

Note sur Popper lecteur de Tarski

Philippe de Rouilhan

Institut d'histoire et de philosophie des sciences et des techniques
(CNRS, Université Paris 1, ENS)

Résumé : 1. Introduction 2. La théorie de la vérité de Tarski est-elle, comme Popper l'affirme à la suite de Tarski lui-même, une réhabilitation de la conception traditionnelle de la vérité comme correspondance aux faits ? — Oui, mais pas pour la raison qu'il donne. 3. La définition explicite de la vérité de Tarski (quand elle est possible) est-elle, comme Popper l'affirme à la suite de Tarski lui-même, purement morphologique (syntaxique) ? — Non. 4. La théorie de Tarski est-elle, comme celui-ci le prétend, « épistémologiquement neutre » ? — La thèse est équivoque, et Tarski peut aussi bien la soutenir que Popper la contester. 5. Popper a-t-il raison pour autant d'aller jusqu'à prétendre que la théorie de la vérité de Tarski fournit un argument en faveur du « réalisme métaphysique » ? — Non. 6. Conclusions.

Abstract: 1. Introduction. 2. Is Tarski's theory of truth, as Popper claims after Tarski himself, a rehabilitation of the traditional view of truth as correspondence to facts? — Yes, but not for the reasons he gives. 3. Is Tarski's explicit definition of truth (when it is possible), as Popper claims after Tarski himself, purely morphological (syntactical)? — No. 4. Is Tarski's theory, as Tarski claims it to be, « epistemologically neutral » ? — This thesis is ambiguous, and Tarski can support it just as Popper can dispute it as well. 5. Is Popper right for all that when he goes as far as to claim that Tarski's theory of truth provides an argument in favour of « metaphysical realism » — No. 6. Conclusions.

1 Introduction

La rencontre de Tarski, la découverte de sa théorie « *sémantique* » de la vérité, fut pour Popper une révélation¹. Le réalisme auquel ce dernier était prêt à souscrire semblait solidaire d'une conception de la vérité — la conception traditionnelle de la vérité comme *correspondance aux faits* — notoirement problématique. Pour Popper, la théorie de Tarski réhabilitait cette conception et offrait ainsi au réaliste ce qui lui avait manqué jusque-là. L'interprétation de Popper de la théorie de Tarski a été vivement, et même violemment, critiquée. Un philosophe français est allé jusqu'à écrire : « L'interprétation que propose Popper de Tarski est pire que fausse : c'est de la bouillie philosophique. » Le but de cet exposé est de rouvrir le procès et de porter un jugement plus serein et probablement plus équitable.

¹Popper se souvient de sa première rencontre de Tarski, en juillet 1934, à un colloque organisé par le Cercle de Vienne : « Je dois souligner qu'à cette époque [...] j'étais mal à l'aise avec la notion de vérité [...] ».

« La raison de ce malaise concernant la notion de vérité était, évidemment, que cette notion était depuis quelque temps attaquée par certains philosophes, et avec de bons arguments. Ce n'était pas tant l'antinomie du Menteur qui m'effrayait, que la difficulté d'expliquer la théorie [de la vérité-]correspondance : que pouvait être la correspondance d'un énoncé avec les faits ? Il y avait en outre une idée que, bien que je l'eusse toujours fermement refusée, je me sentais incapable de combattre effectivement. L'idée à laquelle je fais allusion est que, si nous voulons parler de la vérité, nous devrions être capables de donner un critère de vérité. Je considérais bien qu'il était cependant légitime de parler de vérité. Mais j'étais incapable de défendre mon idée que l'absence de critère de vérité ne pouvait être utilisée comme argument contre la légitimité logique de la notion de vérité.

« Je suis heureux de n'avoir jamais donné expression écrite à ce malaise particulier, qui était complètement injustifié, comme chacun ici aujourd'hui s'en rendra compte. » [Popper 1972, 319–320].

Un peu plus loin, Popper parle de « joie intense et [de] soulagement » à la découverte de la théorie de Tarski [Popper 1972, 322].

Puis, encore un peu plus loin : « Comme je l'ai mentionné auparavant, j'ai rencontré Tarski pour la première fois en juillet 1934 à Prague. C'est au début de l'année 1935 que je l'ai à nouveau rencontré, à Vienne, au colloque de Karl Menger, auquel Tarski et Gödel participaient [...]. C'est à cette époque que j'ai demandé à Tarski de m'expliquer sa théorie de la vérité, ce qu'il a fait dans une conférence de peut-être vingt minutes sur un banc (un banc que je n'oublie pas) au *Volksgarten* de Vienne. Il m'a aussi permis de voir le jeu d'épreuves de la traduction allemande de son grand article sur le concept de vérité, qu'il venait juste de recevoir de l'éditeur de *Studia Philosophica*. Il n'y a pas de mots qui puissent dire combien j'ai appris de tout cela, pas de mots qui puissent exprimer ma gratitude. Bien que Tarski ne fût qu'un peu plus âgé que moi, et bien que nous fussions, à cette époque, en termes très intimes, je le considérais comme l'homme que je pouvais vraiment regarder comme mon professeur de philosophie. Je n'ai jamais autant appris de personne d'autre. » [Popper 1972, 322].

Les pièces au dossier portant signature de nos deux protagonistes ne sont pas nombreuses. Du côté de Tarski, le fameux mémoire sur le concept de vérité — le *Wahrheitsbegriff* — aux trois éditions successives (polonaise [1933], allemande, avec un important *post-scriptum* [1935], anglaise, [1956]), et quelques articles destinés à un public plus large ([1935], [1944], [1969]); du côté de Popper, essentiellement le chapitre 9 d'*Objective Knowledge*, « Philosophical Comments on Tarski's Theory of Truth » [1972, 319–340], avec, en addendum, la reprise d'un petit article plus ancien, [1955], à quoi l'on peut ajouter les multiples passages de l'œuvre où l'auteur redit sa dette à l'égard du fondateur de la sémantique comme science, notamment la deuxième section de l'article publié comme chapitre 10 de *Conjectures and Refutations* [1963, 223–228]. Quant aux commentateurs, laissant à ce papier la légèreté d'une note, je ne m'arrêterai à aucun d'eux en particulier, exception faite pour D. Miller, grand connaisseur de Popper devant l'Éternel et son digne héritier, juste pour dire que son article sur le même sujet, « Popper and Tarski » [1999], devrait être lu en contrepoint du mien, à moins que ce ne soit l'inverse.

L'interprétation de Tarski par Popper tient essentiellement en trois thèses, dont les deux premières ont déjà été évoquées dans ce qui précède. Sur ces deux premières Popper est d'accord avec Tarski ; sur la dernière, qui se divise en deux, il ne l'est pas.

Thèse 1 - *La théorie de la vérité de Tarski est bien, comme celui-ci l'affirme lui-même, une réhabilitation de la conception traditionnelle de la vérité comme correspondance aux faits.*

Thèse 2 - *La définition explicite de la vérité de Tarski (quand elle est possible) est bien, comme celui-ci l'affirme lui-même, purement morphologique (syntaxique).*

Thèse 3 - (a) *La théorie de la vérité de Tarski n'est pas, au contraire de ce que celui-ci prétend, « épistémologiquement neutre » ; (b) elle constitue même, plus précisément, un argument en faveur du « réalisme métaphysique ».*

J'examinerai successivement les trois thèses et soutiendrai ce qui suit.

Sur la thèse 1 - *Popper a raison (avec Tarski), mais sa présentation est trompeuse*, voir § 2.

Sur la thèse 2 - *Popper a tort (avec Tarski)*, voir § 3.

Sur la thèse 3 - *3a est équivoque, et Tarski est aussi bien fondé à la soutenir que Popper à la contester*, voir § 4 ; quant à 3b, *Popper a tort (contre Tarski)*, voir § 5.

Je reprendrai l'essentiel des idées que je défends en conclusion, voir § 6.

2 La théorie de la vérité de Tarski est-elle, comme Popper l'affirme à la suite de Tarski lui-même, une réhabilitation de la conception traditionnelle de la vérité comme correspondance aux faits ? — Oui, mais pas pour la raison qu'il donne.

2.1. Rappelons d'abord l'essentiel de la théorie de la vérité de Tarski : celui-ci retient de la conception traditionnelle de la vérité comme correspondance aux faits l'idée que la définition explicite ou implicite de la notion de vérité relative aux énoncés d'un langage-objet (e.g., un fragment de l'anglais) est *matériellement adéquate* si l'on peut, dans le métalangage, en *déduire logiquement, moyennant les principes de la syntaxe du langage-objet*, toutes les *définitions partielles*, dites aussi *équivalences-T*, relatives aux énoncés du langage-objet. Exemple paradigmatique d'équivalence-T (en supposant la théorie développée en français) :

« Snow is white » est un énoncé vrai ssi la neige est blanche.

Au contraire des logiciens d'aujourd'hui, Tarski, comme Carnap et Church, par exemple, intégrait dans la notion de langage des considérations preuve-théorétiques (*proof-theoretical*). En particulier, la notion de déduction logique mobilisée ci-dessus était intégrée dans le métalangage sous la forme d'axiomes logiques et des règles d'inférence logiques ; et de même les principes de la syntaxe du langage-objet, sous la forme d'axiomes.

2.2. La théorie de la vérité de Tarski est bien, comme Popper l'affirme à la suite de Tarski lui-même, une réhabilitation de la conception traditionnelle de la vérité comme correspondance aux faits, mais ce jugement doit être affiné, et ne saurait justifier la présentation trompeuse que Popper donne de la théorie de Tarski.

Car Popper présente cette théorie en termes lourds et insistants de « correspondance » et de « faits »², sans mot dire de ce qui fait pourtant l'originalité et la simplicité de cette théorie par rapport à la tradition, à

² « Si, comme le suggère la théorie de Tarski, la vérité est la correspondance aux faits, alors abandonnons un instant complètement le mot « vérité » et ne parlons à la place que de « la correspondance des énoncés aux faits qu'ils décrivent ».

savoir qu'elle réussit à faire droit à l'intuition fondamentale de la conception traditionnelle de la vérité comme correspondance avec les faits *tout en faisant l'économie des notions problématiques de correspondance et de fait*. La théorie de la vérité de Tarski ne peut être légitimement interprétée comme *réhabilitation* de la conception traditionnelle de la vérité comme correspondance aux faits que dans la mesure où elle est d'abord interprétée à partir d'elle *dans les termes de cette économie*.

On pourrait imaginer un logicien, appelons-le Popski, qui aurait élaboré une théorie de la vérité fondée sur une ontologie de *faits* et sur une analyse de la notion générale (non exclusivement sémantique) de *correspondance*, théorie selon laquelle la vérité devrait être comprise comme *correspondance aux faits* :

s est vrai ssi_{df} il existe un fait x tel que s correspond à x

« C'est, je pense, l'impossibilité apparente de découvrir ou d'expliquer cette correspondance qui a rendu toutes les théories pré-tarskiennes de la vérité-correspondance si suspectes. [...] »

« Maintenant soyons audacieux et supposons sérieusement qu'il y a des énoncés qui correspondent aux faits. Toute théorie ayant à traiter de cette situation doit être capable de parler (1) des énoncés d'un certain langage, que nous appelons [...] le langage-objet, et (2) des faits et de ce qui se présente comme des faits.

« (1) Pour parler des énoncés, nous devons avoir à notre disposition des *noms* des énoncés. [...] Cela veut dire que toute théorie de la correspondance doit être formulée dans un métalangage [...]. »

« (2) Pour parler d'une relation quelconque entre les énoncés et les faits nous devons avoir à notre disposition des descriptions des faits ; c'est-à-dire que nous devons être capables de décrire dans notre métalangage tous les faits que nous pouvons décrire dans le langage-objet. [...] »

« Ainsi constatons-nous que toute théorie traitant de la correspondance entre les énoncés et les faits et, par conséquent, d'une certaine relation entre les énoncés et les faits, doit être formulée dans un métalangage qui, en dehors des mots logiques usuels, a à sa disposition trois types d'expressions :

« (1) Des noms d'énoncés [...] d'un certain langage-objet [...]. »

« (2) Des énoncés décrivant les faits (y compris les non-faits) en discussion dans ce langage-objet [...]. »

« (3) [Des termes sémantiques, par exemple : « X correspond aux faits ».] » [Popper 1972, 324–325].

Et un peu plus loin : « En enseignant la théorie de la vérité de Tarski, j'ai trouvé que les choses étaient rendues plus faciles, au moins pour certains de mes élèves, si je parlais de cette manière de la *correspondance aux faits* plutôt que de la vérité. [...] »

« [P]arler de la correspondance aux faits (au lieu de la vérité) semble être d'un réel secours pour certains élèves. Cela leur permet de voir plus clairement que, et pourquoi, l'énoncé qui prend la place de [...] « p » [dans « P correspond aux faits si et seulement si p »] est, et pourquoi il *doit* être, un énoncé métalinguistique d'un certain *fait* (ou de quelque chose qui se présente comme un fait), c'est-à-dire la description métalinguistique d'une certain état de choses décrit aussi dans le langage-objet. » [Popper 1972, 326–327].

(où « s » est une variable parcourant la classe des énoncés du langage-objet). Soit dit en passant, pour autant qu'on puisse l'imaginer, une telle théorie aurait sur celle de Tarski l'avantage considérable de pouvoir s'appliquer à un langage-objet variable, indéterminé, inconnu, et non pas seulement, comme celle de Tarski, à des langages-objets explicitement donnés avec leur traduction dans le métalangage. Mais ce que Popski aurait fait, Tarski ne l'a pas fait. Popper non plus, d'ailleurs, ni personne.

2.3. On m'objectera peut-être que le fait même que Tarski ait pu éliminer les notions de correspondance et de fait de la théorie de la vérité montre que l'usage de ces notions, en ce qui concerne la vérité, est inoffensif et donc légitime.

Derrière l'objection, je crois entendre l'adage de Quine : « Définir, c'est éliminer », ou plutôt sa converse. Éliminer, ce serait définir, ou du moins légitimer. Mais il n'en est rien. Tarski élimine les notions de correspondance et de fait, mais il ne les définit nullement, il n'en légitime nullement l'usage.

Bien sûr, au lieu de dire que « s est vrai », on peut toujours dire que « s correspond aux faits », ou quelque chose de ce genre, et l'on peut définir l'un par l'autre :

$$s \text{ correspond aux faits ssi}_{df} s \text{ est vrai,}$$

ou, si l'on préfère :

$$s \text{ correspond à un fait ssi}_{df} s \text{ est vrai,}$$

ou même :

$$\text{il existe un fait } x \text{ tel que } s \text{ correspond à } x \text{ ssi}_{df} s \text{ est vrai,}$$

mais de telles définitions contextuelles des notions en cause n'en autorisent qu'un usage tellement défectif qu'elles sont à peu près dénuées d'intérêt. En particulier elles ne justifient aucun usage séparé de ces notions, ce qu'on pourrait mettre en relief à l'aide de tirets, par exemple :

$$s \text{ correspond-aux-faits ssi}_{df} s \text{ est vrai.}$$

La comparaison avec le cas du calcul infinitésimal sera peut-être éclairante. Dans leur reconstruction du Calcul, Bolzano, Cauchy et Weierstrass éliminaient la notion leibnizienne d'*infinitésimal*, et ils le faisaient par voie de définition contextuelle permettant d'en gouverner l'usage sur un mode *simulatoire*. Le point crucial est que cette définition, mobilisant la notion préalablement définie de *limite*, était assez subtile pour

montrer dans quelle large mesure, suffisante pour les besoins essentiels du Calcul, l'usage de cette notion était inoffensif et donc justifié. Rien de comparable avec les définitions ci-dessus envisagées, dont le pouvoir de simulation est dérisoire. Ni, *a fortiori*, chez Tarski, qui laisse les notions de correspondance et de fait tomber en déshérence, et ne cherche pas plus à les définir et à en innocenter l'usage que ne le faisait Laplace pour la notion de Dieu en élaborant un système dont elle était purement et simplement absente.

2.4. Popper se livre à des remarques non sans rapport avec les considérations précédentes (voir § 2.3), mais bien différentes³. Si la vérité a été définie en termes de correspondance et de faits, dit-il à peu près, il doit être possible, inversement, de définir (en un sens propre, non contextuel, non simulateur) les faits, la réalité, en termes de vérité et de correspondance supposés disponibles :

la réalité =_{df} les faits =_{df} ce à quoi les énoncés vrais correspondent,
 x est un fait ssi_{df} il existe s tel que s est vrai & s correspond à x .

Certes, et je dirais même plus. La vérité ayant en fait été définie par Tarski sans mobiliser aucune notion problématique comme celles de correspondance et de fait, elle *est* disponible, et la notion de fait se trouve ainsi suspendue à la *seule* notion de correspondance.

Mais à aucun moment, ni chez Tarski ni chez Popper, la moindre lumière n'a été jetée sur la notion de correspondance, ce dont — faut-il s'en étonner ? — Popper ne semble avoir cure. Si bien que, quand même cette notion serait formellement disponible et partant les définitions ci-dessus formellement correctes, les notions de fait et de réalité garderaient tout leur mystère.

³« La théorie de Tarski nous permet de *définir la vérité* comme correspondance aux faits ; mais nous pouvons l'utiliser pour *définir la réalité* comme ce à quoi les énoncés vrais correspondent. Par exemple, nous pouvons distinguer les *faits réels*, c'est-à-dire, les faits (prétendus) qui sont réels, des *faits (prétendus) qui ne sont pas réels* (c'est-à-dire, des non-faits). Ou, pour le dire plus explicitement, nous pouvons dire qu'un fait prétendu [...] est un fait réel si et seulement si l'énoncé qui le décrit [...] est *vrai* ; sinon le fait prétendu n'est pas un fait réel (ou, si vous préférez le dire ainsi, ce n'est pas un fait du tout).

« Et tout comme Tarski nous permet de remplacer le terme « vérité » par « l'ensemble des énoncés (ou phrases) vrais », nous pouvons remplacer le terme « réalité » par « l'ensemble des faits réels ».

« Je suggère donc que, si nous pouvons définir le concept de vérité, nous pouvons définir aussi le concept de réalité. [...] Je veux simplement dire que, s'il est possible de définir « vérité » comme « correspondance aux faits », ou, ce qui revient au même, comme « correspondance à la réalité », alors il est également possible de définir « réalité » comme « correspondance à la vérité ». » [Popper 1972, 329]

2.5. Dans le passage de la conception traditionnelle de la vérité à la théorie de Tarski, les notions de correspondance et de fait ont disparu. Popper comprend la nouvelle théorie dans l'horizon de la tradition, en fermant les yeux sur cette disparition. D'autres philosophes, inversement, aveuglés par cette disparition, en sont venus à dénier que la théorie de Tarski eût le moindre lien avec la conception traditionnelle. Au lieu de mettre un énoncé tel que « Snow is white » en correspondance avec le monde, comme le ferait une définition partielle de la vérité fidèle à la tradition :

« Snow is white » est vrai
ssi il existe un fait x tel que « Snow is white » correspond à x ,

la définition partielle de Tarski (déjà mentionnée) :

« Snow is white » est vrai ssi la neige est blanche,

au dire de ces philosophes, se contente de mettre en relation l'énoncé en question (« Snow is white ») avec un autre énoncé (« La neige est blanche »).

L'argument est pour le moins étonnant. La définition partielle de Tarski parle bien de l'énoncé « Snow is white », qu'elle mentionne, pour ce faire, du côté gauche, mais non de l'énoncé « La neige est blanche », qu'elle ne mentionne pas, mais utilise simplement, du côté droit, pour parler de la neige sous le rapport de la blancheur. La définition partielle de Tarski met bien l'énoncé « Snow is white » en relation avec le monde. Elle le fait tout autant, quoique d'une autre manière, que la définition partielle de style traditionnel. Les auteurs de l'argument feraient bien de réfléchir à la différence qu'il y a entre la définition partielle de Tarski et la sortie correspondante d'un manuel de traduction anglais-français :

« Snow is white » se traduit par « La neige est blanche ».

3 La définition explicite de la vérité de Tarski (quand elle est possible) est-elle, comme Popper l'affirme à la suite de Tarski lui-même, purement morphologique (syntaxique) ? — Non.

3.1. Un deuxième rappel, sur la façon dont Tarski définit explicitement la vérité pour un langage-objet (quand c'est possible). A son instar,

lui qui envisage, à des fins pédagogiques, des langages-objets auxquels n'appartiennent qu'un nombre fini de phrases [Tarski 1935 ; 1956, 188 ; 1969], choisissons pour langage-objet, \mathcal{L}_0 , le fragment de l'anglais ne contenant que les deux phrases « Snow is white » et « Grass is blue », et pour métalangage, \mathcal{M}_0 , un fragment convenable du français, contenant entre autres choses les traductions françaises respectives, « La neige est blanche » et « L'herbe est bleue », des deux phrases en question. La définition explicite tarskienne de la vérité pour \mathcal{L}_0 dans \mathcal{M}_0 est alors la suivante :

s est vraie ssi_{df}
 s est « Snow is white » et la neige est blanche ou
 s est « Grass is blue » et l'herbe est bleue.

Cette définition répond au critère d'adéquation formulé par Tarski (voir § 2.1), avec la particularité remarquable que cette définition suffit, à elle seule, pour la déduction logique des deux équivalences-T :

« Snow is white » est vraie ssi la neige est blanche,
 « Grass is blue » est vraie ssi l'herbe est bleue.

En général, la définition explicite tarskienne de la vérité pour un langage-objet, \mathcal{L} , dans un métalangage, \mathcal{M} , ne permet de déduire logiquement toutes les équivalences-T que sur un mode enthymématique, en s'appuyant sans mot dire sur les principes de la syntaxe de \mathcal{L} .

Ainsi, avec une définition explicite matériellement adéquate de la vérité pour \mathcal{L} dans \mathcal{M} , les équivalences-T relatives aux phrases de \mathcal{L} sont ce qu'on pourrait appeler des *théorèmes syntaxiques* de \mathcal{M} , et sont même, dans certains cas, comme on vient de le voir, des *théorèmes logiques* de \mathcal{M} . Si \mathcal{L} est contenu dans \mathcal{M} et que la fonction de traduction est la fonction identique (traduction dite homophonique), et donc que les équivalences-T sont du genre :

« La neige est blanche » est vraie ssi la neige est blanche,

sans doute prendra-t-on acte d'un cœur léger du caractère de théorème logique ou au moins syntaxique des équivalences-T. Mais si \mathcal{L} n'est pas contenu dans \mathcal{M} et que l'on oublie la manière dont la définition en question est construite, il y a de quoi s'étonner. Pour savoir, par exemple, que « Snow is white » est vraie ssi la neige est blanche, à supposer qu'on ait là une équivalence-T relative à \mathcal{L} dans \mathcal{M} , ne faut-il pas savoir que « Snow is white », en tant que phrase de \mathcal{L} , veut dire que la neige est

blanche, ou veut dire quelque chose de matériellement équivalent ? — ou peut-être savoir que « Snow is white », en tant que phrase de \mathcal{L} , se traduit dans \mathcal{M} par « la neige est blanche », ou par une phrase matériellement équivalente ? — en tout cas savoir quelque chose que la syntaxe de \mathcal{L} , à elle seule, ne peut vous apprendre ? Pour assurer le caractère de théorème logique ou au moins syntaxique des équivalences-T, il faudrait que ce *quelque chose* fût de quelque façon intégré dans le concept même de vérité engagé dans les équivalences-T. Mais comment ?

La construction tarskienne réussit ce tour de force. Mais il n'y a pas de miracle, c'est qu'elle n'est pas elle-même construite de façon purement syntaxique. La définition donnée ci-dessus de la vérité pour \mathcal{L}_0 est transparente sur ce point. Le *definiens* de cette définition est pur de tout terme sémantique, comme le voulait Tarski, mais tous ses termes n'appartiennent pas pour autant à la syntaxe de \mathcal{L}_0 . Certes, les termes de mention (avec guillemets), « « Snow is white » » et « « Grass is blue » », des phrases de \mathcal{L}_0 appartiennent à la syntaxe de \mathcal{L}_0 , mais non la traduction (sans guillemets), « la neige est blanche » et « l'herbe est bleue », de ces phrases dans \mathcal{M}_0 . Qui plus est, la place de ces traductions dans le *definiens* est prescrite par le manuel de traduction de \mathcal{L}_0 dans \mathcal{M}_0 , que la syntaxe de \mathcal{L}_0 est bien incapable, à elle seule, d'établir. Bref, *la définition explicite tarskienne de la vérité pour \mathcal{L}_0 dans \mathcal{M}_0 n'est pas purement syntaxique*. Il en va de même, de façon générale, de la définition explicite tarskienne de la vérité pour \mathcal{L} dans \mathcal{M} , si \mathcal{L} contient des termes extra-logiques (par exemple, le terme singulier « Snow » et le prédicat « is white ») : le *definiens* de cette définition contient entre autres choses la traduction de ces termes (« la neige », « est blanche ») à une place prescrite par le manuel de traduction de \mathcal{L} dans \mathcal{M} , et *cette définition n'est pas, ne peut pas être purement syntaxique*. Et il en irait à l'évidence de même de toute définition explicite répondant au critère tarskien d'adéquation.

3.2. Curieusement, Tarski [1936b ; 1956, 406] prétend le contraire : « [L]a sémantique devient une partie de la morphologie du langage[-objet] si cette dernière est entendue en un sens suffisamment large » (voir aussi [1935 ; 1956, 273]). La « morphologie » en question n'est autre que ce que Church, par exemple, appelle « syntaxe » d'un langage, à savoir l'étude des expressions de ce langage selon leur forme, indépendamment de tout contenu ([Tarski 1935 ; 1956, 251], [Church 1956, 58]). Et quand Tarski précise que cette morphologie doit être entendue en un sens suffisamment large, il ne veut certainement pas dire qu'elle doit être entendue en un sens suffisamment large *pour que des termes extra-logiques comme « la neige » et « est blanche » puissent lui appartenir*, il doit simplement

vouloir dire que *sa logique sous-jacente peut être aussi riche que l'on veut*.

De façon non moins étonnante, Church emboîte le pas de Tarski⁴. Il fait certes preuve d'un minimum de prudence quand il écrit que, « *d'une certaine façon*, la sémantique peut être réduite à la syntaxe » (je souligne), mais il n'hésite pas à qualifier la définition explicite de Tarski (quand elle est possible) de « *purement syntaxique* » (je souligne).

L'affirmation de Tarski endossée par Church n'est justifiée que dans le cas des langages-objets dont la traduction métalinguistique appartient au langage de leur morphologie, ou syntaxe, *lato sensu*. C'est le cas des langages logico-mathématiques, et en particulier des langages pris à titre d'exemples dans le *Wahrheitsbegriff*. Mais le contenu du mémoire, et déjà son titre bien compris (à partir de l'édition allemande), « Le concept de vérité dans les langages formalisés », n'impliquaient aucune telle limitation. Encore moins l'article de 1944, où Tarski élargissait le champ de droit de ses investigations aux « langages à structure rigoureusement spécifiée » [Tarski 1944, § 6]. Ce manque de vigilance, tout à coup, de Tarski et de Church, ces parangons, par ailleurs, de rigueur et de précision, ne laisse pas d'étonner⁵.

3.3. On ne peut décidément faire confiance à personne, et Popper aurait dû se méfier, au lieu de suivre la leçon, comme il semble vouloir le faire⁶. Je dis « semble », car son exposé sur ce point est obscur et demande à être revu et corrigé pour devenir l'expression d'une pensée clairement identifiable. Quand Popper parle de « morphologie (ou syntaxe) *d'(of)* un langage », il faut croire, sauf à le convaincre de confusion

⁴ « Comme il ressort du travail de Tarski, en un certain sens, la sémantique peut être réduite à la syntaxe. Tarski a particulièrement insisté sur la possibilité de trouver, pour un langage formalisé donné, une propriété purement syntaxique des formules bien formées qui coïncide en extension avec la propriété sémantique d'être une phrase vraie. » [Church 1956, 65].

⁵ Sur cette question, voir aussi [Rouilhan 1999, 312–315].

⁶ « Les termes [sémantiques du métalangage] en tant que termes du métalangage [...] ont le même caractère morphologique que les [noms d'expression du langage-objet] : c'est-à-dire qu'ils appartiennent à la morphologie développée *dans* le métalangage (même si ce n'est pas à cette partie du métalangage qui contient la morphologie ou syntaxe *du* langage-objet et peut être développée dans le langage-objet lui-même. » [1972, 326, n. 12, italiques de Popper].

« [L]es termes qui sont sémantiques relativement au langage-*objet* peuvent avoir dans le métalangage comme tel le même statut que ses autres termes morphologiques ou syntaxiques. La sémantique d'un langage-objet L_n peut être une partie de la syntaxe *du* métalangage d'ordre supérieur (disons L_{n+1}) : nul terme de caractère non morphologique ou non syntaxique n'entre nécessairement dans L_{n+1} . Cela revient à une réduction de la sémantique de L_n à la syntaxe *de* L_{n+1} . » [1972, 327, je souligne deux fois la préposition « de » (*of*)]. » [1972, 327].

mentale, qu'il ne vise pas la même chose que Tarski (ou Church) sous la même description. Il vise très probablement la partie de la morphologie de ce langage ou d'un autre langage qui s'exprime dans ce langage même. S'il en est ainsi, alors Popper veut dire, en résumé, que les termes sémantiques relatifs à un langage-objet peuvent être explicitement définis en termes purement morphologiques (ou syntaxiques) dans un métalangage d'ordre supérieur (en un sens à préciser, voir § 4.1) à celui de ce langage-objet. Je ne vois pas d'interprétation plus charitable que celle-là, qui clarifie la position de Popper, et lui reconnaît, dans ce qu'il faut bien appeler, pour finir, une erreur, la caution fatale des plus grands.

4 La théorie de Tarski est-elle, comme celui-ci le prétend, « épistémologiquement neutre » ? — La thèse est équivoque, et Tarski peut aussi bien la soutenir que Popper la contester.

4.1. Dernier rappel sur la théorie de la vérité de Tarski. Pour simplifier, limitons-nous désormais, comme Tarski le fait dans le *Wahrheitsbegriff* (hormis le *post-scriptum*), aux langages (en position de langage-objet ou de métalangage) formalisés fondés sur la théorie des types simple. La terminologie ne doit pas nous égarer : un langage *formalisé* n'est pas un langage *formel*, au sens de vide de sens, c'est un langage purifié de toute ambiguïté, ou, pour le dire comme Tarski, « un langage dans lequel le sens de chaque expression est univoquement déterminé par sa forme » [1935 ; 1956, 165–166].

Pour tout langage-objet, \mathcal{L} , d'ordre fini, n , Tarski montre comment il est possible de construire une définition explicite matériellement adéquate d'un prédicat de vérité pour \mathcal{L} dans un métalangage, \mathcal{M} , *pourvu que* \mathcal{M} *soit d'ordre* $> n$; il montre aussi que, si le langage-objet est d'ordre infini ω , on doit se contenter d'une définition implicite dans un métalangage d'ordre ω .

Ce n'est pas ici le lieu de s'abîmer dans la contemplation de l'antique et vénérable paradoxe du Menteur et du grand théorème moderne d'indéfinité que Tarski a su en tirer. Prenons simplement acte du résultat : si une théorie scientifique, \mathcal{T} , de langage \mathcal{L} d'ordre κ ($1 \leq \kappa \leq \omega$), est développée dans le cadre de la théorie des types simple, il est possible, dans ce même cadre, de définir, explicitement ou implicitement, de

façon matériellement adéquate, un prédicat de vérité pour \mathcal{L} , et que le langage de cette définition est *nécessairement d'ordre* $\geq \kappa$. Cette condition sur l'ordre montre dans quelle mesure la théorie de la vérité pour le langage d'une théorie scientifique a, ontologiquement, partie liée avec cette dernière.

Ajoutons que Tarski, à l'instar de Quine, dont il se révèle philosophiquement si proche, dans leurs discussions avec Carnap au début des années quarante (voir [Mancosu 2005]), n'était certainement pas prêt à accepter l'idée d'une philosophie première, indépendante en droit des sciences, particulières ou non, ni, réciproquement, l'idée que ces sciences fussent en droit indépendantes de la philosophie, à supposer que rien de sérieux fût digne de ce nom⁷.

Curieusement, Tarski prétend que sa théorie de la vérité est « épistémologiquement neutre »⁸. Popper le conteste⁹. Ni l'un ni l'autre ne donnent très exactement leurs raisons, ni ne sont très explicites sur le sens en lequel ils entendent la neutralité en question. Or on peut l'entendre en un sens faible, qui donne raison à Tarski, ou en un sens fort, qui donne raison à Popper.

⁷Ce que ne dément pas l'autodérision de cette profession de foi de Tarski lors d'une réunion de l'Association of Symbolic Logic (ASL), à Chicago, le 29 avril 1965 : « Vous savez, il se trouve que je suis un anti-platoniste beaucoup plus extrême. [...] [J]e représente cette espèce très brutale et naïve d'anti-platonisme, une chose que je pourrais décrire comme du matérialisme, ou du nominalisme mâtiné de matérialisme, et il est très difficile pour un homme de vivre toute sa vie avec cette position philosophique, spécialement si c'est un mathématicien, spécialement si, pour certaines raisons, son passe-temps favori est ce qu'on appelle la théorie des ensembles » (cité par [Mancosu 2005, 341]).

⁸« En fait, la définition sémantique de la vérité n'implique rien concernant les conditions sous lesquelles une phrase telle que (1) :

(1) *la neige est blanche*

peut être affirmée. Elle implique seulement que, lorsque nous admettons ou rejetons cette phrase, nous devons être prêts à affirmer ou à rejeter la phrase corrélatrice (2) :

(2) *la phrase « la neige est blanche » est vraie.*

« Aussi pouvons-nous accepter la conception sémantique de la vérité sans abandonner quelque position épistémologique que nous puissions avoir. Nous pouvons demeurer réalistes naïfs, réalistes critiques ou idéalistes, empiristes ou métaphysiciens — ce que nous étions auparavant. La conception sémantique est complètement neutre par rapport à toutes ces positions. » [Tarski 1944, 361-362].

⁹« Comme j'étais un réaliste du sens commun critique, et conscient du fait que je soutenais par conséquent une théorie « métaphysique », j'étais fort intéressé par ce qui m'apparaissait comme un aspect réaliste de la théorie de la vérité de Tarski, un aspect dont je suspecte que Tarski peut dénier la simple existence. » [Popper 1972, 323].

4.2. Tarski l'entend certainement au sens faible, qui lui donne raison. On voit bien ce qu'il a en tête. Non pas l'indépendance de toute science, donc aussi de la sémantique devenue enfin science, et donc enfin de la théorie sémantique de la vérité, à l'égard de la philosophie (voir § 4.1), mais, beaucoup plus modestement, l'idée qu'au delà des disputes possibles autour du concept de vérité, toute personne un tant soit peu éclairée sur le sens en question du mot « vrai », un tant soit peu maîtresse de son bon usage en ce sens, devrait au minimum reconnaître la validité des définitions partielles de la vérité et accepter l'exigence selon laquelle elles doivent être déductibles, en un certain sens que le profane peut laisser au clerc le soin d'articuler, de toute définition prétendue matériellement adéquate de la vérité. Et il semble difficile, en effet, de ne pas souscrire avec Tarski à une version aussi faible de la thèse de neutralité.

4.3. Mais si nous entendons la thèse de neutralité en un sens plus fort, comme Popper sans doute le faisait, il en va tout autrement. Sans chercher ici à savoir ce qu'il en était exactement pour lui, je supposerai que la neutralité en question est relative notamment au débat sur le réalisme scientifique, et qu'elle implique que, dans le passage d'une interprétation réaliste d'une science à une interprétation instrumentaliste, l'explication que Tarski donne de la vérité ne changerait pas. Je montrerai que cette conséquence n'est pas le cas, et donc que la théorie de la vérité de Tarski n'est pas neutre au sens supposé.

Le débat entre le réaliste et l'instrumentaliste peut prendre plusieurs formes, je m'en tiendrai, pour mettre à l'épreuve la thèse de neutralité au sens fort, au cas où s'opposent le nominaliste, pour lequel seuls existent les individus, d'ordre 1, et le réaliste prêt sans état d'âme à conférer à toutes les entités non individuelles, d'ordre > 1 , la même dignité ontologique qu'aux individus. J'imaginerai donc (avec les mêmes notations que ci-dessus, voir § 3.1) la théorie scientifique \mathcal{T} développée dans le cadre de la théorie des types simple, et d'ordre supérieur κ ($1 < \kappa \leq \omega$). La théorie de Tarski nous apprend qu'il est possible de définir de façon matériellement adéquate un prédicat de vérité, « Vr », pour \mathcal{L} , et que le langage de cette définition est nécessairement d'ordre $\geq \kappa$, voir § 4.1. Soit (toujours dans le même cadre) \mathcal{T}^+ , de langage \mathcal{L}^+ , une extension de \mathcal{T} contenant cette définition. Pour le réaliste, pas de problème, « Vr » peut légitimement prétendre au statut de prédicat de vérité pour \mathcal{L} . Mais pour le nominaliste ? La définition en question ne jouit d'aucun privilège par rapport aux autres expressions d'ordre supérieur de \mathcal{L}^+ , elle n'est pas plus réelle ni moins instrumentale qu'elles. La définition de « Vr » n'est donc plus susceptible de valoir comme définition réelle de la vérité

pour \mathcal{L} , c'est maintenant une définition instrumentale, une fausse définition. Pour le dire de façon lapidaire : la vérité réelle est devenue vérité instrumentale ; la vérité, fausse vérité. En ce sens, la théorie de la vérité de Tarski n'est pas neutre. Pas plus que la science en général à l'égard de la philosophie.

5 Popper a-t-il raison pour autant d'aller jusqu'à prétendre que la théorie de la vérité de Tarski fournit un argument en faveur du « réalisme métaphysique » ?

— Non.

Non content de contester la neutralité de la théorie de la vérité de Tarski, Popper va jusqu'à affirmer que cette théorie fournit un argument en faveur du « réalisme métaphysique »¹⁰. Je ne chercherai pas plus à savoir ce que Popper veut dire exactement par cette expression que je n'ai cherché tout à l'heure à savoir en quel sens exactement il entendait, pour la contester, la thèse de neutralité. J'admettrai que l'affirmation de Popper implique que, dans le conflit des interprétations, la théorie de la vérité de Tarski pourrait être invoquée par le réaliste contre l'instrumentaliste.

La théorie de la vérité de Tarski a bien pu sembler au réaliste Popper la meilleure réponse possible à l'objection que croyaient jusque-là pouvoir lui opposer ses adversaires en lui représentant les difficultés inhérentes à la conception traditionnelle de la vérité comme correspondance aux faits. Mais, loin d'être la meilleure, cette réponse avait la faiblesse d'accorder à l'objection une pertinence qu'en réalité elle n'avait pas. En réalité, quoi que Popper ait pu en penser, les difficultés en question ne menaçaient nullement sa position.

Ce n'est pas que la théorie de la vérité de Tarski soit neutre dans le conflit des interprétations au sens où elle jouirait d'un privilège d'extra-territorialité, pour ainsi dire, qui lui éviterait *a priori* d'être emportée dans la tourmente. Au contraire, je crois avoir montré qu'elle ne l'est pas, voir § 4. Pour pouvoir prétendre fournir une explication réelle de la

¹⁰ « La théorie de Tarski, comme vous le savez tous, et comme il l'a d'abord souligné, est une *réhabilitation* et une élaboration de la théorie classique de la vérité selon laquelle la vérité est correspondance aux faits ; et ceci me semble donner appui au réalisme métaphysique. » [Popper 1972, 323].

vérité, la théorie de Tarski a besoin, en effet, d'être *elle-même* interprétée de façon réaliste.

Théorie de la vérité de Tarski et réalisme ont donc bien partie liée. Mais le lien est *asymétrique*. Ce n'est pas la théorie de la vérité de Tarski qui peut fournir un argument en faveur du réalisme (interprétation réaliste de \mathcal{T}), c'est, inversement, le réalisme (interprétation réaliste de \mathcal{T}^+) qui peut seul donner à la théorie et à son objet (la vérité pour \mathcal{L}) une légitimité réelle, et non seulement instrumentale.

Je rejoins finalement les philosophes déflationnistes sur la place qu'il convient d'accorder au problème de la vérité dans la querelle du réalisme : la même, tous privilèges abolis, qu'il convient de lui reconnaître en tout temps et en tout lieu de l'entreprise de connaissance en général (voir, e.g., [Horwich 1990, chap. 4]).

6 Conclusions

En conclusion je dirai que Popper lecteur de Tarski ne fait certes pas excès d'honneur au genre classique *X lecteur de Y*, mais pas au point de mériter l'indignité dont on a pu le frapper. Et je rappellerai les thèses en grande partie négatives et de caractère théorique plutôt qu'exégétique que j'ai soutenues :

(a) Personne n'a jamais construit de définition des concepts de correspondance et de fait invoqués dans la notion traditionnelle de la vérité comme correspondance aux faits, mais l'intuition fondamentale de cette notion ne présuppose pas ces concepts, et la définition sémantique du concept de vérité fait bien droit à cette intuition, voir § 2.

(b) Une définition explicite purement syntaxique du concept de vérité faisant droit à l'intuition fondamentale de la notion traditionnelle de vérité comme correspondance aux faits est en général impossible, voir § 3.

(c) L'explication de la notion de vérité comme correspondance aux faits n'est pas neutre au sens où le passage de l'interprétation réaliste de la science à son interprétation instrumentaliste ne changerait rien à cette explication, voir § 4.

(d) Le concept sémantique de vérité peut être utilisé ou mentionné comme n'importe quel autre concept dans le débat sur le réalisme, mais il n'a aucun rôle privilégié à y jouer, voir § 5¹¹.

¹¹Merci à Alain Boyer, David Miller et Elie Zahar : au premier, pour m'avoir invité à me faire publiquement lecteur de Popper lecteur de Tarski, et à tous les trois, pour

Références

CHURCH, ALONZO

1956 *Introduction to Mathematical Logic*, Princeton : Princeton University Press.

HORWICH, PAUL

1990 *Truth*, Oxford : Blackwell, 1990 (2nd ed, Oxford : Clarendon Press, 1998).

MANCOSU, PAOLO

2005 Harvard 1940-1941 : Tarski, Carnap and Quine on a Finitistic Language of Mathematics for Science, *History and Philosophy of Logic*, 26 (2005) : 327–357.

MILLER, DAVID

1999 Popper and Tarski, in *Popper's Open Society After 50 Years. The Continuing Relevance of Karl Popper* (ed. by I. Jarvie and S. Pralong), London and New York : Routledge, 1999, 56–70.

POPPER, KARL

1955 Note on Tarski's Definition of Truth, *Mind*, 64 (1955) (édition légèrement révisée, [1963, 335–340]).

1963 *Conjectures and Refutations. The Growth of Scientific Knowledge*, London and Henley : Routledge and Kegan Paul, 1963 (4th ed. [revised], 1972).

1972 *Objective Knowledge. An Evolutionary Approach*, Oxford : At the Clarendon Press, 1972 (revised ed., 1979).

RIVENC, FRANÇOIS

1996 Théorie de la vérité et sémantique des conditions de vérité : le projet de Tarski, *Les Études philosophiques*, n°3/1996, 381–402.

ROUILHAN, PHILIPPE DE

1999 Les tableaux de Beth : syntaxe ou sémantique?, *Philosophia Scientiæ*, 3 (4) (1998-1999), 303–322.

TARSKI, ALFRED

1933 *Projecie prawdy w jezykach nauk dedukcyjnych* (Le concept de vérité dans les langages des sciences déductives), Varsovie.

les discussions stimulantes auxquelles cette version de ma lecture fait plus ou moins discrètement écho. Je sais que je ne les ai pas totalement convaincus. Et merci à François Rivenc, le « philosophe français », pour ses précieuses remarques sur une version antérieure de cette « Note », finalement beaucoup plus longue que prévu.

- 1935 Der Wahrheitsbegriff in den formalisierten Sprachen, *Studia Philosophica*, 1 (1936) : 261–405 (tirés à part datés de 1935 ; c'est l'édition allemande de [1933] ; pour l'édition anglaise, *The Concept of Truth in Formalized Languages*, voir [1956, 152–278]).
- 1936a O ugruntowaniu naukowej semantyki (La construction d'une sémantique scientifique), *Przełga Filozoficzny*, 39 (1936) : 50–57.
- 1936b Grundlegung der wissenschaftlichen Semantik, *Actualités Scientifiques et Industrielles*, 39 (1936) : 1–8 (c'est l'édition allemande de [1936a] ; pour la version anglaise, *The Establishment of Scientific Semantics*, voir [1956, 401–408]).
- 1944 The Semantic Conception of Truth and the Foundations of Semantics, *Philosophy and Phenomenological Research* 4 (1944) : 341–376.
- 1956 *Logic, Semantics, Metamathematics — Papers from 1923 to 1938* (ed. and trans. by J. Woodger), Oxford : At the Clarendon Press, 1956 (Second edition edited and introduced by J. Corcoran, Indiana : Hackett Publishing Company, 1983).
- 1969 *Truth and Proof*, *Scientific American*, 220 (6) (1969) : 63–77.