7

Alfred Tarski, *Il concetto di verità nei linguaggi formalizzati* (1933)

di *Carlotta Pavese**

7.1
Cenni biografici

Alfred Tarski nacque a Varsavia nel 1901, da una famiglia ebreo-polacca. Studiò presso l'Università di Varsavia, dove ottenne il dottorato nel 1924, sotto la supervisione del logico Stanisław Leśniewski. Si sposò nel 1929 con Maria Witkowska da cui ebbe due figli. Dal 1924 al 1939 ricoprì posizioni temporanee presso l'Università di Varsavia e insegnò anche presso le scuole superiori per sostenere sé e la sua famiglia, senza riuscire a trovare un'occupazione permanente. All'inizio degli anni Trenta, entrò in contatto con il Circolo di Vienna, dove conobbe Karl Menger, Kurt Gödel e Rudolf Carnap. A questo periodo risale la prima stesura del *Concetto di verità nei linguaggi formalizzati*, pubblicato nella prima versione in polacco nel 1933. Nel 1939 ricevette un invito da Quine a visitare Harvard. Partì per gli Stati Uniti nell'agosto di quell'anno, poco prima dell'invasione tedesca della Polonia e dello scoppio della Seconda guerra mondiale. Occupò posizioni temporanee in diverse università, e nel 1945 ottenne una posizione permanente presso l'Università della California a Berkeley. Morì nel 1983.

7.2
Riferimenti bibliografici dell'opera

EDIZIONE ORIGINALE A. Tarski, *Pojęcie prawdy w językach nauk dedukcyjnych*, Nakładem Towarzystwa Naukowego Warszawskiego, Warszawa 1933; poi nella collezione del 1995 curata da J. Zygmunt,

* Carlotta Pavese insegna Filosofia alla Cornell University.

Pisma Logiczno-Filozoficzne, 1 Prawda, Wydawnictwo Naukowe pwn, Warsaw, pp. 13-172.

TRADUZIONE ITALIANA A. Tarski, *Il concetto di verità nei linguaggi formalizzati*, a cura di F. Rivetti, Vita e Pensiero, Milano 1963 (edizione di riferimento nel corso del capitolo).

TRADUZIONE TEDESCA Appare nel 1935 con il titolo *Der Wahrheitsbegriff in den formalisierten Sprachen* nel primo volume di "Studia Philosophica".

TRADUZIONE INGLESE La prima, ad opera di J. Woodger, risale al 1935, ma viene pubblicata per la prima volta nel 1956 nel volume collettaneo del lavoro di Tarski *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, ed. by J. Corcoran, Hackett Publishing Company, Indianapolis (IN), pp. 152-278.

7.3 Contesto

Ai tempi della pubblicazione dell'articolo di Tarski, il concetto di verità svolgeva un ruolo centrale in logica, comparendo nelle definizioni di validità e di conseguenza logica del tempo e in diversi risultati logici importanti, come i teoremi di Löwenheim (1915) e Skolem (1920) e quelli di completezza e di incompletezza di Gödel (1929; 1931). Sebbene queste e altre teorie logiche matematiche impiegassero la nozione di verità, lo facevano informalmente, senza mai fornirne una definizione.

D'altra parte, il concetto di verità era considerato problematico da molti pensatori del tempo e in particolare da diversi esponenti del Circolo di Vienna, con cui Tarski entra in contatto all'inizio degli anni Trenta. Un aspetto molto noto del pensiero del Circolo di Vienna è la *critica alla metafisica*, che segue dal criterio di significato come verificabilità, secondo cui enunciati dotati di senso sono o gli enunciati dimostrabili della logica e della matematica o gli enunciati empirici verificabili (in qualche senso da specificare). Gli enunciati della metafisica, poiché né analitici né verificabili empiricamente, venivano considerati come privi di significato. Il concetto di verità, come distinto sia dal concetto sintattico di dimostrabilità sia dal concetto empirico di verificabilità, era visto come un concetto *semantico* – cioè come avente a che fare con la relazione tra linguaggio e mondo. Come tale, il concetto di verità era considerato in odore di metafisica dagli esponenti del

Circolo di Vienna e gli enunciati dove appariva venivano rifiutati come privi di senso. Si narra che uno degli esponenti principali del Circolo, Otto Neurath, in contatto epistolare con Tarski fin dall'inizio degli anni Trenta, obiettasse regolarmente all'uso della parola "vero" con l'esclamazione "Metaphysik!" (Cohen, Neurath, 1973, pp. 82-3).

Ad acuire il sospetto nei confronti del concetto di verità c'erano i paradossi semantici come il paradosso del mentitore, noto fin dall'antichità e di cui non si conoscevano soluzioni soddisfacenti. Per un'illustrazione del paradosso, si consideri l'enunciato seguente:

(a) ***s* non è un enunciato vero**.

Se chiamiamo ***s*** il primo enunciato in neretto di questo capitolo, (b) segue da (a):

(b) "***s* non è un enunciato vero**" = ***s***

D'altra parte, il bicondizionale (c) è indipendentemente plausibile:

(c) "***s* non è un enunciato vero**" è vero se e solo se ***s* non è un enunciato vero**.

Ora, dall'identità (b) e dall'apparentemente innocuo bicondizionale (c) segue la contraddizione (d):

(d) ***s*** è vero se e solo se ***s* non è un enunciato vero**.

Tarski vide con più chiarezza di chiunque altro che, in mancanza di un modo adeguato di bloccare il sorgere di tali contraddizioni, i paradossi semantici sollevavano seri dubbi sulla coerenza delle teorie logiche e matematiche che impiegavano il concetto di verità.

Gli stessi teoremi di incompletezza di Gödel, resi pubblici nel 1930, ebbero in un certo senso l'effetto di *aumentare* il sospetto nei confronti del concetto di verità. Nel primo teorema dell'incompletezza dell'aritmetica, Gödel dimostra che ci sono enunciati matematici veri che non sono dimostrabili. Questo risultato di Gödel, che Tarski fu uno dei primi ad apprezzare, provava in maniera definitiva che il concetto di verità matematica non poteva essere completamente ridotto a nozioni logico-sintattiche. In questo senso, il risultato di Gödel rese più urgente il problema di capire se il concetto semantico di verità, come distinto da quello di dimostrabilità, fosse scientificamente accettabile.

In questo panorama filosofico si colloca l'articolo sulla verità di Tarski. Con esso, Tarski intese riabilitare il concetto di verità dai sospetti di molti esponenti del Circolo di Vienna, sospetti che come abbiamo visto erano rafforzati dall'esistenza dei paradossi semantici. L'obiettivo era dimostrare che una definizione formalmente accettabile del concetto di verità fosse possibile senza dover fare appello a nozioni semantiche a loro volta non definite e allo stesso tempo fornire una soluzione di principio ai paradossi semantici, così da eliminare il sospetto sulla natura irriducibilmente metafisica del concetto di verità e sulla sua coerenza. Come Tarski stesso dichiarò in seguito, solo attraverso una definizione matematicamente adeguata del concetto semantico di verità l'uso scientifico del concetto di verità in matematica e in logica poteva essere giustificato e gli stessi risultati di Löwenheim-Skolem e di Gödel potevano ricevere una fondazione adeguata.

7.4
Contenuti

Nell'introduzione, Tarski precisa che l'obiettivo principale dell'articolo è quello di fornire una definizione di verità che sia *materialmente* adeguata – cioè in modo dimostrabile una definizione di *verità*, piuttosto che di qualche altro concetto – e *formalmente* adeguata. Una definizione di verità per un linguaggio interpretato è formalmente adeguata se formulata usando una lista finita di termini legittimi e matematicamente chiari. In particolare, Tarski chiarisce che la definizione sarebbe matematicamente accettabile solo qualora non utilizzi concetti *semantici*, cioè aventi a che fare con la relazione tra linguaggio e mondo, che non siano a loro volta definiti in termini non semantici.

Nella prima sezione, Tarski spiega i problemi che sorgono nel tentativo di fornire una definizione formalmente corretta di verità per linguaggi "ordinari," come l'italiano o l'inglese. Il problema principale è che certe caratteristiche di questi linguaggi li rendono suscettibili ai paradossi semantici. Come abbiamo visto, da enunciati come (a) e (c), insieme con la premessa intuitiva (b), è molto semplice derivare una contraddizione. Secondo Tarski, l'origine della contraddizione risiede nel fatto che è possibile sostituire '*p*' nello Schema T con un enunciato che contiene il predicato "vero" – ad esempio, "***s* non è un enunciato vero**" – e '*N*' con un particolare nome per quell'enunciato – cioè "***s***":

Schema T: *N* è vero se e solo se *p*

Tarski nota che non c'è modo di impedire che questa sostituzione possa avvenire in un linguaggio ordinario, perché i linguaggi ordinari sono *universali*, nel senso che contengono, oltre ai loro enunciati, i loro nomi e predicati semantici come "essere vero" o "denotare". Tarski pensava che proprio questo carattere universale dei linguaggi ordinari fosse alla radice dei paradossi semantici e minasse la possibilità di usare in maniera coerente il predicato di verità per questi linguaggi.

Nella seconda sezione, Tarski introduce il linguaggio della teoria del calcolo delle classi che userà per illustrare il suo metodo di definizione della verità. Questo linguaggio è un linguaggio *formale*, nel senso che è possibile definire la classe dei suoi enunciati sulla base di caratteristiche puramente strutturali. I linguaggi formali per cui Tarski intende fornire una definizione di verità *non* sono universali: non contengono nomi per le loro espressioni né predicati semantici come "è vero". Un linguaggio formale L nel senso di Tarski è distinto dal suo metalinguaggio ML. ML contiene una *traduzione* di ogni espressione di L; inoltre, contiene i nomi *strutturali-descrittivi* delle espressioni di L, cioè che descrivono la struttura delle espressioni che nominano. Un modo di introdurre nomi strutturali-descrittivi – ma non l'unico – è usare le virgolette quadrate ⌜, ⌝: se *e* è un enunciato del nostro linguaggio, ⌜*e*⌝ è un nome strutturale-descrittivo per quell'enunciato. Infine, ML contiene le risorse espressive per parlare delle proprietà semantiche di L. In particolare, L *non* contiene il suo predicato di verità, che è invece un'espressione di ML.

La terza sezione è il cuore dell'articolo, dove Tarski fornisce la definizione di verità per il caso particolare del linguaggio del calcolo delle classi. Come abbiamo visto all'inizio, una definizione è materialmente adeguata qualora sia dimostrabilmente una definizione di verità piuttosto che di qualche altro concetto, come quello di dimostrabilità o di asseribilità. In questa sezione, Tarski chiarisce che quel che garantisce che una definizione di verità sia materialmente adeguata è che soddisfi la Convenzione T:

Convenzione T
Una definizione D formalmente adeguata del predicato "vero in L" in un metalinguaggio per L è una definizione di verità se e solo se

(i) per ogni enunciato *e* di L, un esempio dello Schema T è derivabile da D (cioè per ogni enunciato *e* di L, è derivabile da D un enunciato che risulta dalla sostituzione di '*N*' nello Schema T con un nome strutturale descrittivo di *e* e dalla sostituzione di '*p*' con la traduzione di *e* nel metalinguaggio di L);

(ii) è dimostrabile da quella definizione che solo gli enunciati di L possono soddisfare il predicato "vero in L".

Tarski spiega che, per un linguaggio che possiede solo un numero finito di enunciati, è molto semplice costruire una definizione di verità che soddisfi la Convenzione T. Ad esempio, si consideri un linguaggio che abbia tre enunciati $e1$, $e2$ e $e3$, tradotti nel metalinguaggio come $p1$, $p2$ e $p3$ e con nomi rispettivamente $N1$, $N2$ e $N3$. La seguente è una definizione di verità materialmente adeguata:

Definizione 1
N è vero in L se e solo se
o $N = N1$, e $p1$;
o $N = N2$, e $p2$;
o $N = N3$, e $p3$.

Il problema è che questo metodo non è proponibile per linguaggi con un numero infinito di enunciati. Perché, in questo caso, una simile definizione di verità dovrebbe avere una lunghezza infinita. Quindi è necessario usare un *metodo ricorsivo*. Ad esempio, si consideri un linguaggio senza variabili ma con due costanti individuali '1' e '0', il segno di identità '=', quello di negazione '¬', quello di congiunzione '&'. Tale linguaggio ha un numero infinito di enunciati, ottenuti tramite la ripetuta applicazione della negazione e della congiunzione dei suoi enunciati atomici $\ulcorner 0=0 \urcorner$, $\ulcorner 1=0 \urcorner$, $\ulcorner 1=1 \urcorner$, $\ulcorner 0=1 \urcorner$. La seguente è una definizione "ricorsiva" di verità per questo linguaggio:

Definizione 2
N è vero in L se e solo se
o $N = \ulcorner 1=1 \urcorner$;
o $N = \ulcorner 0=0 \urcorner$;
o $N = \ulcorner \neg e \urcorner$ e $\ulcorner e \urcorner$ non è vero;
o $N = \ulcorner e1 \mathbin{\&} e2 \urcorner$ e $\ulcorner e1 \urcorner$ è vero e $\ulcorner e2 \urcorner$ è vero.

La Definizione 2 è ricorsiva nel senso che in essa la verità di un enunciato complesso è definita nei termini della verità degli enunciati da cui deriva tramite l'applicazione dei connettivi logici '¬' e '&': ad esempio un enunciato ⌜$e1$ & $e2$⌝ è definito come vero se e solo se gli enunciati che lo compongono ⌜$e1$⌝ e ⌜$e2$⌝ sono veri.

Dopo questi preliminari, Tarski ci insegna come definire ricorsivamente la verità per il particolare linguaggio del calcolo delle classi. Questo linguaggio contiene il segno di negazione '¬' (= *non*), il segno di disgiunzione 'v' (= *o*), il quantificatore universale '∀' (= *per ogni*) e il segno di inclusione 'I' (= *è incluso in*), una costante individuale per l'insieme vuoto '∅' e un insieme di variabili. Il metalinguaggio del calcolo delle classi contiene la traduzione di tutte queste espressioni del linguaggio oggetto, oltre che i nomi strutturali descrittivi per queste espressioni. Come nomi per le espressioni del linguaggio del calcolo delle classi, Tarski non usa le virgolette ⌜ ⌝ introdotte prima. Invece, introduce nomi per ogni espressione del linguaggio del calcolo delle classi. ∏ è il nome del quantificatore universale, *i* è il nome del segno di inclusione 'I' e 1, 2 sono nomi delle variabili 'x1' e 'x2'. Inoltre ~ è il nome della negazione ('¬'), ∨ il nome per la disgiunzione ('v') e ∅ è il nome del simbolo dell'insieme vuoto ('∅'). Tali nomi si combinano secondo la sintassi degli enunciati del linguaggio oggetto a formare nomi strutturali-descrittivi per quegli enunciati. Ad esempio, si consideri l'enunciato:

(e) ∀x2∀x1Ix1x2

(f) traduce (e) nel metalinguaggio:

(f) *per ogni classe a e b*, $a \subseteq b$

e (g) è il nome strutturale-descrittivo di (e) nel metalinguaggio:

(g) ∏2∏1*i*1,2

Oltre alla traduzione e ai nomi strutturali-descrittivi, il metalinguaggio del linguaggio del calcolo delle classi contiene le risorse per parlare delle proprietà semantiche del linguaggio del calcolo delle classi e queste risorse includono il predicato di verità per quel linguaggio.

Tarski spiega che una definizione come la Definizione 2 non è estendibile al linguaggio del calcolo delle classi perché in questo linguaggio non tutti gli enunciati complessi sono composti da altri enunciati. Ad esempio, un enunciato quantificato come $\ulcorner \forall x \mathrm{I} \varnothing x1 \urcorner$ ha come componente la funzione enunciativa – $\ulcorner \mathrm{I} \varnothing x1 \urcorner$. Una funzione enunciativa è un enunciato se e solo se non contiene variabili libere. Se contiene variabili libere una funzione enunciativa non è né vera né falsa. Quindi la verità, ad esempio, di $\ulcorner \forall x \mathrm{I} \varnothing x1 \urcorner$ non può essere una funzione della verità della funzione enunciativa $\ulcorner \mathrm{I} \varnothing x1 \urcorner$, perché tale funzione enunciativa non è né vera né falsa. Per definire la verità degli enunciati quantificazionali come $\ulcorner \forall x \mathrm{I} \varnothing x1 \urcorner$, Tarski quindi propone di definire una nozione di *soddisfacimento* di una funzione enunciativa da parte di certi individui. Poiché il linguaggio del calcolo delle classi parla di *classi*, il soddisfacimento di una funzione enunciativa è da parte di classi di individui. Ad esempio, una funzione enunciativa come $\ulcorner \mathrm{I} x1 \varnothing \urcorner$ è soddisfatta da una classe a se e solo se $a \subseteq \varnothing$ (di fatto nessun insieme soddisfa tale funzione enunciativa); una funzione enunciativa come $\ulcorner \forall \mathrm{x}2 \mathrm{I} \mathrm{x}1 \mathrm{x}2 \urcorner$ è soddisfatta da una classe a se e solo se per ogni classe b, $a \subseteq b$.

Una funzione enunciativa può contenere un numero arbitrario di variabili. Per questo motivo, Tarski propone che la relazione di soddisfacimento sia definita non come relazione tra una funzione enunciativa e un individuo ma come una relazione tra una funzione enunciativa e una *sequenza infinita* di individui. Le variabili del linguaggio del calcolo delle classi sono numerate e ognuna corrisponde a un elemento di una sequenza con lo stesso numero. Ad esempio, 'fk' si riferisce all'individuo **k** nella sequenza f e a esso corrisponde la variabile 'k'. Una sequenza f infinita di insiemi soddisfa la funzione enunciativa $\prod 2 i1, 2$ se solo se per ogni classe b, $f1 \subseteq b$.

Che cosa succede quando la funzione enunciativa è un enunciato universale come (e)? Tarski propone che tale enunciato – il cui nome strutturale descrittivo nel metalinguaggio è $\prod 1 \prod 2 i1, 2$ – sia da ritenersi soddisfatto da una sequenza infinita f se e solo se $\prod 2 i1,2$ è soddisfatto da tutte le sequenze g che differiscono da f al massimo nel loro primo elemento, e questo a sua volta richiede che la funzione enunciativa $i1, 2$ sia soddisfatta da tutte le sequenze j che differiscono da g al massimo nel loro secondo elemento. Questo in ultima analisi significa che $\prod 1 \prod 2 i1, 2$ è soddisfatto da una sequenza infinita f se e solo se per ogni classe a, b, $a \subseteq b$.

La Definizione 3 (corrispondente alla definizione 22 dell'originale) fornisce una definizione di soddisfacimento per il linguaggio del calcolo delle classi L*:

Definizione 3
1. *i*n,m è soddisfatto in L* da una sequenza *f* se e solo se *f*n ⊆ *f*m;
2. ~*N* è soddisfatto in L* da *f* se e solo se *N* non è soddisfatto da *f*;
3. *N*1∨*N*2 è soddisfatto in L* da *f* se e solo se o *N*1 è soddisfatto da *f* o *N*2 è soddisfatto da *f*;
4. ∏n*N* è soddisfatto in L* da *f* se e solo *N* è soddisfatto da tutte le sequenze *g* che differiscono da *f* al massimo in n;
5. Nient'altro è soddisfatto in L* da *f*.

Nello schema "*N* è soddisfatto in L* da *f* se e solo se...", *N* sta per un nome strutturale descrittivo di una funzione enunciativa del linguaggio oggetto e "..." per un enunciato nel metalinguaggio che specifica le condizioni di soddisfacimento di quella funzione enunciativa. La Definizione 3 è *ricorsiva* perché definisce il soddisfacimento di qualsiasi funzione enunciativa in termini del soddisfacimento delle funzioni enunciative che la compongono. La definizione di verità (corrispondente alla definizione 23 dell'originale) viene data nei termini di questa nozione di soddisfacimento, come segue:

Definizione 4
Un enunciato N è vero in L* se e solo se è soddisfatto da tutte le sequenze di insiemi ed è falso se e solo se non è soddisfatto da nessuna sequenza.

Nell'articolo, Tarski non tenta di dimostrare che tale definizione di verità soddisfa la Convenzione T perché, come spiega, tale dimostrazione dovrebbe avvenire in un *meta*-metalinguaggio che dovrebbe essere a sua volta introdotto. Invece di fornire una vera e propria dimostrazione che la Convenzione T è soddisfatta dalla sua definizione di verità, Tarski si limita a illustrare questo fatto con alcuni esempi. Si consideri di nuovo l'enunciato (e) e l'esempio corrispondente dello Schema T:

(h) ∏1∏2i1, 2 è vero in L* se e solo se per ogni classe *a*, *b*: *a* ⊆ *b*.

Vogliamo dimostrare che (h) è derivabile dalla Definizione 4. Dalla Definizione 4, si ottiene (i) sostituendo N con il nome strutturale descrittivo di (e):

(i) ∏1∏2i1, 2 è vero in L* se e solo se ogni sequenza infinita di classi lo soddisfa.

Secondo la Definizione 3, (l) segue:

(l) ∏1∏2i1, 2 è soddisfatta da una sequenza *f* se e solo se tutte le sequenze *g* che differiscono da *f* al massimo nel primo elemento soddisfano ∏2i1, 2, – se e solo se tutte le sequenze *j* che differiscono da *g* al massimo nel secondo elemento soddisfano i1, 2 (cioè $g1 \subseteq j2$) – se e solo se per tutti le classi *a* e tutti le classi *b*, $a \subseteq b$.

Ma (i) e (l) a sua volta implicano (h) – che ∏1∏2i1, 2 è vero in L* se e solo se per ogni classe *a*, *b* $a \subseteq b$ – proprio l'esempio dello Schema T che si voleva derivare dalla definizione di verità (Definizione 4). Tramite esempi di questo tipo, Tarski dà un'illustrazione di che cosa significa che la definizione di verità da lui fornita per il linguaggio del calcolo delle classi L* soddisfa la Convenzione T.

Nella sezione quarta, Tarski spiega come estendere il metodo illustrato a tutti i linguaggi di *ordine finito*. L'ordine di un linguaggio dipende dal tipo di variabili che il linguaggio contiene. Variabili su individui sono di tipo 1, variabili su insiemi di individui sono di tipo 2, variabili su insiemi di insiemi di individui sono di tipo 3 ecc. Il linguaggio del calcolo delle classi contiene solo variabili di secondo tipo (su classi di individui) – quindi è un linguaggio di secondo ordine. La definizione di soddisfacimento usata per definire la verità per il linguaggio del calcolo delle classi non può essere usata per qualsiasi linguaggio di ordine finito. La ragione è che la nozione di soddisfacimento per il linguaggio del calcolo delle classi era definita come una relazione tra una funzione enunciativa e una sequenza infinita di oggetti *dello stesso tipo* (cioè, classi di individui). Ma un linguaggio qualsiasi di ordine finito può contenere variabili su insiemi di individui (tipo 2) così come variabili su insiemi di insiemi di individui (tipo 3) o come variabili su insiemi di insiemi di insiemi di individui (tipo 4). Per estendere il metodo di definizione della verità illustrato per il linguaggio del calcolo delle classi occorre dunque una nuova definizione del concetto di soddisfacimento, adeguata per il nuovo linguaggio. Tarski discute due metodi ma solo uno consente di definire la nozione di soddisfacimento per qualsiasi linguaggio di ordine finito: il metodo di *unificazione semantica delle variabili*. Questo metodo consiste nell'associare ogni individuo *a* a una relazione a due posti: la relazione R che vige tra due individui *b* e *c* se e solo se $b = a$ e $c = a$. Tramite questa correlazione ogni classe di individui può essere correlata una a una con una classe di relazioni: a ogni classe A di individui corrisponde la classe A* di tutte quelle relazioni R che sono correlate nel modo descritto con gli elementi *a* di A. Con questa unificazione semantica, un linguaggio qualsiasi di ordine finito viene reinterpretato in modo tale che tutte le sue variabili, di qualsiasi ordine esse siano, stiano uniformemente per relazioni a due posti: le variabili del linguaggio ora appartengono alla *stessa* categoria semantica. In questo modo, la relazione di soddisfacimento può essere definita come prima in termini di una relazione a due posti, ovvero una relazione tra funzioni enunciative e *sequenze di relazioni*.

Nell'ultima sezione, Tarski prova un risultato noto come quello dell'"indefinibilità della verità" per i linguaggi di ordine infinito, usando come esempio il linguaggio della *teoria generale degli insiemi*. Il problema che sorge per questo tipo di linguaggio è, ancora una volta, quello di fornire una definizione adeguata di soddisfacimento. Date le caratteristiche di questo linguaggio, la relazione di soddisfacimento in questo caso dovrebbe essere una relazione tra funzioni enunciative e *un numero infinito di sequenze* (di individui, di insiemi di individui, di insiemi di insiemi di individui ecc.). Una relazione simile non è una relazione nella teoria dei tipi *finiti*, che è la teoria adottata da Tarski per il linguaggio della teoria generale degli insiemi. Come chiarisce nel poscritto, accettando una teoria dei *tipi transfiniti*, è possibile fornire una definizione di verità anche per linguaggi di ordine infinito. In questo senso, il risultato dell'indefinibilità dipende dalla scelta della metateoria. Tuttavia, nel poscritto Tarski precisa che, alla luce di questo fatto, quello dell'indefinibilità della verità può essere riformulato come il risultato che la verità non può essere definita se l'ordine del metalinguaggio *non è superiore* all'ordine del linguaggio oggetto.

7.5 Importanza, effetti e sviluppi

Tarski presentò la sua teoria della verità nel 1935 presso il Congrès internationale de philosophie scientifique di Parigi. La ricezione dell'articolo fu mista. Molti, tra cui Carnap, considerarono riuscito il tentativo di Tarski di riabilitare scientificamente il concetto di verità. Neurath, d'altra parte, vide nella teoria tarskiana un tentativo di risuscitare una teoria corrispondentista della verità e per questo motivo non la considerò una riabilitazione riuscita del concetto di verità dal suo statuto di concetto metafisico. Il dibattito sulle ambizioni riduzioniste della definizione di Tarski non finisce con il Circolo di Vienna. Field (1972) osservò che il progetto di Tarski di definire il concetto di verità in termini che non usassero concetti semantici a loro volta non definiti non può davvero considerarsi riuscito, visto che la definizione di Tarski assume come primitivo il concetto di *traduzione*, che entra sia nella definizione di metalinguaggio che dello Schema T (per una discussione critica in italiano di Field, 1972, cfr. Marconi, 1984).

Indipendentemente dalle sue ambizioni riduzioniste, la teoria della verità di Tarski ha avuto una profonda influenza sulla riflessione filosofica sul significato dei linguaggi naturali. Davidson (1967) propose che una teoria del significato degli enunciati di un linguaggio naturale dovesse prendere le forme di una definizione tarskiana della verità: secondo Davidson, una teoria del significato è adeguata solo se implica gli esempi dello Schema T per tutti gli enunciati del linguaggio di riferimento; e la nozione di significato è da definire sul modello della definizione tarskiana di soddisfacimento, nel senso che deve specificare il significato di un'espressione complessa ricorsivamente tramite la specificazione del significato delle espressioni che la compongono.

Pressoché tutta la discussione contemporanea sui paradossi semantici nasce come reazione alla soluzione dei paradossi suggerita da Tarski. Come abbiamo visto, cardine della soluzione di Tarski è la distinzione tra linguaggio oggetto e metalinguaggio: l'antidoto ai paradossi è imporre la restrizione che un linguaggio *non* possa contenere il suo predicato di verità. Questa restrizione era concepita da Tarski come una vera e propria restrizione *sintattica*: ogni enunciato *e* che sia equivalente a *e non è vero* è sintatticamente malformato. Una conseguenza è che la soluzione di Tarski non si estende ai linguaggi semanticamente chiusi, come i linguaggi naturali, che contengono il loro predicato di verità. La teoria della verità di Kripke (1975) così come le teorie revisioniste di Gupta e Belnap (1993) prendono le mosse da una critica di questo e altri aspetti della teoria tarskiana.

Infine, tutta la riflessione filosofica contemporanea sulla natura della verità parte in un modo o nell'altro dalla teoria della verità di Tarski. Le teorie deflazionistiche della verità (ad esempio Field, 1994) si ispirano esplicitamente alla teoria tarskiana della verità e si definiscono *in termini di* essa: secondo il deflazionismo, *tutto* quel che c'è da dire sulla verità è catturato dallo Schema T.