

Andrea Bottani

Imperscrutabilità del riferimento e teoria del significato

Abstract - "Imperscrutabilità del riferimento" è il nome di una nota tesi di Quine: l'idea che sia impossibile dire, in base alla totalità dell'evidenza accessibile all'interprete, a che cosa i termini singolari di un linguaggio si riferiscano, o di che cosa i suoi predicati siano veri. Accanto all'indeterminatezza di forma logica e di condizioni di verità, l'imperscrutabilità del riferimento è uno degli aspetti o delle forme in cui l'indeterminatezza della traduzione può manifestarsi. Qui, tuttavia, l'imperscrutabilità del riferimento viene discussa a prescindere da altre forme di indeterminatezza (ignorando, ad esempio, casi in cui il riferimento di una parola è banalmente indeterminato per il fatto che un manuale di traduzione tratta la parola come un termine singolare e un altro manuale ugualmente corretto come un predicato). La tesi che voglio discutere è che, anche se la forma logica e le condizioni di verità degli enunciati fossero determinate, ciò non basterebbe a fissare il riferimento. Questa è l'unica forma che la tesi di imperscrutabilità del riferimento può assumere per chiunque neghi l'indeterminatezza di forma logica e di condizioni di verità oppure ritenga, come Davidson, che queste forme di indeterminatezza non possano essere sistematiche.

Cos'è intesa, la tesi di imperscrutabilità del riferimento può essere ottenuta come conseguenza di due assunzioni:

A) L'idea che l'attribuzione di riferimenti alle parole da interpretare sia una costruzione teorica, non esposta a un controllo empirico diretto: i riferimenti delle parole sono qualunque cosa generi una teoria del significato empiricamente adeguata per il linguaggio oggetto (irriducibilità empirica del riferimento).

B) L'idea che una stessa teoria del significato ammetta teorie del riferimento irriducibili e conflittuali. Mentre A) rappresenta in qualche modo una filiazione del principio fregeano di contestualità del significato, B) si fonda sulla possibilità di definire funzioni permutative sul dominio di interpretazione. Prese insieme, le due assunzioni producono una sorta di *reductio ad absurdum* della nozione di riferimento: se una teoria del significato empiricamente adeguata è producibile sia dicendo che un termine *t* si riferisce ad *a* sia dicendo che si riferisce a *b* ($a = b$), allora, di fatto, *t* si riferisce ad *a* e *t* non si riferisce ad *a*. - cfr. A)

Il paradosso può stimolare reazioni molto diverse. 1. Si può provare a riformare la relazione di riferimento in modo da sottrarla alla presa del paradosso, come tenta di fare Quine suggerendo che il riferimento sia relativo o come si può provare a fare con qualche analogo del riferimento relativo, ad es. il riferimento parziale di Field, oppure moltiplicando le relazioni di riferimento; 2. Si può mandare il riferimento in soffitta, come ci consiglia di fare Davidson; 3. Si può provare a negare l'imperscrutabilità del riferimento abbandonando A) o B) o entrambe, nel qual caso il paradosso non avrebbe gambe per camminare.

La relazione comprende tre parti. Una ricostruzione dell'argomentazione che va da A) e B) alla tesi di imperscrutabilità del riferimento. Una discussione delle implicazioni che la tesi, se vera, potrebbe avere per la nozione di riferimento e particolarmente un confronto degli atteggiamenti di Quine e Davidson in proposito. Infine, un paio di osservazioni concernenti la tesi di imperscrutabilità. Anche se nulla viene detto per confutare direttamente la tesi, ci si sforza di mostrare che esistono dei vizi argomentativi nel modo in cui la tesi è sostenuta da coloro che la accettano sicché, nel migliore dei casi, non abbiamo ragioni definitive per sostenere che la tesi è vera. I motivi, fondamentalmente, sono due. In primo luogo, se *L* è un linguaggio abbastanza potente da esprimere una teoria del riferimento per se stesso (come ad es. il linguaggio ordinario che ci è familiare), è impossibile che due diversi schemi di riferimento per *L* generino teorie della verità equivalenti per *L*. In secondo luogo, il fatto che due teorie dell'interpretazione assegnino ad ogni enunciato di un linguaggio *L* condizioni di verità equivalenti non significa ancora che rappresentino esattamente nello stesso modo le relazioni di dipendenza logica valide fra enunciati di *L*. Le condizioni di verità assegnate dalle due teorie agli enunciati di *L* potrebbero essere equivalenti senza essere *logicamente* equivalenti.

I. Le alterne fortune semantiche del riferimento

Uno dei problemi più importanti dello studio del significato, o semantica, riguarda il ruolo della teoria del riferimento in un resoconto del significato. Per Frege, la teoria del riferimento era solo una parte della teoria semantica, un'altra parte essendo la teoria del senso (e un'altra ancora, se Dummett ha ragione, la teoria della forza). Ma

già nel *Tractatus*, come anche nell'atomismo logico di Russell, la semantica perdeva spessore, appiattendosi per intero sulla teoria del riferimento (nel *Tractatus*, una proposizione è una configurazione di parti ognuna delle quali sta per un oggetto e la funzione semantica di ogni parte si esaurisce nel nominare ciò di cui è nome). Una ventina di anni più tardi, Tarski continuava a descrivere la semantica come lo studio di "certe relazioni fra le espressioni di un linguaggio e gli oggetti [...] 'indicati' da tali espressioni" (Tarski 1944, tr.it. p. 32).

Ma proprio in quegli anni, Quine cominciava a denunciare come una "confusione" questa identificazione di significato e riferimento. Quine rielaborava considerazioni di chiara matrice freghiana, ma le sue conclusioni finivano per divergere totalmente da quelle di Frege. Mentre autori come Carnap e Church cercavano - nella prospettiva di Frege - di integrare una semantica referenzialista alla Tarski con una teoria del significato non referenziale cos' da avere una teoria semantica completa, la critica del "mito del museo" tenne Quine fermamente alla larga da questa tentazione. Ma anche abbandonando i significati determinati e il mito del museo, qualcosa continuava a fuoriuscire dalla teoria del riferimento ed era il problema della traduzione: come sia possibile tradurre una lingua in un'altra e che cosa distingua una traduzione corretta da una scorretta. Una buona traduzione infatti, qualunque cosa sia, deve conservare il significato di ciò che viene tradotto e viceversa il significato, qualunque cosa sia, è certo qualcosa che si conserva attraverso la traduzione. Cos', il rifiuto di assorbire il significato nel riferimento e l'abbandono del "mito del museo" produssero in Quine, congiuntamente, la tendenza a formulare e affrontare il problema del significato come problema della traduzione. Ma questa tendenza fin' per ripercuotersi sulla stessa nozione di riferimento.

Fu la tesi di indeterminatezza della traduzione - l'idea che esistano molti modi empiricamente corretti di tradurre un linguaggio in un altro - a generare problemi per il riferimento. Dato che per Quine un'espressione identificata come un termine può sempre essere indifferentemente tradotta con termini aventi riferimenti diversi, risulta insensato chiedersi, in assoluto, a che cosa l'espressione da tradurre si riferisca e il riferimento diventa tanto indeterminato quanto la traduzione: nulla, nel comportamento osservabile del parlante, può rivelare che cosa i suoi termini denotino, proprio come nulla può rivelare quale delle molte traduzioni possibili delle sue parole sia quella corretta. E poiché il significato, qualunque cosa sia, è certo qualcosa di pubblico [1] - intersoggettivamente trasmissibile e comunemente condiviso - il riferimento non può entrare a costituire il significato per motivi analoghi a quelli per cui non possono entrarvi quelle "creature delle tenebre" che sono i sensi freghiani. Cos', dal *Tractatus* a *Word and Object* si comp' un intero percorso e dall'identificazione tout court di significato e riferimento si arrivò all'idea, diametralmente opposta, che il riferimento sia "imperscrutabile" e come tale non possa entrare a costituire il significato neppure come componente parziale. Nello stesso tempo, l'opera di demolizione della teoria freghiana del significato fu completa: mentre la critica del "mito del museo" toglieva di mezzo il sinn, la tesi di imperscrutabilità del riferimento (da qui in avanti, I.R.) eliminava per sempre la bedeutung. Liquidate le due componenti fondamentali del significato freghiano, del significato fin' per non restare più nulla, se si eccettua il fatto che esistono modi legittimi e modi illegittimi di "versare" i linguaggi l'uno nell'altro, cioè appunto di tradurre.

Questo lavoro contiene una breve discussione di I.R. e delle sue conseguenze per la teoria del significato. In effetti, la rilevanza filosofica di I.R. non si ferma certo al ruolo che essa gioca in teoria del significato. E' stato osservato ad esempio (cfr. Wallace 1977) che l'imperscrutabilità del riferimento rende problematica la distinzione fra enunciati al passato ed enunciati al futuro, che rende impossibile associare a un'asserzione il fatto che la rende vera o falsa ed anche che rappresenta

un serio pregiudizio per qualsiasi teoria causalista del riferimento (quest'ultimo punto é discusso approfonditamente da Field 1974, 1975). L'elenco potrebbe continuare ma il fatto essenziale, filosoficamente, é che I.R. sembra interferire con certe assunzioni realistiche basilari incorporate nel nostro linguaggio ordinario: per quanto Davidson 1977 sostenga che I.R. non toglie di mezzo la realtà extralinguistica e i suoi componenti elementari (oggetti, eventi e quant'altro), essa sembra toglierci almeno i mezzi per affermarne l'esistenza, svuotando di senso la stessa opzione realista. Nessuno di questi punti, tuttavia, verrà discusso nel presente lavoro. Qui, intendo semplicemente ricostruire una plausibile argomentazione che abbia come conclusione I.R. e valutare se l'argomentazione funziona. Non cercherò quindi di confutare I.R., o di mostrare che il riferimento é effettivamente componente integrante - quand'anche non integrale - del significato. Mi sforzerò invece di mostrare che esistono dei vizi argomentativi nel modo in cui la tesi é sostenuta da coloro che la accettano cosicché, in definitiva, non abbiamo ragioni veramente cogenti per sostenere che la tesi é vera.

II. Cosa vuol dire che il riferimento é imperscrutabile?

Il mio primo obiettivo é di consegnare I.R. ad una formulazione che sia nello stesso tempo sintetica, precisa ed esegeticamente corretta. Una formulazione classica, la più diretta e immediata, é questa:

(I.R.1) E' impossibile dire, in base alla totalità dell'evidenza accessibile all'interprete, che cosa i termini singolari di un linguaggio denotino, o di che cosa i suoi predicati siano veri*.

(I.R.1) é una cattiva formulazione per almeno due motivi. In primo luogo, Quine non crede che sia impossibile dire a che cosa le parole si riferiscono ma solo che sia impossibile dirlo in assoluto: mentre il riferimento assoluto é imperscrutabile, possiamo rendere scrutabile il riferimento relativizzandolo a manuali di traduzione (in effetti, relativizzandolo doppiamente a manuali di traduzione e a linguaggi di sfondo)[2]. In secondo luogo, I.R.¹ asserisce un'impossibilità di cui non precisa la natura. Se chiamiamo "schema di riferimento" per un linguaggio L un resoconto che specifica quali parole di L siano nomi e che cosa questi nomi denotino, quali siano predicati e di che cosa questi predicati siano veri ecc., I.R.¹ potrebbe asserire una delle seguenti tesi: 1. E' impossibile formulare schemi di riferimento. 2. E' impossibile capire se uno schema di riferimento sia corretto o scorretto. 3. Non esistono schemi di riferimento corretti. 4. Esistono schemi di riferimento incompatibili e ugualmente corretti. Quine non ha mai inteso sostenere né 1. né 2. né 3.. 1. é palesemente falso: non é certo impossibile fare ciò che é già stato fatto ed é difficile negare che Tarski, ad esempio, abbia effettivamente formulato uno schema di riferimento per il calcolo delle classi. D'altra parte, 2. e 3 sono incompatibili con parti importanti della teoria quiniiana della traduzione. Un manuale di traduzione, per Quine, é un sistema completo di equazioni sintattiche fra espressioni della lingua da tradurre ed espressioni della lingua in cui tradurre e come tale non incorpora né uno schema di riferimento né una definizione di verità per la lingua da tradurre. Ma se un manuale di traduzione M traduce "London" con "Londra" e "London is a town" con "Londra é una città", allora l'assegnazione di riferimento "London si riferisce a Londra" e il T-enunciato " 'London is a town' é vero se e solo se Londra é una città" sono entrambi veri relativamente a M. Diciamo che un manuale di traduzione M rende vero uno schema di riferimento S (oppure che S é vero relativamente a M) se e solo se tutte le

assegnazioni di riferimento in S sono vere relativamente a M e che M rende vera una teoria della verità V se e solo se tutti i T-enunciati implicati da V sono veri relativamente a M. Ora, Quine crede che uno schema di riferimento sia corretto se vero relativamente ad almeno un manuale di traduzione corretto, scorretto se falso relativamente ad ogni manuale di traduzione corretto. E pensa che esistano manuali corretti e manuali scorretti, che sia possibile distinguere gli uni dagli altri e che si possa decidere se uno schema di riferimento é vero o falso relativamente a un manuale di traduzione. Nel suo insieme, ciò esclude tanto 2. quanto 3..

Quine ha invece chiaramente asserito non solo 4., ma la sua stessa universalizzazione, vale a dire:

(I.R.2) Per ogni schema di riferimento, ne esiste un altro incompatibile e ugualmente corretto.

(I.R.2) afferma che esistono molti modi incompatibili ma ugualmente corretti di dire a che cosa le parole di un linguaggio si riferiscono e non che non ne esiste nessuno, come (I.R.1) sembrava, almeno in prima battuta, suggerire. Ciò che (I.R.2) implica é che, se esiste uno schema corretto di riferimento secondo il quale "Bobby" si riferisce a Bobby, ne esisterà un altro ugualmente corretto ma incompatibile secondo il quale "Bobby" si riferisce a qualcosa di diverso da Bobby. Poiché entrambi gli schemi sono corretti, sarà corretto dire che "Bobby" si riferisce a Bobby, ma anche che "Bobby" si riferisce a qualcosa di diverso da Bobby, dopodiché sembrerebbe debba essere corretto anche dire che "Bobby" si riferisce e non si riferisce a Bobby (almeno assumendo, come il nostro concetto ordinario di riferimento sembra imporre, che se un nome si riferisce a un oggetto, non si riferisce a nulla che sia diverso da esso). Se questa é la situazione di ogni nome del linguaggio, come Quine sembra pensare, allora per ogni nome n e ogni oggetto a, n si riferisce ad a se e solo se n non si riferisce ad a. Questo chiarisce almeno due cose: a) l'imperscrutabilità del riferimento non é un problema di sottodeterminazione empirica del riferimento ma piuttosto di sovrabbondanza o sovradeterminazione: abbiamo troppi schemi di riferimento empiricamente adeguati o corretti. Non é impossibile dire a che cosa le parole di un linguaggio si riferiscono, é impossibile dirlo in modo univoco e consistente. b) L'imperscrutabilità del riferimento non é, come potrebbe sembrare, una impossibilità di ordine meramente epistemico - una impossibilità di vedere o di capire - e neppure, strettamente parlando, una impossibilità di ordine ontologico ma una impossibilità di ordine semantico: la difficoltà coinvolge la stessa nozione di riferimento, che appare minacciata da una sostanziale *reductio ad absurdum*.

(I.R.2) é abbastanza generale da rendere sensato il tentativo di precisarne meglio il senso e il ruolo all'interno della teoria quiniiana della traduzione. Poiché sappiamo che per Quine uno schema di riferimento é corretto se e solo se é vero relativamente a un manuale di traduzione corretto, possiamo "espandere" (I.R.2) come segue:

(I.R.3) Per ogni manuale di traduzione MT (da un linguaggio L al nostro linguaggio),
esiste un manuale di traduzione MT' (da L al nostro linguaggio) tale che:
(i) MT ed MT' sono ugualmente corretti rispetto alla totalità dell'evidenza empirica accessibile all'interprete;
(ii) MT e MT' rendono veri schemi di riferimento incompatibili (per L).

Cos'è descritta, l'imperscrutabilità del riferimento diventa semplicemente uno degli aspetti o delle forme in cui l'indeterminatezza della traduzione può manifestarsi. Una tassonomia quiniiana dell'indeterminatezza, a volerla tracciare per sommi capi,

dovrebbe distinguere almeno i seguenti taxa [3].

1. Indeterminatezza di condizioni di verità

Per ogni manuale di traduzione M da L a L*, esiste un manuale M' ugualmente corretto tale che M ed M' rendono vere teorie della verità incompatibili per L.

2. Indeterminatezza di forma logica.

Per ogni manuale di traduzione M da L a L*, esiste un manuale M' ugualmente corretto tale che M ed M':

- 2a. Discordano su quali parole siano termini singolari, o predicati, o quantificatori ...
- 2b. Discordano sulla logica profonda del linguaggio oggetto.

3. Imperscrutabilità del riferimento.

Per ogni manuale di traduzione M da L a L*, esiste un manuale M' ugualmente corretto tale che M ed M' rendono veri schemi di riferimento che:

- 3a. Associano alle parole di L diverse entità "pescando" da universi diversi (ontologia totale variante);
- 3b. Associano alle parole di L diverse entità "pescando" dallo stesso universo (ontologia totale fissa).

Una tassonomia come questa, e anche più completa e dettagliata di questa, non dice nulla sulle relazioni fra le forme di indeterminatezza che individua ed elenca, quindi non fornisce una descrizione completa del fenomeno dell'indeterminatezza. La domanda è: può una forma qualsiasi di indeterminatezza presentarsi in assenza delle altre o almeno di qualcuna delle altre? Si direbbe di no. Ad esempio, se due manuali di traduzione assegnano agli enunciati del linguaggio oggetto diverse condizioni di verità, non si vede davvero, prima facie, come potrebbero assegnare ad ogni enunciato la stessa forma logica e ad ogni termine lo stesso riferimento; e se discordano sulla forma logica, pare difficile che possano concordare totalmente sulle condizioni di verità e sui riferimenti. Il motivo è che una definizione tarskiana di verità per un linguaggio L è un ingranaggio composto di due ruote: i) uno schema di riferimento per L; ii) un insieme di "regole di proiezione" incorporate in una teoria logica per L: funzioni capaci di generare le condizioni di verità di ogni enunciato di L quando sia noto che cosa denotino i nomi che vi ricorrono, di che cosa siano veri i predicati che vi ricorrono, ecc., in breve quali siano "i riferimenti" delle parole che lo compongono (un esempio di regola di proiezione, il più semplice, è il seguente: se a è un nome e P è un predicato, allora Pa è vero se e solo l'oggetto denotato da a è uno degli oggetti di cui P è vero). Se lo schema di riferimento è fisso, cambiando le regole di proiezione cambia anche la teoria della verità e se viceversa la teoria della verità cambia, qualcosa deve necessariamente essere mutato nelle regole di proiezione. Cos'è, sembra che indeterminatezze di tipo 1 e di tipo 2 non possano presentarsi in assenza di qualche indeterminatezza di tipo diverso.

Ma Quine crede che il caso del riferimento sia diverso e che indeterminatezze di tipo 3. possano presentarsi in isolamento. Questo configura una tesi più forte di (I.R.³), formulabile come segue:

(I.R.⁴) Per ogni manuale di traduzione M (da un linguaggio L al nostro

linguaggio), esiste un manuale di traduzione M' (da L al nostro linguaggio) tale che:

- (i) M ed M' sono ugualmente corretti rispetto alla totalità dell'evidenza empirica accessibile all'interprete;
- (ii) M e M' concordano sulle condizioni di verità e sulle forme logiche degli enunciati di L ;
- (iii) M e M' rendono veri schemi di riferimento incompatibili (per L).

Le differenze fra (I.R.3) e (I.R.4) sono chiare. Mentre (I.R.3) dice che manuali di traduzione ugualmente corretti possono rendere veri schemi di riferimento conflittuali, (I.R.4) dice qualcosa di più: persino manuali che rendono vere teorie della verità e teorie della forma logica perfettamente equivalenti possono rendere veri schemi di riferimento conflittuali ed essere ugualmente corretti. Questo implica che, anche se la forma logica e le condizioni di verità degli enunciati fossero perfettamente determinate, ciò non basterebbe a fissare il riferimento. (I.R.3), invece, non implica nulla del genere.

(I.R.4) è nello stesso tempo meno generale e più forte di (I.R.3) e infatti (I.R.4) implica (I.R.3) ma non viceversa. D'altra parte, la negazione di (I.R.3) implica la negazione di (I.R.4), ma non viceversa: (I.R.4) potrebbe essere falsa e l'imperscrutabilità del riferimento conseguire dall'indeterminatezza di condizioni di verità e/o di forma logica. Ma assumiamo che queste forme di indeterminatezza non sussistano. Congiuntamente a (I.R.3), questa assunzione implica (I.R.4) e congiuntamente a (I.R.4) implica, banalmente, (I.R.3). Quindi, sotto questa assunzione, (I.R.3) e (I.R.4) sono logicamente equivalenti. Dunque (I.R.3) e (I.R.4) si equivalgono per chiunque neghi che si dia indeterminatezza di forma logica e di condizioni di verità. Ma si equivalgono anche per chiunque ritenga, come Davidson, che queste tipologie di indeterminatezza non possano essere pervasive o sistematiche.

I motivi per cui Davidson esclude che indeterminatezze di tipo 1., 2a., 2b. e anche 3a. possano essere sistematiche dipendono in parte da una applicazione a largo raggio del cosiddetto "principio di carità", in parte dall'idea che le teorie dell'interpretazione debbano assumere la forma di teorie tarskiane della verità anziché quella di manuali di traduzione. Ma ciò impone anche di riformulare (I.R.3), perché dal nuovo punto di vista uno schema di riferimento è corretto se è incorporato in una teoria della verità corretta (e non, come per Quine, se è vero relativamente a un manuale di traduzione corretto). Incorporata questa nozione di schema di riferimento corretto in (I.R.2), otteniamo (I.R.3)', la controparte del quiniano (I.R.3) nella teoria davidsoniana dell'interpretazione radicale:

- (I.R.3)' Per ogni teoria della verità T (per un linguaggio L), esiste una teoria T' (per L) tale che:
- (i) T e T' sono ugualmente corrette rispetto alla totalità dell'evidenza accessibile all'interprete;
 - (ii) T e T' incorporano schemi di riferimento incompatibili (per L).

E aggiungendo a (I.R.3)' la clausola

- (iii) T e T' concordano sulle condizioni di verità e sulle forme logiche degli enunciati di L ,

otteniamo (I.R.4)', la controparte del quiniano (I.R.4) nella teoria davidsoniana dell'interpretazione radicale.

Se assumiamo che una teoria corretta della verità non possa essere resa vera da un manuale scorretto e che un manuale di traduzione corretto non possa rendere vera una teoria della verità scorretta, (I.R.3) diventa equivalente a (I.R.3)' e (I.R.4) a (I.R.4)'. Ma dovrebbe ormai essere chiaro che - sempre che non si introducano ulteriori assunzioni - esiste un certo numero di formulazioni non equivalenti di I.R.. E' possibile dimostrarne qualcuna? E' possibile dimostrarle tutte?

III. Che cosa fa pensare che il riferimento sia imperscrutabile?

Secondo Wallace [4], l'imperscrutabilità del riferimento può essere asserita sulla base di un'argomentazione del genere:

A1. Se due schemi di riferimento determinano le stesse condizioni di verità per tutti gli enunciati di un linguaggio, sono ugualmente corretti (vale a dire: o sono entrambi empiricamente corretti, oppure nessuno dei due lo é).

A2. Per ogni schema di riferimento, ne esiste un altro incompatibile che genera, tramite regole di proiezione fisse, le stesse condizioni di verità per tutti gli enunciati di un linguaggio.

(I.R.2) Per ogni schema di riferimento, ne esiste un'altro incompatibile e ugualmente corretto.

Assumerò che questa argomentazione schematizzi fedelmente le ragioni per cui sia Quine che Davidson pensano che il riferimento sia imperscrutabile. Non discuterò quest'assunzione, non rientrando ciò negli scopi del presente lavoro. In ogni caso, non conosco altre argomentazioni altrettanto forti e convincenti a favore dell'idea che il riferimento sia imperscrutabile. Se ve ne fossero, ciò che dirò sulla provabilità di I.R. non avrà valore generale ma varrà limitatamente a una particolare strategia di prova.

E' facile vedere che l'argomentazione é formalmente corretta. Essa costituisce quindi una prova definitiva di (I.R.2) a meno che una delle sue premesse non sia falsa. Ma prima di chiarire e discutere separatamente A1. e A2., é istruttivo chiedersi che cosa A1. e A2. esattamente implicano. Accertato che implicano (I.R.2), cosa dire di (I.R.3), (I.R.4), (I.R.3)' e (I.R.4)'? Prendiamo A2.: se due schemi di riferimento determinano, tramite regole di proiezione fisse, le stesse condizioni di verità per ogni enunciato del linguaggio, appartengono a teorie della verità equivalenti che incorporano esattamente le stesse regole di proiezione. Poiché teorie della verità equivalenti che incorporano le stesse regole di proiezione (i) sono ugualmente corrette e (ii) concordano sulle condizioni di verità e sulla forma logica di tutti gli enunciati del linguaggio, se A2. é vera, anche (I.R.4)' é vera. Quindi, banalmente, anche (I.R.3)' lo é. E se assumiamo, come pensa Quine, che una teoria della verità corretta non possa essere vera relativamente a manuali di traduzione scorretti, A2. implica anche (I.R.3) e (I.R.4) - dato che, sotto questa assunzione, (I.R.3)' e (I.R.4)' implicano (I.R.3) e (I.R.4).

Questo comporta due cose. In primo luogo, l'argomentazione é perfettamente formulabile e accettabile sia nella teoria della traduzione di Quine che nella teoria della interpretazione di Davidson. In secondo luogo, l'argomentazione é una buona prova (sempre che le sue premesse siano vere) tanto delle versioni deboli di I.R. -

(I.R.³) e (I.R.³)' - quanto delle sue versioni forti - (I.R.⁴) e (I.R.⁴)'. E quindi vale come prova del fatto che, anche se la forma logica e le condizioni di verità degli enunciati fossero perfettamente determinate, ciò non renderebbe scrutabile il riferimento.

Riconosciuto che l'argomentazione è formalmente corretta, e stabilito cosa essa provi, che cosa fa pensare che le sue premesse siano vere?

Irriducibilità empirica del riferimento

A far pensare che A1. sia vera è la difficoltà di dare una caratterizzazione in termini non linguistici, o una riduzione empirica, della nozione di riferimento. L'idea che non si possa verificare o falsificare empiricamente un'assegnazione di riferimento ha una lunga e autorevole tradizione in filosofia del linguaggio. Essa è direttamente suggerita dal fatto che neppure l'ostensione può fissare univocamente il riferimento di un termine. Quando qualcuno indica qualcosa, si chiede Wittgenstein*, chi ci dice che indichi nella direzione che va dalla spalla al dito o non piuttosto nella direzione opposta, o in una bizzarra direzione a zig-zag? E cosa fa, osserva Quine*, uno che dice "quello è un coniglio"? Indica un coniglio e usa un predicato che si applica a conigli oppure indica una parte integrale di coniglio e usa un predicato che si applica a parti integrali di coniglio? O, ancora, indica una fase temporale di coniglio e usa un predicato che si applica a fasi temporali di coniglio? Davidson, a sua volta, osserva che la nozione di riferimento è, nella teoria del significato, un costrutto strettamente teorico. Egli sottolinea l'analogia con la fisica. Nella teoria fisica, spieghiamo i fenomeni macroscopici postulando una struttura fine non osservabile - proprio come, nella teoria semantica, spieghiamo le condizioni di verità degli enunciati postulando relazioni "non osservabili" di riferimento. Ma proprio come controlliamo e interpretiamo la teoria fisica a livello macroscopico, così controlliamo e interpretiamo la teoria semantica a livello dei T-enunciati: ad esempio, se i parlanti inglesi assentono a "it is raining" ogniqualvolta piove e in nessun'altro caso, questo è un buon indizio a favore dell'ipotesi che "it is raining" sia vero in inglese se e solo se sta piovendo (naturalmente questa è una grezza semplificazione ma aiuta ad afferrare l'idea).

Cosa intendiamo dire, esattamente, quando diciamo che quella di riferimento, in semantica, è una nozione teorica? Secondo Church 1956, esistono due modi di concepire le regole semantiche di una lingua formalizzata. Possiamo assumere la nozione di riferimento come preliminarmente nota ed usare le regole semantiche per conferire significato a un sistema logistico non interpretato. Oppure possiamo pensare che le regole semantiche, nel loro complesso, diano una definizione della nozione di riferimento nello stesso modo in cui le regole di buona formazione danno una definizione di "ben formato". Il concetto di riferimento apparterrebbe allora alla sintassi teorica, "non essendo stato usato nulla di semantico nella sua definizione" (Church, 1956, p. 65). Secondo Wallace, chi pensa che quella di riferimento sia una nozione teorica, concepisce le regole semantiche in un terzo modo. Egli tratta il concetto di verità come preliminarmente noto e pensa che le regole semantiche diano una definizione corretta di riferimento se generano (tramite le regole di proiezione) condizioni di verità empiricamente accettabili. Da questo punto di vista, la nozione di riferimento "non può propriamente essere ascritta alla sintassi teorica e tuttavia il suo contenuto semantico non è auto-generato ma è totalmente derivato dal concetto semantico di verità" (Wallace, 1977, p. 152).

In termini più superficiali ma più diretti, l'irriducibilità empirica del riferimento significa che la totalità dell'evidenza empirica accessibile all'interprete (il comportamento osservabile del parlante, linguistico e non linguistico, non importa se descritto in termini estensionali o intensionali) non verifica né falsifica direttamente alcuno schema di riferimento. Quindi se due schemi di riferimento

determinano le stesse condizioni di verità per tutti gli enunciati di un linguaggio, non può essere l'evidenza empirica a renderne uno più o meno corretto dell'altro. Né, apparentemente, può essere qualcosa d'altro: quale altro ruolo teorico può avere uno schema di riferimento se non quello di generare una definizione di verità attraverso delle regole di proiezione? Quindi uno schema di riferimento è corretto esattamente nella misura in cui lo sono i T-enunciati (assegnazioni di condizioni di verità) che esso genera tramite le regole di proiezione. Ecco prodotta la premessa A1..

L'irriducibilità empirica del riferimento è una assunzione discussa (ad esempio, è rifiutata dai sostenitori della teoria causale del riferimento). Quindi, un modo di negare A1. è confutare l'idea che il riferimento sia empiricamente irriducibile, nel qual caso potrebbe essere l'evidenza empirica a discriminare fra due schemi di riferimento che generano le stesse condizioni di verità per tutti gli enunciati del linguaggio. Un altro modo di negare A1. è quello di mostrare che la funzione teorica di uno schema di riferimento non si esaurisce nel generare le condizioni di verità di certi enunciati: potrebbe essere qualcosa di diverso dall'evidenza empirica a discriminare due schemi di riferimento che generano le stesse condizioni di verità per ogni enunciato.

Sottodeterminazione semantica del riferimento

L'esistenza di schemi di riferimento incompatibili in grado di generare le stesse condizioni di verità per tutti gli enunciati del linguaggio è documentata dall'esistenza di permutazioni dell'universo, funzioni uno-uno che mappano ogni individuo dell'universo su un individuo dello stesso. Più precisamente, una funzione f è una permutazione (del dominio delle variabili "x", "y" e "z") se e solo se:

- (i) $(\forall x)(\exists y)(y = f(x))$; f non è una funzione parziale: per ogni individuo ne esiste un altro che è f del primo
- (ii) $(\forall x)(\exists z)(x = f(z))$; ogni individuo è f di qualche altro individuo
- (iii) $(\forall x)(\forall y)(f(x) = f(y) \Leftrightarrow x = y)$ Due individui sono identici se e solo se lo sono gli individui che sono f di essi.

Dato uno schema di riferimento S per un linguaggio L e definita una funzione permutativa f sull'universo di L (il dominio delle sue variabili non ristrette), risulta implicitamente dato anche un nuovo schema di riferimento S^f per L . Se S dice che "cane" si applica ad un oggetto arbitrario a se e solo se a è un cane, S^f dice che "cane" si applica ad un oggetto arbitrario a se e solo se $f(a)$ è un cane, e così via, per qualsiasi predicato di L si voglia considerare al posto di "cane" (assumiamo, per ora, che L non contenga costanti individuali fra i suoi termini primitivi - come non ne conterrebbero l'italiano o l'inglese se Kripke avesse torto e i nomi propri fossero descrizioni definite camuffate - e che S si limiti a specificare quali parole di L sono predicati, e a che cosa si applicano).

E' chiaro che, se f non è la funzione di identità, S e S^f sono incompatibili (assegnano ai predicati di L estensioni diverse). Tuttavia si può mostrare che S e S^f , tramite le familiari regole di proiezione tarskiane, determinano le stesse condizioni di verità per ogni enunciato di L . Un abbozzo di dimostrazione può essere dato come segue.

Se L non comprende costanti individuali fra i suoi termini primitivi, tutti gli enunciati di L avranno la forma " $(\exists x) Px$ " oppure " $(\forall x) Px$ " (dove " P " è un predicato

semplice o complesso quanto si vuole), oppure saranno congiunzioni, disgiunzioni, negazioni (e cos' via) di enunciati di questa forma. Le regole di proiezione potranno dunque essere cos' formulate:

Regole di proiezione per L

1. un enunciato di forma " $(\exists x) Px$ " é vero se e solo se qualcosa soddisfa " Px ";
2. un enunciato di forma " $(\forall x) Px$ " é vero se e solo se ogni cosa soddisfa " Px ";
3. le regole di proiezione per enunciati di forma $A \wedge B$, $A \vee B$ ecc., sono date nel modo familiare.

Schema S

$(\forall y)$ (y soddisfa "x é un cane" se e solo se y é un cane)
 $(\forall y)$ (y soddisfa "x é un gatto" se e solo se y é un gatto)
 ecc.

Schema S ϕ

$(\forall y)$ (y soddisfa "x é un cane" se e solo se $f(y)$ é un cane)
 $(\forall y)$ (y soddisfa "x é un gatto" se e solo se $f(y)$ é un gatto)
 ecc.

Supponiamo ora che un qualsiasi enunciato di forma " $(\exists x) Px$ ", per esempio " $(\exists x) x$ é un gatto" sia vero in base ad S. Allora, per 1., qualcosa soddisfa "x é un gatto". Quindi, per S, qualcosa (diciamo a) é un gatto. Ora, se f^{-1} é l'inversa di f [cioé una funzione tale che, per ogni x e ogni y, $x = f(y)$ se e solo se $y = f^{-1}(x)$] allora $f(f^{-1}(a)) = a$ e quindi $f(f^{-1}(a))$ é un gatto. Quindi, per S^f , $f^{-1}(a)$ soddisfa "x é un gatto" e, per 1., " $(\exists x) x$ é un gatto" é vero. Dunque enunciati di forma " $(\exists x) Px$ " sono veri in accordo a S se e solo se sono veri in accordo a S^f , cioé S e S^f generano, tramite 1.- 3., le stesse condizioni di verità per enunciati di questa forma. Essi generano le stesse condizioni di verità (tramite 1.-3.) anche per enunciati di forma " $(\forall x) Px$ ", considerato che " $(\forall x)$ " é definibile come " $\neg(\exists x)\neg$ ". E, di fatto, generano le stesse condizioni di verità per ogni enunciato di L, non potendo avere diverse condizioni di verità enunciati formati applicando le stesse funzioni di verità (\neg , \vee , \wedge , ecc.) ad enunciati aventi le stesse condizioni di verità.

L'argomentazione funziona sotto l'assunzione che f sia una funzione permutativa: se f non soddisfacesse il requisito (ii), $f^{-1}(a)$ potrebbe non esistere, e se non soddisfacesse il requisito (iii), " $f(f^{-1}(a))$ " potrebbe essere un termine a denotazione ambigua; in entrambi i casi sarebbe falso, o né vero né falso, che $a = f(f^{-1}(a))$. D'altronde, se f non soddisfacesse il requisito (i), non avremmo alcuna garanzia che S^f assegni un riferimento ad ogni predicato di L (la funzione f potrebbe restare priva di valore ogniqualvolta abbia per argomento un membro dell'estensione che S assegna a qualche predicato P di L).

L'argomentazione mostra che, se L é un linguaggio privo di costanti individuali (primitive), e se é possibile definire sull'universo di L una funzione permutativa f, allora schemi di riferimento conflittuali produrranno (tramite regole di proiezione prefissate) identiche condizioni di verità per gli enunciati di L. Wallace ha mostrato [5] che ciò vale anche per linguaggi con nomi non eliminabili. Supponiamo che L' sia una estensione di L ottenuta aggiungendo ad L un certo

numero di costanti individuali primitive. S' e S'^f saranno allora schemi di riferimento per L' in tutto identici (rispettivamente) ad S e S^f se si eccettua il fatto che S' e S'^f conterranno, per ogni nome n di L' , una clausola che dice che cosa n denota. Per ogni n , se S' dice che n denota b , S'^f dirà che n denota non $f(b)$ ma $f^{-1}(b)$, dove f^{-1} è l'inversa di f . S'^f è allora esattamente ciò che si ottiene da S' applicando permutazioni reciprocamente inverse alle clausole di S' che dicono di che cosa i predicati di L' sono veri e a quelle che dicono che cosa i nomi di L' denotano: per ogni espressione e di L' , se S' dice che e è un predicato che si applica ad un oggetto arbitrario a se e solo se a è P , S'^f dirà che e si applica ad a se e solo se $f(a)$ è P ; e se S' dice che e è un nome che denota b , S'^f dirà che e denota $f^{-1}(b)$.

Regole di proiezione per L'

- 1., 2. e 3. Come sopra.
4. Un enunciato di forma " Pa " è vero se e solo se l'oggetto denotato da " a " è uno degli oggetti a cui " P " si applica.

Schema S'

Clausole per predicati
Come in S .

Clausole per nomi

"Bobby" si riferisce a Bobby
"Smoggy" si riferisce a Smoggy
ecc.

Schema S'^f

Clausole per predicati
Come in S^f

Clausole per nomi

"Bobby" si riferisce a $f^{-1}(\text{Bobby})$
"Smoggy" si riferisce a $f^{-1}(\text{Smoggy})$
ecc.

Naturalmente, S' e S'^f genereranno identiche condizioni di verità per enunciati di L in cui non ricorrono costanti individuali (quando non siano in gioco costanti individuali, S' e S'^f si comportano esattamente come S e S^f). Ma prendiamo un enunciato atomico di forma soggetto-predicato come "Smoggy è un gatto". Per S' , "Smoggy" denota Smoggy e "è un gatto" si applica ad un oggetto arbitrario a se e solo se a è un gatto. Quindi S' - tramite 4. - genera il T-enunciato: "'Smoggy è un gatto' è vero se e solo se Smoggy è un gatto". Per S'^f , invece, "Smoggy" denota $f^{-1}(\text{Smoggy})$ e "è un gatto" si applica ad un oggetto arbitrario a se e solo se $f(a)$ è un gatto. Quindi S'^f - tramite 4. - genera il T-enunciato: "'Smoggy è un gatto' è vero se e solo se $f(f^{-1}(\text{Smoggy}))$ è un gatto". Ma questa è una banale riformulazione del T-enunciato generato da S' , dato che $f(f^{-1}(\text{Smoggy})) = \text{Smoggy}$ per definizione di f^{-1} . S'

ed S^f non possono che produrre le stesse condizioni di verità anche per enunciati di forma soggetto-predicato.

Per rendere la cosa più intuitiva, potremmo supporre che "f" abbrevi "la cosa a destra di" e "f⁻¹" la sua inversa, cioè "la cosa a sinistra di". Possiamo dire che "Smoggy é un gatto" é vero se e solo se Smoggy é un gatto oppure che é vero se e solo se la cosa a destra della cosa a sinistra di Smoggy é un gatto. Poco cambia, perché la cosa a destra della cosa a sinistra di Smoggy é proprio Smoggy.

III. I problemi della seconda premessa

Può sembrare che, quando un linguaggio ha i mezzi per esprimere attribuzioni di riferimento alle proprie espressioni, non sia più così facile mostrare che schemi di riferimento incompatibili possono generare le stesse condizioni di verità per ognuno dei suoi enunciati. Supponiamo di aggiungere a L' un certo numero di predicati a due posti fra cui "si riferisce a" e un certo numero di nomi "metateorici" che abbiano come referenti le stesse espressioni di L'. Chiamiamo L* questa estensione di L', in tutto simile a una lingua naturale: L*, infatti, proprio come l'italiano, ha i mezzi per esprimere una teoria del riferimento per se stesso.

*Regole di proiezione per L**

1., 2., 3. e 4. Come sopra.

5. Un enunciato di forma "aRb" é vero se e solo se la coppia ordinata degli oggetti cui si riferiscono, nell'ordine, "a" e "b" é una delle coppie cui "R" si applica.

Schema S*

Clausole per predicati

A 1 posto

Come in S.

A 2 posti

$(\forall z)(\forall j)$ ($\langle z, j \rangle$ soddisfa "x é a destra di y" se e solo se z é a destra di j)

$(\forall z)(\forall j)$ ($\langle z, j \rangle$ soddisfa "x si riferisce a y" se e solo se z si riferisce a j)
ecc.

Clausole per nomi

"Bobby" si riferisce a Bobby

"Smoggy" si riferisce a Smoggy

ecc.

" 'Bobby' " si riferisce a "Bobby"

" 'Smoggy' " si riferisce a "Smoggy"

ecc.

Schema S*f

Clausole per predicati

A 1 posto

Come in S^f

A 2 posti

$(\forall z)(\forall j)$ ($\langle z, j \rangle$ soddisfa "x si riferisce a y" se e solo se f(z) si riferisce a f(j))

$(\forall z)(\forall j) (\langle z, j \rangle \text{ soddisfa "x é a destra di y" se e solo se } f(z) \text{ é a destra di } f(j))$

Clausole per nomi

"Bobby" si riferisce a $f^{-1}(\text{Bobby})$

"Smoggy" si riferisce a $f^{-1}(\text{Smoggy})$

ecc.

" 'Bobby' " si riferisce a $f^{-1}(\text{"Bobby"})$

" 'Smoggy' " si riferisce a $f^{-1}(\text{Smoggy})$

ecc.

I T-enunciati generati per " 'Smoggy' si riferisce a Smoggy" rispettivamente da S^* e da S^{*f} sono i seguenti:

S^* : " 'Smoggy' si riferisce a Smoggy" é vero se e solo se "Smoggy" si riferisce a Smoggy.

S^{*f} : " 'Smoggy' si riferisce a Smoggy" é vero se e solo se $f(f^{-1}(\text{"Smoggy"}))$ si riferisce a $f(f^{-1}(\text{Smoggy}))$

E' facile vedere che i due T-enunciati sono del tutto equivalenti. Ma prendiamo il secondo, quello generato dallo schema permutato S^{*f} . Se le condizioni di verità di " 'Smoggy' si riferisce a Smoggy" sono quelle che S^{*f} determina, " 'Smoggy' si riferisce a Smoggy" contraddice S^{*f} stesso, secondo cui "Smoggy" si riferisce a $f^{-1}(\text{Smoggy})$. Adottare S^{*f} equivale quindi a trattare come universalmente falsi enunciati come " 'Smoggy' si riferisce a Smoggy", " 'Bobby' si riferisce a Bobby", " 'Roma' si riferisce a Roma" ecc. Ma é proprio imparando ad accettare enunciati del genere, enunciati conformi a "paradigmi decitazionali analoghi al paradigma tarskiano di verità (Quine PT, 52), che apprendiamo ad usare espressioni come "si riferisce a", "denota", "é il nome di" ecc. E sembra che qualunque schema ci inviti apertamente a considerare come falsi enunciati del genere sia per ciò stesso evidentemente scorretto e inaccettabile. Per evitare il problema, possiamo provare a riformulare S^{*f} come segue.

Schema S^{*f}

Clausole per predicati

A 1 posto

Come in S^f

A 2 posti

$(\forall z)(\forall j) (\langle z, j \rangle \text{ soddisfa "x é a destra di y" se e solo se } f(z) \text{ é a destra di } f(j))$

ecc.

$(\forall z)(\forall j) (\langle z, j \rangle \text{ soddisfa "x si riferisce a y" se e solo se } f(z) \text{ si riferisce a } j)$

Clausole per nomi

Come in S^f

I T-enunciati generati per " 'Smoggy' si riferisce a Smoggy" rispettivamente da S^* e da S^{*f} sono i seguenti:

S^* : " 'Smoggy' si riferisce a Smoggy" é vero se e solo se "Smoggy" si riferisce a Smoggy.

S^{*f} : " 'Smoggy' si riferisce a Smoggy" é vero se e solo se $f(f^{-1}(\text{"Smoggy"}))$ si

riferisce a $f^{-1}(\text{Smoggy})$.

La situazione di S^*f' é molto diversa da quella di Sf^* . Se le condizioni di verità di " 'Smoggy' si riferisce a Smoggy" sono quelle che S^*f' determina, " 'Smoggy' si riferisce a Smoggy" non contraddice S^*f' stesso, secondo cui "Smoggy" si riferisce a $f^{-1}(\text{Smoggy})$, ma lo conferma. Quindi, possiamo adottare S^*f' senza dover trattare come falsi enunciati decitazionali come "'Smoggy' si riferisce a Smoggy", " 'Bobby' si riferisce a Bobby" o simili - proprio ciò che non accadeva per Sf^* . Ma emerge una difficoltà nuova perché S^* e S^*f' non assegnano all'enunciato di L^* " 'Smoggy' si riferisce a Smoggy" le stesse condizioni di verità. La conclusione é che per linguaggi in grado di esprimere teorie del riferimento per se stessi, come L^* e come l'italiano, indeterminanze di riferimento non possono presentarsi separatamente da indeterminanze di condizioni di verità. Cosí, se abbiamo a che fare con un linguaggio L capace di esprimere attribuzioni di riferimento alle proprie stesse espressioni, per esempio con una lingua naturale come l'italiano o l'inglese, il riferimento delle espressioni di L può essere imperscrutabile solo a patto che le condizioni di verità di tutti gli enunciati di L in cui ricorre "si riferisce a", o espressioni simili, siano indeterminate. Per linguaggi del genere, e quindi in particolare per le lingue naturali, una prova veramente cogente dell'imperscrutabilità del riferimento presuppone che si sia preliminarmente provata l'indeterminatezza delle condizioni di verità. Questa é una conclusione interessante ed istruttiva, ma non rappresenta un non sequitur per l'argomentazione che ha (I.R.2) come conclusione, almeno nella misura in cui i motivi di Quine per affermare che le condizioni di verità degli enunciati sono indeterminate sembrano convincenti. Secondo Davidson, vi sono ragioni per pensare che le condizioni di verità degli enunciati non possano essere sistematicamente indeterminate. Ma questo non esclude che almeno certi enunciati del linguaggio, nella fattispecie quelli che includono "si riferisce a", possano avere condizioni di verità del tutto indeterminate.

IV. I problemi della prima premessa

Fra i termini primitivi di L' , aggiungiamo ora i funtori " f " e " f^{-1} " stessi (o espressioni diverse di cui " f " e " f^{-1} " siano comunque le traduzioni canoniche). Otteniamo una estensione di L' che chiamiamo L^+ . Insieme ad L^+ , sarà facile generare due schemi di riferimento S^+ e S^+f in tutto identici a S' e S'^f eccetto per il fatto di contenere, aggiuntivamente, delle clausole per i funtori " f " e " f^{-1} ". Queste clausole dovranno essere tali da far s" che per lo schema "normale" o omofonico S^+ " $f(\text{Smoggy})$ " si riferisca a $f(\text{Smoggy})$ e " $f^{-1}(\text{Smoggy})$ " si riferisca a $f^{-1}(\text{Smoggy})$ mentre per lo schema f -permutato S^+f , " $f(\text{Smoggy})$ " dovrà riferirsi a $f(f^{-1}(\text{Smoggy}))$ e " $f^{-1}(\text{Smoggy})$ " a $f^{-1}(f^{-1}(\text{Smoggy}))$. E' chiaro che, in caso contrario, S^+ e S^+f non produrrebbero le stesse condizioni di verità per enunciati come " $f(\text{Smoggy})$ non é un gatto", " $f^{-1}(f(\text{Bobby}))$ é un cane" ecc. (si ricordino le clausole per predicati date in S' e S'^f). Le clausole per " f " e " f^{-1} " potrebbero quindi essere, sia in S^+ che in S^+f qualcosa del genere.

Schema S^+ e Schema S^+f

$(\forall x) (\forall n) [("f" \cup n \text{ si riferisce a } x) \text{ se e solo se } (\exists y) (n \text{ si riferisce a } y \wedge x = f(y))]$

$(\forall x) (\forall n) [("f^{-1}" \cup n \text{ si riferisce a } x) \text{ se e solo se } (\exists y) (n \text{ si riferisce a } y \wedge x = f^{-1}(y))]$

(dove " n " varia su termini singolari di L^+ e " f^{-1} " » n é l'espressione composta,

nell'ordine, dalle espressioni " f^{-1} " e n).

Queste clausole dicono che una espressione come " $f^{-1}(\text{Smoggy})$ " si riferisce a f^{-1} di ciò cui si riferisce "Smoggy", e quindi a $f^{-1}(\text{Smoggy})$ per S^+ e a $f^{-1}(f^{-1}(\text{Smoggy}))$ per S^+f (vedi le clausole per nomi date in S' e S'^f). Proviamo ora a generare, partendo da S^+ e S^+f , le condizioni di verità di un enunciato come

$$(e^1) \quad \neg \exists x (x = f(\text{Smoggy}))$$

Assumendo che S^+ e S^+f trattino "=" allo stesso modo di un predicato a due posti, essi genereranno per (e^1) - rispettivamente - le seguenti condizioni di verità.

(e^1) é vero se e solo se:

$$\begin{aligned} (e^{1+}) & \quad \neg \exists x (x = f(\text{Smoggy})) \\ (e^{1+f}) & \quad \neg \exists x (f(x) = f(f^{-1}(\text{Smoggy}))) \end{aligned}$$

dove (e^{1+}) e (e^{1+f}) sono rispettivamente l'interpretazione di (e^1) in S^+ e S^+f . Poiché sappiamo per ipotesi che f e f^{-1} sono permutazioni, (e^{1+}) e (e^{1+f}) hanno esattamente le stesse condizioni di verità - vedi la definizione di funzione permutativa data sopra. Per lo stesso motivo, sappiamo anche che (e^{1+}) e (e^{1+f}) sono entrambi falsi (entrambi implicano che f^{-1} non sia una permutazione, perché implicano che f^{-1} non soddisfi la clausola (i) della definizione di funzione permutativa data sopra). Ma ora vorrei disinteressarmi del valore di verità di (e^{1+}) e (e^{1+f}) , e considerare invece cosa possiamo inferire da essi. Poniamo ad esempio che " f " sia "a destra di" e " f^{-1} " sia "a sinistra di". Da (e^{1+}) si potrebbe inferire che non esiste nulla che sia a destra di Smoggy. Ma da (e^{1+f}) non lo si potrebbe inferire: potrebbe esservi qualcosa a destra di Smoggy ma non esistere nulla che stia a sinistra di Smoggy; in tal caso, non esisterebbe nulla che stia a destra di ciò che sta a sinistra di Smoggy e quindi neppure qualcosa che stia a destra di ciò che sta a destra di ciò che sta a sinistra di Smoggy.

Per chiarire meglio la situazione, proviamo a interpretare un altro enunciato di L^+ , e precisamente:

$$(e^2) \quad \neg \exists x (x = f^{-1}(f(\text{Smoggy}))).$$

S^+ e S^+f generano, rispettivamente, le seguenti interpretazioni di (e^2) :

$$\begin{aligned} (e^{2+}) & \quad \neg \exists x (x = f^{-1}(f(\text{Smoggy}))) \\ (e^{2+f}) & \quad \neg \exists x (f(x) = f(f^{-1}(f^{-1}(\text{Smoggy})))) \end{aligned}$$

E' chiaro che (e^{2+}) é una conseguenza logica di (e^{1+}) : se non c'è nulla a destra di Smoggy, certo non può esservi nulla che stia a sinistra di ciò che sta a destra di Smoggy. Ma (e^{2+f}) , al contrario, non é una conseguenza logica di (e^{1+f}) : (e^{1+f}) potrebbe essere vero non perché non esista qualcosa a sinistra di Smoggy ma perché non esiste nulla alla sua destra. In questo caso, (e^{2+f}) sarebbe falso anche se (e^{1+f}) fosse vero. Dunque $(e^{1+}) \div (e^{2+})$ ma $\neg [(e^{1+f}) \div (e^{2+f})]$. In altri termini, secondo S^+ $(e^1) \sim (e^2)$, mentre secondo $S^+f \neg [(e^1) \div (e^2)]$. Secondo S^+ , la congiunzione di (e^1) con la negazione di (e^2) é una contraddizione in termini, secondo S^+f é un enunciato del tutto consistente - anche se falso, assumendo che f sia una permutazione. S^+ e

S^+f , cioè, danno diverse immagini della relazione di conseguenza logica fra enunciati di L^+ , anche se assegnano condizioni di verità equivalenti ad ogni enunciato di L^+ .

VI. Scrutabilità del riferimento?

Torniamo ora all'argomentazione che, muovendo da A1. e A2., arriva all'imperscrutabilità del riferimento. Ho già notato ciò che appare anche ad occhio nudo e cioè che l'argomentazione é formalmente corretta. Ma il fatto che diversi schemi di riferimento possano rappresentare diversamente i nessi di conseguenza logica fra gli enunciati di un linguaggio pur generando per ogni enunciato - tramite regole di proiezione fisse - le stesse condizioni di verità sembra gettare qualcosa più che un'ombra sulla prima premessa dell'argomentazione. A1. asseriva che due schemi di riferimento sono ugualmente corretti se determinano le stesse condizioni di verità per tutti gli enunciati di un linguaggio. La premessa era resa verosimile dall'apparente irriducibilità empirica del riferimento e dall'impressione che, a parte il ruolo che giocano nel produrre una definizione di verità, gli schemi di riferimento non svolgano nessuna altra funzione teorica. Così, sembrava che né l'evidenza empirica né qualcosa d'altro potesse discriminare due schemi capaci di generare le stesse condizioni di verità per ogni enunciato di un linguaggio.

Ho rinunciato a discutere il problema della riducibilità empirica del riferimento, con tutto ciò che il problema coinvolge, ad esempio l'accettabilità di una teoria causalista dei nomi propri, o dei nomi di tipi naturali. Ma anche assumendo che il riferimento sia empiricamente irriducibile, sembra che due schemi che generano definizioni di verità equivalenti possano essere discriminati empiricamente dalla loro disposizione a produrre diverse mappature dei nessi di conseguenza logica fra enunciati. L'attività linguistica del ragionare, dell'inferire o del passare da asserzioni ad altre asserzioni sembra infatti non meno costitutiva di ciò che potremmo chiamare "comportamento linguistico manifesto" di quanto non lo sia l'attività di produrre enunciati in presenza di particolari stimolazioni sensorie, o di specifiche "parti" di mondo (oggetti ed eventi). Alla domanda di Dummett: "come si può manifestare la conoscenza delle condizioni di verità di un enunciato teorico di una teoria scientifica?", Bilgrami risponde: "usando l'enunciato in spiegazioni riuscite" (Bilgrami 1986, p. 104). Alla domanda: "come si può manifestare la conoscenza del riferimento dei nostri termini?" noi diamo una risposta simile: usando certi enunciati, e non altri, dentro a dimostrazioni riuscite (il fatto che (e^1) venga consensualmente trattato come una buona prova di (e^2), ad esempio, configura un pregiudizio negativo a carico di S^+f).

Davidson ha scritto: "Vedere la struttura di un enunciato attraverso gli occhi di una teoria della verità equivale a vederla come costruita a partire da meccanismi bastanti, in numero finito, a costruire ogni enunciato; la struttura dell'enunciato determina cos' le sue relazioni con gli altri enunciati. E in effetti, non si possono dare le condizioni di verità di tutti gli enunciati senza mostrare che alcuni di essi sono conseguenze logiche di altri" (Davidson, 1984, p. 61). Nulla di quanto ho detto sopra contraddice questo assunto. Ma, anche cos', resta il fatto che teorie della verità del tutto equivalenti possono caratterizzare in modi sottilmente diversi la relazione di conseguenza logica. Per sostenere che il riferimento é imperscrutabile non basta mostrare che schemi incompatibili di riferimento generano definizioni di verità equivalenti. E' necessario mostrare che queste equivalenti teorie della verità generano mappature equivalenti del nesso di conseguenza logica. Credo di avere mostrato che questo non si può fare: non, almeno, con il metodo delle permutazioni.

In estrema sintesi: é vero che schemi incompatibili di riferimento possono generare le stesse condizioni di verità per tutti gli enunciati di un linguaggio, ma può

darsi che non generino la stessa teoria del significato per quel linguaggio quando si includa nella teoria del significato - come sembra si debba fare - una rappresentazione dei nessi di dimostrabilità/provabilità fra enunciati. Questo è un vecchio punto, che richiama alla asimmetria-complementarità delle nozioni di verità e asseribilità nella spiegazione della nozione di significato, una tematica su cui manca ovviamente lo spazio per fare qualcosa di più che questo cenno.

Ciò che sorprende è la levità ontologica di questa difesa della nozione di riferimento dal cono d'ombra dell'impercipiabilità. Perché, se segmentiamo la competenza linguistica in competenza referenziale, lessicale e inferenziale (cfr. Marconi 1997), si vede che ciò che salva la relazione di riferimento dalla reductio ad absurdum generata da I.R. è il richiamo alla competenza inferenziale: quanto di più lontano si potrebbe immaginare da quell'area di competenza referenziale in cui sembrerebbe localizzarsi e svolgersi la vicenda del contatto/attrito fra linguaggio e mondo.

Note

1. Rammenta la critica wittgensteiniana del "linguaggio privato".
2. Vedi "Ontological Relativity" in Quine 1969.
3. Questa tassonomia dell'indeterminatezza in Quine è essenzialmente quella tracciata da Davidson in 1979, p. 228.
4. Cfr. Wallace 1977, pp. 146-9.
5. Cfr. Wallace 1977, pp. 149-50.

Bibliografia

- Bilgrami A. (1986) "Meaning, Holism and Truth", in E. Lepore (a cura di) *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson*, Oxford 1986.
- Church A. (1956) *Introduction to mathematical Logic*, Princeton 1956
- Davidson D. (1968) "Semantics for Natural Languages", *Linguaggi nella società e nella tecnica*, Milano, 1970.
- Davidson D. (1977) "Reality without Reference", *Dialectica*, 31, 1977
- Davidson D. (1979) "The Inscrutability of Reference", *The Southwestern Journal of Philosophy*, 10, '79
- Field H. (1974) "Quine and the Correspondence Theory", *Philosophical Review*, 83, 1974.
- Field H. (1975) "Conventionalism and Instrumentalism in Semantics", *Nous*, 9, 1975.
- Marconi D. (1997) *Semantic Competence*, MIT Press (forthcoming)
- Quine W.V.O. (1960) *Word and Object*, Cambridge Mass. 1960
- Quine W.V.O. (1969) *Ontological Relativity and Other Essays*, New York 1969
- Wallace J. (1977) "Only in the Context of a Sentence do Words have any Meaning", *Midwest Studies in Philosophy*, II, 1977
- Wittgenstein L. (1921) *Tractatus Logico Philosophicus*, London 1921