⁹. On peut ainsi s'y objecter en demandant par exemple en quel sens la conséquence logique d'énoncés vrais *doit (must)* elle-même être vraie, sur la simple base de la définition tarskienne, puisqu'on pourrait penser qu'il y a seulement une coïncidence extensionnelle sans qu'il y ait de modalité en jeu⁴⁰. De la même manière, il n'est pas difficile de refuser que la conséquence logique soit *entièrement (completely)* indépendante des constantes extra-logiques, puisque notre reconnaissance de la relation de conséquence logique dépend elle-même de notre connaissance de la signification des termes logiques⁴¹.

Pour montrer que sa définition est correcte et adéquate, Tarski compare ensuite celle-ci à celle de Carnap. En particulier, il vise à montrer que, moyennant quelques efforts de traduction ou d'interprétation des termes de l'une dans l'autre, les deux définitions sont en fait équivalentes. Pour ce faire, Tarski montre qu'on peut comprendre les énoncés contradictoires de Carnap comme des énoncés qui n'ont aucun modèle, et que les énoncés analytiques sont des énoncés satisfaits par n'importe

39. Etchemendy propose une reconstruction du raisonnement de Tarski, que nous étudierons dans le prochain chapitre.

40. Etchemendy (Etchemendy, 1983; 1990), comme nous le verrons au prochain chapitre, propose une critique qui va dans cette direction.

41. C'est le chemin que semble prendre Gila Sher (Sher, 1991; 1996)

quelle séquence (Tarski, 2002, p.187). Par contre, comme nous l'avons vu, il semble que cette équivalence ne soit pas si directe, surtout puisqu'elle suppose un déplacement du traitement substitutionnel des variables et de la quantification vers un traitement objectuel.

1.3.3.2 Démarcation des termes logiques et extra-logiques

Bien que la définition de Tarski soit sur plusieurs plans plus satisfaisante que celle de Carnap, l'appel à la notion de modèle ne spécifie pas ce qui relève proprement du domaine de la *logique* et ce qui compte plutôt comme *extra-logique*. Autrement dit, cette approche ne permet pas de statuer sur la question de la démarcation. Bien au contraire, la capture de la notion de conséquence logique en termes de modèles semble exiger qu'un choix de termes considérés *logiques* soit effectué *préalablement* à la définition. L'article de 1936 se termine par des remarques importantes sur la portée du résultat, qui semble être minoré en regard notamment de l'importance de la question des constantes logiques. Sans réponse à cette question cruciale, Tarski croit qu'il n'a pas solutionné de manière définitive le problème de la définition de la conséquence logique. Puisque cette question, la question des constantes logiques ou de la démarcation, est conceptuellement liée à la définition de la conséquence logique, il convient, avant de conclure ce chapitre, de faire quelques remarques qui aideront à mieux comprendre les mérites et les limites de la définition tarskienne.

On peut regrouper en (au moins) trois groupes les approches ou positions disponibles concernant la question des constantes logiques⁴². Dans un premier groupe, très hétérogène, se trouvent les approches cherchant à déterminer un *critère* sur la base duquel démarquer les constantes logiques des constantes extra-logiques. On peut ici penser au critère technique d'invariance sous les permutations (Tarski, 1986; Sher, 1991), voulant que les constantes logiques soient celles qui sont insensibles aux transformations du domaine d'objets sur lui-même⁴³, ou encore aux

42. Notons que les groupements présentés ici ne se veulent ni mutuellement exclusifs, ni collectivement exhaustifs.

43. Catarina Dutilh Novaes (Dutilh Novaes, 2011; 2014) fait une bonne présentation critique

caractérisations inférentielles des constantes logiques en termes de règles d'introduction et d'élimination (Prawitz, 1985; Dummett, 1991). Dans le premier cas, l'idée est que les constantes logiques sont *formelles*, par opposition à *matérielles*, en ce sens qu'elles sont indifférentes aux contenus particuliers; l'idée de constante logique vient avec une complète « indifference to individual objects » (Dutilh Novaes, 2011). Dans le deuxième cas, ce qui distingue les constantes logiques est le fait que celles-ci peuvent être caractérisées, en respectant certaines contraintes, par des règles d'inférence.

Deuxièmement, on peut vouloir déterminer ce que sont les constantes logiques à partir de considérations *pragmatiques*. C'est l'approche que semblent proposer des auteurs comme William Hanson (Hanson, 1997), Stewart Shapiro (Shapiro, 2014) et, de manière encore plus saillante, Robert Brandom (Brandom, 1994). Pour ce dernier, par exemple, il faut expliquer la signification des constantes logiques en termes de *rôle inférentiel* dans des contextes normatifs. On doit ainsi déterminer ce qu'une personne *fait* lorsqu'elle *utilise* les constantes logiques, ce qui permet en retour de déterminer ce que ces dernières signifient dans le cadre de pratiques ou d'activités inférentielles⁴⁴. Ces approches pragmatiques ont, contrairement aux approches par critères, tendance à être *minimalistes*, en ce sens qu'elles n'ont pas à inclure des constantes logiques simplement parce que certaines d'entre elles tombent sous le critère privilégié, et *holistes*, en ce sens qu'elles s'intéressent au rôle de la logique de manière globale (MacFarlane, 2015).

Enfin, plutôt que de rechercher à déterminer par des critères spécifiques ou des considérations pragmatiques quelles sont les constantes logiques, on peut refuser que la logique porte principalement sur les constantes logiques. Cette position est adoptée par Etchemendy (Etchemendy, 1990, Chapitre 9) et est intimement liée à sa critique de la définition de la conséquence logique de Tarski. Nous n'entrons donc pas dans la présente section dans les détails de cette approche, mais l'étudierons plutôt dans le chapitre suivant, où nous verrons qu'Etchemendy est très sceptique face à l'idée même de constante logique.

de cette position.

44. L'approche de Shapiro est très différente de celle de Brandom et est plutôt axée sur la nature et le rôle de la logique.

Comme Tarski a proposé, dans un article plus tardif (Tarski, 1986), un critère d'invariance sous les permutations permettant de distinguer les constantes logiques et extra-logiques, on pourrait penser que l'article de 1936 fait lui aussi appel à ce critère. Cependant, étudier l'article de 1936 avec les lunettes de l'article de 1966 me semble constituer une erreur d'interprétation. Gila Sher développe sa propre position en procédant à cette espèce d'*anachronisme*, bien qu'elle n'attribue pas cette position à Tarski lui-même. Mais, de manière plus importante, lire l'article de 1936 avec ce critère en tête semble surtout aller en sens contraire des remarques faites par Tarski dans l'article.

Dans l'article de 1936, Tarski souligne que la démarcation n'est pas arbitraire (Tarski, 2002, p.188). En effet, si on exclut des signes comme l'implication ou les quantificateurs, qui comptent parmi les signes par excellence que les logiciens et les mathématiciens considèrent comme des constantes logiques, on arrive à des conséquences qui contredisent les intuitions. Autrement dit, considérer trop peu de constantes comme des constantes logiques rend impossible une définition conforme au concept intuitif de conséquence. Par contre, Tarski reconnaît ne pas pouvoir lui-même fournir un critère précis de démarcation (Tarski, 2002, p.188) ; il n'adopte officiellement aucune des positions du premier groupement que nous avons identifié. De surcroît, Tarski reconnaît que, dans le cas limite, pour un langage strictement formel où tous les signes seraient considérés comme des signes *logiques*, la notion de conséquence logique coïnciderait avec la notion de conséquence matérielle (Tarski, 2002, p.188).

Malgré ce constat d'échec, c'est-à-dire malgré qu'il ne dispose d'aucun critère de démarcation précis, Tarski demeure optimiste sur la possibilité d'élaborer des moyens permettant de distinguer les constantes logiques et extra-logiques. D'un autre côté, en l'absence d'un tel critère, il semble normal de devoir considérer seulement *relativement* le concept de conséquence, puisqu'on ne dispose d'aucune ligne démarquant de manière absolue les constantes logiques et extra-logiques. Or cette démarcation, comme nous l'avons vu, est préalable et essentielle à la définition du concept de conséquence. Ceci expliquerait pourquoi on peut voir une fluctuation plus ou moins grande dans l'utilisation du concept de conséquence

logique par les logiciens et mathématiciens (Tarski, 2002, p.190). Il faut toutefois souligner, comme le fait Shapiro (Shapiro, 2011, p.530), que l'on peut être *tolérant* et accepter une forme de relativisation sans *tout* accepter ; on peut demeurer à la fois ouvert et accepter (ou à tout le moins *considérer*) différentes démarcations, mais ferme et refuser que *n'importe quelle* démarcation soit bonne.

Le moindre qu'on puisse dire, c'est que Tarski, dans son article de 1936 présente une position sur les constantes logiques très (trop ?) prudente. Faute de disposer d'un critère ou d'une manière de trancher la question, il se replie dans une forme de *relativisme* sans se commettre explicitement à l'une ou l'autre des options disponibles sur le marché des idées (invariance, inférentialisme, etc.). Ce relativisme de Tarski est d'ailleurs l'un des angles d'attaques d'Etchemendy à son endroit. Comme j'essaierai de le démontrer dans le troisième chapitre, Etchemendy semble toutefois procéder à une sorte de distorsion du propos de Tarski sur cette question. Il m'apparaît plus juste de ranger l'article de 1936 dans le second groupement, sur les considérations pragmatiques, bien que l'article en question soit loin de faire la présentation d'un agenda défini comme celui de Brandom. En l'absence d'un critère de démarcation précis ou en la présence d'une variété de critères distincts, la décision de procéder à telle ou telle démarcation peut reposer entre les mains du logicien ou du mathématicien, qui peut procéder à des découpages *plus ou moins arbitraires* selon l'utilisation qu'il compte faire de la logique. Pour le dire autrement, faute de disposer d'un critère de démarcation, le logicien jouit d'une certaine latitude qui repose sur ce qu'il *fait* avec la logique.

1.4 Conclusion

Dans ce chapitre, nous avons vu que la définition par Tarski du concept de conséquence logique s'inscrivait dans un projet de capture de notions intuitives à l'intérieur de théories déductives, ce pour quoi on peut considérer les travaux de Tarski comme s'inscrivant dans la tradition formaliste, et que ces objectifs agissaient à titre de contraintes sur les définitions. Nous avons également vu que la définition de Tarski s'inscrivait en réponse aux tentatives syntaxiques, où l'on risque

de rencontrer les problèmes d'incomplétude et de ne pas capturer l'ensemble des conséquences intuitives, et aux travaux de Carnap, où une place importante est accordée à des notions sémantiques (en germe), mais où l'utilisation explicite de ces notions n'est pas encore adéquatement assumée. Enfin, nous avons vu que la définition de Tarski du concept de conséquence visait à corriger les lacunes de ces conceptions par le recours à la notion de satisfaction. Ainsi, Tarski peut procéder à une définition qui est « matériellement adéquate » et « formellement correcte » du concept de conséquence en termes de modèles, en évitant les écueils de ses prédécesseurs et en remplissant les réquisits de l'intuitionnisme formaliste. Mais cette définition repose sur et présuppose donc une démarcation entre les constantes logiques et extra-logiques, sans laquelle la définition ne peut fonctionner.

Par contre, la définition de Tarski soulève à son tour de nombreuses questions en lien avec les principaux problèmes que nous avons identifiés pour la famille d'approches modèle-théorétiques⁴⁵. Nous préciserons et tenterons de donner réponse à un certain nombre de ces questions dans les prochains chapitres, où nous présenterons les principaux éléments de la critique d'Etchemendy et en donnerons une évaluation.

Parmi les questions auxquelles les travaux de Tarski donnent lieu, on peut par exemple se demander en quoi consiste une *interprétation*. S'agit-il d'une représentation du monde (ou d'une partie de celui-ci)? Une interprétation est-elle contrainte par les possibilités du monde, c'est-à-dire les différentes manières que pourrait avoir le monde de se réaliser? Nous verrons que la critique d'Etchemendy repose en grande partie sur la distinction entre sémantique représentationnelle et sémantique interprétationnelle (§2.1.1), mais que cette distinction est incomplète et n'épuise pas les points de vue sur la sémantique (§3.1.1).

On pourrait également demander si la définition tarskienne identifie les *bonnes* caractéristiques du concept de conséquence. Par exemple, la conséquence logique

45. Rappelons que ces approches doivent expliquer pourquoi et justifier comment la nécessité de la conséquence peut-être réduite en termes non modaux, spécifier ce que sont les interprétations et indiquer la manière de quantifier sur les domaines d'interprétation, puis indiquer comment sélectionner le vocabulaire à faire varier d'une interprétation à l'autre.

doit-elle se caractériser par la préservation de vérité, des prémisses à la conclusion ? Est-elle forcément une relation formelle entre les prémisses et la conclusion ? La conséquence logique est-elle une relation que l'on peut connaître *a priori*, indépendamment de toute considération empirique ? Nous répondrons en partie à ces questions, mais nous nous concentrerons surtout sur le fait de savoir si Tarski, dans sa définition, parvient à capturer adéquatement la nécessité associée au concept de conséquence, c'est-à-dire le fait que la conclusion découle *nécessairement* des prémisses. Nous verrons qu'Etchemendy donne une réponse négative à cette question (§2.2), mais que son argument repose en quelque sorte sur une distorsion du propos de Tarski (§3.1.3) et qu'il est possible d'y répondre (§3.2).

Enfin, on pourrait demander quel est exactement le concept *intuitif* de conséquence qu'il s'agit de capturer, et quel rôle épistémologique joue la notion d'intuition dans la justification de Tarski. On voit bien que l'appel aux intuitions peut être jugé problématique, dans le cas notamment des ω -règles. Il n'est en effet pas évident de voir en quoi on passe toujours du vrai au vrai dans le cas des ω -inférences, puisqu'on pourrait refuser que la conclusion *découle* des prémisses⁴⁶.

46. Faute d'espace, nous n'exploiterons pas ce filon en détail dans le mémoire, mais il s'agit d'une autre voie permettant d'éclairer la définition de Tarski et de donner suite aux critiques d'Etchemendy.

CHAPITRE II

LA CRITIQUE D'ETCHEMENDY

Dans le chapitre précédent, nous avons conclu sur un ensemble de questions auxquelles menait l'article de 1936 de Tarski. Dans ce chapitre, nous tenterons de fournir des réponses à certaines de ces questions, en particulier à celles concernant la nature de la notion d'interprétation et la nécessité associée au concept de conséquence, à partir de la critique formulée par Etchemendy à l'endroit de Tarski dans son livre de 1990. Nous présenterons deux problèmes de nature conceptuelle identifiés par Etchemendy, à savoir la confusion entre sémantiques représentationnelle et interprétationnelle (§2.1) et le « sophisme de Tarski » (§2.2). Puis, nous présenterons un problème de nature extensionnelle, à savoir la *sur-génération* de la définition tarskienne (§2.3), avant de conclure sur quelques implications de cette critique (§2.4).

2.1 Qu'est-ce que la sémantique ?

Dans la présentation et l'évaluation qu'il fait des idées de Tarski, Etchemendy traite le concept de vérité logique comme un cas limite du concept de conséquence logique. Autrement dit, les vérités logiques sont considérées comme les conséquences logiques d'un ensemble de prémisses *vide*, ou encore comme les conséquences logiques de *n'importe quel* ensemble de prémisses¹. Ainsi, en dé-

1. De l'hypothèse p , on peut par exemple tirer $p \vee q$, mais on peut aussi tirer $p \supset p$, qui est une loi logique ; dans ce cas, la conclusion ne dépend pas de la nature de l'hypothèse, qui peut

pit du fait que le concept de conséquence soit plus général mais aussi techniquement plus important et plus intéressant que le concept de vérité logique, la plupart des arguments d'Etchemendy ciblent, par souci de simplicité et pour faciliter l'exposé (Etchemendy, 1990, p.11), cette dernière notion et non le concept de conséquence².

La spécificité du projet de Tarski, selon Etchemendy, peut être saisie à partir de la distinction entre les conceptions *interprétationnelle* et *représentationnelle* de la sémantique. Pour bien comprendre cette distinction, il convient d'abord de présenter deux différents points de vue qu'il est possible d'adopter concernant le concept de vérité lui-même. Selon Etchemendy, à la suite par exemple de Donald Davidson (Davidson, 1973), on peut considérer la notion de vérité d'une part comme une *propriété* des énoncés et d'autre part comme une *relation*. Dans le premier cas, on considère que la vérité s'applique aux énoncés à la manière d'un prédicat monadique, comme une fonction avec un seul argument. On aura par exemple le schéma suivant : l'énoncé p est vrai. Au contraire, si l'on considère la vérité comme une relation, c'est seulement relativement à *autre chose* qu'on peut spécifier le concept de vérité. On aura ainsi une relation de la forme « l'énoncé p est vrai dans θ », comme « l'énoncé p est vrai dans un modèle » ou encore « p est vrai dans un monde possible ». Un énoncé est alors vrai *relativement*, et non pas vrai *tout court* ou *absolument*.

Présentés ainsi, les points de vue prédicatif et relationnel ne semblent pas être indépendants l'un de l'autre. Il semble en effet plutôt naturel de considérer que le point de vue relationnel est une *généralisation* du point de vue prédicatif (monadique). Comme le souligne Etchemendy, « " x is true in y " stands to " x is true" as " x is a brother of y " stands to " x is Fred's brother" » (Etchemendy, 1990, p.14). Ainsi, pour obtenir le concept de *fraternité-d'une-personne-en-particulier*, qui est monadique, on n'aurait qu'à spécifier par rapport à *qui* ou à *quelle per-*

être *déchargée*. Autrement dit, on peut tirer $p \supset p$ d'hypothèses quelconques, mais pas $p \vee q$ puisqu'il faut assumer la vérité d'au moins une prémisse pour tirer cette conclusion.

2. Ceci est pour le moins surprenant étant donné le titre de son ouvrage : *The Concept of Logical Consequence*. Etchemendy reconnaît toutefois que Tarski ne vise pas à *réduire* le concept de conséquence logique au concept de vérité logique (Etchemendy, 1990, p.47).

sonne en particulier, par exemple *Fred, Tom, Carl*, et ainsi de suite, on entend le concept relationnel de fraternité. Autrement dit, on peut voir la relation de fraternité entre deux personnes comme la propriété d'être-le-frère-de-quelqu'un (Etchemendy, 1990, p.15). De même, pour obtenir le (ou un) concept de vérité monadique, on n'aurait qu'à spécifier à partir du concept relationnel ce par rapport à *quoi* on définit le concept, par exemple un modèle, un monde possible, etc.

Là où cette distinction devient intéressante, c'est dans l'importance qu'elle revêt lorsqu'il s'agit de spécifier ce par rapport à quoi l'on relativise le concept de vérité. C'est cette technique qui est à l'oeuvre dans la technique consistant à construire des tables de vérité³. Pour construire de telles tables, l'on doit d'abord établir une colonne de référence. Pour un énoncé p quelconque, on peut dresser la colonne de référence ainsi :

p
vrai
faux

Les tables de vérité s'allongeront à mesure qu'on ajoutera des énoncés, puisqu'il faudra tenir compte de l'ensemble des combinaisons possibles entre les valeurs de vérité des différents énoncés. Par exemple, pour quatre énoncés p, q, r et s quelconques, la table de vérité comptera 16 rangées : une rangée où les quatre énoncés sont vrais, une rangée où les quatre énoncés sont faux, quatre rangées où trois énoncés sont vrais et où un énoncé est faux, etc. :

3. Dans ce cas, c'est par rapport aux rangées de la table de vérité que le concept est relativisé.

p	q	r	s
vrai	vrai	vrai	vrai
vrai	vrai	vrai	faux
vrai	vrai	faux	vrai
vrai	vrai	faux	faux
vrai	faux	vrai	vrai
vrai	faux	vrai	faux
vrai	faux	faux	vrai
vrai	faux	faux	faux
faux	vrai	vrai	vrai
faux	vrai	vrai	faux
faux	vrai	faux	vrai
faux	vrai	faux	faux
faux	faux	vrai	vrai
faux	faux	vrai	faux
faux	faux	faux	vrai
faux	faux	faux	faux

Puis, à mesure que la complexité croît, il faut maîtriser l'application de règles récursives propres à chacun des connecteurs extensionnels, c'est-à-dire que pour des énoncés complexes obtenus à partir des énoncés simples, on utilise les colonnes de référence auxquelles on applique les transformations sur les valeurs de vérité, conformément aux différentes règles des connecteurs. Pour des énoncés p et q quelconques, la négation (\neg) et la conjonction ($\&$), on a par exemple :

p	q	$p\&\neg q$
vrai	vrai	faux
vrai	faux	vrai
faux	vrai	faux
faux	faux	faux

Comme le remarque Etchemendy, ces tables n'indiquent toutefois pas *laquelle* (ou

lesquelles) des rangées de la table de vérité sont *correctes*, c'est-à-dire dans quelle situation on se retrouve. Par exemple, si p et q étaient respectivement les énoncés « Robert Brandom est barbu » et « Martha Nussbaum est barbue », ce n'est pas simplement en regardant la table de vérité qu'on pourrait déterminer ce qui est le cas ; les tables de vérité indiquent seulement les rangées dans lesquelles les énoncés complexes comme « $p \& \neg q$ », obtenus par l'application de règles récursives sur les énoncés simples, sont vrais ou faux. Pour déterminer la ou les rangées qui sont adéquates, on peut, selon Etchemendy, choisir entre deux alternatives. D'une part, une rangée pourrait être considérée comme correcte en vertu de l'état du monde, ou, d'autre part, en vertu du langage pour lequel on construit une table de vérité. Dans le premier cas, un énoncé de la langue française est vrai (« Robert Brandom est barbu ») ou faux (« Martha Nussbaum est barbue ») parce que le monde est tel qu'il est⁴. Dans le second cas, la vérité ou la fausseté des énoncés est non plus due au monde, mais plutôt à la *signification* des énoncés et des expressions du langage. Ainsi, c'est parce que les expressions « Robert Brandom », « être », « barbu », etc., signifient ce qu'elles signifient en français qu'on peut obtenir des énoncés vrais ou faux, dans la langue française, à partir de ceux-ci.

Dans les deux cas de notre exemple, c'est la deuxième rangée de notre table de vérité qui est adéquate, mais pour des raisons différentes. Avec l'approche centrée sur le monde, on obtient une explication relativisant la notion de vérité au monde ; on obtient ainsi des formulations comme « $p \& \neg q$ est vrai dans le monde actuel » ou plus généralement « p est vrai dans W », où W tient lieu de variable pour les mondes possibles. Dans le second cas, centré sur la langue française, on obtient des formulations comme « $p \& \neg q$ est vrai dans la langue française » ou plus généralement « p est vrai dans L », où L tient lieu de variable pour les langages possibles. Ces deux alternatives, drastiquement opposées selon Etchemendy, donnent lieu à des explications contrefactuelles différentes. Pour l'alternative reposant sur le monde, on a par exemple :

Si le monde avait été différent (par exemple si les femmes avaient été barbues et les hommes imberbes), alors l'énoncé « Robert Brandom est barbu » aurait été faux, et ce, en dépit de la signification qu'ont

4. Aux dernières nouvelles, Robert Brandom porte la barbe, mais pas Martha Nussbaum.

les expressions « Robert Brandom », « barbu », etc., en français.

Pour l'alternative reposant au contraire sur la signification des expressions dans une langue, on a l'explication contrefactuelle suivante :

Si la langue française avait été différente (par exemple si l'expression « Robert Brandom » avait désigné Martha Nussbaum) alors l'énoncé « Robert Brandom est barbu » aurait été faux, et ce, en dépit du fait que Robert Brandom a une barbe.

Dans un cas, on commence avec une théorie de la signification pour les énoncés du langage ; des rangées dans les tables de vérité sont alors correctes en vertu du monde, et un changement de rangée équivaut à un changement dans l'état du monde. Dans l'autre cas, les faits *extra-linguistiques*, comme le fait que Robert Brandom a une barbe ou que Martha Nussbaum n'en a pas, sont fixés ; des rangées dans les tables de vérité sont alors correctes en vertu du langage, et un changement de rangée dans les tables de vérité équivaut à un changement dans la signification des énoncés (Etchemendy, 1990, p.18-19). Ces deux approches, reposant sur le monde et sur la signification linguistique des énoncés, donnent respectivement lieu, on l'aura deviné, aux approches représentationnelle et interprétationnelle de la sémantique.

2.1.1 L'approche représentationnelle et l'approche interprétationnelle

2.1.1.1 L'approche représentationnelle

L'approche *représentationnelle* de la sémantique vise à expliquer comment le monde exerce son influence sur les valeurs de vérité des énoncés d'un langage particulier, envisagé comme ne subissant aucune modification dans le temps (Etchemendy, 1990, p.20). Mais les tables de vérité ne peuvent pas encore être appelées des structures *sémantiques* :

« From this “representational” standpoint, our truth table gives us valuable information about *truth*, but certainly not about how truth would be affected by changes in [rows of the truth tables]. Rather,

it tells us how truth would be affected by changes in the world, by changes that are *represented* or *depicted* by changes in row. » (Etchemendy, 1990, p.20-21)

L'approche représentationnelle doit ainsi indiquer comment la vérité des énoncés se trouve affectée par des changements ayant lieu dans le monde, alors que les tables de vérité peuvent être vues comme des simples taches d'encre n'ayant pas à *représenter* quoi que ce soit.

Selon Etchemendy, pour que la technique des tables de vérité devienne une véritable méthode de *représentation*, il faut procéder de la manière suivante. On commence d'abord par établir un ensemble de modèles servant à représenter toutes les configurations possibles du monde influençant les valeurs de vérités des énoncés du langage. Selon l'approche représentationnelle, les modèles M sont des structures formées d'un domaine de discours (non vide), ainsi que d'une fonction assignant une extension du domaine aux expressions extra-logiques de manière à assigner une valeur de vérité aux énoncés (McKeon, 2010, p.40). L'un des rôles du domaine de discours, en plus de déterminer la portée des quantificateurs⁵, est d'imposer une limite sur les valeurs sémantiques que peuvent prendre les expressions du langage. Quant à elles, les *fonctions* de nos modèles assignent par exemple des individus du domaine aux noms d'individus et des sous-ensembles ordonnés du domaine aux prédicats et aux relations de manière à lier les expressions linguistiques et le monde, dans les limites établies par le domaine.

Pour un langage (servant à représenter une partie très limitée du monde) ne comportant que deux noms d'individus, par exemple « Robert Brandom » et « Martha Nussbaum », et deux prédicats monadiques, par exemple les propriétés d'être barbu et d'être philosophe politique, on aura les quatre énoncés atomiques suivants : « Robert Brandom est barbu », « Robert Brandom est philosophe politique », « Martha Nussbaum est barbue » et « Martha Nussbaum est philosophe politique ». Puisque chacun de ces énoncés est susceptible d'être vrai ou faux, on a au total 16 combinaisons distinctes de valeurs de vérité. De manière générale, pour un langage bivalent comportant n énoncés atomiques, on aura 2^n rangées

5. Nous reviendrons sur cette question dans les sections suivantes.

dans notre table de vérité, et donc autant de modèles pour représenter l'ensemble des combinaisons possibles.

Sur la base de ces modèles, on fournit ensuite une définition récursive, c'est-à-dire par ordre de complexité croissante, de « p est vrai dans M », pour des énoncés et des modèles arbitraires. Cette relation sert à caractériser indirectement « p est vrai dans W » et doit donc assurer que les énoncés de notre langage soient vrais dans les modèles représentant des mondes où les énoncés sont effectivement vrais. Par exemple, si un modèle représente un monde où Robert Brandom porte la barbe, la définition devrait assurer que l'énoncé « Robert Brandom est barbu », si celui-ci signifie ce qu'il signifie en français, est *vrai* dans ce modèle, alors qu'un modèle représentant un monde où Robert Brandom n'a *pas* de barbe devrait rendre *faux* un tel énoncé. Autrement dit, l'énoncé p est vrai dans le modèle M seulement si le modèle assigne à cet énoncé la valeur « vrai » (Etchemendy, 1990, p.21).

Avec une telle approche de la sémantique, on a ce qu'il faut pour développer une théorie de la signification pour les langues naturelles. C'est par exemple ce qu'a voulu faire Donald Davidson, en s'inspirant et en prolongeant les travaux de Tarski (Davidson, 1967). Afin de pouvoir appliquer la sémantique telle que développée par Tarski aux langues naturelles (et non aux langages formalisés), Davidson doit faire l'hypothèse qu'on peut réduire la signification des énoncés à leurs *conditions de vérité*. Mais comme nous l'avons vu dans le chapitre précédent, et comme le remarque Etchemendy, ce n'est pas là l'objectif de Tarski ; Tarski ne veut pas donner une caractérisation et une justification de « p est vrai dans W », étant donné des langages considérés fixes, mais plutôt de « p est vrai dans L », en particulier pour les langages formalisés (Etchemendy, 1990, p.23). Selon Etchemendy, en supposant que notre objectif soit du côté représentationnel, la construction de notre théorie des modèles devrait toutefois satisfaire deux critères :

1. Pour juger notre théorie adéquate, l'ensemble des modèles devrait contenir *toutes et seulement* les configurations intuitivement possibles du monde. Si par exemple on excluait des rangées de nos tables de vérité, notre représentation serait incomplète (Etchemendy, 1990, p.23).
2. Comme notre définition récursive repose sur des intuitions concernant l'in-

fluence du monde sur les valeurs de vérité, notre théorie ne peut être adéquate qu'à la condition suivante : un énoncé est vrai dans un modèle si et seulement si l'énoncé est vrai dans un modèle M_i s'avérant adéquat dans la représentation d'un monde W_i . (Etchemendy, 1990, p.24).

Ces critères, pris conjointement, permettent d'arriver aux affirmations suivantes. Premièrement, les énoncés qui sont vrais dans *tous* les modèles d'une sémantique représentationnelle sont des énoncés représentant des vérités *nécessaires*⁶. Si ce n'est pas le cas, c'est que nous avons ou bien négligé des configurations possibles, ou bien que nous avons déclaré un énoncé vrai dans un modèle alors qu'il devrait en fait être déclaré faux. Deuxièmement, lorsqu'un énoncé est vrai dans tous les modèles où un autre énoncé est vrai, alors la vérité du premier énoncé doit être une condition nécessaire de la vérité du second. On dit alors que le premier énoncé est une *conséquence* du second énoncé.

Les critères que nous venons d'énoncer permettent de voir pourquoi l'approche représentationnelle de la sémantique ne permet pas de caractériser adéquatement la notion de vérité logique, ni non plus le concept plus général de conséquence logique, comme Tarski souhaite le faire. En effet, des vérités qui sont nécessaires mais qui ne sont toutefois pas des vérités logiques, comme les vérités mathématiques non réductibles aux vérités logiques⁷, seront vraies dans tous les modèles de notre sémantique. Autrement dit, une sémantique représentationnelle ne permet pas de distinguer adéquatement différents aspects de certaines notions comme la nécessité. Un autre problème des sémantiques représentationnelles surgit lorsqu'on se demande si nos modèles représentent bien *toutes* et *seulement* les configurations possibles du monde. Le problème, plus précisément, est que la notion de configuration possible repose elle-même sur la notion de modèle. Autrement dit, la notion de configuration possible est présupposée et on tombe dans un cercle vicieux dès qu'il s'agit de déterminer si notre représentation du monde est exhaustive. Ainsi, l'approche représentationnelle de la sémantique présuppose les notions que Tarski souhaite élucider (comme la vérité logique) ; Etchemendy en conclut que Tarski ne

6. On voit ainsi en quoi l'aspect modal se trouve *réduit* en des termes non modaux, c'est-à-dire que la nécessité est associée à l'obtention d'énoncés vrais dans un nombre de modèles.

7. Voir par exemple le théorème d'incomplétude de Gödel (Gödel, 1967), ayant signé l'échec du logicisme.

peut donc pas approcher la sémantique de manière *représentationnelle*. Selon lui, l'objectif de la sémantique dans le projet tarskien est alors d'illuminer les règles de fonctionnement de notre langage, et non de fournir une analyse de la notion de vérité nécessaire (Etchemendy, 1990, p.25).

2.1.1.2 L'approche interprétationnelle

Ayant écarté l'approche représentationnelle du projet tarskien, Etchemendy présente ensuite l'approche de la sémantique qu'il attribue à Tarski, à savoir l'approche *interprétationnelle*; cette présentation se fait en caractérisant de manière plus précise la notion de satisfaction que nous avons présentée dans le chapitre précédent. Selon Etchemendy, l'approche interprétationnelle fait varier les significations (*meanings*) d'un ensemble d'expressions plutôt que les possibilités du monde. Autrement dit, on donne des interprétations différentes du langage pour lequel on veut établir une sémantique. Ainsi, pour cette approche, la distinction entre constantes logiques et extra-logiques revêt une importance cruciale.

En supposant qu'on dispose d'une manière de distinguer les constantes logiques et extra-logiques, il faut d'abord faire en sorte qu'à chaque énoncé S du langage corresponde une unique fonction propositionnelle S^* ; c'est ce qu'Etchemendy appelle une fonction propositionnelle distinguée (*distinguished sentential function*) (Etchemendy, 1990, p.52). Ceci est fait en assignant des variables spécifiques à chaque constante extra-logique du langage (Etchemendy, 1990, p.52). On assignera par exemple la variable « x_1 » à l'expression « Robert Brandom », la variable « x_2 » à l'expression « Martha Nussbaum », et la variable « y_1 » à l'expression désignant la propriété d'être barbu⁸. Des énoncés comme « Robert Brandom est barbu » et « Martha Nussbaum est barbue » deviendront ainsi les fonctions propositionnelles distinguées « x_1y_1 » et « x_2y_1 », où des variables différentes sont assignées aux expressions différentes de manière à pouvoir distinguer les énoncés différents les uns des autres. Ce recours à des variables spécifiques permet en

8. Plus spécifiquement, c'est aux *occurrences* de ces expressions qu'on assignera *une* variable spécifique.

quelque sorte de simplifier la « mécanique » du concept de satisfaction développé par Tarski ; les séquences nous intéressant seront limitées à celles assignant des objets du domaine de satisfaction approprié aux variables apparaissant dans nos fonctions propositionnelles, et seulement à celles-là. Etchemendy appelle d'ailleurs ces séquences des *séquences limitées* (Etchemendy, 1990, p.53). Autrement dit, le domaine des séquences sera limité par les variables spécifiques apparaissant dans nos fonctions propositionnelles.

Puisqu'il y a une correspondance entre les constantes extra-logiques et les variables ainsi qu'entre les énoncés et les fonctions propositionnelles, c'est-à-dire puisqu'il y a exactement autant de variables qu'il y a d'expressions extra-logiques et autant de fonctions propositionnelles que d'énoncés, il y a une manière, selon Etchemendy, de construire notre sémantique sans passer par les variables et les fonctions propositionnelles. Pour ce faire, on doit d'abord introduire des séquences dont le domaine sera constitué des constantes extra-logiques elles-mêmes, et non pas de variables par lesquelles nous aurons remplacé celles-ci. Pour le dire autrement, on évite ainsi le « détour » par les variables spécifiques (comme « x_1 », « y_2 », etc.) en considérant *directement* les expressions (comme « Robert Brandom », « Martha Nussbaum », etc.) comme pouvant faire l'objet d'une variation. Etchemendy appelle *séquences directes* (*d-sequences*) celles dont le domaine est constitué des constantes extra-logiques ; elles sont en fait des fonctions qui assignent des objets du domaine de satisfaction approprié aux expressions à faire varier (Etchemendy, 1990, p.53). On ajuste ensuite notre notion de satisfaction pour qu'elle puisse tenir compte de ces séquences directes : une séquence directe satisfait *directement* (*d-satisfies*) un énoncé S si la séquence limitée correspondante satisfait la fonction propositionnelle S^* (Etchemendy, 1990, p.54).

Illustrons ceci à l'aide d'un exemple. Imaginons une séquence directe, appelons-la J , qui assigne Martha Nussbaum à l'expression « Robert Brandom » et la propriété d'être philosophe politique à l'expression « est barbu ». La séquence limitée correspondante, appelons-la J^* , satisfait alors la fonction propositionnelle « x_1y_1 » si et seulement si Martha Nussbaum est philosophe politique. Cette fonction propositionnelle étant associée à l'énoncé « Robert Brandom est barbu », on obtient

alors, en vertu de la définition donnée de la satisfaction directe, l'équivalence suivante : la séquence directe J satisfait directement « Robert Brandom est barbu » si et seulement si la séquence limitée J^* satisfait « x_1y_1 ». Corollairement, on a donc que la séquence directe J satisfait directement « Robert Brandom est barbu » si et seulement si Martha Nussbaum est philosophe politique.

En procédant de cette manière avec la notion de satisfaction, on assure un traitement contrefactuel convenable et on évite ainsi que des séquences satisfassent *trivialement* des énoncés ne comptant aucune variable simplement parce que ceux-ci sont vrais (Etchemendy, 1990, p.54). On ne considère plus qu'un énoncé comme « Robert Brandom est barbu », en vertu du fait qu'il ne compte aucune variable et qu'il est vrai en français, est satisfait par *toutes* les séquences. De même, on évite que des séquences échouent trivialement à satisfaire des énoncés ne comptant aucune variable simplement en vertu de la fausseté de ceux-ci. La satisfaction directe est pour ainsi dire plus stricte que la satisfaction tout court. Elle permet donc de déterminer *si* 1) la valeur de vérité d'un énoncé serait différente *si* 2) les expressions extra-logiques de ces énoncés étaient interprétées différemment, c'est-à-dire conformément aux assignations déterminées par la séquence directe. Ainsi, bien que dans le monde actuel l'énoncé « Robert Brandom est barbu » soit vrai, si l'expression « Robert Brandom » avait désigné une personne différente, par exemple Martha Nussbaum, alors cet énoncé aurait été vrai à condition que Martha Nussbaum soit barbue.

Avec ces notions en main, on peut maintenant affirmer qu'un énoncé est une *vérité logique* dans le cas où il est directement satisfait par toutes les séquences directes. Un argument est *logiquement valide*, c'est-à-dire que la conclusion est une *conséquence logique* des prémisses, si, et seulement si, il préserve directement la satisfaction par *toutes* les séquences directes. Selon Etchemendy, on obtient la version tarskienne de la sémantique interprétationnelle si l'on considère les séquences directes comme les *modèles* tarskiens et si l'on considère la satisfaction directe comme la vérité relative à ces séquences (Etchemendy, 1990, p.55), au sens où nous l'avons présentée précédemment.

On peut donc considérer les différentes séquences directes, ou les différents modèles, comme des fonctions fournissant des réinterprétations possibles des expressions de notre langage. Notre théorie des modèles, envisagée de manière interprétationnelle, permet ainsi de donner une caractérisation de « p est vrai dans L » pour un ensemble de langages. Toutes les réinterprétations ne sont cependant pas acceptées ; dans les mots d'Etchemendy, les réinterprétations qui nous intéressent sont celles qui sont « *semantically well-behaved* » (Etchemendy, 1990, p.55). Les réinterprétations admissibles sont celles qui respectent l'idée de catégorie sémantique, à savoir les réinterprétations respectant certaines conditions concernant les domaines de satisfaction. Ainsi, les séquences ne pourront par exemple substituer aux noms d'individus que des noms de membres (ou de membres potentiels) du domaine d'individus, aux prédicats que des propriétés faisant partie (ou faisant potentiellement partie) du domaine des prédicats, et ainsi de suite. On n'arrivera donc pas à des énoncés comme « Est barbu est barbu » à partir de l'énoncé « Robert Brandom est barbu », puisque les séquences permettant la réinterprétation de l'expression « Robert Brandom » sont contraintes par la catégorie à laquelle cette expression appartient⁹.

Les modèles de l'approche interprétationnelle ne visent pas à représenter les différentes configurations possibles du monde, comme c'était le cas pour l'approche représentationnelle. Plutôt, il faut maintenant comprendre ces modèles comme indiquant quelles réinterprétations des énoncés d'un langage L sont acceptables (*semantically well-behaved*) (Etchemendy, 1990, p.56). Selon la perspective représentationnelle, un énoncé comme « Robert Brandom est barbu ou Robert Brandom n'est pas barbu » serait vrai même si Robert Brandom était différent. Selon la perspective interprétationnelle, on ne s'intéresse pas à savoir ce qui arriverait à la vérité de l'énoncé si Robert Brandom se rasait la barbe ; ce qui nous intéresse plutôt est le fait de savoir que l'énoncé serait vrai même si l'énoncé « Robert Brandom est barbu » était interprété différemment conformément aux catégories sémantiques, par exemple comme signifiant « Martha Nussbaum est philosophe

9. Un poids important reposant sur cette notion de catégorie sémantique, on a donc ici un angle d'attaque à partir duquel critiquer l'approche interprétationnelle de la sémantique telle que présentée par Etchemendy. Nous verrons un problème de cette approche dans la section suivante.

politique ».

2.1.2 Les deux approches coïncident-elles ?

La distinction effectuée entre les approches représentationnelle et interprétationnelle de la sémantique permet à Etchemendy de soutenir que Tarski ne doit pas être vu comme un théoricien des modèles au sens contemporain du terme, mais aussi qu'il existe un défaut majeur à l'approche interprétationnelle. Pour défendre ces affirmations, Etchemendy montre que les sémantiques modèle-théorétiques interprétationnelles ne peuvent pas toutes être envisagées de manière représentationnelle, et, inversement, que les sémantiques représentationnelles ne peuvent pas toutes être envisagées de manière interprétationnelle.

D'abord, certaines théories respectent les exigences interprétationnelles sans respecter les exigences représentationnelles :

« [T]his becomes apparent when we consider different theories that emerge from different selections of [...] [logical constants]. With other choices of [logical constants] we encounter one of two problems : either the resulting class of models, when seen representationally, omits depictions of genuinely possible configurations of the world, or there is simply no way to view the class of models *as* representations. »
(Etchemendy, 1990, p.61)

Si l'on considère l'énoncé « Robert Brandom est barbu » comme une constante logique, il n'y a plus de modèle représentant la possibilité que Robert Brandom *ne soit pas* barbu et on fait ainsi face à un problème du premier type (Etchemendy, 1990, p.61). Autrement dit, on peut omettre des représentations en *incluant* par exemple un surplus d'expressions dans notre ensemble de constantes logiques. Bien que ceci ne pose pas en soi problème pour la perspective interprétationnelle, ceci enfreint le premier critère que nous avons présenté pour les sémantiques représentationnelles, à savoir que l'ensemble des représentations soit complet.

Pour illustrer le second type de problème, à savoir le fait qu'il soit impossible pour une interprétation particulière d'être vue comme une représentation du monde, on

peut par exemple considérer certaines expressions, typiquement admises comme des constantes logiques, comme faisant partie de notre ensemble d'expressions variables (Etchemendy, 1990, p.62). Autrement dit, l'on peut *exclure* à outrance certaines expressions de notre ensemble de constantes logiques. Ainsi, si l'on considère que le connecteur « ou » peut faire l'objet d'une variation et se voir assigner les valeurs du connecteur « et », on obtiendra un modèle qui n'a aucun équivalent dans une sémantique représentationnelle (et non pas un modèle représentant une configuration possible mais *inusitée* du monde).

Les théories interprétationnelles ne peuvent pas toutes être envisagées selon la perspective représentationnelle, mais l'inverse est aussi vrai, c'est-à-dire qu'on ne peut pas toujours voir de manière interprétationnelle les sémantiques représentationnelles. Un premier problème est que bien que deux modèles puissent assigner des séquences d'objets différents (personnes, propriétés, etc.) aux expressions du langage, des modèles *différents* peuvent très bien ne pas représenter des configurations du monde elles-mêmes différentes les unes des autres (Etchemendy, 1990, p.63). Par exemple, les expressions « Boris Vian » et « Vernon Sullivan » dénotent le même individu, puisque cette dernière est un pseudonyme utilisé par l'auteur Boris Vian dans les années 1940 et 1950. Ces expressions, si elles conservent leur signification, comme c'est le cas lorsque l'on vise à établir une sémantique représentationnelle, donnent donc lieu à des modèles isomorphes, c'est-à-dire que ces expressions sont intersubstituables (dans les contextes extensionnels). On peut ainsi considérer qu'une sémantique omettant toutes les occurrences de l'expression « Vernon Sullivan » est adéquate, en ce sens qu'une telle théorie n'omet aucune configuration possible du monde. Toutefois, cette sémantique n'est pas adéquate du point de vue interprétationnel puisqu'elle restreint de manière arbitraire les modèles en restreignant les réinterprétations admissibles.

Un autre problème empêchant l'approche représentationnelle d'être vue de manière interprétationnelle concerne des restrictions pouvant être appliquées sur les représentations. Il est en effet acceptable d'exclure certains modèles lorsqu'on souhaite donner une *représentation* du monde, sans pour autant que ces exclusions soient acceptables d'un point de vue interprétationnel. Pour illustrer ceci, imagi-

nous un langage simple composé des énoncés suivants : « L'herbe est blanche », « L'herbe est rouge » et « L'herbe est verte »¹⁰. Comme ces énoncés sont, pris individuellement, susceptibles d'être vrais ou faux, une table de vérité exhaustive compterait 8 rangées :

L'herbe est blanche	L'herbe est rouge	L'herbe est verte
Vrai	Vrai	Vrai
Vrai	Vrai	Faux
Vrai	Faux	Vrai
Vrai	Faux	Faux
Faux	Vrai	Vrai
Faux	Vrai	Faux
Faux	Faux	Vrai
Faux	Faux	Faux

Par contre, comme les attributions de couleur sont mutuellement exclusives, c'est-à-dire que le fait d'être coloré d'une certaine couleur exclut la possibilité d'être coloré d'une autre couleur, on n'a pas à inclure *toutes* ces représentations dans notre théorie. Autrement dit, comme on vise à donner une représentation du monde, on peut tenir compte de la possibilité *empirique* plutôt que de la possibilité *logique*. Ainsi, la table de vérité suivante, excluant les rangées où plusieurs attributions de couleurs sont vraies simultanément, serait tout à fait acceptable comme base d'une sémantique représentationnelle pour notre langage simple :

L'herbe est blanche	L'herbe est rouge	L'herbe est verte
Vrai	Faux	Faux
Faux	Vrai	Faux
Faux	Faux	Vrai
Faux	Faux	Faux

Les sémantiques représentationnelles visent à capturer les effets des variations du monde en fixant les significations des énoncés du langage. Au contraire, les

10. Il s'agit d'une adaptation d'un exemple fourni par Etchemendy (Etchemendy, 1990, p.63).

sémantiques interprétationnelles visent à capturer les effets des variations dans la signification des expressions du langage :

« By including [models in an interpretational theory] that assign true to ["L'herbe est rouge"], we acknowledge that this sentence, though [actually] false, could be assigned a different meaning, perhaps that Lincoln had a beard, and thereby say something true. [...] *[T]his restriction from eight to four models though easily motivated from the representational standpoint, would make little sense in interpretational semantics.* » (Etchemendy, 1990, p.63, je souligne)

On voit ainsi pourquoi, selon Etchemendy, les deux approches de la sémantique ne peuvent pas être réconciliées.

Les approches représentationnelle et interprétationnelle ne sont donc pas simplement deux *perspectives* différentes sur la sémantique qu'on pourrait adopter à gré : « Clearly, representational and interpretational semantics are entirely different enterprises, governed by entirely different standards » (Etchemendy, 1990, p.63). Il semble que, selon Etchemendy, ce point n'ait pas été relevé, ou du moins n'ait pas été souligné avec une force suffisante par la plupart des interprètes de Tarski. On a ainsi pu mal comprendre le projet de Tarski dans l'article de 1936 sur la base d'une confusion entre les deux approches. On a par exemple pu penser que Tarski voulait réduire les mondes possibles de l'approche représentationnelle aux modèles tarskiens de l'approche interprétationnelle¹¹ (Popper, 2002, p.453). Ce qu'il est important de noter, et ce vers quoi la critique d'Etchemendy pointe, est toutefois que l'approche interprétationnelle, s'affairant à caractériser « *p* est vrai dans *L* » pour un monde fixé *W*, ne présente qu'un intérêt limité :

« One simple reason for this is that with a sufficiently broad range of languages to choose from, all sentences are precisely on a par. That is, for any true sentence of (say) English, we can devise *some* languages in which it is false; similarly, any false sentence can always find a home in which it happens to be true. This in spite of any logical or semantic properties the sentence may originally have had. *Sentences*,

11. Etchemendy ne cible pas Popper comme un auteur ayant pu faire cette erreur, mais on peut penser qu'il serait d'accord pour dire que celui-ci ne distingue pas adéquatement les approches représentationnelle et interprétationnelle de la sémantique.

at least in the sense in which these are things that can wander from language to language, do not carry with them the semantic characteristics necessary to ensure *any* truth value. » (Etchemendy, 1990, p.78-79)

Ceci devient encore plus évident lorsqu'on introduit dans notre langage une nouvelle catégorie d'expressions, celle des quantificateurs. C'est le second problème conceptuel que nous tenterons de mettre en évidence dans la section suivante, ainsi que ses implications sur la définition tarskienne de la conséquence logique.

2.2 Quantification et nécessité

Une fois établi le fait que les approches représentationnelle et interprétationnelle de la sémantique ne coïncident pas, Etchemendy en vient à souligner un autre défaut conceptuel de l'approche interprétationnelle. Ce défaut réside dans la conception de la logique du premier ordre avec quantificateurs et a pour conséquence l'impossibilité de caractériser adéquatement la nécessité associée au concept de conséquence logique. C'est d'ailleurs la raison pour laquelle, selon Etchemendy, les approches modèles-théorétiques standards ont divergé des principes interprétationnels tels que développés par Tarski.

2.2.1 Logique du premier ordre, quantificateurs et restrictions

De l'avis d'Etchemendy, un problème surgit pour l'approche interprétationnelle avec ce qu'il appelle les *cross-term restrictions*, restrictions liées à l'utilisation des quantificateurs et à la difficulté de déterminer leur statut par rapport à la démarcation des termes logiques et extra-logiques. Selon lui, lorsqu'on enrichit notre langage avec un quantificateur comme « quelque chose » (*something*), l'on n'arrive plus à donner une caractérisation interprétationnelle satisfaisante. Le problème est qu'il ne semble pas acceptable de traiter le quantificateur comme une constante logique, comme le voudrait Tarski, puisque, selon Etchemendy, « its contribution to the truth values of sentences still differs as radically from model to model as

that of 'Abe Lincoln' or 'was president' » (Etchemendy, 1990, p,76). Autrement dit, contrairement à un connecteur comme la conjonction, dont la contribution sémantique semble identique peu importe les séquences ou les modèles, l'expression « quelque chose » est intimement liée au domaine d'interprétation, qui n'est pas fixe de modèle en modèle.

Pour déterminer le statut de cette expression, au moins deux autres alternatives, toutes deux problématiques, s'offrent à nous (Etchemendy, 1990, p.66). D'abord, on pourrait considérer le quantificateur introduit comme un terme variable faisant partie de la catégorie sémantique des quantificateurs *existentiels*. Ainsi, on pourrait l'interpréter comme signifiant « quelqu'un » (un être humain), « quelque chien », et ainsi de suite. Ceci revient en fait à modifier la portée du quantificateur. Si le quantificateur en question porte sur l'ensemble des êtres humains, on a par exemple l'équivalence suivante : une séquence assignant des êtres humains à *E* satisfait « *E* est barbu » si et seulement si *quelqu'un* est barbu. On ne pourrait cependant pas réinterpréter l'expression « quelque chose » comme signifiant « toute chose » ou « la chose », par exemple, puisque ces réinterprétations seraient considérées comme inadmissibles (*ill-behaved*) par notre sémantique, c'est-à-dire qu'elles ne comptent pas comme des quantificateurs *existentiels*.

Mais une autre alternative s'offre à nous ; l'on pourrait considérer que seule une partie de l'expression « quelque chose » doit varier, tandis que l'autre partie doit être traitée comme une constante logique. Ainsi, on pourrait considérer la partie « quelque » du quantificateur comme ayant une signification fixe et interpréter de différentes manières la partie « chose » du quantificateur, comme signifiant par exemple, « humain », « chien », et ainsi de suite. Autrement dit, on pourrait permettre à nos modèles d'interpréter différemment une partie du quantificateur, c'est-à-dire de lui assigner un domaine différent, mais on ne permettrait pas l'interprétation *universelle* du quantificateur en vertu du statut constant de la partie « quelque » ; interpréter « quelque chose » comme signifiant « toute chose » serait dans ce cas contraire à notre sélection de constantes logiques.

Peu importe laquelle de ces deux alternatives est choisie — et il semble qu'elles

ne soient finalement pas si différentes —, on aboutira au problème suivant. Selon Etchemendy, on ne peut pas accepter qu'un modèle assigne à la fois l'individu Martha Nussbaum à l'expression « Robert Brandom » et l'ensemble des chiens à l'expression « quelque chose » (ou à « chose »), puisque l'individu assigné à l'expression « Robert Brandom » doit faire partie du domaine de discours de notre modèle, à savoir à l'ensemble associé à l'expression « quelque chose ». Le problème est que l'interprétation qu'on fait d'une expression entre en conflit avec l'interprétation qu'on fait d'une expression différente. Dans notre exemple, ce sont les interprétations des expressions « Robert Brandom » et « quelque chose » qui sont problématiques.

Pour éviter ceci, c'est-à-dire pour éviter les interprétations problématiques, il faut donc recourir à des restrictions sur les quantificateurs. C'est ce qu'Etchemendy appelle des *cross-term restrictions*, qui imposent en fait des « constraints on the simultaneous interpretation of two or more expressions », peu importe que ces expressions appartiennent à la même catégorie sémantique ou non (Etchemendy, 1990, p.71). Certaines séquences directes sont ainsi exclues de l'ensemble de nos modèles non pas parce que l'interprétation d'une expression est inadmissible (*ill-behaved*) en elle-même, mais plutôt parce que l'interprétation d'une expression particulière est incompatible avec celle d'une autre expression.

L'effet de telles restrictions est de réduire l'ensemble des modèles de notre sémantique. En retour, ceci a pour effet d'augmenter le nombre d'énoncés qui sont vrais dans tous les modèles et ainsi l'ensemble d'arguments préservant la vérité dans tous les modèles (Etchemendy, 1990, p.71). Ultimement, les *cross-term restrictions* ont donc pour effet de modifier les énoncés qu'on considère logiquement vrais et les arguments qu'on considère logiquement valides. Or, il semble ne pas y avoir de raison liée à notre sémantique interprétationnelle d'imposer des *cross-term restrictions* sur l'interprétation de nos expressions :

« our decision [...] [to impose restrictions] is not guided by anything over and above our intuitions about which inferences should be valid and which sentences should be logically true. Any such restriction stands in violation of Tarski's account of the logical properties. » (Et-

chemendy, 1990, p.73)

Si par exemple on exclut de notre ensemble de modèles la séquence (directe) où « chose » a pour domaine l'ensemble des chiens et les autres interprétations jugées problématiques du même genre, alors des arguments de la forme suivante seront considérés comme valides : Robert Brandom est barbu, donc quelque chose est barbu. Sans une telle restriction sur l'interprétation des expressions, on doit choisir entre considérer l'argument valide *malgré* le fait qu'il existe des séquences ne préservant pas la vérité des prémisses à la conclusion (par exemple si on assigne l'ensemble des chiens à « chose »), ou bien déclarer l'argument invalide même s'il est considéré intuitivement valide¹².

Etchemendy croit qu'il y a une apparente circularité dans le recours aux *restrictions* en général, puisqu'on risque de déclarer des énoncés comme logiquement vrais s'ils sont vrais sous toutes les interprétations qui ne falsifient pas les énoncés nous apparaissant, *intuitivement*, logiquement vrais. Il en va de même pour les restrictions sur les quantificateurs : « If we must [...] impose cross-term restrictions to “fine tune” the semantics, but can motivate the restrictions only by noting that the unmodified semantics gets things wrong, then our analysis is in serious trouble » (Etchemendy, 1990, p.73). En d'autres mots, il manque un support indépendant permettant de justifier le recours aux restrictions.

Selon Etchemendy, face à ce problème, on peut tenter de réviser l'approche de Tarski de manière à rendre acceptable le recours occasionnel aux *cross-term restrictions*. Mais la façon de faire ceci n'est pas évidente à première vue, dans la mesure où la manière la plus directe d'imposer les restrictions est en fait circulaire : « We can hardly say that a sentence is logically true in L just in case it is true in all interpretations that remain after imposing those cross-term restrictions needed to produce the proper collection of logical truths for L » (Etchemendy, 1990, p.75). L'un des défauts majeurs de l'approche de Tarski est donc le suivant :

12. Plus spécifiquement, Etchemendy croit qu'il y a alors une violation d'un critère de *persistence* : « If the property of *not being logically true* should persist through *contractions* of the language, the property of *being logically true* should persist through *expansions* » (Etchemendy, 1990, p.31). Dans ce cas, on voudrait que l'ajout de l'ensemble des chiens à notre domaine ne change rien au statut de l'argument.

il n'y a pas de sélection de constantes logiques permettant d'arriver à un résultat adéquat concernant la notion de vérité logique, et, par extension, de conséquence logique, pour les langages avec quantificateurs. Etchemendy croit d'ailleurs qu'il s'agit d'une indication conceptuelle montrant que l'approche interprétative de la sémantique est inadéquate. L'explication de ce qu'Etchemendy appelle le « sophisme de Tarski » permettra de voir ceci de manière plus évidente.

2.2.2 Le « sophisme de Tarski » et la question des constantes logiques

Etchemendy soutient que la définition tarskienne de la conséquence logique échoue à justifier adéquatement la nécessité de la conclusion dans un argument valide. En d'autres mots, selon Etchemendy, la définition de Tarski ne capture pas adéquatement le fait que les prémisses d'un argument valide impliquent nécessairement la conclusion. La caractéristique modale de la définition de Tarski se trouve dans l'exigence suivante : un argument peut être déclaré valide si et seulement si il est *impossible* que les prémisses soient vraies et la conclusion fausse (Tarski, 1983, p.414 ; Tarski 2002, p.183).

Certains commentateurs ont défendu l'idée qu'on ne devrait pas comprendre la formulation de Tarski comme visant à « injecter » de la nécessité dans le concept de conséquence logique, ou, en d'autres termes, qu'on ne devrait pas rechercher une modalité forte dans la définition de Tarski (Ray, 1996; Gomez-Torrente, 1998). Une telle position est défendue en soutenant que l'utilisation par Tarski du mot « doit » (*must*) dans sa formulation ne vise pas à capturer une notion de nécessité, mais devrait plutôt être comprise comme visant une sorte de *généralité*. On obtiendrait ainsi des formulations comme « la conclusion d'un argument valide est vraie *toutes les fois* que les prémisses sont elles-mêmes vraies » ou « la conclusion d'un argument valide est vraie *chaque fois* que les prémisses sont elles-mêmes vraies », formulations visiblement plus faibles que celle de Tarski. Néanmoins, cette ligne d'argumentation anti-modale apparaît défectueuse, en ce qu'elle présuppose que l'idée de généralité exclut *de facto* l'idée de modalité, alors que les modalités peuvent être comprises de manière quantificationnelle, par exemple comme

quantification sur les mondes (Kripke, 1959). D'autres auteurs, comme Patterson, défendent l'idée que Tarski devrait être compris comme mettant l'accent sur la notion de vérité et disant ainsi que les « consequences of *true* sentences must be *true* » (Patterson, 2008, p.206). Bien que cette lecture soit plus attrayante que la précédente, en ce qu'elle tente d'explicitier un lien entre la définition du concept de conséquence de Tarski et celle de Carnap, rien ne semble y disqualifier la présence d'une modalité. Ainsi, les arguments à l'effet qu'on ne devrait voir aucune modalité dans la définition de Tarski devraient être rejetés.

L'objectif de Tarski avec sa définition étant de capturer un concept intuitif d'une manière formellement correcte pour les systèmes déductifs, la coïncidence entre les deux concepts — c'est-à-dire intuitif et formel — peut être établie à l'aide de la clause suivante : σ est une conséquence intuitive de Γ si et seulement si σ est une conséquence tarskienne de Γ pour une sélection de constantes logiques (Etchemendy, 1990, p.86). Le problème, selon Etchemendy, est que la définition de Tarski supporte au plus le conditionnel suivant : si σ est une conséquence intuitive de Γ , alors σ est une conséquence tarskienne de Γ pour une sélection de constantes logiques. En d'autres termes, un argument intuitivement valide $\langle \Gamma, \sigma \rangle$ est aussi valide au sens de Tarski pour une certaine sélection de constantes logiques (Etchemendy, 1990, p.86). Bien entendu, l'équivalence ne peut être obtenue que si la *converse* du conditionnel peut également être obtenue, à savoir si les conséquences tarskiennes sont également des conséquences intuitives. Mais, Etchemendy croit qu'il n'est pas possible de prouver cette équivalence puisque le dernier conditionnel est faux. Remarquons toutefois que Etchemendy croit que *si* l'équivalence pouvait être établie, alors l'analyse de Tarski serait satisfaisante :

« Needless to say, *if* [the] equivalence could somehow be shown, *then* Tarski's definition of consequence could hardly be faulted. But in order to show that the equivalence holds, we must show that if σ is a Tarskian consequence of Γ , then it is a consequence "in the ordinary sense". That is, we must show that if all the members of Γ are true, σ must be true as well. » (Etchemendy, 1990, p.86, je souligne)

C'est ici que se trouve, selon Etchemendy, le lien modal erroné entre les prémisses et la conclusion d'un argument dans l'analyse tarskienne : c'est seulement en

commettant un « sophisme modal » qu'on arrive au bi-conditionnel recherché. Ainsi, c'est seulement sur la base d'une erreur de raisonnement qu'on a cru (et qu'on continue de croire) que l'analyse de Tarski est adéquate. Voici le passage, que nous avons déjà évoqué au chapitre précédent, de l'article de 1936 où Etchemendy voit un problème dans le raisonnement de Tarski¹³ :

« It seems to me that everyone who understands the content of the above definition [the formal/technical definition] must admit that it agrees quite well with common usage [the intuitive concept]. This becomes still clearer from its various consequences. In particular, it can be proved, on the basis of this definition, that *every consequence of true sentences must be true*, and also that the consequence relation which holds between given sentences is completely independent of the sense of the extra-logical constants which occur in these sentences. »
(Tarski, 1983, p.417, je souligne)

Ce passage, d'après Etchemendy, vise à fournir une *justification* de la deuxième partie du bi-conditionnel. Mais Etchemendy juge que cette justification est incomplète et que l'équivalence est ainsi fautive (Etchemendy, 1990, p.87). Pour défendre son point, Etchemendy fournit un contre-exemple montrant qu'on ne peut pas (et qu'on ne devrait pas) accepter la ligne d'argument de Tarski.

Selon Etchemendy, « Lincoln avait une barbe » est une conséquence au sens de Tarski de l'énoncé « Washington était président » lorsque *toutes* les expressions de l'argument sont fixées, puisque la substitution d'expressions variables — ici, l'ensemble vide — par des séquences d'objets des catégories appropriées préserve la vérité de la conclusion ou falsifie au moins l'une des prémisses de l'argument. Avec un tel exemple, Etchemendy semble saisir une possibilité soulevée par Tarski lui-même d'un « cas limite » (*extreme case*), où l'on pourrait regarder tous les termes du langage comme des termes logiques (Tarski, 1983, p.420). Toutefois, comme le soulignent des auteurs comme Manuel Sanchez-Miguel, « in the extreme case Tarski contemplates [...] the first sentence counts as logically true and the

13. Dans cette section, nous utilisons la traduction de Woodger (Tarski, 1983), puisque c'est à elle qu'Etchemendy fait référence. Bien qu'Etchemendy propose dans cette section du texte une *reconstruction* du « raisonnement de Tarski » (*Tarski's reasoning*), il y a des raisons de douter que Tarski offre une *preuve* de la nécessité de la conséquence logique. Ce n'est toutefois pas *cet* enjeu spécifique que nous souhaitons examiner ici.

argument as logically valid » (Sanchez-Miguel, 2003, p.164)¹⁴. On peut donc voir que le cas limite présenté par Etchemendy n'est pas conforme avec notre saisie intuitive du concept de conséquence logique¹⁵. Même si la vérité est préservée pour toutes les séquences et même si, conséquemment, l'argument est valide au sens de Tarski, « Lincoln avait une barbe » n'est clairement pas une *conséquence intuitive* de « Washington était président », car on n'accepterait certainement pas que celle-là *doit être vraie* étant donné que celle-ci est vraie¹⁶.

Etchemendy a certainement raison de soutenir qu'une conséquence tarskienne ne respecte pas le sens intuitif du concept de conséquence logique, puisqu'une conséquence au sens ordinaire n'est pas le résultat d'un *choix* de termes à maintenir fixes :

« It is perfectly clear that with many selections of [logical constants], there are Tarskian consequences that are not genuine consequences, and hence that [the equivalence] is simply false. Yet our proof that every Tarskian consequence of true sentences must be true is perfectly correct. The problem is not with our proof, but with thinking that this proof shows that any modal relation holds between the premises and conclusion of the argument $\langle \Gamma, \sigma \rangle$. » (Etchemendy, 1990, p.87)

Selon Etchemendy, l'argument de Tarski échoue ainsi à établir la modalité qu'on recherche dans le concept de conséquence logique, et l'impression de succès qu'on tire de cette analyse est seulement imputable à un « sophisme modal » :

« To show that all Tarskian consequences are consequences in the ordinary sense, we would need to prove a theorem with an embedded modality. Specifically, we would have to show that, for any Γ and σ , if (1) σ is a Tarskian consequence of Γ (for some [logical constants]) then the following are jointly incompatible : (2) All the members of Γ are true [and] (3) σ is false. But, of course, all we can show is that for any

14. Voir aussi (Ray, 1996, p.648-651).

15. Notons que le fait de garder fixes tous les termes a pour effet de bloquer le problème des *cross-term restrictions* sur les quantificateurs, qui, comme nous l'avons vu, est un aspect important de l'argument d'Etchemendy.

16. Ceci est bien sûr dû au fait que l'argument est satisfait de manière *triviale* par *n'importe quelle* séquence. Comme nous l'avons déjà mentionné, un examen plus approfondi de l'idée d'intuition en jeu dans la définition de Tarski serait requis, mais nous n'avons pas l'espace ici pour examiner cette question.

Γ and σ , the three conditions are jointly incompatible. » (Etchemendy, 1990, p.87)

Notons que, en l'absence de modalités, il n'y aurait aucun problème en tant que tel avec ce raisonnement, puisque le connecteur d'incompatibilité ($|$) est un connecteur logiquement équivalent à la négation d'une conjonction, c'est-à-dire qu'on peut sans problème établir l'équivalence suivante :

$$\neg(P\&Q\&R) \equiv (P \rightarrow (Q | R)) \equiv (P \rightarrow \neg(Q\&R))$$

Le « sophisme modal » de Tarski résiderait ainsi dans le fait de penser que (2) la vérité des prémisses de l'ensemble Γ et (3) la fausseté de la conclusion σ , sont *nécessairement* incompatibles si (1) σ est une conséquence tarskienne de Γ , alors que tout ce qu'on peut dire dans de telles conditions est que (2) et (3) ne peuvent pas être vrais simultanément. Le lien modal recherché entre les prémisses Γ et la conclusion σ dans le concept ordinaire est ainsi absent de la définition tarskienne :

« Now, it should be clear from a purely abstract point of view that the joint incompatibility of (1), (2), and (3), plus the truth of (1), does not entail joint incompatibility of (2) and (3). Here, we need only note the fallaciousness of any inference from Necessarily (if P and Q then not R) to If P then necessarily (if Q then not R). » (Etchemendy, 1990, p.87)

Formellement, on a l'inférence suivante, qui est invalide :

$$\frac{\Box((P\&Q) \rightarrow \neg R)}{P \rightarrow \Box(Q \rightarrow \neg R)}$$

La « preuve » de Tarski situe la nécessité à *l'intérieur* du conditionnel requis pour prouver l'équivalence, affectant seulement son conséquent, alors qu'elle devrait être située à *l'extérieur* du conditionnel principal, de manière à être englobante et à affecter à la fois l'antécédent et le conséquent. Au final, selon Etchemendy, on se retrouve donc avec un sophisme de portée modale (Etchemendy, 1990, p.87)¹⁷. À

17. Il est intéressant de remarquer que ce sophisme est similaire par sa forme au sophisme *de re / de dicto*.

partir des formulations qu'on retrouve dans l'article de Tarski, Etchemendy croit qu'on ne peut pas expliquer le sophisme par une mauvaise compréhension des intentions de Tarski. Ceci, comme nous le verrons dans le prochain chapitre, n'est pas aussi transparent qu'Etchemendy l'affirme. On ne peut pas non plus, selon Etchemendy, penser qu'un logicien aussi réputé que Tarski aurait pu tomber dans un piège aussi bête (*foolish*) qu'une *faute* de raisonnement concernant la portée d'un opérateur modal. Au contraire, Etchemendy croit que Tarski aurait pu juger satisfaisant le fait d'*injecter* la nécessité logique dans sa définition formelle du concept de conséquence alors que le concept intuitif qu'il s'efforçait de capturer était plus ou moins ambigu (Etchemendy, 1990, p.92). Bien que je trouve que cette lecture de Tarski ne lui soit pas très charitable, je crois qu'il y a en fait une manière de comprendre l'argument qui n'est pas victime du sophisme (§3.2). Mais avant de proposer une réponse à Etchemendy et de réhabiliter Tarski sur ce point, il convient de présenter un troisième problème, cette fois de nature extensionnelle, que celui-ci soulève dans la définition tarskienne.

2.3 Le problème de la *sur-génération*

2.3.1 La variation du domaine d'interprétation

Le passage à la logique du premier ordre n'est pas seulement problématique sur le plan conceptuel, mais, selon Etchemendy, rend aussi l'extension de la définition inadéquate. Plus précisément, l'extension de la définition tarskienne est incorrecte hormis pour la logique du premier ordre, mais, selon Etchemendy, l'extension de cette dernière est correcte pour de mauvaises raisons. Le problème est que la définition tarskienne est influencée par des faits qui n'ont rien à voir avec la logique. Ainsi, on peut par exemple montrer que la vérité de certains énoncés du premier ordre dépend du nombre d'objets dans l'univers. Voici une généralisation des énoncés σ d'Etchemendy, qui expriment qu'il y a au moins n objets différents (Etchemendy, 1990, Chapitre 8) :

$$\sigma_n = \exists x_1 \exists x_2 \dots \exists x_{n-1} \exists x_n [(x_1 \neq x_2) \& (x_1 \neq x_3) \& \dots \& (x_{n-1} \neq x_n)]$$

Ce qui intéresse Etchemendy est que la vérité (ou la fausseté) de ces énoncés dépend du nombre d'objets dans les domaines de nos modèles : un énoncé σ_n est vrai s'il y a au moins n objets dans le domaine. De même, un énoncé $\neg\sigma_n$ est vrai s'il y a moins que n objets dans le domaine. Etchemendy défend que des énoncés de cette forme sont déclarés logiquement vrais par l'analyse tarskienne. Par exemple, les énoncés $\neg\sigma_n$ seront des vérités logiques si le domaine de notre modèle est fini et compte exactement k éléments (où $k < n$). Dans les mots d'Etchemendy :

« Clearly, some of these sentences are true. Exactly how many depends, of course, on the size of the universe — that is, on how many objects there happen, in fact, to be. If the universe is infinite then all of the sentences are true, and so will be mistakenly judged to be logically true. If it is finite, then only a finite number of them will be judged logically true. But the important point is not how many of these sentences the definition gets wrong, but rather the fact that the assessments here are clearly dependent on a nonlogical state of affairs. » (Etchemendy, 1990, p.111)

L'ensemble de nos expressions logiques semble donc être suffisamment expressif pour exprimer le fait que le monde compte au moins n objets¹⁸. Mais les énoncés de cette généralisation qui sont vrais ne semblent pas être des vérités *logiques*, en ce sens que le nombre d'objets dans le domaine ne relève pas lui-même de la logique. Pourtant, selon Etchemendy, l'analyse tarskienne semble les déclarer comme tels¹⁹. Ainsi, si l'on fait des énoncés de cette généralisation qui sont vrais la conclusion d'arguments, ceux-ci devraient en principe être considérés comme valides.

Pour préciser le propos d'Etchemendy, il faut souligner qu'en montrant que la définition tarskienne repose sur des faits extra-logiques, on ne montre pas encore

18. Ceci est vrai pour peu qu'on accepte — ce que tous ne font pas — que l'identité fait partie de l'ensemble des constantes logiques, au même titre que les quantificateurs, la négation et la conjonction.

19. Comme le soulignent Mario Gomez-Torrente (Gomez-Torrente, 1999, p.379) et Scott Soames (Soames, 1999, p.119), la notion de vérité logique à laquelle Etchemendy fait appel est construite de manière très (trop?) étroite en ce qu'elle associe vérité logique et analyticit  (Etchemendy, 1990, p.126).

que l'analyse tarskienne *sur-génère*, c'est-à-dire qu'elle compte plus de vérités logiques (ou d'arguments valides) qu'il n'en faut. On montre plutôt qu'elle sur-générait *si* le monde était différent, ou *si* l'on incluait certaines expressions dans notre ensemble de constantes logiques.

Mais alors, si notre ensemble de constantes logiques est *trop expressif*, c'est-à-dire s'il nous permet d'exprimer des énoncés *comme* des vérités logiques alors qu'ils n'en sont pas, peut-on décider de laisser de côté certaines expressions logiques, pour éviter le problème? La réponse à cette question est négative, puisque si on laisse de côté certaines constantes logiques, comme le signe d'identité, on aura le problème inverse. Autrement dit, laisser de côté l'identité fait *sous-générer* notre définition en la faisant déclarer invalides des arguments comme le suivant (alors qu'ils devraient être jugés valides) : Si a est mathématicien et si $a = b$, alors b est mathématicien. L'argument est déclaré invalide puisqu'il ne préserve pas la vérité sous toutes les interprétations des expressions extra-logiques, à savoir le prédicat « mathématicien », les constantes d'individus (a et b) et la relation d'identité ($=$). En effet, si l'on interprétait par exemple $a = b$ comme signifiant « a est le fils de b » et les constantes a et b comme référant respectivement à Karl Menger et à Carl Menger, on aurait des prémisses vraies et une conclusion fausse²⁰. On voit donc que l'inclusion de la relation d'identité dans notre ensemble de constantes logiques a pour résultat de faire sur-générer notre définition du concept de conséquence (on considère certains énoncés σ_n comme étant des vérités logiques, et donc trop d'arguments comme étant valides). À l'inverse, son exclusion a pour résultat de faire sous-générer la définition (on déclare invalides certains arguments qui ne le sont pas). Dans tous les cas, l'extension de notre définition est inadéquate.

Il convient ici de noter que les problèmes extensionnels pourraient être évités en faisant varier les domaines de nos modèles. Ainsi l'énoncé $\exists x_1 \exists x_2 (x_1 \neq x_2)$ n'est par exemple pas considéré comme une vérité logique par les théoriciens des modèles contemporains, *puisque'il est faux dans les domaines ne comptant qu'un seul élément*. Etchemendy croit cependant que la sémantique tarskienne, puisqu'elle

20. Carl est bien le père de Karl, l'algébriste autrichien, mais il était économiste et non pas mathématicien.

ne permet pas un tel traitement des domaines de quantification (Etchemendy, 1988a), n'arrive donc pas au verdict souhaité pour les énoncés σ . Elle les classe incorrectement comme des vérités (ou des faussetés) logiques, ce qui constitue une lacune la différenciant de la théorie des modèles contemporaine²¹. Selon Etchemendy, l'erreur de la sémantique tarskienne s'explique ultimement par le fait qu'elle repose sur un principe défectueux : le principe de réduction. Voyons maintenant en quoi consiste ce principe et les problèmes qu'il entraîne.

2.3.2 Le principe de réduction

Selon Etchemendy, les défaillances extensionnelles dans l'approche de Tarski sont imputables à la nature quantificationnelle de celle-ci, et plus spécifiquement au recours à certains principes gouvernant l'utilisation des quantificateurs. Parmi ces principes, certains ne posent pas problème (Etchemendy, 1990, p.98). Ainsi, on peut accepter le *principe d'instanciation*, qui stipule que si un énoncé quantifié universellement est vrai, alors chacune de ses instances est également vraie. De même, on peut accepter un *principe de clôture* de la logique stipulant que si un énoncé quantifié universellement est logiquement vrai, alors toutes ses instances sont également logiquement vraies.

Par contre, l'approche de Tarski, tout comme celle de Bolzano, fait appel à un troisième principe, le *principe de réduction*, jugé problématique par Etchemendy (Etchemendy, 1990, Chapitre 7). Ce principe stipule que si un énoncé quantifié universellement est vrai, alors toutes ses instances sont logiquement vraies. En d'autres mots, ce principe revient à réduire la vérité logique à la vérité (tout court) d'un ensemble d'énoncés. De toute évidence, il s'agit d'un principe qu'on ne peut pas accepter puisque les généralisations universelles ne sont pas nécessairement des vérités logiques. En effet, la vérité de ces énoncés, comme celle d'un ensemble d'autres énoncés, peut ne découler que du hasard ou des contin-

21. Rappelons qu'Etchemendy, en traitant les quantificateurs comme des expressions extralogiques (§2.2.1), tente de « sauver » Tarski en examinant comment celui-ci pourrait accommoder la relativisation des domaines, mais qu'il juge finalement ces tentatives infructueuses en vertu des problèmes associés au recours aux *cross-term restrictions*.

gences du monde. Pour sauver Tarski, il semble nécessaire de corriger ce principe. Etchemendy propose deux modifications, faisant toutes deux du chemin sur des possibilités ouvertes par Tarski.

2.3.2.1 Première tentative de modification : la conséquence comme *relative* à des sélections de constantes logiques

Comme la notion de vérité logique repose sur la démarcation entre expressions logiques et extra-logiques, le principe que l'on recherche devrait être une version modifiée du principe de réduction qui tient compte de la *relativisation* à la sélection de constantes logiques. Donc, puisque Tarski ouvre la porte à une caractérisation *relationnelle* de la notion de vérité logique, *relative* à une sélection arbitraire de constantes logiques, Etchemendy propose comme première modification du principe de réduction le principe suivant : Si un énoncé quantifié universellement est vrai, alors toutes ses instances sont logiquement vraies *par rapport aux* (*with respect to*) expressions qui ne sont pas liées par un quantificateur universel (Etchemendy, 1990, p.101). Cette formulation du principe est motivée par le fait que la sélection des constantes logiques détermine quelles expressions ne varieront pas et ne tomberont donc pas sous la portée de quantificateurs dans notre généralisation. Si la généralisation d'un énoncé est vraie, alors la définition tarskienne devrait déclarer l'énoncé comme logiquement vrai par rapport aux constantes logiques sélectionnées.

On peut se demander s'il s'agit d'une version plausible du principe, ce à quoi Etchemendy répond par la négative. Il semble en effet naturel, comme le stipule cette modification du principe, de croire qu'une vérité logique est vraie en vertu du sens ou de la signification de certaines expressions constituantes, à savoir des constantes logiques. Ainsi, on pourrait penser qu'une instanciation du principe du tiers exclu, comme l'énoncé « Lincoln était président ou Lincoln n'était pas président », est une vérité logique en vertu de la signification des expressions « ou » et « n' _ pas »²², c'est-à-dire de la disjonction et de la négation, et pas en

22. Je laisse de côté les problèmes de la *grammaire* de la négation française.

vertu d'un quelconque état du monde (Etchemendy, 1990, p.101).

Mais, selon Etchemendy, le statut de *vérités logiques* de certaines instanciations du tiers exclu ne dépend pas *uniquement* des connecteurs de disjonction et de négation. C'est le cas par exemple de certains énoncés faisant appel aux quantificateurs, comme l'énoncé « Tout le monde est heureux ou quelqu'un n'est pas heureux » (Etchemendy, 1990, p.102)²³. Dans ce cas, la vérité de l'énoncé dépend de la signification des expressions « ou » et « n'__ pas », mais également de la signification des expressions « tout le monde » et « quelqu'un ». Ainsi, si l'expression « tout le monde » avait signifié « tous les chiens », l'énoncé aurait été faux. Ceci semble donc confirmer notre intuition à l'effet que la vérité logique d'un énoncé dépend de la signification de certaines de ses expressions constituantes.

On pourrait aussi généraliser cette idée et penser que l'énoncé « Leslie est un homme ou Leslie n'est pas un *bachelor* » est une vérité logique, pour peu qu'on ne fasse varier que le nom « Leslie » (Etchemendy, 1990, p.102). Autrement dit, cet énoncé pourrait lui aussi être considéré comme une vérité logique si, en plus de la disjonction et de la négation, on considérait comme constantes les expressions « est un homme » et « est un *bachelor* ». On a ainsi une manière de comprendre ce que veut dire le fait que la vérité logique d'un énoncé dépend de la signification de certaines expressions, comme le stipule la première modification du principe.

Mais en inspectant un peu plus minutieusement le principe ainsi modifié, on peut conclure qu'il doit être rejeté, puisqu'il permet que des vérités logiques soient considérées comme telles pour des raisons qui n'ont absolument *rien* à voir avec la logique. Ainsi, tous les êtres humains satisfont (pour le moment) la généralisation suivante : pour tout individu x , si x était président des États-Unis, alors x était un homme (Etchemendy, 1990, p.104). La reformulation du principe est ainsi respectée, bien que la généralisation soit vraie pour des raisons *historiques* purement contingentes, et non pas pour des raisons *logiques*. Le problème avec cette version du principe est qu'elle énonce une condition *nécessaire* pour établir la vérité logique d'un énoncé, mais pas une condition *suffisante*; les vérités

23. On joue ici sur la traduction entre les quantificateurs : le quantificateur universel est par définition équivalent à la négation du quantificateur existentiel.

logiques exigent davantage que le simple fait de reposer sur la signification de certaines expressions (Etchemendy, 1990, p.105-106). Autrement dit, on n'a pas de garantie que la vérité logique d'un énoncé repose *uniquement* sur le sens des expressions, et penser le contraire est une confusion entre conditions nécessaires et conditions suffisantes.

2.3.2.2 Deuxième tentative de modification : restrictions sur la sélection de constantes logiques

Pour espérer corriger le principe, il faut ainsi déterminer ce qu'il a de problématique et quel problème il doit éviter. Le coeur du problème du principe de réduction est qu'il associe la vérité logique d'un énoncé à la vérité d'une généralisation universelle. Ceci a pour effet d'introduire une influence extra-logique dans notre définition du concept de conséquence logique (ou à tout le moins de rendre *possible* l'introduction d'une telle influence). Or il semble d'une part que les vérités logiques ne devraient pas dépendre de faits extra-logiques, et d'autre part qu'associer le statut logique des vérités logiques à la simple vérité (ou fausseté) matérielle d'autres énoncés rend impossible l'obtention d'une *extension adéquate* pour notre définition. Si le statut logique d'un énoncé repose sur des faits extra-logiques — dans les mots d'Etchemendy : des « *substantive extra-logical facts* » —, alors notre caractérisation des notions logiques est erronée (ou le serait si les faits en question étaient changés).

Introduire des noms ou des prédicats dans notre sélection de constantes logiques expose notre caractérisation des notions logiques à une influence extra-logique. Pour éviter ceci, et pour modifier la formulation originale du principe de réduction, on peut alors se demander s'il est possible d'imposer des restrictions sur notre sélection de constantes, c'est-à-dire de déterminer un critère de *logicalité* (*logicality*). On pourrait par exemple, selon Etchemendy, penser, comme deuxième tentative de modification du principe, à la formulation suivante : Si un énoncé quantifié universellement est vrai, et si toutes les constantes logiques apparaissant dans les fonctions propositionnelles satisfont un critère de logicalité, alors toutes

ses instances sont logiquement vraies (Etchemendy, 1990, p.110).

Mais, même si cette caractérisation semble plausible — du moins plus plausible que la première —, elle demeure sujette à l'influence des faits extra-logiques. Autrement dit, elle a le même défaut, qui selon Etchemendy est impossible à éviter : elle ne nous immunise pas contre l'influence extra-logique. Ce point est passé sous le radar de Tarski puisque celui-ci a présupposé, à tort selon Etchemendy, que les « facts of a nonlogical sort can [exert their] influence [...] only if expressions of a nonlogical sort are included in the set of fixed terms » (Etchemendy, 1990, p.111). Comme nous le verrons, Etchemendy juge que cette hypothèse est erronée puisqu'il semble que les faits extra-logiques ne concernent pas seulement les individus ou les propriétés, mais aussi, par exemple, le nombre d'individus qui existent dans le monde.

2.3.2.3 Peut-on éviter les « accidents logiques » ?

Le problème ne résulte pas d'une sélection incorrecte de constantes logiques ou de critère de démarcation, mais plutôt du fait qu'on assume qu'il est possible d'*exclure* l'influence de faits extra-logiques de notre caractérisation du concept de conséquence. Or Etchemendy affirme qu'il y a, en plus de la taille des domaines, des assomptions extra-logiques dont l'influence sur notre caractérisation du concept de conséquence est importante : on n'a qu'à penser, selon lui, à l'axiome de la paire ou à l'hypothèse du continu (Etchemendy, 1990, p.122-124).

L'axiome de la paire stipule qu'on peut obtenir un ensemble C à partir de deux ensembles A et B . Cet axiome, bien qu'il puisse être obtenu à partir d'autres axiomes de la théorie des ensembles de Zermelo-Fraenkel et ne soit donc pas indépendant, est généralement admis afin de définir la notion de couple et de n -tuple ordonnés (Potter, 2004, p.291). L'hypothèse du continu énonce qu'il n'y a pas d'ensemble dont la cardinalité soit plus petite que 2^{\aleph_0} , la cardinalité des nombres réels, mais plus grande que \aleph_0 , la cardinalité des entiers naturels (Potter, 2004,

p.268)²⁴. Sans aller dans le détail des arguments d'Etchemendy à ce sujet, qui sont plutôt techniques, on peut retenir que l'axiome de la paire et l'hypothèse du continu sont importants pour les fondements de la théorie des ensembles, mais surtout que le fait de les accepter nous commet à une influence extra-logique, ce à quoi, selon Etchemendy, on peut et on doit s'objecter. Après tout, « why should the *logical* truth of a sentence depend on such highly abstract *set-theoretic* claims, claims that are not, intuitively, a matter of logic at all? » (Etchemendy, 1990, p.123). Puisque la définition tarskienne ne permet pas d'exclure une telle influence, elle est donc défailante sur le plan extensionnel.

Il y a au moins deux leçons à retenir de ceci, l'une portant sur ce qu'on recherche et l'autre sur la futilité de le rechercher. Premièrement, pour éviter le problème de la sur-génération, il faudrait inclure une garantie que des faits extra-logiques, comme ceux concernant la théorie des ensembles, n'influencent pas notre caractérisation du concept de conséquence logique. Il s'agit d'une formulation du principe de clôture que nous avons formulé ci-dessus, à l'effet que seules les instances de généralisations logiquement vraies sont également logiquement vraies. Une autre leçon, liée à la précédente, est que l'on recherche en quelque sorte un remplacement du principe de réduction, qui est un principe défectueux, par le principe de clôture (Etchemendy, 1990, p.143). Mais pour utiliser le principe de clôture, il faut une spécification de ce que veut dire le fait qu'une généralisation soit *logiquement* vraie (et pas simplement historiquement vraie, ou physiquement vraie, etc.). Autrement dit, on est à la recherche de ce qu'on recherchait déjà par ailleurs.

2.4 Conclusion

La critique d'Etchemendy vise à montrer que la définition tarskienne est conceptuellement incorrecte, c'est-à-dire qu'elle ne rend pas compte adéquatement du concept de conséquence, et ce pour deux raisons principales. D'abord, elle repose sur une conception interprétationnelle de la sémantique, et Etchemendy croit que si l'on « prend les modèles au sérieux », alors il faut qu'on « acknowledge the re-

24. L'hypothèse est parfois présentée sous la forme $2^{\aleph_0} = \aleph_1$

presentational nature of model-theoretic semantics » (Etchemendy, 1988b, p.104). Puis, la définition tarskienne échoue à garantir la nécessité du concept de conséquence : « [it] just omits the guarantee [the necessity of truth preservation of valid arguments], attempting to replace it with that which the guarantee is a guarantee of » (Etchemendy, 2008, p.270). La définition tarskienne ne caractérise donc pas adéquatement le concept de conséquence logique.

De plus, la définition tarskienne est extensionnellement inadéquate. Elle déclare logiquement valides les arguments qui sont intuitivement considérés comme tels, mais seulement pour la logique du premier ordre et, de surcroît, elle le fait seulement pour les *mauvaises raisons*. Pour reprendre l'exemple de Willard V.O. Quine, il serait inadéquat de définir « a un coeur » en termes de possession d'un foie, puisqu'il est en quelque sorte accidentel que tous les animaux ayant un coeur aient *aussi* un foie. Selon Etchemendy, on peut expliquer les raisons pour lesquelles l'analyse de Tarski donne une extension adéquate, pour les langages du premier ordre, par le fait que ceux-ci sont relativement faibles et ont des ressources expressives relativement limitées. Les théories du premier ordre reposent également sur des hypothèses (*assumptions*) de la théorie des ensembles relativement fortes, qui passent inaperçues. Ainsi, le monde compense pour les défauts dans l'analyse tarskienne des propriétés logiques. Si l'on supprime certaines hypothèses sur lesquelles elle repose, ou si l'on augmente les ressources expressives en passant par exemple à une logique du second ordre, alors l'analyse des propriétés logiques ne peut plus tenir (Etchemendy, 1990, p.125-126).

Etchemendy tire au moins deux conséquences des problèmes qu'il identifie dans la définition tarskienne. Premièrement, la question de la démarcation entre expressions logiques et extra-logiques est un faux problème ; c'est le *mythe* des constantes logiques (Etchemendy, 1990, Chapitre 9). Ce n'est pas à cause de caractéristiques particulières des « constantes logiques » que l'extension de l'analyse, pour la logique du premier ordre, est adéquate, mais à cause d'hypothèses qui passent pour ainsi dire inaperçues, et à cause de la relative faiblesse des langages évalués.

« [I]n some cases we may be fortunate : if the relevant portions of the world are sufficiently varied, if none of the substantive generalizations

come out true, then the account will not issue any faulty declarations of logical truth. For the only sentences that will be judged logically true will then be instances of *generalizations* that are logically true, and these instances will, by the closure principle, be logically true as well. In these cases, but only in these cases, Tarski's definition will yield a reasonable assortment of logical truths. But we succeed here not because [the principle of reduction] is correct, however modified, or because we have chosen the right "logical constants". Our success is due to [the closure principle] and simple good fortune. » (Etchemendy, 1990, p.130)

Autrement dit, déplacer l'analyse vers la question des constantes logiques est une confusion : si aucune généralisation extra-logique n'est vraie, le succès extensionnel de l'analyse ne peut pas être dû au principe de réduction, qui est problématique et ne peut être corrigé, mais seulement aux contingences du monde, qui font en sorte qu'on n'obtient pas plus de vérités *logiques* qu'il n'en faut, et au principe de clôture. La seule manière de s'assurer qu'on ne sous-génère pas, c'est-à-dire qu'on omette de considérer comme valides des conséquences intuitives, est de laisser tomber la garantie qu'on ne sur-génère pas, c'est-à-dire qu'on accepte qu'il puisse y avoir des conséquences non intuitives considérées valides. Il y a donc une tension entre d'un côté l'inclusivité de notre *ensemble* de constantes logiques et de l'autre l'étroitesse du même ensemble (Etchemendy, 1990, p.134).

La seconde conséquence tirée par Etchemendy concerne la conception qu'on se fait de la logique du premier ordre, et en particulier le statut accordé aux théorèmes de complétude et de fiabilité. Etchemendy reconnaît que l'appareillage déductif de la logique du premier ordre est jugé extensionnellement correct, même si aucun critère de démarcation logique ne peut *garantir* le caractère adéquat de cet appareillage (Etchemendy, 1990, Chapitre 11). Mais, selon lui, les théorèmes de complétude et de fiabilité pour la logique du premier ordre, établissant respectivement que les lois logiques sont toutes démontrables par le système, et que tout ce qui est démontré par le système est une loi logique, sont généralement interprétés comme apportant une telle garantie. D'une part, le théorème de fiabilité est vu comme assurant qu'on ne sur-génère pas, en ce sens que les déductions sont valides (Etchemendy, 1990, p.144). D'autre part, le théorème de complé-

tude est vu comme assurant que nos systèmes déductifs ne sous-génèrent pas, en ce sens que les arguments valides peuvent tous être déduits (Etchemendy, 1990, p.144). Or cette interprétation des théorèmes *présuppose* que la caractérisation modèle-théorique du concept de conséquence est conceptuellement et extensionnellement adéquate (Shapiro, 2011, p.533-534). Etchemendy croit donc que si l'on montre que la définition tarskienne du concept de conséquence ne sur-génère pas, c'est parce qu'on présuppose en fait la fiabilité du système déductif. À la lumière de la critique d'Etchemendy, il semble donc qu'on ne peut pas conserver cette interprétation des théorèmes.

Dans le prochain chapitre, nous tenterons de présenter certaines limites de la critique d'Etchemendy en montrant que certains de ses arguments reposent en quelque sorte sur des distorsions du propos de Tarski (§3.1). Nous tenterons aussi de donner suite à la critique d'Etchemendy sur la question du « sophisme de Tarski », non pas en nous opposant directement à son propos — l'argument attribué à Tarski par Etchemendy, tel que celui-ci le présente, est bel et bien fallacieux —, mais plutôt en nous intéressant à la manière dont il faudrait lire Tarski pour éviter le sophisme. Autrement dit, je ne m'intéresserai pas vraiment à savoir si la reconstruction que donne Etchemendy de l'argument de Tarski est adéquate, mais tenterai plutôt de montrer comment éviter le sophisme (§3.2).

CHAPITRE III

LE DÉPASSEMENT DE LA CRITIQUE D'ETCHEMENDY

Dans les chapitres précédents, nous avons présenté la définition sémantique du concept de conséquence logique donnée par Tarski (Tarski, 2002) puis sa critique par Etchemendy, en particulier telle qu'il la développe dans son livre de 1990 (Etchemendy, 1990). Nous avons vu que cette critique suivait trois axes principaux. Le premier axe porte sur un problème conceptuel dans la définition de Tarski, qui repose sur une approche interprétationnelle de la sémantique plutôt que sur une approche représentationnelle (§2.1). Le second axe, lui aussi de nature conceptuelle, vise à montrer que la définition échoue à capturer l'aspect modal du concept de conséquence (§2.2). Enfin, le troisième axe de la critique d'Etchemendy, cette fois de nature extensionnelle, concerne le fait que la définition tarskienne sur-génère ou sous-génère, c'est-à-dire qu'elle déclare ou bien trop ou bien trop peu d'arguments comme étant valides, selon la manière de démarquer les expressions logiques et extra-logiques (§2.3). Dans ce chapitre, je procéderai à une critique qui montrera la limite des arguments d'Etchemendy contre la définition tarskienne du concept de conséquence logique.

Ma critique se déroulera en deux temps. Dans un premier temps, je procéderai négativement en montrant des défauts présents sur chacun des trois axes de la critique d'Etchemendy (§3.1). Ceci permettra en quelque sorte de voir comment il serait possible de donner suite aux arguments qu'il développe contre la définition tarskienne. Dans un deuxième temps, à partir des défauts identifiés dans sa critique, je tenterai de donner une réponse à l'un de ces arguments d'Etchemendy, à

savoir l'argument du « sophisme de Tarski », de manière à réhabiliter la définition tarskienne (§3.2).

3.1 Les distorsions d'Etchemendy

3.1.1 Le faux dilemme d'Etchemendy ?

Nous avons vu qu'Etchemendy effectuait une distinction entre les approches représentationnelle et interprétationnelle de la sémantique (§2.1.1) et qu'il jugeait que ces approches ne pouvaient pas être réconciliées (§2.1.2). Mais le portrait brossé par Etchemendy d'une sémantique reposant *ou bien* sur le monde *ou bien* sur le langage est plutôt réducteur. Etchemendy semble penser que l'approche représentationnelle est *la* sémantique reposant sur le monde, et que l'approche interprétationnelle, qu'il attribue à Tarski, est *la* sémantique reposant sur le langage. Il n'indique toutefois pas en quoi ces approches sont les seules possibles pour qui veut donner une définition sémantique du concept de conséquence logique. Comme le souligne Gila Sher, « [Etchemendy's] reasoning involves an uncritical uniqueness assumption : there is exactly one theory of logic as based on language and exactly one theory of logic as based on the world » (Sher, 1996, p.683). Ainsi, si l'on accepte l'alternative qu'Etchemendy propose et si l'on juge que l'approche interprétationnelle, étant problématique, doit être rejetée, il faut alors adopter l'approche représentationnelle.

Pourtant, Etchemendy semble lui-même reconnaître qu'il existe différentes moultures de l'approche interprétationnelle, puisqu'il juge que la *théorie de la variation* de Bolzano, faisant appel à la notion de substitution, se distingue de la sémantique tarskienne reposant sur la notion plus générale de satisfaction (Etchemendy, 1990, Chapitre 3). Puisqu'il les présente toutes les deux comme des théories linguistiques, Etchemendy doit en quelque sorte accepter qu'il y a plusieurs manières pour une sémantique de reposer sur des variations au niveau du langage, et doit ainsi accepter qu'il n'y a pas *une seule* théorie interprétationnelle, mais

bien *plusieurs*¹. De même, on peut penser que l'approche représentationnelle de la sémantique donne lieu à des théories différentes, par exemple selon notre critère d'individuation des objets (McKeon, 2010, p.57-59), selon notre critère pour l'identité des propriétés (Lowe, 1997) ou encore selon qu'on accepte l'existence de propriétés essentielles ou non (Forbes, 1997), puisqu'on admettra alors des variations du monde plus ou moins importantes. On peut donc penser qu'Etchemendy commet éventuellement un sophisme du *faux dilemme* lorsqu'il propose un choix entre deux approches alternatives, l'approche représentationnelle et l'approche interprétationnelle, qui ne sont pourtant pas *exhaustives*. Pour répondre à ce problème, il faudrait qu'Etchemendy démontre que ces deux alternatives sont les seules valables pour la caractérisation du concept de conséquence, ce dont on peut douter.

Il y a toutefois un autre problème associé au faux dilemme d'Etchemendy, sans doute plus pressant que le précédent. Ce problème est lui aussi lié à la distinction entre les approches représentationnelle et interprétationnelle de la sémantique, mais concerne plutôt le fait que celles-ci ne semblent pas *s'exclure mutuellement*. Etchemendy suppose, ce qui n'est pas invraisemblable pour les langages extensionnels qu'il considère, que la valeur sémantique des expressions dépend, d'un côté, de leur signification linguistique et, de l'autre, du monde. C'est pourquoi il propose de faire varier ou bien le monde (auquel cas on considère le langage comme étant fixé) ou bien le langage (auquel cas on considère le monde comme étant fixé). Il n'est néanmoins pas farfelu (et il est même probablement nécessaire) d'envisager une *combinaison* de variations. Plus précisément, on pourrait considérer des variations possibles et simultanées du monde *et* du langage, de sorte que notre sémantique tienne compte de la complexité de leur interrelation. Des auteurs ont ainsi pu refuser que la théorie des modèles devait être comprise tel que le stipule Etchemendy : « [In] standard model theory [...] both 'world' and 'language' get varied, in an interrelated, controlled manner » (Shapiro, 1998, p.149).

1. Rappelons que Bolzano s'intéresse d'abord aux *représentations en soi* (*Vorstellung an sich*) et aux *propositions en soi* (*Satz an sich*), qui ne sont pas des entités linguistiques (Bolzano, 2011, §19) ; c'est seulement moyennant une adaptation, au prix d'une simplification de ses idées (Rusnock et Burke, 2011), qu'on peut associer Bolzano à Tarski. Etchemendy reconnaît ce point (Etchemendy, 1990, n.2 et n.5, p.162-163), mais semble lui-même en négliger la portée.

On comprendra mieux l'importance de ce point, pour une sémantique digne de ce nom, à l'aide d'une analogie offerte par John MacFarlane (MacFarlane, 2000).

Supposons qu'on s'intéresse aux systèmes de coordonnées terrestres, à savoir aux systèmes représentant la position de points sur la surface de la Terre. Ces systèmes font en général appel à des principes spécifiant le lien entre coordonnées et positions. On peut par exemple penser à un principe comme le suivant : la latitude et la longitude d'un bateau dépendent du positionnement de celui-ci sur la surface de la Terre. Étant donné un système particulier de coordonnées déterminant une échelle de latitudes et de longitudes, on pourra par exemple affirmer que c'est le positionnement d'un objet qui détermine sa latitude et sa longitude par rapport au système de coordonnées en question. Par contre, sans une connaissance du fonctionnement de *ce* système, on ne pourrait pas déterminer ce que sont les latitudes et les longitudes possibles, pour *ce* système, en considérant seulement les positionnements possibles du bateau à la surface de la Terre. C'est qu'en effet, le principe liant coordonnées et positionnement *présuppose* un ensemble de latitudes et de longitudes possibles, à savoir celles qui sont déterminées par le système de coordonnées lui-même (MacFarlane, 2000, p.10). Autrement dit, le système de coordonnées détermine des latitudes et des longitudes possibles qu'il est nécessaire de connaître pour déterminer la latitude et la longitude d'un bateau à partir de sa position. Les variations possibles de coordonnées sont donc fonction de ce système de coordonnées, et plus précisément de la manière dont celui-ci est élaboré.

Il en va de même pour le fonctionnement de notre sémantique. En effet, il semble que la valeur sémantique d'une expression est déterminée à la fois par sa signification et par l'état du monde, mais seulement par rapport à un ensemble de variations possibles, c'est-à-dire par rapport à la spécification préalable de la catégorie sémantique de cette expression. La spécification de la catégorie sémantique d'une expression revient en fait à déterminer la contribution potentielle des expressions de cette catégorie à la valeur de vérité des énoncés. On pourra ainsi différencier les groupes verbaux, les noms propres, les noms communs, et ainsi de suite, par leurs contributions sémantiques. Rappelons que c'est la raison pour laquelle on a pu interdire la construction de formules comme « Est barbu est

barbu » dans l'élaboration d'une sémantique interprétationnelle (§2.1.1.2); ceci a été rendu possible par la spécification d'un ensemble de valeurs sémantiques possibles pour la catégorie des noms propres.

Lorsqu'Etchemendy présente les approches représentationnelle et interprétationnelle de la sémantique, il omet d'indiquer cette étape préalable à la détermination des variations, tant pour les modèles envisagés comme configurations du monde que pour les modèles envisagés comme réinterprétations des expressions du langage. Ainsi, pour l'approche représentationnelle telle que la présente Etchemendy, les modèles représentent les configurations possibles du monde en assignant des éléments du domaine aux expressions du langage selon leur signification actuelle, puisque le langage est considéré fixe (§2.1.1.1). Mais on pourrait penser que la signification actuelle des expressions extra-logiques n'a qu'une importance secondaire, et que ce qui importe véritablement pour notre sémantique est l'ensemble de valeurs *possibles* que peuvent prendre les expressions extra-logiques. Comme le dit MacFarlane :

« In asking about the range of semantic values generated by variations in possible meanings and possible worlds, we presuppose a range of possible semantic values that depends on neither factor. That is, we *presuppose* a well-defined semantic category. » (MacFarlane, 2000, p.11)

Etchemendy passe sous silence le fait que les variations possibles, tant pour l'approche représentationnelle que pour l'approche interprétationnelle, ne dépendent pas directement du monde ou du langage, mais plutôt de la manière dont on élabore notre sémantique. En ce sens, on peut affirmer que les notions de configuration possible et de réinterprétation possible, pour les approches représentationnelle et interprétationnelle, ne sont pas neutres, mais déjà chargées théoriquement. On peut aussi dire qu'elles semblent aller de pair, comme les notions de latitude et de longitude. Mais ce fait est pour ainsi dire dissimulé par l'exposé d'Etchemendy, un peu comme s'il choisissait d'élaborer deux systèmes de coordonnées exclusifs, l'un se limitant à la latitude et l'autre se limitant à la longitude. On peut donc penser qu'Etchemendy est coupable de commettre un sophisme du faux dilemme à un deuxième titre, puisqu'il ne mentionne pas le fait qu'une sémantique peut

tenir compte de variations *à la fois* au niveau du monde *et* au niveau du langage.

3.1.2 Le problème de la sur-génération

Comme nous l'avons vu, Etchemendy juge que les problèmes conceptuels de la définition tarskienne se répercutent sur l'extension de celle-ci. Selon lui, le problème de la sur-génération de la définition tarskienne est dû au fait que celle-ci repose sur le principe de réduction, lequel stipule que les vérités logiques peuvent être réduites à la vérité d'énoncés quantifiés universellement (§2.3.2). Le résultat de cette réduction est que certains énoncés sont déclarés incorrectement comme des vérités (ou des faussetés) logiques. Par exemple, si l'on traite comme constantes les expressions logiques habituelles, et si notre domaine compte plus de 2 objets, la définition tarskienne du concept de conséquence semble déclarer des énoncés comme $\exists x_1 \exists x_2 (x_1 \neq x_2)$ ou $\exists x_1 \exists x_2 \exists x_3 [(x_1 \neq x_2) \& (x_1 \neq x_3) \& (x_2 \neq x_3)]$ comme des vérités logiques (Etchemendy, 1990, p.111). Les théoriciens des modèles contemporains évitent ce problème en assurant que le domaine de quantification varie d'interprétation en interprétation, de sorte que l'énoncé $\exists x_1 \exists x_2 (x_1 \neq x_2)$ soit déclaré faux dans les modèles à un seul élément, que l'énoncé $\exists x_1 \exists x_2 \exists x_3 [(x_1 \neq x_2) \& (x_1 \neq x_3) \& (x_2 \neq x_3)]$ soit déclaré faux dans les modèles à moins de 3 éléments, et ainsi de suite (§2.3.1). Ainsi, ces énoncés ne sont pas des vérités (ni des faussetés) logiques.

Certains auteurs refusent le verdict rendu par Etchemendy, affirmant que la variation des domaines d'interprétation n'est pas étrangère à Tarski (Hodges, 1986; Gomez-Torrente, 1996; Ray, 1996). Il y a en effet plusieurs articles rédigés avant 1936 dans lesquels Tarski mentionne l'importance des domaines variables. Par exemple, dans l'article de 1933 sur le concept de vérité, Tarski réfère aux travaux de Hilbert en disant la chose suivante :

« [A]nother concept of a relative character plays a much greater part than the absolute concept of truth and includes it as a special case. This is the concept of *correct* or *true sentence in an individual domain a*. By this is meant [...] every sentence which would be true in the

usual sense if we restricted the extension of the individuals considered to a given class *a*. » (Tarski, 1983, p.199)

On a ainsi suggéré que Tarski, dans son article de 1936, reprenait cette idée et acceptait, sans toutefois le mentionner explicitement, que le domaine puisse varier d'interprétation en interprétation². Puisque l'argument d'Etchemendy à l'effet que la définition tarskienne sur-génère dépend essentiellement d'une conception fixe du domaine d'interprétation, montrer que la définition tient compte de la relativisation du domaine suffit à immuniser Tarski contre cette critique. Mais une telle lecture de Tarski est plutôt faible, en ce qu'elle suppose que la variation du domaine est présente dans la définition de Tarski alors qu'on n'en trouve absolument aucune trace. Or, si la variation du domaine joue un rôle déterminant dans la définition, on est en droit de s'attendre à ce que Tarski en explique le fonctionnement, d'autant que Tarski est un logicien reconnu pour sa rigueur et sa précision. De plus, comme le souligne Timothy Bays, la notion de modèle est elle-même introduite pour résoudre un écueil spécifique, à savoir un problème concernant la richesse expressive des langages (§1.3.2); on voit par contre mal pourquoi Tarski aurait pu vouloir relativiser les domaines d'interprétation dans l'article de 1936 sans avoir en tête un problème à régler, problème qu'il n'évoque même pas.

La variation du domaine d'interprétation en interprétation n'est toutefois pas la seule stratégie pour répondre à l'argument d'Etchemendy; on pourrait aussi défendre que la critique sur ce point n'est pas décisive puisqu'elle n'entraîne pas les conséquences alléguées par Etchemendy. Autrement dit, on pourrait soutenir que la conception fixe des domaines adoptée par Tarski dans l'article de 1936 ne cause pas forcément de problèmes fâcheux pour la définition, comme le prétend Etchemendy. Il y a en effet une manière d'obtenir l'équivalent de la variation des domaines d'interprétation à partir d'une conception fixe des domaines : étant donné un ensemble d'énoncés *S*, on n'a qu'à introduire un nouveau prédicat *D*,

2. Mario Gomez-Torrente (Gomez-Torrente, 1996; 2009) défend que l'absence de mention explicite à la relativisation des domaines dans l'article de 1936 s'explique par le fait qu'il s'agissait de la conception *standard* des modèles à l'époque. Timothy Bays (Bays, 2001) et Paolo Mancosu (Mancosu, 2006; 2010) ont toutefois montré que cette question était loin de faire l'unanimité parmi les contemporains de Tarski.

tenant lieu de domaine, et à ensuite relativiser chacun des énoncés quantifiés de l'ensemble S à ce prédicat. L'énoncé $\exists x_1 \exists x_2 (x_1 \neq x_2)$, une fois le prédicat introduit, deviendra par exemple $\exists x_1 \exists x_2 [Dx_1 \& Dx_2 \& (x_1 \neq x_2)]$. On obtiendra ainsi une correspondance entre les domaines relativisés et les domaines fixes. Comme le souligne Bays, ce résultat est important puisque « every theorem concerning the collection of variable-domain models [...] can be translated into an equally interesting (and, indeed, essentially *identical*) theorem concerning the collection of fixed-domain models » (Bays, 2001, p.1711-1712). Ceci permet aussi d'éviter les problèmes associés aux *cross-term restrictions* des quantificateurs de l'approche interprétationnelle (§2.2.1), dont Etchemendy fait grand cas³. Ainsi, la conception fixe des domaines ne rencontre pas nécessairement les problèmes soulevés par Etchemendy.

Ce qu'il est intéressant de noter, en ce qui nous concerne, c'est que Tarski a lui-même recours à cette technique *avant* 1936. Par exemple, Tarski utilise le prédicat B , qu'il appelle « universe of discourse » (Tarski, 1983, p.320), dans un article de 1935 sur les fondements de l'algèbre de Boole. Plus spécifiquement, B est un ensemble primitif dont la cardinalité, c'est-à-dire le nombre d'éléments, peut varier. On a donc une manière d'axiomatiser une structure — l'algèbre de Boole — en relativisant les axiomes à un prédicat particulier (§1.1.1).

En guise de second exemple de relativisation des domaines à des prédicats, on peut aussi penser à un texte sur les limitations des théories déductives, où Tarski conclut de la manière suivante :

« In conclusion we note the following : when applying Th. 1 to special deductive theories and to general metamathematics we must restrict ourselves to those axiom systems from which it follows that there are no individuals outside the domain of discourse of the theory discussed. We could remove this restriction by means of an appropriate generalization of Th. 1. To obtain this generalization we should introduce a new notion, in fact the notion of a *sentential function intrinsic under a given variable (or constant) 'a'* or, in other words, the notion of a

3. Comme nous le verrons, cette approche du problème de la sur-génération permet aussi de voir sous un nouveau jour l'argument du « sophisme de tarski ».

property intrinsic for a given class a. Roughly speaking, a property is intrinsic for a class a if it involves exclusively elements of a , subsets of a , relations between elements of a , etc. (and not, for example, individuals outside of a). » (Tarski, 1983, p.392)

Le point sur lequel j'aimerais ici attirer l'attention n'est pas tant que Tarski a *effectivement* recours à la méthode relativisant les énoncés à un prédicat de domaine dans l'article de 1936. Après tout, il n'en est pas fait mention, pas plus que de la variation des domaines d'interprétation. Plutôt, il faut voir que le pas à franchir pour éviter le problème soulevé par Etchemendy n'est pas si grand, et, surtout, qu'il est à la portée de Tarski. On peut ainsi penser que le problème extensionnel qu'identifie Etchemendy dans la définition tarskienne n'est pas un problème robuste, dans la mesure où celui-ci est relativement facile à éviter⁴.

3.1.3 Comprendre l'argument de Tarski

Les problèmes que nous avons identifiés jusqu'ici dans la critique d'Etchemendy permettent de jeter un nouvel éclairage sur le second problème conceptuel de la définition tarskienne, à savoir que celle-ci échouerait à capturer la nécessité du concept de conséquence. Nous avons vu que l'introduction de quantificateurs posait problème pour l'approche interprétationnelle telle que présentée par Etchemendy (§2.2.1). C'est ce qui ne permettait pas, selon lui, à la définition tarskienne de capturer adéquatement la nécessité du concept de conséquence. Selon Etchemendy, c'est seulement en commettant un sophisme, le « sophisme de Tarski », que Tarski a pu affirmer que sa définition était conceptuellement adéquate (§2.2.2). Ce sophisme modal imputé à Tarski repose sur le lien entre les trois conditions suivantes (Etchemendy, 1990, p.87) :

1. La conclusion σ est une conséquence tarskienne de l'ensemble de prémisses

4. Il y a beaucoup plus à dire à ce sujet, notamment parce qu'une telle réponse à l'argument d'Etchemendy a des implications sur notre compréhension de l'exemple des ω -inférences de Tarski (§1.2.1) et sur la question des logiques d'ordre supérieur. Ce débat devient cependant rapidement technique, et je n'ai pas l'espace ici pour développer les arguments en jeu. On consultera avec profit les articles de Paolo Mancosu (Mancosu, 2006; 2010) et Mario Gomez-Torrente (Gomez-Torrente, 1996; 2009), qui se répondent en quelque sorte et offrent des points de vue opposés sur la question des domaines d'interprétation.

Γ ;

2. Tous les énoncés de l'ensemble de prémisses Γ sont vrais;
3. La conclusion σ est fausse.

Plus spécifiquement, le sophisme consiste à croire que la vérité de (1) entraîne nécessairement l'incompatibilité entre (2) et (3). Une solution au problème soulevé par Etchemendy serait de refuser la condition (1) telle que formulée en proposant qu'il s'agit plutôt d'un conditionnel comme le suivant :

4. Si la conclusion σ est une conséquence tarskienne de l'ensemble de prémisses Γ alors (2) et (3) sont nécessairement incompatibles.

Pour être accepté, un tel conditionnel devrait être défendu, ce qui semble une solution plutôt *ad hoc* au problème. Une autre solution, pour rendre l'argument valide, et ainsi pour le rendre résistant à la critique d'Etchemendy, serait de lire le raisonnement de Tarski selon l'inférence (valide) suivante :

$$\frac{\Box((P \& Q) \rightarrow \neg R)}{\Box P \rightarrow \Box(Q \rightarrow \neg R)}$$

Pour le dire simplement, nous voudrions que P , l'antécédent de la conclusion de l'inférence, soit lui-même nécessaire, de manière à éviter le saut illégitime de modalité. Voici une preuve, en déduction naturelle, de l'inférence⁵ :

$$\frac{\frac{\frac{\Box((P \& Q) \rightarrow \neg R)}{(P \& Q) \rightarrow \neg R} \Box E}{\frac{\frac{\frac{\Box P}{P} \Box E \quad \frac{\Box Q}{Q} \Box E}{P \& Q} \& I}{\frac{\frac{\neg R}{Q \rightarrow \neg R} \rightarrow I}{\Box(Q \rightarrow \neg R)} \Box I} \rightarrow I} \rightarrow E}{\Box P \rightarrow \Box(Q \rightarrow \neg R)} \rightarrow I$$

Ce qu'il faut est donc que, du fait que σ est une conséquence tarskienne de Γ pour une sélection de constantes logiques, σ est *nécessairement* une conséquence tarskienne de Γ pour une sélection de constantes logiques. Ceci donne l'implication

5. Dale Jacquette (Jacquette, 2006) donne une preuve similaire en utilisant la méthode de Fitch.

suivante : $\models P \rightarrow \models \Box P$. On a ici une règle de nécessité qu'on retrouve dans les logiques modales standards, essentielle à des systèmes comme T , S_4 et S_5 , par exemple. Cette règle ne veut pas dire, bien entendu, que $\Box P$ peut être obtenu à partir de la simple assertion de P , mais plutôt que tout théorème, pouvant être dérivé de nos hypothèses initiales par l'application correcte de règles d'inférences, est lui-même nécessaire. Autrement dit, la règle de nécessité rend compte du fait que les théorèmes sont une conséquence nécessaire d'un système logique. Afin de sauver Tarski de la critique d'Etchemendy, il faut montrer que le conditionnel requis pour éviter le « sophisme de Tarski » peut être considéré comme un théorème. Comme le sophisme modal qu'il impute à Tarski est une partie charnière de la critique d'Etchemendy, éviter le sophisme est un résultat positif : cela élimine une raison de rejeter l'analyse tarskienne.

Pour Tarski, les arguments intuitivement valides sont nécessairement valides, c'est-à-dire qu'il y a une caractéristique modale associée au concept de conséquence. On pourrait dire que les arguments valides doivent l'être pour toutes les interprétations ou dans tous les mondes possibles⁶. Dans la présentation qu'Etchemendy fait du raisonnement de Tarski, les conclusions tarskiennes ne sont pas valides dans tous les mondes possibles ; sa critique met en lumière deux problèmes convergents et reliés dans la définition de Tarski. Premièrement, en sélectionnant arbitrairement les constantes logiques et les expressions à faire varier dans un argument, on peut produire des contre-exemples à la définition tarskienne. Deuxièmement, le sophisme peut être évité si les conséquences tarskiennes sont elles-mêmes nécessaires, mais on ne peut pas obtenir cette nécessité tant qu'il n'y a pas de restrictions adéquates sur la sélection de constantes logiques. Pour résoudre ce problème, il semble qu'on doit se tourner vers un aspect jugé important par Tarski, et vers lequel la lecture d'Etchemendy semble elle-même pointer par moments : il faut envisager l'application de restrictions adéquates sur la sélection de constantes logiques. Nous nous pencherons sur cette question dans la section suivante.

6. On s'intéressera aux interprétations si l'on adopte une perspective interprétationnelle de la sémantique et aux mondes possibles si l'on adopte une perspective représentationnelle. On s'intéressera à la fois aux interprétations et aux mondes possibles si l'on refuse, comme nous l'avons fait dans la section précédente, que ces approches s'excluent mutuellement.

3.2 (Comment) Peut-on sauver Tarski du « sophisme de Tarski » ?

Etchemendy a raison d'affirmer que le sophisme qu'il attribue à Tarski est, en un sens, indépendant du choix de constantes logiques :

« This fallacy becomes quite apparent when we consider the arguments that come out valid when we include all expressions in [the set of fixed terms]. But it is crucial to recognize that the inference remains fallacious, and for exactly the same reasons, regardless of our choice of fixed terms. The fallacy may be easier to spot when we include names and predicates in [the set of fixed terms], but the inference is no less fallacious when we only hold fixed (say) the truth-functional connectives. The argument does not depend on [the set of fixed terms], and it does not get better or worse according to what we suppose the [fixed terms] to be. » (Etchemendy, 1990, p.88)

Cependant, la reconstruction qu'Etchemendy fournit de l'argument de Tarski est trompeuse, en ce qu'elle suggère que le présumé sophisme n'a *absolument* rien à voir avec la sélection de constantes logiques. D'une part, il est vrai que le passage de $\Box((P \& Q) \rightarrow \neg R)$ à $P \rightarrow \Box(Q \rightarrow \neg R)$ est invalide, peu importe quelles expressions sont considérées comme constantes. En d'autres termes, même si l'on change l'ensemble des constantes logiques, la reconstruction de l'argument demeure fallacieuse. Mais d'autre part, si l'ensemble des constantes logiques est adéquatement restreint — par exemple si, contrairement à Etchemendy, on ne le garde pas *vide* — de manière à rendre les conséquences tarskiennes nécessaires, alors il semble que l'on peut éviter le piège tendu par Etchemendy. On éliminerait ainsi l'une des objections principales formulées par Etchemendy à l'endroit de l'analyse tarskienne.

Dans l'exemple choisi par Etchemendy pour mettre en évidence les défauts de la définition tarskienne, l'ensemble des constantes logiques est *maximal*, c'est-à-dire qu'aucune expression n'y étant pas déjà comprise ne pourrait être ajoutée à l'ensemble des expressions logiques. Ceci fait en sorte que la validité de l'argument dépend seulement de la vérité *matérielle* des énoncés constituants, et non de sa *forme*. L'exemple en question est censé montrer que certaines conséquences

tarskiennes ne sont pas des conséquences intuitives, c'est-à-dire qu'on n'accepterait pas qu'elles sont des conséquences au sens ordinaire (Tarski, 2002, p.176). L'exemple sert donc à rejeter le conditionnel recherché en illustrant l'un de deux problèmes possibles pour la validité d'un argument, à savoir que 1) certaines expressions extra-logiques des prémisses Γ d'un argument $\langle \Gamma, \sigma \rangle$ sont considérées comme des constantes logiques et certaines prémisses sont rendues fausses par les substitutions, mais la conclusion σ n'est pas une vérité logique, ou 2) certaines expressions extra-logiques de la conclusion σ sont gardées fixes et celle-ci est rendue vraie par les substitutions, mais les prémisses ne sont pas logiquement fausses.

En dépit de ce contre-exemple à la définition tarskienne, Etchemendy ne produit pas un contre-exemple dans lequel les expressions fixées sont *toutes, et exclusivement*, des constantes *logiques*. Mais on peut prouver par une réduction à l'absurde, en supposant sa fausseté pour en dériver une contradiction, qu'un principe de nécessité $\models P \rightarrow \models \Box P$ doit être accepté dans l'analyse de Tarski (si les conséquences tarskiennes doivent être vues comme étant elles-mêmes nécessaires). Faisons donc une telle supposition. Il existe alors un argument de la forme $\langle \Gamma, \sigma \rangle$, valide au sens de Tarski dans un monde logiquement possible W_i , mais qui n'est pas valide dans un monde possible différent W_j . Autrement dit, on a $P \& \neg \Box P$.

Puisqu'au moins deux contraintes classiques pèsent sur les mondes possibles, à savoir qu'ils soient consistants et que chaque monde possible puisse être différencié de tous les autres mondes possibles, il doit y avoir au moins une différence entre les mondes W_i et W_j . Ces différences sont généralement comprises comme des *événements* différents ayant lieu dans les mondes et permettant de distinguer *n'importe quels* mondes entre eux, mais elles peuvent aussi être vues comme reposant sur l'*instanciation de prédicats*, c'est-à-dire que les mondes peuvent être distingués par les propriétés et relations entre individus⁷. Il suit de ceci que si le principe de nécessité est faux, alors la validité au sens de Tarski dépend de l'instanciation d'au moins un prédicat différenciant entre eux les mondes possibles.

7. La base d'individuation des mondes possibles (événements, prédicats, situations, etc.) ne change pas la nature de l'argument que je souhaite développer.

Ceci, toutefois, n'est pas la manière de déterminer la validité d'un argument, au sens de Tarski. Un argument $\langle \Gamma, \sigma \rangle$ est valide dans un monde possible W_i si et seulement si toutes les séquences (avec toutes les expressions logiques fixées, et seulement ces expressions) préservent la satisfaction pour les fonctions propositionnelles du schéma d'argument $\langle \Gamma^*, \sigma^* \rangle$ dans le monde W_i . Ainsi, on s'assure que la définition ne sur-génère pas et ne sous-génère pas. Si la préservation de la satisfaction dépendait d'un prédicat instancié dans ce monde, il y aurait alors une séquence pour laquelle toutes les prémisses seraient satisfaites dans le monde W_j mais pour laquelle la conclusion ne serait *pas* satisfaite.

Pour voir que ce n'est pas le cas, supposons que la validité, au sens de Tarski, d'un argument $\langle \Gamma, \sigma \rangle$ dans un monde W_i dépend d'un prédicat P_i , et considérons la séquence substituant aux expressions variables de l'argument le même prédicat, non instancié dans W_j . Si la validité dépendait de P_i , alors la séquence rendrait l'argument invalide. Mais, en vertu de notre hypothèse initiale, un argument est valide si, et seulement si, il préserve la satisfaction pour toutes les séquences, y compris la séquence qui substitue P_i , prédicat qui sert à distinguer les mondes possibles W_i et W_j . On est donc face à une contradiction. Il en découle que la validité au sens de Tarski ne peut pas être déterminée par l'instanciation de prédicats particuliers dans des mondes logiquement possibles différents. Il en découle aussi qu'il ne devrait pas y avoir d'argument $\langle \Gamma, \sigma \rangle$ valide dans certains mondes possibles sans être également valide dans tous les mondes possibles, et donc que les conséquences tarskiennes sont elles-mêmes nécessaires, en ce qu'un argument valide dans un monde possible est valide dans tous les mondes possibles.

On a donc une preuve du principe de nécessité $\models P \rightarrow \Box \models P$ et une raison de croire que les conséquences tarskiennes sont nécessaires. Rappelons que dans la reconstruction d'Etchemendy, P signifie que σ est une conséquence tarskienne de Γ pour une sélection de constantes logiques. Pris avec l'inférence modale (valide) de $\Box((P \& Q) \rightarrow \neg R)$ à $\Box P \rightarrow \Box(Q \rightarrow \neg R)$, le principe de nécessité suffit à contrer l'objection d'Etchemendy à l'effet que Tarski commet un sophisme modal. On a ainsi la modalité recherchée entre Q (*All the sentences of Γ are true*) et $\neg R$ (*σ is not false*) dans la définition de Tarski, exactement là où Etchemendy croit

qu'elle devrait être pour que l'argument soit accepté (Etchemendy, 1990, p.86).

Il est clair, en lisant l'article de 1936, que Tarski ne cautionne en aucun cas les sélections arbitraires de constantes logiques dans l'application du critère de préservation de satisfaction. Au contraire, la définition du concept de conséquence est présentée en termes de *modèles*, compris comme des séquences arbitraires d'objets satisfaisant des fonctions propositionnelles. Ces fonctions propositionnelles sont obtenues à partir d'énoncés dans des arguments, en substituant *uniformément et exclusivement* les expressions extra-logiques par des variables⁸. Loin de prétendre avoir présenté une définition absolument adéquate du concept de conséquence, Tarski reconnaît qu'il puisse y avoir des difficultés avec l'extension de son analyse (Tarski, 2002, p.188). Ces remarques sont d'ailleurs en conformité avec les doutes qu'il soulève lui-même sur la possibilité d'une analyse formelle entièrement satisfaisante du concept intuitif.

Le problème le plus important souligné par Tarski est celui de la démarcation entre expressions logiques et extra-logiques (§1.3.3.2). Sans cette distinction, la définition ne peut pas être adéquate. Ceci veut donc dire que l'intention de Tarski était de spécifier et de fixer *exclusivement* les expressions logiques, et *pas* les expressions extra-logiques, même si force est d'admettre qu'une démarcation tranchée est difficile, sinon impossible, à opérer (Tarski, 2002, p.188). Tarski reconnaît qu'il ne dispose pas d'un critère permettant de distinguer entre ces deux ensembles d'expressions. Il concède même, comme on l'a vu, que dans le cas limite, on pourrait imaginer que toutes les expressions d'un langage, par exemple pour un langage purement formel, soient considérées comme des constantes logiques. Dans ce cas, l'analyse en termes de préservation de la satisfaction ferait coïncider extensionnellement conséquences formelle et matérielle (Tarski, 2002, p.188).

Pour Tarski, dans son article de 1936, la question de (la possibilité de) la démarcation des expressions logiques et extra-logiques et sa justification philosophique demeure ouverte (Tarski, 2002, p.188). *En un sens*, ceci a pour effet de relativiser la conséquence logique à une sélection arbitraire d'expressions logiques. Toute-

8. La variante utilisant la *satisfaction directe*, telle que nous l'avons présentée dans le deuxième chapitre, arrive essentiellement au même résultat.

fois, Tarski n'accepte *jamais* que la conséquence logique repose sur une sélection *arbitraire* de constantes logiques. Ceci enlève beaucoup de mordant à la critique d'Etchemendy, puisque celle-ci repose sur une lecture quelque peu distordue de la définition tarskienne et de son application : en aucun cas Tarski ne dit que la conséquence logique est relative à une sélection d'expressions fixées, mais seulement relative à des expressions *logiques*, peu importe la manière de définir ou de déterminer celles-ci. L'une des objections principales d'Etchemendy, à savoir que Tarski aurait commis un « sophisme modal », repose ainsi sur une présentation problématique des arguments de Tarski.

CONCLUSION

Maintenant que nous avons présenté la définition tarskienne du concept de conséquence (Chapitre 1), les arguments d'Etchemendy contre celle-ci (Chapitre 2) et quelques limites de ces arguments (Chapitre 3), il est temps de dresser un bilan de l'argumentation que nous avons développée au cours de ce mémoire.

Dans le premier chapitre, j'ai tenté de défendre l'idée que l'article de 1936 de Tarski devait être compris dans le cadre d'un projet plus vaste dans le domaine des métamathématiques, projet que Tarski appelle l'*intuitionnisme formaliste*. Ce projet se caractérise notamment par le fait qu'il vise à capturer des concepts, dont on a une saisie intuitive préalable, en en donnant une définition à l'intérieur de théories déductives, qui agissent à titre de contraintes formelles sur les définitions. On a donc minimalement deux conditions à respecter, qui sont des conditions importantes dans l'ensemble des travaux de Tarski, à savoir que la définition qu'on donne des concepts soit « matériellement adéquate » et « formellement correcte ». Mettre l'accent sur l'intuitionnisme formaliste constitue une lecture originale de Tarski, dans la mesure où peu d'auteurs ont explicitement présenté l'oeuvre de Tarski en ces termes. Cette lecture a l'avantage de faire la lumière sur le type de définition qui intéresse Tarski, soit la définition de termes au sein de systèmes déductifs, et aussi de mieux situer ses travaux dans leur contexte historique, en les opposant par exemple à la position formaliste de Hilbert.

J'ai aussi montré, dans ce chapitre, en quoi la définition tarskienne du concept de conséquence est une définition *sémantique*, ce qui en fait l'originalité conceptuelle et historique, en ce qu'elle repose sur la notion de satisfaction et sur la distinction

entre niveaux de langages. Ceci a notamment permis de montrer comment cette définition est une réponse à la fois aux problèmes rencontrés par les tentatives syntaxiques, qui n'arrivent pas à compter les ω -inférences parmi la classe des conséquences logiques et qui font face au problème de l'incomplétude démontré par Gödel, et par la tentative « sémantique » de Carnap, qui rencontre les limites d'une conception substitutionnelle de la quantification. Enfin, j'ai montré que la définition de Tarski repose en grande partie sur la démarcation entre deux catégories d'expressions, les termes logiques et les termes extra-logiques, et que l'article de 1936 ne propose aucun critère de démarcation entre ces deux types de termes. J'ai plutôt soutenu, en me basant sur les données textuelles présentes dans l'article, que Tarski fait appel à des considérations pragmatiques pour établir cette démarcation.

Dans le second chapitre, j'ai tenté de montrer comment la définition tarskienne prête flanc à la critique, tant sur le plan intensionnel que sur le plan extensionnel, en examinant les principaux arguments de John Etchemendy. La critique d'Etchemendy est une critique *interne*, comme nous l'avons exposé dans ce chapitre ; elle vise à critiquer la définition tarskienne *à partir de* la théorie des modèles, et non à partir d'une perspective *externe*, comme peuvent le faire les auteurs travaillant en théorie de la preuve.

J'ai d'abord montré que la définition tarskienne devait être complétée en précisant le sens de la notion de sémantique, et plus particulièrement le sens de la notion d'interprétation. C'est d'ailleurs là que se situe le premier problème conceptuel identifié par Etchemendy, qui juge problématique la définition tarskienne puisqu'elle ne ferait pas appel à une approche représentationnelle de la sémantique. J'ai aussi montré en quoi l'approche interprétationnelle (ou tarskienne) de la sémantique débouche sur le « sophisme de Tarski », seconde erreur conceptuelle selon Etchemendy, puisqu'elle n'arriverait pas à donner une caractérisation adéquate d'une catégorie d'expressions de la logique du premier ordre, celle des quantificateurs. Ainsi, comme l'ensemble des constantes logiques sur lequel la définition doit reposer ne peut pas être caractérisé adéquatement d'une manière unique, la définition n'arrive pas à caractériser adéquatement la nécessité associée au concept

de conséquence.

Enfin, j'ai montré, toujours selon Etchemendy, que ces problèmes conceptuels sont couplés à un problème de nature extensionnelle, à savoir que la définition tarskienne n'arriverait pas à capturer adéquatement la classe des conséquences logiques, un peu à la manière dont les approches syntaxiques échouaient à compter les ω -inférences comme des arguments valides. Le double problème de sur-génération et de sous-génération de la définition tarskienne serait imputable au fait que celle-ci repose sur le principe de réduction, qui consiste à affirmer qu'on peut réduire la vérité logique d'énoncés à la vérité de généralisations universelles. Ceci a pour effet d'introduire une influence *extra-logique* dans notre définition du concept de conséquence logique.

La critique formulée par Etchemendy à l'endroit de la définition tarskienne n'est toutefois pas sans failles, comme j'ai essayé de le montrer dans le troisième chapitre. Premièrement, la distinction entre l'approche représentationnelle de la sémantique, que préconise Etchemendy, et l'approche interprétationnelle, qu'il attribue à Tarski, est en fait un faux dilemme. En effet, cette distinction n'épuise pas l'ensemble des possibilités de théories sémantiques, c'est-à-dire qu'il existe à la fois plusieurs variantes représentationnelles et plusieurs variantes interprétationnelles de la sémantique. La présentation que fait Etchemendy de cette distinction tend aussi à rendre inconciliables les variations simultanées du monde et du langage, alors qu'on peut faire varier les deux dimensions simultanément.

Deuxièmement, lorsqu'il argumente à l'effet que la définition tarskienne est inapte à capturer adéquatement l'extension du concept de conséquence logique, Etchemendy insiste sur le fait que la définition tarskienne omet, contrairement à la théorie des modèles contemporaine, de faire varier le domaine d'interprétation des modèles. Il ne mentionne toutefois pas que Tarski peut obtenir l'effet de la variation dans sa définition en introduisant un prédicat tenant lieu de domaine, c'est-à-dire en relativisant les domaines à ce prédicat. Autrement dit, Etchemendy exagère l'importance de la variation du domaine et ses conséquences sur la définition tarskienne.

Troisièmement, dans la reconstruction qu'il donne du « raisonnement de Tarski », Etchemendy omet de mentionner qu'il existe une manière simple de lire l'argument de Tarski, d'une manière qui n'est pas fallacieuse, à savoir en introduisant une modalité portant sur l'antécédent du conditionnel voulant que la conclusion σ d'un argument soit une conséquence tarskienne d'un ensemble de prémisses Γ . À partir de notre critique sur chacun de ces trois axes, on a ainsi pu voir que les arguments d'Etchemendy reposaient en quelque sorte sur des distorsions du propos de Tarski, ou à tout le moins sur une lecture ne lui étant pas suffisamment charitable.

Bien entendu, nous nous sommes concentrés, dans ce mémoire, sur quelques aspects centraux de la critique d'Etchemendy ; nous en avons laissé certains autres de côté, lesquels occupent une place moins importante dans les débats contemporains, tel que le recours par Etchemendy au « squeezing argument » de Georg Kreisel¹. En vertu de ce choix méthodologique, nous n'avons pu qu'effleurer, par exemple, la question des logiques d'ordre supérieur ou la reprise que fait Etchemendy de la notion d'intuition.

La critique d'Etchemendy, malgré les défauts que nous avons tenté de mettre en évidence, a toutefois le mérite de nous forcer à revoir ou à faire une nouvelle lecture de la définition tarskienne. Elle nous oblige ainsi à préciser conceptuellement la définition, tout comme elle nous oblige à nous questionner sur son développement historique et sur sa place dans la logique et dans la philosophie de la logique contemporaines. Il s'agit donc d'une contribution inestimable à la philosophie de la logique, comme en témoigne la littérature à laquelle elle a donné lieu, que ce soit pour venir à la défense de Tarski ou pour ajouter à la critique. C'est pour en souligner l'importance que j'ai tenté, dans le troisième chapitre, de montrer une manière de donner suite aux arguments d'Etchemendy et, pour ainsi dire, de réhabiliter Tarski.

Ce travail est loin d'épuiser ce qui peut être dit au sujet du concept de conséquence

1. On consultera avec profit l'article de Peter Smith (Smith, 2011) pour une présentation générale des « squeezing arguments » et l'article d'Owen Griffiths (Griffiths, 2014) pour une critique de l'utilisation qu'en fait Etchemendy.

logique, comme j'espère l'avoir montré dans les paragraphes qui précèdent. Les recherches menées dans le cadre de mon mémoire m'ont aussi conduit à m'intéresser à des questions apparentées, comme la question du pluralisme logique. Le pluralisme logique est un ensemble de thèses à l'effet qu'il peut exister plusieurs systèmes logiques également admissibles ou corrects. Une porte d'entrée au pluralisme consisterait par exemple à penser qu'il existe plusieurs caractérisations adéquates du concept de conséquence logique. En particulier, les travaux de Tarski semblent eux-mêmes déboucher sur une forme de pluralisme, en ce sens notamment que Tarski n'offre pas lui-même un critère permettant la démarcation stricte des expressions logiques et extra-logiques.

Un pluralisme d'inspiration formaliste, reprenant en la généralisant la définition sémantique du concept de conséquence donnée par Tarski, a récemment été défendu par Jc Beall et Greg Restall (Beall et Restall, 2006). Plusieurs problèmes se posent cependant pour le pluralisme que ces auteurs développent. On peut notamment juger problématique le fait que Beall et Restall négligent les conséquences non transitives et non réflexives (Caret et Hjortland, 2015), puisque de telles conséquences ont des applications dans la résolution de paradoxes (Ripley, 2013), ou encore le fait qu'ils négligent la logique linéaire, pourtant centrale d'un point de vue sous-structural (Girard, 1989). Un autre problème du pluralisme de Beall et Restall est qu'il est limité à des conséquences *sémantiques*². J'aimerais, dans mes recherches futures, tenter de résoudre en partie ces problèmes du pluralisme logique, en étudiant par exemple la possibilité de considérer l'idée de normalisation des preuves (Gentzen, 1955), qui est en quelque sorte une contrainte sur les manipulations syntaxiques et qui est liée aux propriétés des systèmes logiques (décidabilité, consistance), comme critère permettant la discussion du pluralisme.

J'espère que ce mémoire, par un examen de la définition sémantique donnée par Tarski du concept de conséquence logique, a pu faire voir la richesse des débats en philosophie de la logique, et la contribution philosophique que peuvent avoir des résultats techniques. Jaakko Hintikka, à mon avis, a vu juste en affirmant la

2. Restall a adopté une perspective preuve-théorique pour prolonger ce pluralisme (Restall, 2014).

chose suivante :

« Quite often, developments in the more technical parts of logic have fully as much philosophical interest as the work that is done in the soi-disant “philosophical logic”, even though most of such philosophical implications have not caught the eye of the majority of “philosophical” logicians. » (Hintikka, 1988, p.20)

RÉFÉRENCES

- Aristote (2014). *Premiers Analytiques*. Paris : Flammarion.
- Bays, T. (2001). On Tarski on Models. *Journal of Symbolic Logic*, 66 : 1701–1726.
- Beall, J. et Restall, G. (2006). *Logical Pluralism*. Oxford : Clarendon Press.
- Belnap, N. (1962). Tonk, Plonk, and Plink. *Analysis*, 22 : 130–134.
- Belnap, N. (1993). On Rigorous Definitions. *Philosophical Studies*, 72 : 115–146.
- Bolzano, B. (2011). *Théorie de la science*. Paris : Gallimard.
- Brandom, R. (1994). *Making it Explicit : Reasoning, Representing, and Discursive Commitment*. Cambridge, MA : Harvard University Press.
- Caret, C. R. et Hjortland, O. T. (2015). *Foundations of Logical Consequence*. Oxford : Oxford University Press.
- Carnap, R. (1934/1937). *The Logical Syntax of Language [Logische Syntax der Sprache]*. London : Kegan Paul.
- Davidson, D. (1967). Truth and Meaning. *Synthese*, 17 : 304–323.
- Davidson, D. (1973). In Defense of Convention T. Dans Leblanc, H. (dir.), *Truth, Syntax and Modality*. Amsterdam : North-Holland.
- Detlefsen, M. (2005). Formalism. Dans Shapiro, S. (dir.), *The Oxford Handbook of Philosophy of Mathematics and Logic (p. 236–317)*. Oxford University Press.

- Dummett, M. (1991). *The Logical Basis of Metaphysics*. Cambridge, MA : Harvard University Press.
- Dutilh Novaes, C. (2011). The Different Ways in which Logic is (said to be) Formal. *History and Philosophy of Logic*, 32(4) : 303–332.
- Dutilh Novaes, C. (2012). *Formal Languages in Logic : A Philosophical and Cognitive Analysis*. Cambridge : Cambridge University Press.
- Dutilh Novaes, C. (2014). The Undergeneration of Permutation Invariance as a Criterion for Logicality. *Erkenntnis*, 79 : 81–97.
- Etchemendy, J. (1983). The Doctrine of Logic as Form. *Linguistics and Philosophy*, 6 : 319–334.
- Etchemendy, J. (1988a). Tarski on Truth and Logical Consequence. *Journal of Symbolic Logic*, 53 : 51–79.
- Etchemendy, J. (1988b). Models, Semantics and Logical Truth. *Linguistics and Philosophy*, 11(1) : 91–106.
- Etchemendy, J. (1990). *The Concept of Logical Consequence*. Cambridge, MA : Harvard University Press.
- Etchemendy, J. (2008). Reflections on Consequence. Dans Patterson, D. (dir.), *New Essays on Tarski and Philosophy*, chap. 11 (p. 263–299). New York : Oxford University Press.
- Field, H. (1972). Tarski's Theory of Truth. *The Journal of Philosophy*, 64(12) : 347–375.
- Forbes, G. (1997). Essentialism. Dans Hale, B. et Wright, C. (dir.), *A Companion to the Philosophy of Language* (p. 515–533). Oxford : Blackwell.
- Frege, G. (1879/1999). *L'Idéographie [Begriffsschrift]*. Paris : Librairie Philosophique J. Vrin.
- Gentzen, G. (1934-1935/1955). *Recherches sur la déduction logique [Untersuchungen über das Logische Schliessen]*. Paris : Presses Universitaires de France.

- Girard, J.-Y. (1989). *Proofs and Types*. Cambridge : Cambridge University Press.
- Gödel, K. (1931/1967). On formally undecidable propositions of Principia Mathematica and related systems. Dans van Heijenoort, J. (dir.), *From Frege to Gödel : A Source Book in Mathematical logic, 1979-1931*. Cambridge, MA : Harvard University Press.
- Gomez-Torrente, M. (1996). Tarski on Logical Consequence. *Notre Dame Journal of Formal Logic*, 37 : 125–151.
- Gomez-Torrente, M. (1998). On a Fallacy Attributed to Tarski. *History and Philosophy of Logic*, 19 : 227–234.
- Gomez-Torrente, M. (1999). Logical Truth and Tarskian Logical Truth. *Synthese*, 117(3) : 375–408.
- Gomez-Torrente, M. (2009). Rereading Tarski on Logical Consequence. *The Review of Symbolic Logic*, 2(2) : 249–297.
- Griffiths, O. (2014). Formal and Informal Consequence. *Thought*, 3(1) : 9–20.
- Hand, M. (2007). Objectual and Substitutional Interpretations of the Quantifiers. Dans Jacquette, D. (dir.), *Philosophy of Logic (p. 649–674)*. Amsterdam : North-Holland Elsevier.
- Hanson, W. H. (1997). The Concept of Logical Consequence. *The Philosophical Review*, 106 : 365–409.
- Hart, H. L. (1961). *The Concept of Law*. Oxford : Oxford University Press.
- Hintikka, J. (1988). On the Development of the Model-Theoretic Viewpoint in Logical Theory. *Synthese*, 77(1) : 1–36.
- Hodges, W. (1986). Truth in a Structure. *Proceedings of the Aristotelian Society*, 86 : 135–151.
- Jacquette, D. (2006). Tarski's Analysis of Logical Consequence and Etchemendy's Criticism of Tarski's Modal Fallacy. Dans Jadacki, J. et Pasniczek, J. (dir.),

- The Lvov-Warsaw School : The New Generation* (p. 345–368). Amsterdam/New York : Rodopi.
- Jané, I. (2006). What is Tarski's Common Concept of Consequence? *The Bulletin of Symbolic Logic*, 12 : 1–42.
- Kirkham, R. L. (1992). *Theories of Truth : A Critical Introduction*. Cambridge, MA : MIT Press.
- Kripke, S. (1959). A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24 : 1–14.
- Kripke, S. (1975). Outline of a Theory of Truth. *The Journal of Philosophy*, 72 : 690–716.
- Lowe, E. J. (1997). Objects and Criteria of Identity. Dans Hale, B. et Wright, C. (dir.), *A Companion to the Philosophy of Language* (p. 613–633). Oxford : Blackwell.
- MacFarlane, J. (2000). What is Modeled by Truth in All Models? [Document non publié].
- MacFarlane, J. (2015). Logical Constants. Dans Zalta, E. N. (dir.), *The Stanford Encyclopedia of Philosophy*. Récupéré de <http://plato.stanford.edu/archives/fall2015/entries/logical-constants>.
- Mancosu, P. (2006). Tarski on Models and Logical Consequence. Dans Ferreiros, J. et Gray, J. (dir.), *The Architecture of Modern Mathematics : Essays in History and Philosophy* (p. 209–238). Oxford : Oxford University Press.
- Mancosu, P. (2010). Fixed- versus Variable-domain Interpretations of Tarski's Account of Logical Consequence. *Philosophy Compass*, 5(9) : 745–749.
- Mancosu, P., Zach, R., et Badesa, C. (2010). The Development of Mathematical Logic from Russell to Tarski, 1900-1935. Dans *The Adventure of Reason : Interplay between Philosophy of Mathematics and Mathematical Logic, 1900-1940* (p. 5–119). New York : Oxford University Press.

- Manzano, M. (1990/1999). *Model Theory*. Oxford : Clarendon Press.
- McKeon, M. (2010). *The Concept of Logical Consequence : An Introduction to Philosophical Logic*. New York : Peter Lang Publishing.
- Murzi, J. et Carrara, M. (2014). More Reflections on Consequence. *Logique et Analyse*, 227 : 223–258.
- Patterson, D. (dir.) (2008). *New Essays on Tarski and Philosophy*. Oxford : Oxford University Press.
- Patterson, D. (2012). *Alfred Tarski : Philosophy of Language and Logic*. Basingstoke : Palgrave Macmillan.
- Peregrin, J. (2014). *Inferentialism : Why Rules Matter*. Basingstoke : Palgrave Macmillan.
- Pitkin, H. F. (1967). *The Concept of Representation*. Berkeley : University of California Press.
- Popper, K. (1935/2002). *The Logic of Scientific Discovery*. London and New York : Routledge.
- Potter, M. (2004). *Set Theory and its Philosophy Set Theory and its Philosophy : A Critical Introduction*. New York : Oxford University Press.
- Prawitz, D. (1985). Remarks on some Approaches to the Concept of Logical Consequence. *Synthese*, 62 : 153–171.
- Prawitz, D. (2005). Logical Consequence from a Constructive Point of View. Dans Shapiro, S. (dir.), *The Oxford Handbook of Philosophy of Mathematics and Logic* (p. 671–695). Oxford : Oxford University Press.
- Prior, A. N. (1960). The Runabout Inference-Ticket. *Analysis*, 21 : 38–39.
- Ray, G. (1996). Logical Consequence : A Defense of Tarski. *The Journal of Philosophical Logic*, 25 : 617–677.

- Read, S. (1988). *Relevant Logic : A Philosophical Examination of Inference*. Oxford : Blackwell.
- Restall, G. (2000). *An Introduction to Substructural Logics*. New York : Routledge.
- Restall, G. (2005). Multiple Conclusions. Dans Hajek, P., L.Valdes-Villanueva, et Westertahl, D. (dir.), *Logic, Methodology and Philosophy of Science : Proceedings of the Twelfth International Congress* : 189–205. London : KCL Publications.
- Restall, G. (2014). Pluralism and Proofs. *Erkenntnis*, 79(2) : 279–291.
- Ripley, D. (2013). Paradoxes and failures of cut. *Australian Journal of Philosophy*, 91 : 139–164.
- Rusnock, P. et Burke, M. (2011). Etchemendy and Bolzano on Logical Consequence. *History and Philosophy of Logic*, 31(1) : 3–29.
- Ryle, G. (1949). *The Concept of Mind*. London : Hutchinson.
- Sanchez-Miguel, M. G.-C. (2003). Gomez-Torrente on Modality and Tarskian Logical Consequence. *Theoria*, 18(47) : 159–170.
- Shapiro, S. (1998). Logical Consequence : Models and Modality. Dans Schirn, M. (dir.), *The Philosophy of Mathematics Today (p. 131–156)*. Oxford : Clarendon Press.
- Shapiro, S. (2005). Logical Consequence, Proof Theory and Model Theory. Dans Shapiro, S. (dir.), *The Oxford Handbook of Philosophy of Mathematics and Logic (p. 651–670)*. Oxford : Oxford University Press.
- Shapiro, S. (2011). Varieties of Pluralism and Relativism for Logic. Dans Hales, S. D. (dir.), *A Companion to Relativism*. Wiley-Blackwell.
- Shapiro, S. (2014). *Varieties of Logics*. Oxford : Oxford University Press.
- Sher, G. (1991). *The Bounds of Logic : A Generalized Viewpoint*. Cambridge, MA : MIT Press.

- Sher, G. (1996). Did Tarski Commit 'Tarski's Fallacy'? *Journal of Symbolic Logic*, 61(2) : 653–686.
- Smith, P. (2011). Squeezing Arguments. *Analysis*, 71(1) : 22–30.
- Soames, S. (1999). *Understanding Truth*. New York : Oxford University Press.
- Sundholm, G. (2002). A Century of Inference : 1837-1936. Dans Gärdenfors, P., Wolenski, J., et Kijania-Placek, K. (dir.), *In the Scope of Logic, Methodology and Philosophy of Science*, volume II (p. 565–580). Kluwer Academic Publishers.
- Sundholm, G. (2012). 'Inference Versus Consequence' Revisited : Inference, Consequence, Conditional, Implication. *Synthese*, 187(3) : 943–956.
- Tarski, A. (1931). Sur les Ensembles Définissables de Nombres Réels. *Fundamenta Mathematicae*, 17(1) : 210–239.
- Tarski, A. (1956/1983). *Logic, Semantics, Metamathematics : papers from 1923 to 1938*. Hackett Publishing.
- Tarski, A. (1986). What are Logical Notions? *History and Philosophy of Logic*, 7 : 143–154.
- Tarski, A. (2002). On the Concept of Following Logically. *History and Philosophy of Logic*, 23 : 155–196.
- Tarski, A. (2009). Du Concept de Conséquence Logique. Dans Bonnay, D. et Cozik, M. (dir.), *Philosophie de la Logique : Conséquence, Preuve et Vérité* (p. 83–97). Paris : Librairie Philosophique J. Vrin.
- Wagner, P. (dir.) (2009). *Carnap's Logical Syntax of Language*. Basingstoke : Palgrave Macmillan.
- Wansing, H. (2000). The Idea of a Proof-Theoretic Semantics and the Meaning of the Logical Operations. *Studia Logica*, 64(1) : 3–20.
- Wittgenstein, L. (1921/1961). *Tractatus Logico-Philosophicus*. London : Routledge.