

Hilary PUTNAM
ZNAČENJE I MENTALIZAM¹

Danas bi neko ko čita recenziju neke filozofske knjige u *New York Review of Books* mogao naići na riječ “intencionalnost”. Ipak, mali broj recenzenata nikada ne kaže što znači ova riječ. Ne zbog toga što je riječ besmislena; radije zato što je ona postala riječ za naziv poglavlja: riječ koja označava cijeli niz tema i pitanja, a ne jednu određenu temu. Konkretno, slijedeće činjenice se obično navode kao primjeri “intencionalnosti”: (1) činjenica da riječi, rečenice i druge “reprezentacije” imaju *značenje*; (2) činjenica da reprezentacije mogu *referirati* na (tj. biti istinite za) neke aktualno postojeće stvari ili na svaku od aktualno postojećih stvari; (3) činjenica da reprezentacije mogu biti *o* nečemu što *ne postoji*; i (4) činjenica da stanje uma može imati “stanje stvari” kao svoj predmet, kao kad neko kaže “ona vjeruje da *on je pouzdan*”, “on se nada da će *njegov šef dobiti otkaz*”, “ona strahuje da *neće biti hrane u kući*”.

Kada je kompjuterska revolucija izbila u svijetu, rasprostranjeno se očekivalo (i tvrdilo se) da će kompjuterski modeli objasniti prirodu tih različitih pojava. Ukratko, ljudi su očekivali da će se dati reduktivno objašnjenje različitih tema uključenih u poglavlje pod naslovom “intencionalnost”. Sada kada se to pokazalo ne tako jednostavnim, brojni mislioci počinju sugerirati da nije tako loše ako se to ne može učiniti; intencionalnost je ionako samo obilježje “narodne psihologije”. Ako se ne može dati prvoklasno znanstveno objašnjenje intencionalnih činjenica i fenomena, to nije zato što znanstveni redukcionizam nije prava mjera za metafiziku, već je to zato što tamo, da tako kažemo, nema ničega što bi se moglo reducirati. Želim tvrditi da su oba stava pogrešna; da intencionalnost neće biti reducirana i neće nestati.

Ta tvrdnja – tvrdnja da se “intencionalnost neće reducirati i da neće nestati” – ponekad se naziva “Brentanova teza”, prema filozofu kojem se (ne sasvim tačno)¹ pripisuje to da je brani u drugom dijelu devetnaestog stoljeća. Ponekad se to stajalište navodi kao pozitivna tvrdnja: tvrdnja da je intencionalnost *primitivni fenomen*, zapravo fenomen koji povezuje misao i stvar, umove i vanjski svijet.

Na određeni način, ovo pozitivno gledište slijedi odmah iz jednog negativnog, međutim, postoji i jedan džoker u paketu karata. Taj džoker

¹ Tekst “Značenje i mentalizam” preveden je ovdje iz knjige Hilary Putnam: *Representation and Reality*. Chapter 1. *Meaning and Mentalism*. The MIT Press, 1998. pp.1-18.

je stari filozofski problem o Jednom i Mnogom. Ako pretpostavimo da, kad god imamo različite fenomene okupljene pod jednim imenom, Mora Postojati Nešto Što Svi Imaju Zajedničko, dakle, doista će slijediti da postoji jedan pojedinačni fenomen (i, ako se ne može reducirati, mora biti “primitivan”) koji odgovara intencionalnosti.

Da biste vidjeli poteškoće, razmotrite svojstvo “crveno”. Intuitivno, sve “crvene” stvari imaju “nešto zajedničko”. Ali u znanstvenom smislu one to nemaju, osim ako je to “refleksnost” – dispozicija da selektivno emitiraju i apsorbiraju određene valne duljine svjetlosti. Takva dispozicija – dispozicija koja utječe na stvari na određeni način (stvari koje nisu ljudska mentalna stanja) – bi se u sedamnaestom stoljeću nazvala “tercijarnim svojstvom”. (Sekundarna svojstva bila su dispozicije koje utječu na naše umove, a primarna svojstva su bila “u samoj stvari” baš kao što ih zamišljamo.) Ako se ograničimo na nedispozicijska (ili “strukturna”) svojstva koja se mogu opisati u fizičkoj znanosti, onda ne postoji znanstveno opisivo svojstvo koje je zajedničko svim crvenim stvarima, ne postoji strukturno svojstvo koje konstituira “crvenilo” crvene zvijezde, crvene svjetlosti, crvene jabuke itd., osim ako smo spremni razmotriti veliku (možda beskonačnu) disjunkciju strukturnih svojstava biti jedno “fizičko svojstvo”. Ipak, postoji smisao u kojem one još uvijek imaju “nešto zajedničko” – nešto nedispozicijsko i nedisjunktivno. U običnom jezičkom smislu *izraza* crveno sve su one “crvene”. Naravno, hoće li neko priznati da je to *stvarno* “nešto zajedničko”, ovisit će o tome vjeruje li neko da je zdravorazumska verzija svijeta jednako legitimna kao i znanstvena verzija. Filozof (poput mene) ne treba odustati od tvrdnje da crvene stvari “imaju nešto zajedničko”. Ali on mora to pitanje razdvojiti od pitanja: “Imaju li one nešto zajedničko što se može opisati u nedispozicijskim terminima na razini egzaktnosti?” Stvari mogu imati nešto zajedničko u jednom opisu svijeta, a ne u drugom.

Pokušat ću pokazati da nema znanstveno opisivog svojstva koje svi slučajevi nekog određenog intencionalnog fenomena imaju zajednički. Ovom tezom mislim zaniijekati da postoji neka znanstveno opisiva “priroda” koju svi slučajevi “referencije” općenito, ili “značenja” općenito, ili “intencionalnosti” općenito, imaju; također mislim poricati da postoji bilo kakvo znanstveno opisivo svojstvo (ili “priroda”) koju svi slučajevi bilo kojeg specifičnog intencionalnog fenomena, recimo, “misliti da ima puno mačaka u susjedstvu”, imaju kao zajedničko. Ali te se fenomene ne može odbaciti kao puku narodnu psihologiju, osim ako se sama ideja *da postoje stvari i da mislimo o njima* može odbaciti kao narodna psihologija.

Međutim, usporedba intencionalnosti sa “crvenim” je pogrešna. Bolja usporedba – koju je predložio Wittgenstein² – je s izrazom “igra”. Čak i na nivou običnog jezika, čudno je reći da sve igre “imaju nešto zajedničko”, a to je *da su igre*. Za neke igre podrazumijevaju se pobjeđiva-nje i gubljenje, dok za druge (“ringe ringe raja”) ne; neke igre se igraju za zabavu igrača, druge (gladijatorske igre, profesionalne igre) ne; neke igre imaju više od jednog igrača, druge nemaju; i tako dalje. Na isti način, kada pažljivo ispitamo sve slučajeve u kojima bismo htjeli da je neko “referirao” na nešto (ili čak sve slučajeve u kojima je neko “referirao” na neku određenu stvar), ne nalazimo bilo *koju* relaciju između riječi i onoga na šta ona referira.

U svakom slučaju, ako Wittgenstein kaže da riječ “igra” ne stoji za neko svojstvo, postoji očita kritika koju neko može iznijeti o njegovoj tvrdnji. “Hume je već razlikovao obični (ili, kako je rekao, ‘prirodni’) smisao riječi kao što su ‘relacija’ i ‘svojstvo’ i logički (ili, kako je rekao, ‘filozofski’) smisao tih riječi. Zašto bismo trebali smatrati da je posebno važno da ne postoje svojstva *u običnom jezičkom smislu* izraza koja sve igre imaju zajednički? U logičkom smislu pojma, *postoji* svojstvo koje sve igre imaju zajedničko, to jest *disjunkciju* raznih kriterija koje koristimo da kažemo za stvari da su ‘igre’.” Wittgenstein nudi ono što bi moglo izgledati kao neuvjerljiv odgovor: predstaviti pojam “porodična sličnost” kao što je “igra” kao disjunkciju egzaktnih pojmova znači pogrešno predstaviti njegov karakter. Riječi poput “igra” imaju nejasnoće, fleksibilnost, “otvorenu teksturu” (kako ju je Waismann nazvao), koje se ne mogu reproducirati bez određenog razdvajanja potpuno određenih svojstava.

Ovo izgleda neuvjerljivo, antifilozofski jer, naposljetku, nije li cijela *svrha* racionalne rekonstrukcije “izoštavanje” naših nejasnih pojmova zdravog razuma? Ne želimo da “racionalna rekonstrukcija” pojma dijeli nejasnoću samog predanalitičkog pojma. Ako se sjetimo da Wittgensteina zapravo ne zanimaju riječi poput “igra”, nego riječi poput “referencija”, “jezik”, “značenje”, onda je situacija vrlo različita. Ovdje ću pokušati pokazati da fenomen otvorene tekture daleko nadilazi puku *labavost* konvencionalne primjene koju nalazimo u slučaju riječi “igra”. (Zapravo, to daleko nadilazi puku labavost primjene u slučaju “igre” također.) I to je upravo otvorena tekstura referencije koja pobjeđuje klasične filozofske slike.

Ako je to ispravno, moramo naučiti vidjeti da se nalazimo u položaju koji ne odgovara niti filozofskoj slici intencionalnosti kao fenomenu

koji se reducira na fizičke (ili možda kompjuterske) izraze, niti slici intencionalnosti kao mita, čak ni slici intencionalnosti kao pojedinačnog “ireducibilnog fenomena”.

Dio mog zadatka je ilustrirati (primjenjujući ga na određeni problem) filozofski stav koji napušta mnoge tradicionalne pretpostavke o Pojavi i Stvarnosti; stav koji odustaje, naprimjer, od pretpostavke da je ono što je stvarno ono što je “ispod” ili “iza” ili “temeljnije” nego naše svakodnevne pojave, stav koji odustaje od pretpostavke o Jednom u Mnogštvenom – i također odustaje od pretpostavka da svaki fenomen ima “posljednju prirodu” za koju moramo dati (metafizički reducirano) objašnjenje. Zbog toga je snaga mog argumenta negativna. Tvrdim da je određeni način razmišljanja o značenju i prirodi uma u osnovi pogrešan. Uvijek je manje uzbudljivo čuti da neko kritikuje pokušaj rješenja problema nego čuti kako objavljuje da je pronašao rješenje. Ali mislim da možemo naučiti nešto o prirodi značenja i, možda, nešto o prirodi psihologije, posmatrajući zašto određene ideje o značenju i njegovom mjestu u umu ne funkcioniraju.

Fodor i Chomsky

Da bi objasnili što nije u redu s načinom na koji su filozofi i kognitivni znanstvenici općenito pristupali pitanjima o značenju, bit će potrebno ispitati brojne različite načine na koje se standardizirani pristup očitovao, niz različitih (iako blisko povezanih) načina razmišljanja. Način razmišljanja o kojem ću najprije raspravljati objašnjava Jerry Fodor u djelu *The Language of Thought*.³ Fodor priznaje da mnogo od svoje inspiracije duhuje radu Noama Chomskog. Međutim, Chomsky se nikada nije posvetio mogućnosti pronalaženja “psihološki stvarnih” entiteta koji imaju dovoljno svojstava koja smo preanalitički dodijelili “značenjima” kako bismo opravdali identifikaciju. “Reprezentacije” i “urođene ideje” o kojima Chomsky piše duboke su sintaktičke strukture i sintaktičke univerzalijske. Fodorov program stoga nije istovjetan s Chomskyjevim, već više njegovo odvažno proširenje.

U svakom slučaju, postoji široko rasprostranjeno očekivanje da će se Chomskyjeve ideje prije ili kasnije proširiti na područje semantike, očekivanje koje je odgovorno za velik dio pozornosti koju njegovim idejama poklanjaju francuski neostrukturalisti, američki kognitivni znanstvenici i drugi. Chomsky je poznat po tome što je predložio teoriju prema

kojoj je gramatika “urođena” u umu. Prema Chomskom, postoji univerzalna gramatika – struktura i skup kategorija koje su univerzalne, i to ne samo zato što su ljudska okruženja u određenim aspektima podjednaka već zato što je ova Univerzalna Gramatika ugrađena u osnovnu strukturu samog uma.⁴ Chomsky nadalje sugerira da ova urođena jezična struktura ne karakterizira čitav um, već način funkcioniranja određenog “modula” u umu, takozvanog “jezičkog organa”.⁵ Čini se da Chomsky jezički organ shvaća kao relativno “glup” organ, neovisan o općoj inteligenciji (“ako postoji takva stvar”, rekao bi Chomsky). Čini se da je taj naglasak na gluposti jezičnog organa oštar zaokret od Chomskyjevog ranijeg modela uma kao učenja njegovog maternjeg jezika – uz pomoć, naravno, njegovog poznavanja univerzalne gramatike – stvaranjem hipoteza. Novija pisanja Chomskoga i Fodora prikazuju um kao skup automatski funkcionirajućih “modula”,⁶ a ti napisi naglašavaju “odozdo-prema-gore” za razliku od “odozgor-prema-dolje” obradu – to jest automatsku obradu nasuprot obrade koja se oslanja na opću povezanost i opće informacije.

U svakom slučaju, s obzirom na to da su ključne ideje Chomskog teoretiziranja (1) ideja lingvističkih univerzalija, (2) hipoteza o urođenosti i (3) nedavna ideja modularnosti, forma koja se može očekivati da je dobije Chomskyjeva teorija o semantičkom nivou je relativno jasna (i Fodorova teorija dobija očekivanu formu), čak i ako detalji mogu imati različite oblike. Chomskyjeva teorija semantičkog nivoa reći će da postoje “semantičke reprezentacije” u umu/mozgu; da su one urođene i univerzalne; i da su svi naši koncepti razloživi u takve semantičke reprezentacije. To je teorija koju se nadam da ću uništiti.

Također sam skeptičan glede ideje univerzalne gramatike,⁷ ali to neću raspravljati u ovom radu. Chomskyjev rad, a osobito njegovo oživljavanje “mentalizma” i njegov govor o univerzalnosti u jeziku, pobudio je svjetsku pozornost, a to nije zato što ljudi imaju ogroman interes za gramatiku. Te ideje privukle su pažnju ljudi vrlo daleko od bilo kakve brige o tehničkoj lingvistici: lakanovskim psihoanalitičarima, antropolozima, dječjim psiholozima, filozofima svih vrsta. Očigledno je da ljudi predviđaju da će Chomskyjeva ideja imati implikacije u odnosu na pitanja veća od načina na koji stječemo sintaksu.

Ne bih pokušao uništiti teoriju urođenih semantičkih reprezentacija ako ne mislim da se mnogo toga može naučiti iz proučavanja pitanja koja ona postavlja i odgovora koje ona predlaže, i ako ne mislim da su briljantni mislioci koji predlažu takve teorije u rukama intelektualne čežnje koju samu po sebi vrijedi uzeti ozbiljno. Čežnja je ona koja se objašnjava

dvjema činjenicama o nedavnom razmišljanju o umu.

Jedna je činjenica robusnost najstarijeg obrasca objašnjenja naših mentalnih aktivnosti: objašnjenje u smislu vjerovanja i želja. Bez obzira na to koliko snažno vladala plima bihevizma, mi nikada nismo prestali objašnjavati svoje ponašanje i ponašanje drugih u smislu vjerovanja i želja. Kažemo: “Danas sam otišao u školu jer sam znao da moram pođavati razred.”; ili “Otišao sam u trgovinu jer sam znao da nemamo mlijeka, i želio sam da dam mlijeko u kavu.”

Bihevizam u svom radikalnom obliku sugerirao je da mi ne trebamo ništa od ovoga, jer sve o čemu zapravo govorimo su uvjetovani odgovori itd. Možda se može bez “vjerovanje – želje” govora – kad se radi sa štakorima u vrlo kontroliranim situacijama, ali je čak i tako veliki znanstvenik kao Skinner naletio na nevolje kada je pokušao upotrijebiti “stimulacija – odgovor” jezik kako bi opisao ljudsko verbalno ponašanje. Ono što je Skinner morao učiniti, u osnovi, bilo je proširiti pojmove stimulacija i odgovora tako da (kao što je Chomsky istaknuo u poznatom pregledu prije mnogo godina)⁸ oni postaju prazni. Naprimjer, tokom pokušaja da se analizira govor o Drugom svjetskom ratu, Skinner je rat nazvao stimulacijom. Chomsky je istaknuo da, kada pojam stimulacije postane toliko širok, *Drugi svjetski rat* je “stimulacija” (a odgovor se događa dvadeset godina kasnije), govor o “stimulaciji – odgovoru” postao je puki žargon bez prave kontrole. Dakle, sigurno postoje neki dobri razlozi za odbranu objašnjenja “vjerovanja – želje”.

Druga činjenica je sve veća tendencija da se mozak smatra kompjuterom, a naša psihološka stanja kao softverski aspekt kompjutera. U istraživanjima koja se temelje na takvom pristupu (naprimjer, rad umjetne inteligencije) često se pretpostavlja da kompjuter ima ugrađeni (i time “urođeni”) formalizirani jezik koji može koristiti i kao medij reprezentacije i kao medij računanja. (Ideja *lingua mentisa*, jezika uma, uistinu je drevna ideja koja se pojavila, pomalo nalik ideji Početka Svemira.)

Ako identificiramo kompjuterski *lingua mentis* s Chomskyjevim “semantičkim reprezentacijama”, dolazimo do poznate slike: slike uma kao kriptografa. Um misli svoje misli na mentalnom jeziku (Mentalese), kodira ih u lokalnom prirodnom jeziku, a zatim ih prenosi (recimo, govoreći ih naglas) slušatelju. Slušatelj također ima kriptograf u njegovoj glavi, naravno, koji će potom dešifrirati “poruku”. U ovoj slici prirodni jezik, daleko od toga da je bitan za misao, samo je sredstvo za komunikaciju misli. Ideja o oživljavanju “želja – vjerovanje” psihologije i ideja o kompjuterskom modelu uma može se prizivati iz više razloga. Ako je

Chomsky u pravu, cijelo čovječanstvo ima jednu prirodu, kao što su vjerovali mislioci iz osamnaestog stoljeća. Chomsky je naglasio tu povezanost s prosvjetiteljstvom, te s političkim idealima slobode, jednakosti i bratstva.⁹ Ali čak i osim ovog odjeka, razumljivo je da bi mnogi mislioci trebali biti privučeni programom koji okuplja “želja – vjerovanje” psihologiju i kompjutersko modeliranje.

Želja da se ovo dvoje zajedno uzme i mene je dugo držala. To su dominantne antibiheviorističke tendencije i ja sam vjerovao da će se ujediniti. Želja koja privlači Fodora, a zatim, kao što je mene nekoć obuzimala,¹⁰ to je želja da se “želja – vjerovanje” psihologija učini “znanstvenom” jednostavnim identificiranjem u potpunosti s kompjuterskom psihologijom. Kada sam predložio ovaj program (pod nazivom “funkcionalizam”), mislio sam da je način na koji se to da učiniti jednostavan: jednostavno pretpostavljamo da su želje i vjerovanja “funkcionalna stanja” mozga (tj. karakteristike definirane u smislu kompjuterskih parametara plus odnosa s biološki karakteriziranim unosima i rezultatima). Naprimjer, može se pretpostaviti da *vjerovanje da ima mlijeka u supermarketu* pokazuje jednu od formula¹¹ u *lingua mentisu*, čiji je prijevod “ima mlijeka u supermarketu” u posebnom “registru vjerovanja”. Prikazivanje druge formule u “registru želja” moglo bi biti *željenje mlijeka za sutrašnji doručak*. I prelazak iz tih dvaju kompjuterskih stanja u rezultantu, tj. akcija odlaska u supermarket i kupovina mlijeka, može biti rezultat određenog algoritamskog postupka na ovim prikazanim formulama (kao i na drugim). U takvoj slici mentalna psihologija običnog jezika, narodna psihologija, je gruba aproksimacija idealnom kompjuterskom modelu onoga što se događa u mozgu. *Idealna* “želja – vjerovanje” psihologija bila bi izomorfna kompjuterskom opisu onoga što se događa u mozgu. Napravite tu pretpostavku i imate *mentalizam* u njegovom najnovijem obliku. Mentalizam je samo najnoviji oblik opće tendencije u povijesti misli, tendencija da se pojmovi smatraju znanstveno opisivim (“psihološki stvarnim”) entitetima u umu ili mozgu. I to je cijela ta tendencija koja je, kako ja tvrdim, pogrešna.

Tri razloga zašto mentalizam ne može biti upravu

1. Značenje je holističko

Doktrina nazvana “holizam značenje” nastala je kao reakcija na logički pozitivizam; ona je ponudila dokaze koji pobijaju pozitivističke pokušaje da pokažu kako se svaki izraz koji možemo razumjeti može definirati u smislu jedne ograničene grupe izraza (“opservacijski izrazi”). Argumenti na koje upućujem bili su djelo W. v. Quinea.¹² Ove argumente je Fodor¹³ uvelike prihvatio, ali čini se da ne cijeni njihov značaj za njegovo poduzeće.

Holizam je, dakle, u prvom redu suprotstavljen pozitivizmu. Pozitivistički pogled na jezik insistira na tome da sve smislene deskriptivne riječi u našem jeziku moraju imati definicije u smislu riječi u “osnovnom” vokabularu, rječniku koji se sastoji od riječi koje stoje za pojmove što su *epistemološki primitivniji* od, recimo, teoretskih pojmova znanosti. Omišljeni kandidat pozitivista bio je vokabular koji se sastoji od izraza za senzacije, ili, u svakom slučaju, izraza za ono što bi trebalo biti “opazivo”. Ako formuliramo pozitivizam kao tezu o istinosnim uvjetima za rečenice, a ne kao tezu o odredivosti izraza, možemo reći da su, kao povijesna činjenica, pozitivisti izvorno insistirali na tome da značenje rečenice bude dano (ili da bude sposobno da bude dano) pravilom koje određuje u kojim se iskustvenim situacijama rečenica može tvrditi.

Sada, glavnina filozofije znanosti dvadesetog stoljeća sastojala se u postepenom odbacivanju tog stajališta. Sami logički pozitivisti su se prebacili od zagovaranja gledišta ka njegovom kritiziranju. U osnovi, ono što se ostvarilo (čak i od samih pozitivista) bilo je da teorije ne mogu biti testirane rečenicu po rečenicu. Ako rečenice od kojih se teorija sastoji imaju vlastita nezavisna iskustvena značenja ili proizvode toliko puno tvrdnji koje se mogu zasebno testirati o tome kakvo će iskustvo biti, onda bi se mogla testirati znanstvena teorija testiranjem rečenice 1 i testiranjem rečenice 2 i testiranjem rečenice 3 i tako dalje. Ali, zapravo, pojedinačni postulati teorije općenito nemaju (ili vrlo malo) iskustvenih posljedica kada ih razmatramo izolirano od drugih tvrdnji teorije. Naprimjer, Newtonova teorija univerzalne gravitacije (bez ikakvih dodatnih navoda o graničnim uvjetima) kompatibilna je s bilo kojom orbitom. Čak bi se mogla pomiriti *kvadratna* orbita s teorijom univerzalne gravitacije govoreći: “Pa, to znači da na sistem djeluju ne-gravitacijske sile.” Jedino u prisutnosti velikog broja tvrdnji proizlaze sve njegove takozvane “poslje-

dice” iz znanstvene teorije. Kao što Quine kaže, rečenice odgovaraju testu iskustva “kao korporativno tijelo”, a ne jedna po jedna. (Otuda i izraz “holizam”.)

Isto vrijedi i za jezik svakodnevnog života. Ako vam neko, primjerice, kaže da je *lopov ušao kroz taj prozor, i na tlu ispod prozora je blatno tlo*, “zaključit ćete” da *u blatu ima otisaka stopala*. Ali to nije *logična posljedica* navedenih činjenica, jer ste očito upotrijebili nepotvrđenu pomoćnu hipotezu u smislu da, *ako je lopov ušao kroz taj prozor, hodao je po tlu kako bi došao do prozora*, i druge stvari opće informacije. Ako vaš doušnik kaže: “Ne, on je nosio čizme”, umjesto da očekujete otiske cipela u blatu, sada ćete očekivati da ćete pronaći otiske drugačijeg oblika. Ono što ima iskustveni unos je korporativno tijelo izjava i taj unos nije jednostavna suma iskustvenog unosa pojedinačnih izjava.

U običnom jeziku, za razliku od formaliziranog jezika, ovaj fenomen postaje još više prožimajući onim što se ponekad naziva “nemonotoničnost” logike svakodnevnog diskursa. U formaliziranom jeziku, ako neko kaže: “Sve ptice lete” i također kaže: “Nojevi su ptice”, može se zaključiti “Nojevi lete”. Ali običan jezik nije takav. Ako kažem “Sokolovi lete” ne namjeravam da moj slušatelj zaključi da će soko s polomljenim krilom letjeti. Ono što očekujemo ovisi o cijeloj mreži uvjerenja. Ako jezik opisuje iskustvo, to čini kao mreža, a ne kao rečenica po rečenica.

Holizam značenja također je u suprotnosti s velikom sklonošću naglašavanja *definicije* kao sredstva kojim se značenje riječi treba objasniti ili fiksirati, tj. suprotstaviti onoj poznatoj zagonetki: “Definirajte svoje pojmove!” Ona ima taj aspekt (koji je Quine naglasio) zbog sugeriranja da iz holizma odmah proizlazi to da se većina pojmova *ne može* definirati – ili, barem, ne može se definirati ako neko pod “definicijom” misli nešto što je fiksirano jednom zauvijek, nešto što apsolutno obuhvaća značenje izraza.

Zašto holizam sugerira ovo? Jer, kad se cijeli korpus vjerovanja suoči s neposlušnim iskustvima, “revizija može udariti bilo gdje”, kao što je to Quine izrazio. Čak i ako je izraz izvorno uveden u znanost putem izričito formulirane definicije, status rezultirajuće istine nije zauvijek privilegiran, kao što bi morao biti ako je izraz jednostavno sinonim za definicije.

Primjer iz povijesti fizike može pomoći razjasniti ovu važnu tačku. U Newtonovoj fizici izraz *momentum* definiran je kao “masa puta ubrzanje”. (Zamislite, ako želite, da je taj izraz izvorno izjednačen s ovim *definiensom* odlukom konvencije Newtonovih fizičara.) Ubrzo je postalo

jasno da je inercija konzervirana količina (kako je Leibniz već mislio). S razvojem vektorske analize stereotip o inerciji kao količini koja je konzervirana i koja ima skalarnu vrijednost i smjer – smjer kretanja čestice – postao je univerzalan među fizičarima. No, prihvaćanjem Einsteinove Posebne teorije relativnosti pojavila se poteškoća.

Einstein nije osporavao ideju da objekti imaju inerciju, ili da je konzervirana, ili da je u smjeru kretanja čestice. Ali on je pokazao da bi načelo posebne relativnosti bilo narušeno ako bi inercija bila tačno jednaka (ostatku) masi/e puta ubrzanje.

Što učiniti? Einstein je proučavao slučaj “bilijarskih kugli” (čestice u elastičnom sudaru). Budući da Newtonova fizika “djeluje”, njihova inercija mora se dati pomoću formule “masa puta ubrzanje” *gotovo* tačno, barem kada su ubrzanja “nerelativistička” (mala u odnosu na c , brzinu svjetlosti). Može li postojati količina s osobinama da je (1) konzervirana u elastičnim sudarima, (2) da je bliže i bliže “masa puta ubrzanje” kako brzina postaje mala i (3) njezin smjer je smjer kretanja čestice? Einstein je pokazao da postoji takva količina te je i on (i svi ostali) zaključio da je ta količina zapravo inercija. Revidirana je tvrdnja da je inercija egzaktno jednaka masi puta ubrzanje. Ali ovo je tvrdnja koja je izvorno bila “definicija”! I bilo je razumno revidirati ovu tvrdnju; jer zašto tvrdnja da je inercija očuvana nije imala barem jednako veliko pravo da se održi kao tvrdnja “inercija je masa puta ubrzanje” kada je otkriven sukob?

Filozof tradicionalnog stila bi mogao odgovoriti na ovo posljednje pitanje govoreći: “Zato što ‘Inercija je masa puta ubrzanje’ daje samo značenje riječi ‘inercija’. Ne možete revidirati analitičku istinu.” Ali takav filozof nameće skup kategorija – ideja fiksnih definicija izraza i analitičkih istina koje nemaju realiteta za aktualnu znanstvenu praksu. Zapravo, on se bavi poviješću – kako je izraz prvi put došao u znanost – kao da određuje buduće izbore koji su znanstvenicima dopušteni. Kao što Quine kaže, istina stipulacijom nije trajna osobina rečenica. Kada se tvrdnje u našoj mreži vjerovanja moraju mijenjati, moramo napraviti “kompromise”; i ono što je najbolji kompromis u danom kontekstu ne može se odrediti konzumiranjem tradicionalnih “definicija” izraza.

Još jedan tradicionalni potez će reći: “Pa znanstvenici su odlučili promijenite značenje ‘inercije’.” Ako ovo objašnjava promjenu u istinosnoj vrijednosti – znanstvenici dodijeliti rečenici “Inercija je masa puta ubrzanje” nakon usvajanja relativiteta, onda mora slijediti da sada govorimo o drugačijoj fizičkoj veličini. Ali ne, još uvijek pričamo o istoj dobroj staroj inerciji – veličinama koje se čuvaju u elastičnim sudarima. To

je fizička veličina “inercija” na koju se uvijek referira ako referira na bilo šta. Pokazalo se da ta veličina, sama inercija, nije potpuno jednaka masi puta ubrzanje. Ako se to čini čudno, to je zato što nismo navikli misliti da su značenja povijesni entiteti u smislu u kojem su osobe ili nacije povijesni entiteti. Ja, Hilary Putnam, imao sam kovrčavu plavu kosu kad sam bio mali. Nisam govorio engleski, nego samo francuski, a svoje ime nisam mislio kao “Hilary Putnam”, već kao “Hilaire Poot-nomm”. Sada imam prorijeđenu sijedu kosu, koja uopće nije kovrčava, govorim engleski više nego francuski i zovem sebe “Hilary Putnam”. Ipak, ja sam ista osoba. Postoje prakse koje nam pomažu odlučiti kada postoji dovoljan kontinuitet kroz promjenu kako bi se opravdala tvrdnja da ista osoba još uvijek postoji. Na isti način tretiramo “inerciju” kao da referira na istu količinu na koju je uvijek referirala, a postoje i prakse koje nam pomažu odlučiti da postoji dovoljno kontinuiteta kroz promjenu kako bi se opravdalo to. Značenje ima neki identitet kroz vrijeme, ali bez suštine.

2. Značenje je dijelom normativni pojam

Tvrdio sam drugdje¹⁴ da se pojmovi o utemeljenom ili opravdanom ili razumnoj vjerovanju ne mogu svoditi na fizikalne pojmove. Neki od argumenata će se pojaviti u kasnijim poglavljima ove knjige. Ali čak i ako bismo mogli dati reduktivnu analizu pojma opravdanog vjerovanja, recimo, identificiranjem “biti opravdano” sa “biti ishod takvih i takvih metoda”, ili “takvog i takvog algoritma”, ili “takvog i takvog kompjuterskog programa”, taj bi algoritam morao biti jednako složen kao i opis “opće inteligencije” idealiziranog induktivnog suđenja. Vidjeli smo, iz naše kratke rasprave o holizmu značenja, da testiranje znanstvene teorije nije nešto što se može učiniti samo traženjem operativnih definicija svih pojmova i testiranjem rečenica koje sadrže teoriju jednu po jednu. Umjesto toga, to uključuje vrlo neopipljive stvari, kao što je procjenjivanje jednostavnosti (koja sama po sebi nije jedan faktor, već različite stvari u različitim situacijama) i vaganje jednostavnosti protiv naše želje za uspješnim predviđanjem i protiv naše želje da sačuvamo određenu količinu prošlosti doktrina. To uključuje imati nos za “ispravan” kompromis između tih vrijednosti. Sposobnost da napravi ove procjene i kompromise je ono što Fodor naziva “općom inteligencijom” i ne očekuje da se opća inteligencija objasni u terminima “modula” u doglednoj budućnosti, ako ikada. Opisivanje prirode opće inteligencije je beznadan problem, prema Fo-

doru, a cijela tačka Fodorove “hipoteze modularnosti” je odvojiti problem razumijevanja “jezičnog organa” od problema razumijevanja opće inteligencije.

Sada želim reći da su pojmovi prikupljeni pod naslovom poglavlja “značenje” (ili “intencionalnost”), naprimjer ključni pojmovi “istog značenja” i “iste referencije”, jednako složeni kao i pojmovi prikupljeni u ovom poglavlju pod naslovom “opća inteligencija”. To ne znači da uvijek treba mnogo inteligencije da bi se moglo reći da dva pojma imaju isto značenje ili istu referenciju. Međutim, postoje mnogi slučajevi u kojima nije potrebno mnogo inteligencije za rješavanje problema u induktivnom ili deduktivnom zaključivanju. Da bi se utvrdila intrinzična složenost podatka, potrebno je pitati: Koliko to može biti teško u najtežem slučaju?

Dakle, teorija sinonimnosti bi bila neka teorija koja odlučuje o pitanjima interpretacije. Razmotrite, međutim, koliko suptilna pitanja mogu biti pitanja o interpretaciji, čak i kada se bavimo tekstovima koji nisu osobito “književni”. Već je spomenuta činjenica da su znanstvenici koji su upotrijebili riječ “inercija” koristili je kao ime za konzerviranu količinu, a ne kao sinonim za “masa puta ubrzanje” (čak i ako su to nazvali “definicija inercije”). Drugi primjer je naše poznavanje činjenice da je Bohr 1934. godine, kada je upotrijebio riječ “elektron” (electron), govorio o istim česticama koje je 1900. godine nazvao “elektronima” (electrons). To ne određujemo uspoređujući teorije i opise elektrona koje je Bohr dao u ova dva različita vremena i vidio da su oni isti, jer nisu. Teorija iz 1900. kaže da elektroni kruže oko jezgre baš kao što planeti idu oko Sunca, tj. elektroni imaju trajektorije, dok teorija iz 1934. (koja je u suštini sadašnja kvantna teorija) kaže da elektron nikad nema putanju – činjenica je da nikada nije imala poziciju i zamah u isto vrijeme. Ipak, fizičar bi mogao opisati razvoj kasnije teorije iz prethodnog na ovaj način: u devetnaestom stoljeću otkrili smo da elektroni imaju određeni omjer masenog naboja odbijajući elektronske zrake u magnetskom polju; kasnije smo drugim eksperimentom otkrili što je naboj elektrona (a time i vrijednost mase elektrona); otkrili smo da je električna struja struja elektrona; otkrili smo da svaki atom vodika sadrži jedan elektron i jedan proton; neko smo vrijeme mislili da elektroni imaju trajektorije, ali onda smo otkrili Princip komplementarnosti i tako dalje.

Ukratko, priču bi ispričao kao priču o uzastopnim promjenama vjerovanja o istim objektima, a ne kao priču o uzastopnim “promjenama značenja”. I ista vrsta “opće inteligencije” uključena je u njegovu odluku da sve ove pojave “elektrona” tretira kao sinonimne koliko je uključena u

njegovu odluku da kasnije istraživačke programe u priči tretira kao produžetak ranijih; vrsta odluke koja ima središnju ulogu u teorijskoj evaluaciji. U stvari, tretiranje “elektrona” kao očuvanje barem njegove reference netaknutom kroz sve te teorijske promjene i tretiranje Bohorove teorije iz 1934. godine kao istinski nasljednik njegove teorije iz 1900. godine gotovo su iste odluke: jednom je odluka opisana kao odluka o značenju ili referenca nekog pojma i jednom kao odluka o obiteljskim odnosima istraživačkih programa.

Ova odluka ilustrira ono što se u interpretaciji naziva “dobronamjernost”/naklonjenost ili “korist od sumnje”.¹⁵ Kada 1900. godine tumačimo Bohra prema onome što nazivamo “elektroni”, naša svjetla stvaraju barem neka njegova uvjerenja iz 1900. godine, a tumačiti ga kao nepostojeće objekte bilo bi odbaciti sve njegova uvjerenja iz 1900. godine kao potpuno pogrešna. I, naravno, Bohr je 1934. godine proširio isti “dobronamjerni” odnos prema svom jastvu iz 1900. godine (zbog čega je i dalje koristio riječ “elektron” u svim tim radovima).

Sva interpretacija ovisi o “dobronamjernosti/naklonjenosti” jer uvijek moramo umanjiti barem neke razlike u vjerovanju kada interpretiramo. Naprimjer: pretpostavimo da čitamo roman napisan prije dvjesto godina na engleskom i nailazimo na imenicu “biljka”. U normalnom kontekstu mi se ne ustručavamo identificirati ovu “biljku” s našim sadašnjim engleskim izrazom “biljka”; ipak, pri tome ignoriramo mnoštvo razlika u vjerovanju. Naprimjer, vjerujemo da biljke sadrže hlorofil, znamo za fotosintezu i ciklus ugljičnog dioksida i kisika i tako dalje. Te su stvari središnje u našem današnjem pojmu o tome što je biljka. Sve su to stvari bile nepoznate prije dvjesto godina. Ipak, (osim ako nismo filozofi ili filozofski gledano povjesničari nauke) ne kažemo da su ljudi prije dvjesto godina “živjeli u drugom svijetu” ili da su njihovi pojmovi “nesamjerljivi” sa pojmovima koje sada imamo,¹⁶ što bi, ako bi se uzelo doslovno (naravno, nikad nije!), podrazumijevalo da ne možemo interpretirati obično pismo koje je neko napisao prije dvjesto godina. Ukratko, pojam biljka tretiramo kao nešto što posjeduje identitet kroz vrijeme, ali bez suštine, i pojam elektron tretiramo kao nešto što posjeduje identitet kroz vrijeme, ali bez suštine.

Pa ipak, ne interpretiramo uvijek riječi na takav način da maksimiziramo broj istinskih vjerovanja koje bi govornik imao (pod našim svjetlima) ako je tumačenje bilo ispravno, suprotno gruboj verziji ideje o “dobronamjernosti u interpretaciji”. Ovdje je primjer ove grube verzije: veliki metalurg Cyril Stanley Smith mi je jednom predložio (kao šalu, ali

jednu s ozbiljnom smislom) da doista postoji takva stvar kao što je flogiston (supstanca koja se pretpostavljala, prije nego što je otkrivena uloga kisika, da uzme u obzir izgaranje ostavljajući gorivo i postupno zasićenje – ili “flogisticiranje” – zraka). Flogiston, predložio je Smith, to su valentni elektroni. Ono što ovo čini šalom je da, kao što Smith dobro zna, mi ne govorimo onako kako je on “predložio”; nismo spremni reći: “Teoretičari flogistona govorili su o valentnim elektronima, ali su imali neka svojstva pogrešna.” To bi bila pretjerana “dobronamjernost/naklonjenost”. Znanje da je jedna stvar razumna “dobronamjernost/naklonjenost”, dok bi drugo bilo pretjerano, pokazuje našu punu moć razumijevanja, bilo da je kontekst interpretacija ili “stvarni život”. Nema nade za teoriju istovjetnosti značenja ili reference koja se primjenjuje na takve teške slučajeve i koja je neovisna o našem računu “opće inteligencije”.

Ono što visi na tim teškim odlukama za nas je iznimno važno. Važno nam je ako čitamo roman, jer bi odluka da se riječi romana tretiraju kao izvanzemaljske (“nesumjerljive”), da to učinimo, potpuno promijenila naš odnos prema književnom djelu. I za nas je važno ako pokušavamo razumjeti povijest znanosti, jer će tumačenje koje damo znanstvenim riječima odigrati veliku ulogu u uspostavljanju znanstvenog postignuća ili nedostatka postignuća, njegove racionalnosti ili nedostatka racionalnosti. Na taj način, odlučivanje da se neko interpretira na jedan način, a ne na drugi, intimno je povezano s normativnim prosudbama.

Ako razmišljamo o ulozi pojma istosti značenja u logici,¹⁷ možda se neće činiti tako iznenađujućim da se ispostavi kako taj pojam ima normativnu dimenziju. U logici, ekvivokacija (homonimija, istoimenovanost), tj. upotreba izraza u jednom smislu na jednom mjestu u argumentu i u različitom smislu na drugom mjestu u istom argumentu, je pogreška bilo da li je argument induktivan ili deduktivan. Ali pojam “smisla” ili “značenja” (Fodorov “sadržaj”) nije mogao odigrati tu ulogu u kritici ako nismo protumačili jedan drugoga na takav način da se “značenja” čuvaju pod uobičajenim postupcima fiksiranja i opravdanja vjerovanja. Ako usvojimo prijedloge za značenje operacionalista ili pozitivista prema kojima modificiranje znanstvene teorije gotovo uvijek proizvodi “promjenu u značenju” teorijskih izraza, onda bismo morali reći da svaki znanstvenik koji modificira postojeću teoriju kako bi riješio problem koji neko posjeduje je kriv za dvosmislenost. Bez sumnje, brzo bismo uveli – ili bolje rečeno, ponovno uveli – tradicionalni pojam “promjene značenja” kako bismo mogli razlikovati slučajeve u kojima je znanstvenik počinio pravu “pogrešku dvosmislenosti” u odgovaranju na pitanje i slučajeve u što je

samo u Pickwickovom smislu riječ o “promjeni značenja” između izvornog postavljanja pitanja i davanja odgovora.

Ukratko, bitno je za ulogu koju igraju pojmovi kao “ekvivokacija/ istoimenovanost” – to se pretpostavlja putem epistemičke uloge koju ti pojmovi imaju – da mi interpretiramo jedni druge na takav način da je “značenje” riječi isto (u smislu da se s njima postupa jednako) pod normalnim procedurama fiksiranja i opravdanja vjerovanja. (Zato što interpretativna praksa duguje odanost tom ograničenju da se istovjetnost i razlika značenja ne mogu podudarati s prisutnošću i odsutnošću bilo kojeg lokalnog kompjuterskog odnosa između naših “mentalnih reprezentacija”. Kao što prethodni primjeri ilustriraju, kompjuterska relacija koja se podudarala sa sinonimijom ne može biti “modularna” u Fodorovom smislu, to jest ne može biti psihološki elementarnija od “opće inteligencije”.)

3. Naši pojmovi ovise o našoj fizičkoj i društvenoj okolini na način koji evolucija (koja je završena za naše mozgove prije oko 30.000 godina) nije mogla predvidjeti

Da bi nam dali urođenu zalihu pojmova koji uključuju riječi karburator, birokrat, kvantni potencijal itd., kako to zahtijeva Fodorova verzija hipoteze o urođenosti, evolucija bi morala biti u stanju predvidjeti sve moguće okolnosti budućeg fizičkog i kulturnog okruženja. Očito nije to učinila i nije to mogla učiniti.

Veze između 1, 2 i 3

Mentalisti koji slijede Fodorovo vodstvo predani su ideji da postoje urođene zalihe semantičkih reprezentacija u kojima se svi naši pojmovi mogu eksplicitno definirati. Tačka 3 postavlja očiglednu poteškoću: Kako bi takvi pojmovi kao “karburator” mogli biti urođeni? Primitivni narodi koji nisu poznavali motore s unutarnjim sagorijevanjem ne pokazuju poteškoće u stjecanju takvih pojmova. Prema Fodorovom objašnjenju, to znači da je njihov “jezik mišljenja” sadržavao pojam “karburator” prije nego što su dobili riječ za taj pojam, iako ništa u njihovoj evolucijskoj povijesti nije moglo objasniti kako je pojam “došao tamo”.

Mentalist (poput mog bivšeg ja) koji nije MIT-ove sorte može, naravno, vjerovati u “kompjutersku psihologiju” bez prihvaćanja hipoteze o urođenosti. Ali on se i dalje suočava s ozbiljnim poteškoćama. To je dio

samog pojma kompjuterske psihologije da sve reprezentacije moraju biti opisane sintaktički ili proceduralno, ili kombinacijom sintaktičkih i proceduralnih karaktera. U isto vrijeme, veliki dio argumenta za holizam značenja je da se promjene u zajedničkim “procedurama” za korištenje leksičkog stavka obično ne računaju kao promjena značenja stavka.

Ako je naše razmišljanje konačno provedeno u internalnom mišljenju (samo ne urođenom), ista stvar vrijedi i za stavke u *lingua mentis* što odgovaraju riječima u javnom jeziku koje smo koristili kao primjere (“inercija”, “elektron”, “biljka” itd.). Štoviše, ako *lingua mentis* nije urođen, tada mentalna reprezentacija koja odgovara određenoj stavci u javnom jeziku može varirati, barem u sintaktičkim aspektima. Nijemci ne samo da mogu koristiti drugačiju riječ za biljke nego i govornici engleskog; oni također mogu koristiti drugačiju “mentalnu riječ” ako mentalni vokabular nije urođen. Riječi u *lingua mentis* različitih govornika koji imaju različite sintaktičke “oblike” (različite “pravopise”, da tako kažemo) i različite “procedure” povezane s njima zapravo mogu imati isto značenje i oznaku. (Ako samo ponovimo bučno “značenje” tako da svaka razlika u procedurama koje jedan govornik povezuje s riječju “mačka” i procedure koje drugi govornik povezuje s “mačkom” računaju kao razliku u “značenju” riječi, onda nećemo imati teoriju značenja, već samo potpunu promjenu teme.) Osim toga, ako su semantičke reprezentacije u mozgu razvijene iz iskustva, kao što su riječi u javnom jeziku, a ne sastavljene od jednog urođenog skupa semantičkih primitiva, nema razloga misliti da određena reprezentacija (opisana sintaktički) neće imati različita značenja različitih skupina ljudi. (To znači, “različita značenja” po kriterijima koje koristi dobar interpretator.) Riječ u *lingua mentis* može, da tako kažem, imati jedno značenje za francuske govornike i drugačije značenje za govornike engleskog jezika, kao što ponekad ima pisana riječ jedno značenje na francuskom i drugačije značenje na engleskom jeziku. “Prevođenje” našeg javnog jezika na *lingua mentis* neće riješiti problem konceptualnog sadržaja, nego ga samo premjestiti s jednog jezika u drugi. Sada se više neću baviti tim problemima, ali oni će se ponovno pojaviti u budućim poglavljima.

Postoje veze između 1, 2 i 3 koje je važno vidjeti. Argument protiv pozitivizma i protiv mogućnosti definiranja svih naših pojmova iz neke osnovne zalihe “epistemološki primitivnih pojmova” sažeto je pod tačkom 1. Srce argumenta bilo je da se usvoji pojam “značenja” prema kojemu – nečemu znanstvena otkrića (otkriće da je voda H₂O, da inercija nije egzaktno proizvod mase i ubrzanja, da elektroni poštuju načelo kom-

plementarnosti, ili da biljke sadrže hlorofil i da prave fotosintezu) *mijenjaju značenje* relevantnih pojmova kršeći načelo spomenuto pod tačkom 2, načelo da se značenja tretiraju kao ista pod uobičajenim procesima fiksiranja i opravdanja vjerovanja. Reći da smo promijenili značenje riječi “voda”, kad smo odlučili da je voda H₂O, ne bi samo išlo protiv naših intuicija sinonimije; to bi prekršilo ovo interpretativno načelo, koje je središnje za epistemičku funkciju pojma “promjena značenja”.

Osim toga, postoji veza između tački 1 i 3: ako su rana shvatanja logičkih pozitivista bili u pravu (tj. ako je tačka 1 bila netačna), onda nam evolucija ne bi morala dati tako nevjerovatne “urođene” pojmove, kao što su karburator ili pozitivni naboj, čak i ako su hipoteze o innatnosti bile istinite; to bi nam samo trebalo dati neke zalihe osnovnih pojmova (izrazi opažanja) iz kojih bi se mogli definirati. Ali (kao što Fodor prepoznaje) naši se pojmovi ne mogu definirati iz skupa pojmova koji su mnogo manji i biološki primitivniji od cijelog leksikona. Ukratko, istina holizma značenja blokira jedini način ispunjavanja prigovora 3 koji ima biološki smisao.¹⁸ (U *The Language of Thought* Fodor ne pokušava odgovoriti na prigovor 3; umjesto toga on se jednostavno divi činjenici da svi ti nevjerovatni pojmovi moraju biti urođeni – budući da to zahtijevaju činjenice, prema njegovoj teoriji). Ukratko, sofisticirani mentalizam MIT raznolikosti nije blokiran od bilo koje od ovih tački odvojeno, već od 1, 2 i 3 koje djeluju zajedno.

Zapravo, moj pravi razlog za početak naše rasprave o kompjuterskim (i fizikalističkim) teorijama značenja uz ispitivanje teorije *The Language of Thought* bio je upravo da ilustrira način na koji će teorije vjerovatno zagovarati načelo holizma značenja i suprotstaviti se raznim “principima dobronamjernosti/ naklonjenosti” (osobito principu da se značenja zadržavaju pod normalnim fiksiranjem vjerovanja). Ne mislim da su Fodorijanci i Chomskyjanci većina među kognitivnim znanstvenicima ili filozofima koji favoriziraju kompjuterske i fizikalističke teorije značenja; ali argumenti koje sam postavio protiv Fodora, posebno onih koji se temelje na holizmu značenju i na interpretativnoj maksimi da značenja nisu izmijenjena uobičajenim procedurama promjene vjerovanja i opravdanja, ponovno će se pojaviti kada uzmemo u obzir teorije koje nisu posvećene Chomskyjeovoj hipotezi o urođenosti.

U sljedećim poglavljljima bit će i drugih pitanja o kojima ćemo također morati razgovarati. Da bismo predstavili neka od ovih pitanja, moramo razmotriti važan aspekt značenja koji sam do sada namjerno zanemarivao u ovoj raspravi. To je način na koji je značenje “interak-

tivno”, to jest način na koji to ovisi ne samo o onome što je u našim glavama nego i o onome što je u našoj okolini i kako smo u interakciji s tim okruženjem. To će biti predmet sljedećeg poglavlja.

BILJEŠKE

1. Lloyd Carr je istakao netačnost u ovom pripisivanju meni. Brentano je smatrao da su mentalni fenomeni okarakterizirani usmjerenjem na “sadržaj”: stoga su teme 1, 3 i 4 na mom popisu tema, uključene u poglavlje “intencionalnost”, središnji dio onoga što je Brentano razmatrao. Smatrao je da je “Brentanova teza” poslužila kao način prikazivanja autonomije mentalističke psihologije (“akt-psihologija”), pokazujući da je mentalno odvojeno od stvarnog (vanjskog) svijeta. Koliko znam, Brentano nikada nije koristio riječ “intencionalnost”, niti se izrazima “intencionalno nepostojanje” i “intencionalno postojanje” koristio da bi referirao na odnos uma i stvarnog svijeta, kao što su filozofi počeli koristiti riječ “intencionalnost” nakon Husserla. Bio je to Husserl, a ne Brentano, koji je u intencionalnosti mentalnog shvatio da pruža način razumijevanja povezanosti uma i svijeta i kako to da u aktima svijesti dolazimo do predmeta.
2. Jedna od nekoliko pouzdanih rasprava o ovom jako pogrešno shvaćenom aspektu Wittgensteinovog mišljenja postoji u Rush Rheesovoj knjizi *Discussions with Wittgenstein* (New York: Schocken, 1970). Vidi posebno 46ff.
3. *The Language of Thought* (New York: Thomas Y. Crowell, 1975).
4. Vidi, naprimjer, Chomskyjevu knjigu *Reflections on Language* (New York: Pantheon, 1975).
5. Chomsky govori o “podsistemu (za jezik) koji ima specifično integrirani karakter i koji je ustvari genetski program za specifični organ” u raspravi sa Piagetom, Pappertom i drugim ponovo odštampanoj u *Language and Learning*, ed. Massimo Piatelli (Cambridge, Mass.: Harvard University Press, 1980). Vidi također *Reflections on Language*.
6. Ili “informacijsko enkapsulirani ulazno-izlazni uređaji”.

7. Vidi poglavlje 5. moje knjige *Mind, Language and Reality*, vol. 2. *Philosophical Papers* (Cambridge: Cambridge University Press, 1983).
8. Recenzija Skinnerovog teksta *Verbal Behavior*, *Language* 35, (1955): 26-58.
9. Vidi njegovu knjigu *Problems of Knowledge and Freedom* (New York: Panthenon, 1971).
10. Ono što me je obuzelo, naravno, bila je ideja spajanja psihologije želje i vjerovanja i kompjuterske psihologije; hipoteza o urođenosti me nikada nije privlačila kao način da se to uradi.
11. Naravno, urođeni “jezik” mozga, ako postoji, nije doslovno napisan. Vidi “What Is Innate and Why”, poglavlje 14. u: *Mind, Language and Reality*.
12. “Two Dogmas of Empiricism”, izvorno objavljen u *Philosophical Review* u januaru 1951; ponovo odštampan u Quineovom *From Logical Point of View* (Cambridge, Massachusetts: Harvard University Press, 1961); “Carnap on Logical Truth”, izvorno objavljen u *The Philosophy of Rudolf Carnap* (La Salle, Ill.: Open Court, 1963); ponovno odštampan u Quineovim *Ways of Paradox* (Cambridge, Mass.: Harvard University Press, 2d izd. 1976); *Word and Object* (Cambridge, Mass.: MIT Press, 1960).
13. Vidjeti, naprimjer, njegovo *Psychological Explanation* (New York: Random House, 1968) i *The Modularity of Mind* (Cambridge, Mass.: MIT Press, 1983), posebno posljednjih 45 stranica.
14. *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981); nekoliko radova u: *Realism and Reason*, sv. 3, moje *Philosophical Papers* (Cambridge: Cambridge University Press, 1983). Vidi također moje *Reflexive Reflections*, *Erkenntnis* 22, (1985).
15. Vidi poglavlje 13. moje knjige *Mind, Language and Reality*.
16. Ovdje, dakako, upućujem na relativističke poglede Thomasa Kuhna, čija je knjiga *The Structure of Scientific Revolutions* (2d ed., enlarged, Chicago: University of Chicago Press, 1983) postala klasika.
17. U metateoriji obične “ekstenzionalne” logike – iskaznog računa, teorije kvantifikacije, teorije skupova – ne treba, zapravo, pojam istovjetnosti *smisla*, kažu većina autora, već pojmovi

ekstenzivne semantike – denotacija ili referencija – da igraju važnu ulogu u svim postupanjima. U sljedećem poglavlju ću raspravljati o tome da svaka teorija istovjetnosti “smisla” mora pretpostaviti teoriju istovjetnosti referencije.

18. Zanimljivu odbranu protiv ovog prigovora iznio je Massimo Piattelli-Palmarini u *The Rise of Selective Theories: A Case Study and Some lessons from Immunology*, poglavlje 6. u: *Language Learning and Concept Acquisition: Foundational Issues*, ed. William Demopoulos and Ausonio Marras (Norwood, N. J.: Ablex, 1986). Piattelli ističe da, kada je Niels Kai Jerne prvi put predložio da tijelo proizvodi protutijela u svim probnim postavkama (1955), postojala je ogromna nevoljkost da se to povjeruje. Ipak je ta hipoteza naknadno potvrđena. “Niels Jerne postavlja stvarnu prisutnost u svakom organizmu ‘unutarnje slike bilo kojeg molekularnog oblika’ (str. 127).” Teško je naći argumente da slučaj mora biti jednostavniji u kognitivnim stvarima (str. 129). Međutim, nisam sklon prihvatiti taj intrigantni prijedlog. Pretpostavimo da neko napiše “kratku priču” (ili bolje rečeno plan za kratku priču) na način da na svakom od 200 mjesta postoji mogućnost izvođenja dvije različite akcije: svaka od priča koje generira nacrt može biti samo dvije ili tri stranice dovoljno kratke da bi ih ljudsko biće moglo zapamtiti. Ipak, ukupan broj kratkih priča koje generira nacrt je prevelik – veći od broja elementarnih čestica u svemiru – da bi se sve kratke priče mogle pojedinačno ispisati bilo gdje. Postoji enormna razlika u broju “sadržaja” koje ljudsko biće može naučiti i broj protutijela u ljudskom krvotoku. (Broj protutijela u ljudskom krvotoku je svakako manji od 1017.) Da budemo sigurni, “koncepti” nisu kratke priče. Ali oni često proizlaze iz teorija, i (sličnim argumentom) broj mogućih teorija (i, sumnjam, tipova teorija) – čak i teorije koje su relativno “kratke” – uključuje eksponencijalnu eksploziju koja čini jako nevjerovatnom ideju da je evolucija unaprijed iscrpila sve mogućnosti.

Prijevod s engleskog: Nijaz Ibrulj