

Biblio theca Philoso phica

4(2019)

Janusz Maciaszek

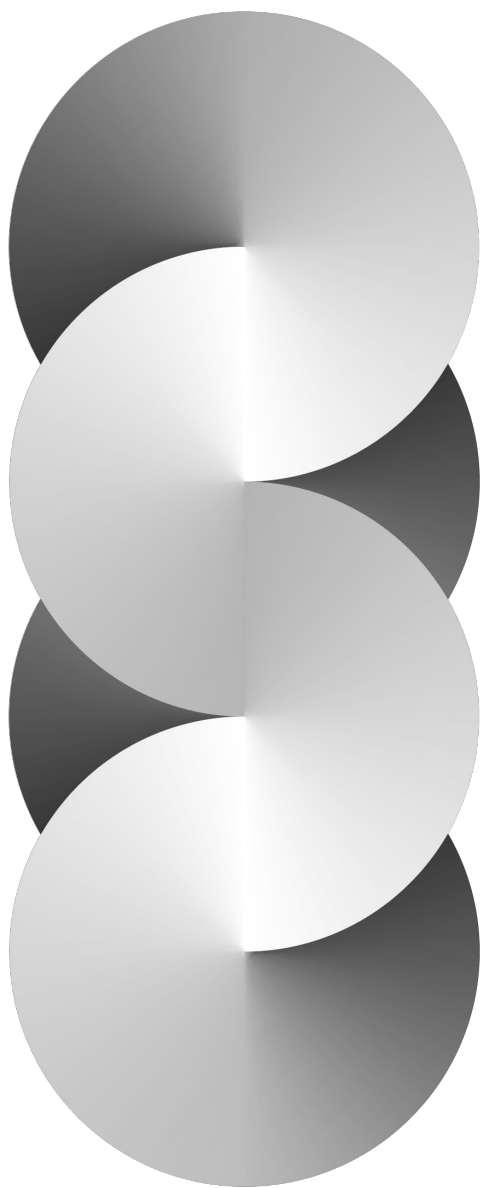
Filozofia
Donalda Davidsona
Podstawy semantyki



WYDAWNICTWO
UNIwersytetu
ŁÓDZKIEGO

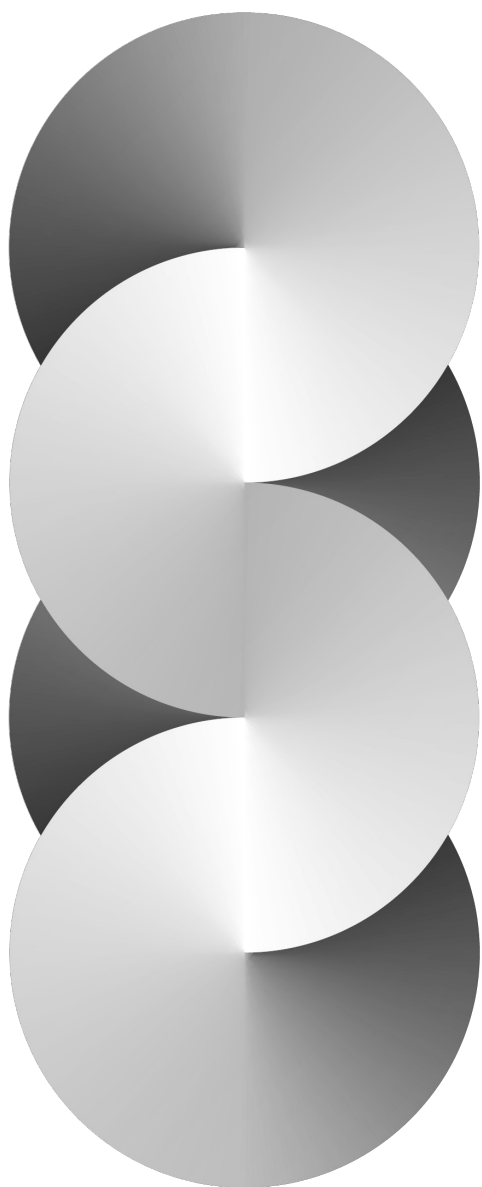
Biblio theca Philoso phica

4(2019)





WYDAWNICTWO
UNIwersytetu
Łódzkiego



Biblio theca Philoso phica

4(2019)

Janusz Maciaszek

Filozofia
Donalda Davidsona
Podstawy semantyki

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Łódź 2019

Janusz Maciaszek – Uniwersytet Łódzki, Wydział Filozoficzno-Historyczny, Instytut Filozofii
Katedra Logiki i Metodologii Nauk, 90-131 Łódź, ul. Lindleya 3/5

RADA NAUKOWA SERII *BIBLIOTHECA PHILOSOPHICA*
Marek Gensler, Andrzej Indrzejczak, Elżbieta Jung, Andrzej Maciej Kaniowski
Krzysztof Kędziora, Ryszard Kleszcz, Janusz Maciaszek, Paweł Pieniążek

RECENZENT

Urszula Żegleń

REDAKTOR INICJUJĄCY

Magdalena Skoneczna

OPRACOWANIE REDAKCYJNE

Monika Grucza

SKŁAD I ŁAMANIE

Janusz Maciaszek

KOREKTA TECHNICZNA

Leonora Gralka

PROJEKT OKŁADKI

Katarzyna Turkowska

Zdjęcie wykorzystane na okładce: © Depositphotos.com/vladystock

Wydrukowano z gotowych materiałów dostarczonych do Wydawnictwa UŁ

© Copyright by Janusz Maciaszek, Łódź 2019

© Copyright for this edition by Uniwersytet Łódzki, Łódź 2019

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego

Wydanie I. W.08870.18.0.M

Ark. druk. 17,125

ISBN 978-83-8142-529-2

e-ISBN 978-83-8142-530-8

<https://doi.org/10.18778/8142-529-2>

Wydawnictwo Uniwersytetu Łódzkiego

90-131 Łódź, ul. Lindleya 8

www.wydawnictwo.uni.lodz.pl

e-mail: ksiegarnia@uni.lodz.pl

tel. (42) 665 58 63

WSTĘP

Donald Davidson (1917–2002) był – w opinii większości badaczy – jednym z najbardziej wpływowych filozofów XX wieku¹. Jego bogata twórczość wzbudza do dzisiaj wiele kontrowersji, których źródłem jest często to, że w czasach postępującej specjalizacji w filozofii Davidson starał się stworzyć spójny system filozoficzny obejmujący metafizykę, teorię poznania, teorię działania, filozofię umysłu, filozofię języka oraz filozofię nauki. Pod względem rozmachu programu filozoficznego Davidson nawiązywał do klasycznych systemów filozoficznych, w szczególności do systemu Arystotelesa. Dla każdej z wymienionych wyżej dyscyplin filozoficznych zaproponował oryginalne i niebanalne rozwiązania. W wielu przypadkach da się oczywiście przedstawić rozstrzygnięcia bardziej atrakcyjne i satysfakcjonujące. Prezentując i oceniając dokonania Donalda Davidsona należy jednak uwzględnić całość zaproponowanego przezeń systemu, który przypomina bardzo skomplikowane, lecz subtelne puzzle, w których każdy element znajduje się dokładnie na swoim miejscu – co stanowi spore wyzwanie dla autora monografii poświęconej jego filozofii. Aby bowiem zrozumieć i docenić całość systemu należy znać jego części. Aby zrozumieć i docenić jego części należy znać całość. Sam Davidson nigdy nie dokonał syntezy własnego systemu filozoficznego, który został przedstawiony w ok. 200 artykułach opublikowanych na przestrzeni niemal 40 lat. Najważniejsze z nich zostały przedrukowane w pięciu zbiorach wydanych przez Oxford University Press. Pierwszy z nich, *Essays on Actions and Events* (2001), wydany po raz pierwszy w roku 1980, obejmuje prace z zakresu filozofii działania oraz filozofii umysłu. W *Inquiries into Truth and Interpretation* (1984) znalazły miejsce prace z zakresu semantyki, zaś w *Subjective, Intersubjective, Objective* (2001a) – głównie z zakresu teorii poznania. Tom *Problems of Rationality* (2004) obejmuje teorię racjonalności, która jednak łączy się ściśle z filozofią umysłu. Ostatni zbiór *Truth,*

¹ Ciekawe informacje na temat życia i twórczości filozoficznej Donalda Davidsona można znaleźć w wywiadzie, którego udzielił Ernestowi Lepore'owi. Wywiad ten opublikowany został w (Davidson 2004).

Language, and History (2005) zawiera, oprócz artykułów poświęconych teorii prawdy i filozofii języka wraz z jej zastosowaniami w teorii literatury, teksty z zakresu historii filozofii, w których Davidson proponuje m.in. oryginalną analizę poglądów Platona, Arystotelesa, Spinozy oraz Hume'a. Na gruncie polskim, poza kilkoma przekładami pojedynczych artykułów, wydany został wybór artykułów Davidsona *Eseje o prawdzie, języku i umyśle* (1992).

Filozofii Donalda Davidsona poświęcono bardzo wiele publikacji, na które składają się zarówno zbiory artykułów, jak i monografie poświęcone różnym fragmentom jego systemu. Z pierwszej grupy należy wymienić przede wszystkim *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson* pod redakcją E. Lepore'a (1986), *Dyskusje z Donaldem Davidsonem o prawdzie, języku i umyśle* pod redakcją U. Żegleń (1996), zmienioną wersję angielską tego zbioru (Żegleń 1999), *Philosophy of Donald Davidson* pod redakcją L. D. Hahna (1999), *Donald Davidson* pod redakcją K. Ludwiga (2003) oraz *A Companion to Donald Davidson* pod redakcją E. Lepore'a oraz K. Ludwiga (2013). Do drugiej grupy należą w pierwszej kolejności dwie monografie również autorstwa E. Lepore'a i K. Ludwiga: *Donald Davidson. Meaning, Truth, Language, and Reality* (2005) oraz *Donald Davidson's Truth-Theoretic Semantics* (2007).

Przedstawiona tu niepełna lista ważniejszych publikacji poświęconych filozofii Davidsona stanowi, obok jego tekstów, bardzo bogaty materiał do studiów dla każdego, kto interesuje się jego twórczością. Prezentowana czytelnikowi monografia nie ma zatem na celu konkutowania z wymienionymi publikacjami. Moim podstawowym celem było zwrócenie uwagi na filozoficzne wątki obecne *explicite* lub jedynie *implicite* w publikacjach Davidsona, który to aspekt jest – moim zdaniem – nie do końca doceniany we wspomnianych opracowaniach, skupiających się przede wszystkim na rekonstrukcji oraz na technicznym aspekcie jego poglądów. Niniejsza praca ma również – co stanowi sprawę niebagatelną – uzupełnić braki w publikacjach na temat Davidsona w języku polskim i jest adresowana do polskiego czytelnika.

Właściwe przedstawienie filozoficznego systemu Donalda Davidsona jest zadaniem stanowiącym spore wyzwanie. Dlatego też prezentowana monografia stanowi jedynie pierwszą część planowanej większej całości. W kolejnych tomach chciałbym poruszyć przede wszystkim zagadnienie zastosowania semantyki Davidsona oraz dokonać analizy jego teorii działania. Nie sposób jednak w tomie poświęconym podstawom semantyki nie wspomnieć choćby krótko o zastosowaniach semantycznej definicji prawdy w realizacji Davidsonowskiego projektu semantycznego (analiza parataktyczna), o przyczynowej teorii metafory, która – wedle Davidsona – wykracza poza zakres semantyki, oraz o Davidsonowskiej filozofii działania – wszak mówienie jest szczegól-

nym rodzajem działania. Nie sposób również nie wspomnieć, choćby w zarysie, o jego poglądach metafizycznych oraz o Davidsonowskiej filozofii umysłu i epistemologii. Wszystkie te tematy wymagają jednak oddzielnego opracowania.

W prezentowanej pracy starałem się nie tylko analizować poglądy Davidsona na tle głównych nurtów filozoficznych, ale wskazywać – tam gdzie jest to możliwe – na jego związki z filozofią polską. Nie chodzi tutaj jedynie o podkreślenie wykorzystania przez Davidsona sformułowanej przez Alfreda Tarskiego semantycznej definicji prawdy, gdyż jest to sprawa oczywista i powszechnie znana. Związki te widoczne są również w przypadku analizy zdań psychologicznych Tadeusza Kotarbińskiego oraz Kazimierza Ajdukiewicza, którego dyrektywalna teoria znaczenia zawiera pewne idee pojawiające się *implicite* w Davidsonowskiej teorii interpretacji.

W tym miejscu nasuwa się zasadnicze pytanie. Skoro Davidson miał ambicje stworzyć system filozoficzny, to dlaczego pierwszy tom projektowanego cyklu poświęcony jest właśnie podstawom jego filozofii języka? Odpowiedź na to pytanie leży w istocie Davidsonowskiego systemu filozoficznego. Wśród problemów poruszanych przez Donalda Davidsona filozofia języka pełni bowiem rolę szczególną, gdyż to ona właśnie łączy różne obszary jego filozofii w jeden spójny system. Powód tego jest oczywisty – człowiek jest istotą obdarzoną mową, która służy do komunikowania się z innymi ludźmi zarówno w sprawach faktów, odczuć, jak i głoszonych teorii – również teorii filozoficznych. Komunikacja językowa nie jest jednak dla Davidsona, jak zazwyczaj zakładano w tradycyjnej filozofii przed tzw. zwrotem językowym, całkowicie neutralna, zaś język nie jest przezroczysty. Język jest środkiem za pomocą którego artykułujemy nasze poznanie, a zarazem wyznacza granice wszelkiego poznania. Widać to wyraźnie z perspektywy ontogenezy języka. W pierwszych miesiącach życia nauczyliśmy się reagować językowo na pewne powtarzające się bodźce. W nauce tej niezbędni byli opiekunowie, którzy już znali język i nauczyli nas konwencjonalnych reakcji na bodźce. Po tym wstępnym okresie nastąpił etap najważniejszy – etap konceptualizacji rzeczywistości, etap kategoryzacji rzeczy i zdarzeń. Na tym etapie również niezbędni są inni ludzie, z którymi nieustannie „uzgadniamy” nasze przekonania na temat rzeczywistości, zgadzając się co do faktów lub spierając się o nie. Czynimy to jednak zawsze przy użyciu języka. Język pełni nie tylko niezwykle istotną rolę w naszych przyczynowych interakcjach z realnie istniejącym światem oraz innymi ludźmi, ale – co równie ważne – owe interakcje wpływają na język. Język jest zatem, jak twierdził Donald Davidson, wystarczająco obiektywny, aby można było twierdzić, że ma się rację – nawet w filozofii.

Davidson uważany jest za autora bardzo wymagającego wobec swoich czytelników. Stosowane przez niego argumenty mają zazwyczaj charakter niezwykle eliptyczny, niekiedy tajemniczy (*criptic*). W jego publikacjach spotykamy wiele powtórzeń oraz sformułowań budzących liczne wątpliwości, gdyż jego poglądy semantyczne stopniowo ewoluowały. Nie były to zmiany radykalne, lecz raczej przesunięcia akcentów będące rezultatem coraz szerszej perspektywy wyznaczanej przez powoli uzupełniany system. Opis ewolucji poglądów Davidsona na semantykę stanowi kolejny cel, jaki postawiłem przed sobą pisząc tę monografię. Oczywiście żadna monografia nie zastąpi lektury tekstów Davidsona – moim celem była jedynie pomoc w ich zrozumieniu oraz zaproponowanie własnej interpretacji tych tekstów. Dlatego też kilka wybranych artykułów, które są – moim zdaniem – kluczowe dla zrozumienia podstaw jego semantyki, omawiam bardzo dokładnie, rozwijając i egzemplifikując myśli Davidsona, zestawiając je często z poglądami innych filozofów języka.

Pierwszy rozdział poświęcony jest semantycznej definicji prawdy, stanowiącej podstawę, na której Davidson oparł swą teorię znaczenia, a następnie teorię interpretacji. Przeprowadzona w tym rozdziale analiza jest bardzo drobiazgowa, gdyż w większości publikacji na temat filozofii Davidsona prezentację teorii Tarskiego traktuje się w sposób bardzo skrótowy, zakładając, że jest powszechnie znana. Moje doświadczenie wskazuje jednak, że nie do końca odpowiada to stanowi faktycznemu. Dotyczy to zarówno filozofów, jak i matematyków.

Powstanie semantycznej definicji prawdy było motywowane zastosowaniami w metamatematyce. Z czasem poglądy Tarskiego ewoluowały w stronę teorii modeli, pozostawiając tym samym semantyczną definicję prawdy filozofom. Nie ulega wątpliwości, że na jej temat powstało wiele wartościowych opracowań – na kilka z nich powołuję się w tej pracy, ale dość często zdarzają się raczej powierzchowne odwołania do tej teorii. Sprzyja temu fakt, że wynikające z niej T-zdania przybierają niekiedy postać trywialną, o czym świadczy pochodzący od Tarskiego przykład „Zdanie *Śnieg jest biały* jest prawdziwe wtedy i tylko wtedy, gdy śnieg jest biały”. Dlatego też jestem przekonany, że nie da się w pełni zrozumieć poglądów semantycznych Donalda Davidsona, nie dostrzegając problemów, które towarzyszyły Tarskiemu podczas formułowania swej definicji. W związku z tym odtwarzam ewolucję poglądów Alfreda Tarskiego oraz kontekst filozoficzny, w którym się owe poglądy kształtowały. Chodzi tu głównie o dominujący w pierwszej połowie XX w. fizykalizm, który wywarł znaczny wpływ na postawę Tarskiego. Wskazuję również powody, dla których wcześniejsza definicja prawdy w języku, na której oparł swą semantykę Davidson, przekształciła się z czasem

w definicję prawdy w modelu, które to podejście jest Davidsonowi nie tylko obce, ale stanowi przedmiot jego zdecydowanej krytyki².

Drugi rozdział poświęcony jest przede wszystkim rozwinięciu i wyjaśnieniu idei, które Davidson zawarł, w sposób niezwykle eliptyczny, w *Theories of Meaning and Learnable Languages* (1966) – swym pierwszym artykule poświęconym filozofii języka. Przede wszystkim odtwarzam i rozwijam w nim Davidsonowską krytykę empirystycznych teorii znaczenia, które można wywieść z Arystotelesa teorii znaku językowego. Donald Davidson poświęcił temu zagadnieniu jedynie kilka słów, chociaż problem ten jest niezwykle ważny nie tylko dla uzasadnienia słuszności podstaw semantyki Davidsona, ale również z powodu nawiązania do konstruktywnej krytyki empiryzmu podjętej przez Quine'a w *Dwóch dogmatach empiryzmu* (1951), a następnie kontynuowanej przez samego Davidsona w *O pojęciu schematu pojęciowego* (1974). Z tego punktu widzenia poglądy filozoficzne Donalda Davidsona wpisują się doskonale w przemiany zachodzące na gruncie filozofii analitycznej w drugiej połowie wieku XX.

Rozdział trzeci poświęcony jest Davidsonowskiej teorii znaczenia opartej na semantycznej definicji prawdy, która została określona przez jej autora nieco oksymoronicznym mianem „teorii znaczenia bez znaczeń”. Wybór odpowiednio zmodyfikowanej semantycznej definicji prawdy – zwanej od tego momentu teorią prawdy – na podstawę teorii znaczenia znajduje swe uzasadnienie w rozdziale drugim, zaś jej nieodniesieniowa interpretacja nawiązuje do rozważań rozdziału pierwszego. W rozdziale tym dokładnie omawiam i rozwijam argumenty, które Davidson wysuwa przeciw semantynom odniesieniowym w swym kanonicznym tekście *Prawda i znaczenie* (1967). Rozważania trzeciego rozdziału kończy dokładne omówienie krytycznego artykułu autorstwa J. A. Fostera *Meaning and Truth Theory* (1976) oraz odpowiedzi, których Davidson udzielił Fosterowi. Tekst Fostera jest niewątpliwie trudny, jednak zdecydowałem się na jego dokładne skomentowanie, gdyż odpowiedzi, których udzielił Davidson, kończą pierwszy okres jego poglądów semantycznych, które można opatrzyć etykietą „teoria znaczenia”, zapoczątkowując jednocześnie okres teorii interpretacji.

Teorii interpretacji poświęcony jest ostatni, a zarazem najdłuższy rozdział monografii. W rozdziale tym nie ograniczam się do teorii interpretacji *sensu stricto*, lecz nawiązuję do poglądów epistemologicznych Davidsona. Uznałem bowiem, że nie da się w pełni zrozumieć czym jest interpretacja, nie znając poglądów Davidsona na temat poznania. W rozdziale tym zawarłem również analizę Davidsonowskiego antykonwencjonalizmu, który znajduje swój wyraz

² Chodzi tu o krytykę tzw. teoriomodelowych semantyk języka naturalnego.

w paradoksalnie brzmiącym stwierdzeniu, że nie istnieje nic takiego jak język, oraz omówiłem warunki możliwości interpretacji, czyli Davidsonowską krytykę schematu pojęciowego oraz tzw. Zasadę życzliwości. Sporo miejsca poświęciłem również Quine'a przekładowi radykalnemu oraz Davidsonowskiej interpretacji radykalnej. Za szczególnie ważny i reprezentatywny dla późniejszych poglądów Donalda Davidsona uznałem artykuł *A Nice Derangements of Epitaphs* (1986), którego dokładnemu omówieniu również poświęciłem sporo miejsca. Monografię kończy paragraf poświęcony problemowi niezdeterminowania interpretacji. Pewne tematy związane z interpretacją, np. teorię decyzji, potraktowałem w sposób bardziej skrótowy, gdyż uznałem, że zagadnienia te zostały wyczerpująco omówione w dostępnej literaturze oraz – moim zdaniem – nie pełnią roli kluczowej w zrozumieniu głównych idei Davidsona.

W niniejszej publikacji nawiązuję w pewnym stopniu do napisanej przeze mnie 10 lat temu monografii *Znaczenie, prawda, przekonania. Problematyka znaczenia w filozofii języka* (2008), ale jest to tekst zasadniczo nowy. W przeciągu owych lat zmieniłem bowiem częściowo rozumienie Davidsonowskiej filozofii języka. W szczególności w większym stopniu niż kiedyś dostrzegam filozoficzną doniosłość Tarskiego semantycznej definicji prawdy oraz różnicę między wczesnymi a późniejszymi poglądami Davidsona. Inaczej również oceniam rolę przekładu w teorii prawdy.

Na treść prezentowanej monografii wpłynęły niewątpliwie moje wystąpienia seminaryjne i konferencyjne oraz prowadzone przeze mnie zajęcia dydaktyczne, podczas których próbowałem wielokrotnie wyjaśnić słuchaczom filozofii niebanalność i filozoficzną wagę poglądów Donalda Davidsona. Uznałem wówczas, że podnoszone przez niego zagadnienia semantyczne należy mocno osadzić w kontekście jego filozofii oraz opatrzyć – tam gdzie jest to tylko możliwe – licznymi i odpowiednio dobranymi przykładami. Książka przeznaczona jest przede wszystkim dla filozofów interesujących się językiem oraz filologów, którzy pragną zapoznać się z filozoficzną refleksją nad językiem. Davidsona teoria interpretacji znalazła bowiem poczesne miejsce również w teorii literatury.

ROZDZIAŁ 1

ALFREDA TARSKIEGO SEMANTYCZNA DEFINICJA PRAWDY

Sądzę, że zarówno językoznawcy jak i filozofowie zainteresowani językami naturalnymi przeoczyli kluczowe znaczenie teorii prawdy (...) dlatego, iż przecenili oni trudności, na jakie napotyka próba podania formalnej teorii prawdy danego języka naturalnego
(Davidson 1992/1970, s. 60–61)¹.

Wokół Tarskiego semantycznej definicji prawdy narosło wiele nieporozumień. Dotyczą one głównie filozoficznych aspektów definicji, gdyż jej waga dla metamatematyki nie jest kwestionowana. Ze względu na cele tej monografii, wspomniany aspekt semantycznej definicji prawdy nie będzie nas jednak interesował. Trudności interpretacyjne sprowokowane zostały przez samego Tarskiego, który w publikacjach poświęconych semantycznej definicji² prawdy pisze o niej w sposób nie do końca przejrzysty, a czasami wręcz nie do końca konsekwentny. Pośród filozofów posiada ona wielu zagorzałych zwolenników, np. Donalda Davidsona, którzy twierdzą, że stanowi ona źródło wielu inspiracji filozoficznych i pomimo swej formalnej prostoty charakteryzuje się filozoficzną głębią. Dla przeciwników definicja prawdy zbudowana przez Tarskiego jest z filozoficznego punktu widzenia absolutnie nieciekawa i stanowi jedynie kolejną nieudaną próbę zdefiniowania tego pojęcia. Są również tacy, którzy uważają, że wartość semantycznej definicji prawdy

¹ Tłumaczenia cytatów pochodzących z pozycji, dla których nie podaję w bibliografii polskich przekładów pochodzą ode mnie. Tam, gdzie jest to niezbędne, w odniesieniu bibliograficznym podane są dwa lata oddzielone ukośnikiem. Pierwszy to rok wydania, z którego pochodzi cytat, zaś drugi to rok pierwszego wydania, który pozwala na znalezienie danej pozycji w bibliografii.

² Teksty poświęcone semantycznej definicji prawdy zostały wydane w latach 1930–1944.

polega przede wszystkim na zdemaskowaniu prawdy jako pojęcia nieciekawego i trywialnego. Prowadzi to wprost do minimalistycznych, np. deflacyjnych, koncepcji prawdy.

Celem tego rozdziału jest przybliżenie czytelnikowi filozoficznych problemów związanych z semantyczną definicją prawdy i przedstawienie tych jej zastosowań na polu filozofii języka, które są istotne z punktu widzenia celów monografii poświęconej podstawom stworzonej przez Donalda Davidsona semantyki języka naturalnego. Ponieważ Davidson wykorzystał semantyczną definicję prawdy jako teorię znaczenia, nie jest możliwe właściwe zrozumienie podstaw jego semantyki bez znajomości tych aspektów semantycznej definicji prawdy, które pozwoliły Davidsonowi na zaadaptowanie jej do potrzeb filozofii języka. Co więcej, wiele wątpliwości dotyczących Davidsonowskiej teorii znaczenia ma swoje źródło w analogicznych wątpliwościach związanych z semantyczną definicją prawdy. Davidson nie tylko bowiem adaptował teorię prawdy do potrzeb teorii znaczenia, ale był jednym z najważniejszych komentatorów myśli filozoficznej Alfreda Tarskiego. Prezentując semantyczną definicję prawdy starałem się unikać zbędnych formalizmów oraz odwołań do logiki formalnej. Jednak czytelnik, który uzna, że rozważania przedstawione w tym rozdziale są zbyt drobiazgowe, może poprzestać na przeczytaniu jedynie pierwszych paragrafów rozdziału.

1. PROSTY PRZYKŁAD SEMANTYCZNEJ DEFINICJI PRAWDY

W jednej z pierwszych prac poświęconych semantycznej definicji prawdy Tarski deklaruwał, że jego celem – a zarazem problemem, który zamierzał rozwiązać – jest:

(...) *konstrukcja metodologicznie poprawnej i merytorycznie trafnej definicji pojęcia zdania prawdziwego*. Definicja ta ma uwzględniać intuicje, które są zawarte w tak zwanym klasycznym ujęciu pojęcia prawdy, według którego *prawdziwy* [o zdaniu] znaczy tyle, co „zgodny z rzeczywistością” (1932, s. 9)³.

³ W rozdziale tym konsekwentnie pomijam nazwisko Tarskiego w odniesieniach bibliograficznych, podając – tam gdzie jest to konieczne – jedynie rok pierwszego wydania oraz stronę ze zbiorowego wydania dzieł Tarskiego (1995). W całej monografii – w miejscach, w których nie będzie prowadziło to do nieporozumień – będę pomijał w odniesieniach bibliograficznych nazwisko autora oraz rok wydania.

Zanim wyjaśnimy na czym polega metodologiczna poprawność oraz merytoryczna trafność warto rozpatrzyć prosty przykład semantycznej definicji prawdy.

Semantyczną definicję prawdy podaje się zawsze dla konkretnego języka⁴, który będziemy nazywali dalej językiem przedmiotowym, sama zaś definicja sformułowana jest w metajęzyku. Język przedmiotowy, nazwijmy go L , zawiera nieskończenie wiele indeksowanych⁵ zmiennych x_1, x_2, \dots oraz skończenie wiele wyrażeń stałych: symbol \subset (interpretowany w metajęzyku jako bycie częścią właściwą), symbol \sim (interpretowany w metajęzyku jako negacja zdaniowa, czyli zaprzeczenie zdania) oraz symbol \forall (interpretowany w metajęzyku jako „każdy”).

Metajęzyk, oznaczmy go M , jest językiem polskim wzbogaconym o dodatkowe pojęcia semantyczne: „nieskończony ciąg obiektów”, „przyporządkowanie terminowi jednostkowemu (np. zmiennej) obiektu przez ciąg obiektów”, „spełnianie formuły przez ciąg obiektów”. W języku M występują również meta-zmienne a, b, \dots oznaczające obiekty oraz niewystępujący w języku L jednoargumentowy predykat „prawdziwy”, którego argumentami są zdania języka przedmiotowego. Integralnymi składnikami M są środki służące do tworzenia z wyrażeń języka L ich metajęzykowych nazw, np. cudzysłowy, które będą konsekwentnie stosował w poniższym przykładzie, będącym uproszczeniem przykładu podanego przez Tarskiego w krótkim komunikacie *Pojęcie prawdy a sformalizowane nauki dedukcyjne* (1930/31).

Semantyczna definicja prawdy ma charakter rekurencyjny. W jej skład wchodzi zapewniająca ową rekurencyjność aksjomaty. W naszym przykładzie tzw. aksjomatem odniesienia jest⁶:

(A) Nieskończony ciąg obiektów C przyporządkowuje zmiennej „ x_i ” i -ty wyraz tego ciągu (oznaczymy ten wyraz jako i_C)⁷.

⁴ Mówiąc ogólnie o semantycznej definicji prawdy, bez odwołania do konkretnego języka, mamy na myśli – podobnie jak w przypadku np. definicji klasycznej – jej ogólny schemat lub sposób budowy. Należy pamiętać, że dla każdego języka budujemy zupełnie inną definicję prawdy. Dlatego też semantyczna definicja prawdy jest zrelatywizowana do języka.

⁵ Indeksowanie zmiennych nie jest kluczowe w języku przedmiotowym, lecz stanowi niewątpliwie ułatwienie w sformułowaniu aksjomatu odniesienia dla zmiennych.

⁶ Przytoczone tu nazwy typów aksjomatów nie pochodzą bezpośrednio od Tarskiego. Aksjomat odniesienia stanowi odpowiednik niejasnego – zdaniem Tarskiego – pojęcia odnoszenia się lub oznaczania. W przypadku stałych indywidualnych, które nie występują w naszym przykładzie, aksjomat odniesienia ma postać następującą:

Dla każdego ciągu C , C przyporządkowuje stałej „ c_i ” ten sam obiekt a_i .

⁷ Ponieważ rozróżnienie między językiem przedmiotowym a metajęzykiem jest bardzo istotne w przypadku aksjomatów semantycznej definicji prawdy oraz wynikających z niej T-zdań,

zaś aksjomatami spełniania:

(B) Ciąg C spełnia formułę „ $x_i \subset x_j$ ” wtw, gdy obiekt i_C jest częścią właściwą obiektu j_C .

(C) Ciąg C spełnia formułę „ $\forall x_i \varphi$ ” wtw, gdy każdy ciąg powstały z ciągu C przez dowolne zastąpienie jego i -tego wyrazu spełnia formułę „ φ ” (innymi słowy: dla każdego obiektu a , $C^{i=a}$ spełnia „ φ ”).

(D) Ciąg C spełnia formułę „ $\sim \varphi$ ” wtw, gdy ciąg C nie spełnia „ φ ”.

Stałym elementem każdej definicji prawdy jest aksjomat prawdziwości, który łączy pojęcie spełniania z prawdziwością:

(E) Formuła jest prawdziwa wtw, gdy jest spełniona przez każdy ciąg⁸.

W naszym przykładzie semantyczna definicja prawdy składa się z aksjomatów (A) – (E).

Z semantycznej definicji prawdy wynikają tzw. T-zdania, zwane w literaturze polskiej również P-równoważnościami, odgrywające kluczową rolę w spełnianiu przez definicję wymienionego na początku paragrafu wymogu merytorycznej trafności. Ponadto jeśli semantyczną definicję prawdy potraktujemy jako teorię, to T-zdania pełnią funkcję wynikających z aksjomatów twierdzeń tej teorii⁹. W jednym z kilku podanych przez Tarskiego sformułowań T-zdania mają postać:

(1.1) Zdanie „ p ” jest prawdziwe wtw, gdy Z .

W powyższym schemacie p jest zdaniem języka przedmiotowego, zaś zdanie Z jest jego przekładem na metajęzyk. Poniżej znajduje się wyprowadzenie przykładowego T-zdania dla podanej wyżej teorii prawdy:

(1.2) Zdanie „ $\forall x_i \forall x_j (x_i \subset x_j)$ ” jest prawdziwe wtw, gdy dla każdego a i b , a jest częścią właściwą b .

(Zauważmy, że formuła $\forall x_i \forall x_j (x_i \subset x_j)$ jest – przy standardowych znaczeniach odpowiednich wyrażen M – intuicyjnie fałszywa, gdyż nie dla każdej pary obiektów pierwszy z nich jest częścią właściwą drugiego).

1. Zdanie „ $\forall x_i \forall x_j (x_i \subset x_j)$ ” jest prawdziwe.

2. Dla każdego ciągu C , C spełnia „ $\forall x_i \forall x_j (x_i \subset x_j)$ ” (z aksjomatu E).

przyczaczone wyrażenia będą konsekwentnie brał w cudzysłów nawet wówczas, gdy będą pisane kursywą.

⁸ Formuła, którą spełnia każdy ciąg, jest po prostu zdaniem.

⁹ Zakłada się tutaj *implicite* klasyczne reguły wnioskowania.

3. Dla każdego C i dla każdego obiektu a , $C^{i=a}$ spełnia „ $\forall x_i(x_i \subset x_j)$ ” (z aksjomatów A oraz C).
4. Dla każdego C i dla każdych obiektów a i b , $C^{i=a, j=b}$ spełnia „ $x_i \subset x_j$ ” (z aksjomatów A oraz C).
5. Dla każdego a i b , a jest częścią właściwą b (z aksjomatu B).

Jak łatwo zauważyć, w każdym kroku dowodu wynikanie zachodzi również w drugą stronę:

1. Dla każdego a i b , a jest częścią właściwą b .
2. Dla każdego C i dla każdego a i b , $C^{i=a, j=b}$ spełnia „ $x_i \subset x_j$ ” (z aksjomatu B).
3. Dla każdego C oraz dla każdego a , $C^{i=a}$ spełnia „ $\forall x_j(x_i \subset x_j)$ ” (z aksjomatów A oraz C).
4. Dla każdego ciągu C , C spełnia „ $\forall x_i \forall x_j(x_i \subset x_j)$ ” (z aksjomatów A oraz C).
5. Zdanie „ $\forall x_i \forall x_j(x_i \subset x_j)$ ” jest prawdziwe (z aksjomatu E).

Z podanej teorii prawdy wynika również następujące T-zdanie (dowód pozostawiam czytelnikowi):

(1.3) Zdanie „ $\forall x_i \sim(x_i \subset x_i)$ ” jest prawdziwe wtw, gdy dla każdego obiektu a , a nie jest częścią właściwą a (żaden obiekt nie jest swoją własną częścią właściwą).

Jeżeli uznamy, że jest tak, jak głosi prawa strona równoważności (1.3), to zdanie języka przedmiotowego $\forall x_i \sim(x_i \subset x_i)$ uznajemy za zdanie prawdziwe. A zatem prawa strona T-zdania stanowi warunek prawdziwości zdania po jej lewej stronie, chociaż – jak łatwo zauważyć – nie podaje dającego się zastosować kryterium prawdziwości.

2. WYMÓG MERYTORYCZNEJ TRAFNOŚCI

Przez merytoryczną trafność definicji rozumie się jej adekwatność, co oznacza, że definicja poprawnie wyznacza zakres definiowanego terminu. W przypadku semantycznej definicji prawdy oznacza to, że powinna ona utożsamić zakres predykatu „prawdziwy” ze zbiorem zdań prawdziwych w języku L . Wymóg merytorycznej trafności wyraża konwencja T głosząca, że z merytorycznie trafnej semantycznej definicji prawdy mają wynikać wszystkie T-zdania.

Przykładem schematu T-zdania jest podana w poprzednim paragrafie formuła (1.1). Jeśli jednak metajęzyk obejmuje język przedmiotowy, wówczas T-zdanie przybiera postać następującą:

(2.1) X jest prawdziwe wtw, gdy p ,

gdzie pod p można podstawić dowolne zdanie rozważanego języka (języka przedmiotowego), zaś pod X – dowolną nazwę indywidualną tego zdania (por. 1932, s. 10). W szczególności (2.1) może przybrać formę, która sprawia wrażenie trywialnej¹⁰:

(2.2) Zdanie „ p ” jest prawdziwe wtw, gdy p .

W podanym wyżej przykładzie skorzystaliśmy z najprostszego sposobu utworzenia nazwy jednostkowej zdania p , czyli wzięcie tego zdania w cudzośłów. Można jednak utworzyć metajęzykową nazwę zdania również w inny sposób. Jeżeli w książce X trzecie zdanie na osiemnastej stronie brzmi *Paris is beautiful*, to można wyrazić warunek prawdziwości tego zdania w postaci następującej:

(2.3) Trzecie zdanie na osiemnastej stronie książki K jest prawdziwe wtw, gdy Paryż jest piękny.

W powyższym sformułowaniu pojawia się jednak poważny problem. Wprawdzie (2.3) wyraża warunek prawdziwości zapisanego w książce K zdania, nie jest to jednak T-zdanie, gdyż nie wynika z żadnej teorii prawdy. Teoria taka – a będzie o tym jeszcze mowa głównie w rozdziale drugim – musi bowiem liczyć skończoną ilość aksjomatów, które odpowiadają skończonej liczbie słów (lub stałych) w języku L . Dzięki ich rekurencyjnemu charakterowi z aksjomatów tych powinno wynikać potencjalnie nieskończenie wiele T-zdań, które odpowiadają potencjalnie nieskończeniu wielu zdaniom języka L . Próba zbudowania teorii prawdy w postaci aksjomatów odwołujących się do pozycji zdań w książkach musi zakończyć się niepowodzeniem, gdyż jej aksjomaty nie odpowiadałyby słowom, lecz zdaniom – tych zaś, nawet zdań prostych gramatycznie, jest nieskończenie wiele, co prowadziłoby do konieczności sformułowania nieskończenia wielu aksjomatów¹¹. Ponieważ

¹⁰ W obejmowaniu języka przedmiotowego przez metajęzyk nie chodzi jedynie o kształty fizyczne zdań. Oznacza to również, że zdanie p wyraża w języku przedmiotowym oraz w metajęzyku te same sądy (tj. znaczy to samo w obu językach). Dlatego też wynikające z teorii prawdy T-zdania nie są faktycznie trywialne, chociaż sprawiają takie wrażenie.

¹¹ Aksjomaty dla zdań prostych musiałyby mieć postać:

Dla każdego ciągu C , C spełnia formułę „*Paris is beautiful*” wtw, gdy Paryż jest piękny.

Stosując aksjomat prawdziwości z podanego aksjomatu uzyskujemy natychmiast T-zdanie (2.3).

nieskończenie wielu aksjomatów podać się nie da, lecz można podać jedynie skończoną ich liczbę, z „teorii prawdy” odwołującej się do położenia zdań w tekście książki, dałoby się wyprowadzić jedynie skończoną ilość T-zdań, co stoi w sprzeczności z konwencją T, która wymaga, aby z teorii prawdy wynikały wszystkie spośród nieskończonej ilości T-zdań¹². A zatem zdanie po lewej stronie równoważności nie może być całkowicie dowolną nazwą języka przedmiotowego. Na problem ten zwrócił uwagę Tarski w swym klasycznym tekście *Pojęcie prawdy w językach nauk dedukcyjnych* z roku 1933, gdzie w sformułowaniu konwencji T pojawia się jednocześnie warunek, aby metajęzykowa nazwa zdania była nazwą strukturalnoopisową:

Poprawną formalnie definicję symbolu „Vr” [oznaczającego klasę wszystkich zdań prawdziwych] sformułowaną w terminach metajęzyka, będziemy nazywać trafną definicją prawdy, o ile pociąga za sobą następujące konsekwencje:

(a) wszystkie zdania dające się uzyskać z wyrażenia „ $x \in Vr$ wtedy i tylko wtedy, gdy p” przez zastąpienie symbolu „x” nazwą strukturalnoopisową dowolnego zdania rozważanego języka, zaś symbolu „p” – wyrażeniem, stanowiącym przekład tego zdania na metajęzyk.

(b) $Vr \subseteq S$ [gdzie S jest językiem przedmiotowym] (s. 60–61).

Powyższe sformułowanie konwencji T Tarski poprzedził dyskusją na temat możliwości zbudowania definicji prawdy w języku naturalnym (potocznym) (s. 17–31), pokazując, że stosowanie takich nazw zdań jak w (2.3) prowadzi do poważnych trudności. Część owych trudności można uniknąć budując definicję strukturalną, czyli po prostu rekurencyjną:

Wobec takiego obrotu rzeczy nasuwa się myśl sięgnięcia przy rozwiązywaniu postawionego zagadnienia do innych metod postępowania. Zwróć tu uwagę na jedną tylko próbę w tym kierunku, mianowicie na próbę zbudowania *definicji strukturalnej*. Ogólny schemat tego rodzaju definicji byłby w przybliżeniu następujący: *zdanie prawdziwe to jest wyrażenie posiadające takie a takie strukturalne własności* (tj. własności, dotyczące kształtu i kolejnego następstwa poszczególnych części składowych wyrażenia) *lub też dające się uzyskać z takich*

¹² W takiej „definicji prawdy” nazwami zdań języka *L* byłyby pary złożone ze „współrzędnej” zdania w książce oraz tytułu książki. Można jednak stworzyć teorię prawdy dla języka, który „szyfruje” w książkach nie zdania, lecz wyrazy. Załóżmy dla uproszczenia, że książka *K* zawiera wszystkie wyrazy języka angielskiego (lecz z oczywistych powodów nie zawiera wszystkich zdań tego języka). Niech *Paris* będzie 753-tym słowem pojawiającym się w *K*, *is* 24-tym słowem, zaś *beautiful* słowem 1856-tym. Nie jest trudno sformułować aksjomaty, z których wynika:

Zdanie będące ciągiem 753-go, 24-tego i 1856-tego wyrazu książki *X* jest prawdziwe wtw, gdy Paryż jest piękny.

a takich strukturalnie opisanych wyrażen przy pomocy takich, a takich strukturalnych przekształceń (s. 28).

Każda semantyczna definicja prawdy jest definicją strukturalną (rekurencyjną) w podanym wyżej sensie, zaś w przykładzie z poprzedniego paragrafu metajęzykowe nazwy zdań po lewej stronie (1.2) i (1.3) są nazwami strukturalnoopisowymi. Wymóg ten jest spełniony, gdyż równoważności te zostały wyprowadzone z odpowiednich aksjomatów semantycznej definicji prawdy. Przyniesiony w poprzednim paragrafie dowód wynikania (1.2) z aksjomatów definicji prawdy pokazuje, że zdanie posiadające „takie a takie strukturalne własności”, czyli w naszym przypadku zdanie $\forall x_i \forall x_j (x_i \subset x_j)$, rozumiane jako ciąg składających się nań symboli, jest prawdziwe wtedy i tylko wtedy, gdy dla każdego a oraz b , a jest częścią właściwą b . Aby podać warunek prawdziwości zdania nie wystarczy zatem wziąć tego zdania w cudzysłów, zaś po prawej stronie podać jego przekład na metajęzyk – należy równoważność w odpowiedni sposób wyprowadzić z aksjomatów teorii prawdy.

W związku z powyższymi uwagami na temat konwencji T pojawia się pierwsze zastrzeżenie Alfreda Tarskiego wobec możliwości ścisłego zdefiniowania prawdy dla języka naturalnego. Wedle Tarskiego jest bardzo wątpliwe, czy z powodu braku precyzyjnych reguł gramatycznych da się sformułować dla tego języka aksjomaty semantycznej definicji prawdy. Drugie zastrzeżenie wiąże się z tym, że w języku naturalnym występują pojęcia semantyczne, np. predykat „prawdziwy”, co może prowadzić do paradoksów¹³. W podanym powyżej przykładzie paradoks powstanie, gdy trzecie zdanie na osiemnastej stronie książki K ma postać: „Trzecie zdanie na osiemnastej stronie książki K jest fałszywe”. W przypadku strukturalnoopisowych nazw zdań podobnego paradoksu można uniknąć, jeżeli wykluczmy z języka przedmiotowego predykat „prawdziwy”. Stanowi to oczywiste ograniczenie w stosowaniu semantycznej definicji prawdy do języków naturalnych, w których predykat ten faktycznie występuje. Wprawdzie Tarski kończy początkowe rozważania konkluzją, że przy próbie budowania semantycznej definicji prawdy dla języków naturalnych pojawiają się nieprzezwyciężone trudności, ale pod koniec tego samego artykułu dopuszcza taką możliwość pod warunkiem poddania języka naturalnego odpowiedniej formalizacji (s. 158). Podobną opinię wygłasza w pracy *Semantyczna koncepcja prawdy i podstawy semantyki* (1944, s. 244).

¹³ Paradoks powstanie, gdy trzecie zdanie na osiemnastej stronie książki K będzie brzmiało: „Trzecie zdanie na osiemnastej stronie książki K jest fałszywe”. W przypadku nazw strukturalnoopisowych paradoksu można uniknąć, gdy wykluczmy z języka przedmiotowego predykat „prawdziwy”. Stanowi to oczywiste ograniczenie w stosowaniu semantycznej definicji prawdy do języków naturalnych, w których predykat ten faktycznie występuje.

Przyjrzyjmy się teraz, w jaki sposób konwencja-T zapewnia merytoryczną trafność semantycznej definicji prawdy. Po lewej stronie T-zdania występuje definiowany metajęzykowy predykat „prawdziwy”, którego argumentem jest metajęzykowa strukturalnoopisowa nazwa zdania *p* języka przedmiotowego, zaś po prawej stronie występuje przekład zdania *p* na metajęzyk (jeśli metajęzyk obejmuje język przedmiotowy, to przekład jest identycznościowy, tj. dla każdego zdania *p* jego przekładem na metajęzyk jest to samo zdanie). Prawą stronę równoważności można niezbyt precyzyjnie rozumieć jako warunek tego, że jest tak, jak głosi zdanie *p*. Sama definicja prawdy nie rozstrzyga, czy faktycznie tak jest, natomiast osoba znająca metajęzyk (a zarazem rozumiejąca tłumaczenie zdania *p* na metajęzyk) może niekiedy dowiedzieć się, czy „rzeczy mają się tak, jak głosi tłumaczenie *p* na metajęzyk”, czyli uznać prawdziwość zdania *p*. Daje temu wyraz sam Tarski:

Rzeczywiście, semantyczna definicja prawdy nie implikuje niczego odnośnie do warunków uznawania zdań takich jak (1)¹⁴

(1) *śnieg jest biały*.

Implikuje jedynie to, że ilekroć uznajemy lub odrzucamy takie zdanie, musimy być gotowi uznać bądź odrzucić odpowiadające mu zdanie (2)

(2) zdanie „śnieg jest biały” jest prawdziwe (s. 268).

Ponieważ zdania uznajemy lub odrzucamy zgodnie z naszymi przekonaniami, często możemy się mylić i w konsekwencji błędnie uznawać jakieś zdanie za prawdziwe¹⁵. Warto również zauważyć, że większość znanych definicji prawdy podaje pewne kryteria prawdziwości. Należy do nich korespondencyjna definicja prawdy, która operuje kryterium zgodności zdania, sądu lub myśli z rzeczywistością, koherencyjna definicja prawdy, która operuje kryterium niesprzeczności sądu z innymi sądami przyjętymi wcześniej, definicja oparta na kryterium zgody powszechnej oraz definicja pragmatystyczna, która operuje kryterium użyteczności sądu w działaniu. Semantyczna definicja prawdy nie jest z pewnością definicją tego typu. Co więcej, we wspomnianych definicjach prawdy¹⁶ kryteria prawdziwości nadają terminowi „prawda” pewien sens. Semantyczna definicja prawdy jest definicją, która podaje jedynie zakres predykatu „prawdziwy” w języku przed-

¹⁴ Dotyczy to przypadku, gdy metajęzyk obejmuje język przedmiotowy.

¹⁵ Pod tym względem semantyczna definicja prawdy nie różni się zasadniczo od definicji wielu pojęć. Znajomość definicji nazwy gatunkowej nie gwarantuje trafnego zaliczenia pewnego obiektu do zakresu tego terminu.

¹⁶ Nie jest to, rzecz jasna, lista kompletna, lecz jedynie zawierająca przykłady. Ponadto każda z wymienionych przeze mnie definicji występuje w różnych wariantach.

miotowym, czyli pokazuje, w jaki sposób – przy pewnych zastrzeżeniach, o których będzie mowa dalej – można generować zdania prawdziwe.

Tarski, podobnie jak większość filozofów Szkoły lwowsko-warszawskiej, utrzymywał bardzo bliskie związki z Kołem wiedeńskim. Jednym z wyznaczników tego neopozytywistycznego nurtu było eliminowanie z filozoficznego dyskursu wszelkich terminów, które nie posiadały jasnych, najlepiej empirycznych, kryteriów stosowności. Takimi terminami były m.in. „znaczenie”, „treść” i „pojęcie”, rozumiane jako treść terminu. Pomimo że Tarski często używał terminu „pojęcie”, to formułując semantyczną definicję prawdy nie chodziło mu o definicję podającą znaczenia lub treści terminu „prawda”, czyli pojęcie prawdy w wymienionym wyżej sensie, lecz o to, aby podana definicja wyznaczała zakres, czyli ekstensję, terminu „prawdziwy” dla zdań ustalonego języka. Pojęcie prawdy było przez Tarskiego rozumiane jako termin „prawdziwy” lub predykat „prawdziwy”. Zauważmy również, że skoro semantyczna definicja prawdy jest zrelatywizowana do języka przedmiotowego, to nie może być definicją tzw. prawdy absolutnej, gdyż taka definicja musiałaby podać treść predykatu „prawdziwy” niezależnie od języka. Jeżeli bowiem pragniemy wyznaczyć zakres predykatu „prawdziwy” w innym języku, musimy zbudować od początku nową definicję.

3. METODOLOGICZNA POPRAWNOŚĆ

Wymóg metodologicznej poprawności semantycznej definicji prawdy ma przynajmniej dwa związane ze sobą aspekty. Pierwszy z nich wiąże się z charakterystyczną dla lat 30. i 40. XX w. niechęcią do semantyki, która była uważana za źródło paradoksów, oraz charakterystyczną dla Koła wiedeńskiego i panującego wówczas fizykalizmu tendencją redukcjonistyczną. Semantyczna definicja prawdy miała pozwolić na uniknięcie paradoksów semantycznych, a tym samym umożliwić zbudowanie wolnej od nich „naukowej semantyki”.

W *O ugruntowaniu naukowej semantyki* Tarski dostrzega rolę pojęć semantycznych w filozofii, wskazując jednak na te trudności w ich użyciu, które prowadzą do paradoksów. Autor wymienia w tym kontekście powszechnie znany paradoks kłamcy, paradoks wyrazu „heterologiczny”¹⁷ oraz paradoks

¹⁷ Zamiast terminu „heterologiczny” używa się również terminu „heterosemantyczny”. Paradoks związany z tym terminem można przedstawić w sposób następujący. Nazwijmy heterosemantycznym wyraz (najlepiej przymiotnik), który nie posiada cechy, jaką oznacza. Przykładem wyrazu heterosemantycznego byłby wyraz „jednosylabowy”, gdyż wyraz ten nie jest jednosylabowy. Paradoks pojawia się, gdy kryterium heterosemantyczności zastosujemy do samego wyrazu „heterosemantyczny” – okazuje się bowiem, że wyraz „heterosemantyczny” jest heterosemantyczny, gdyż nie jest heterosemantyczny.

definiowalności Richarda¹⁸. Trudności te były – zdaniem Tarskiego – spowodowane tym, że nie zdawano sobie sprawy z relatywnego charakteru pojęć semantycznych, tj. „z tego, że pojęcia te odnoszą się zawsze do wyrażen jakiegoś określonego języka” (1936, s. 174) i nie rozróżniano języka przedmiotowego od metajęzyka. Wprawdzie rozróżnienie na język przedmiotowy i metajęzyk nie było odkryciem Tarskiego, ale zastosowanie tego rozróżnienia w semantycznej definicji prawdy, która nie generuje paradoksów tego typu, stanowi niewątpliwie jego zasługę. Jak już wspomniałem w poprzednim paragrafie, cel ten został osiągnięty w ten sposób, że język przedmiotowy nie zawierał żadnych pojęć semantycznych, które stosuje się do jego wyrażen, gdyż pojęcia te występują jedynie w metajęzyku¹⁹.

Sprawą nieco mniej znaną był fakt, że celem Tarskiego było również budowanie definicji fizykalistycznie akceptowalnej. Oznaczało to, że pragnął zdefiniować prawdę w terminach fizycznych i matematycznych. Sprawa ta wymaga dokładniejszego omówienia, gdyż wiąże się z nią pewna interpretacja semantycznej definicji prawdy, która jest niezwykle istotna z punktu widzenia dokonanej przez Davidsona próby jej adaptacji jako teorii znaczenia. Nie jest bowiem do końca jasne, czy podstawowe pojęcia semantyczne występujące w semantycznej definicji prawdy są pojęciami *sensu stricto* semantycznymi (tylko precyzyjnie zdefiniowanymi w sposób fizykalistycznie akceptowalny), czy też są pojęciami *de facto* syntaktycznymi, które w języku nauk dedukcyjnych są również „fizykalistycznie dopuszczalne”. W aksjomatach semantycznej definicji prawdy występują bowiem pojęcia na pierwszy rzut oka semantyczne: przyporządkowanie przez ciąg obiektu zmiennej (w aksjomatach odniesienia) oraz spełniania funkcji zdaniowej przez ciąg obiektów (w aksjomatach spełniania). Sprawie tej poświęcę więcej miejsca w kolejnych paragrafach rozdziału. W tym momencie zwróćmy

¹⁸ Paradoks Richarda polega na spostrzeżeniu, że za pomocą pewnych (skończonych) wyrażen języka naturalnego da się jednoznacznie definiować liczby rzeczywiste. Następnie buduje się nieskończony ciąg wszystkich (skończonych) wyrażen języka naturalnego, które jednoznacznie definiują te liczby. Kolejność wyrażen w ciągu zależy od długości wyrażen, a w przypadku wyrażen o tej samej długości decyduje kolejność alfabetyczna. Ciąg ten wyznacza jednoznacznie ciąg wszystkich liczb rzeczywistych definiowalnych za pomocą (skończonych) wyrażen języka naturalnego. Następnie, odwołując się do tego ciągu liczb, definiuje się (w języku naturalnym) liczbę rzeczywistą, która do tego ciągu nie należy, co przeczy założeniu, że wyjściowy ciąg liczb jest ciągiem wszystkich liczb rzeczywistych definiowanych za pomocą skończonych wyrażen języka naturalnego.

¹⁹ Stanowi to niewątpliwie wyzwanie dla Donalda Davidsona i innych filozofów, którzy starali się dostosować semantyczną definicję prawdy do języków naturalnych, gdyż w językach tych termin „prawdziwy” faktycznie występuje. Davidson do pewnego stopnia bagatelizował ten fakt.

jedynie uwagę, że na początku lat 30. XX w. Tarski skłaniał się ku drugiej możliwości, o czym zdaje się świadczyć następujący cytat:

[w definicji prawdy] unikam stosowania pojęć o treści semantycznej [semantycznej], których sprecyzowanie stwarza przynajmniej takie same trudności jak pojęcie prawdy (1932, s. 10).

Podobnie we Wstępie do *Pojęcia prawdy* Tarski pisze:

(...) nie zamierzam użyć w tej konstrukcji [definicji prawdy] żadnego pojęcia natury semantycznej, o ile uprzednio nie uda mi się sprowadzić go do innych pojęć (1933, s. 14).

Przytoczone cytaty można interpretować jako wymóg, aby w definicji prawdy unikać wszelkich pojęć semantycznych. Zauważmy również, że pojawiające się w aksjomatach semantycznej definicji prawdy pojęcia *prima facie* semantyczne nie występują w T-zdaniach (poza predykatem „prawdziwy”). Być może zatem wymienione pojęcia są *de facto* pojęciami syntaktycznymi, ale w „semantycznym przebraniu”. Na przykład ciąg obiektów byłby faktycznie ciągiem nazw metajęzykowych, zaś spełnianie formuły przez ciąg obiektów byłoby spełnianiem przez ciąg nazw metajęzykowych.

W późniejszym okresie Tarski zwrócił się w stronę innej możliwości, co zauważył – w nieco innym kontekście – Adam Nowaczyk w serii swych publikacji o Tarskiego semantycznej teorii prawdy zebranych w zbiorze *Polawianie sensu w filozoficznej głębi* (2006, s. 146). Definicję prawdy Tarski zaczął bowiem traktować jako definicję prawdy w modelu, a zatem do modelu ją relatywizował. Przy takim rozumieniu prawdy pojęcie to jest, podobnie jak pozostałe pojęcia, o których była mowa, *sensu stricto* semantyczne, tj. ciąg obiektów jest faktycznie ciągiem obiektu dziedziny modelu, zaś spełnianie funkcji zdaniowej przez ciąg obiektów sprowadza się do spełniania jej przez pewne obiekty z dziedziny interpretacji, pary takich obiektów, trójki obiektów itd. Z kolei Davidson wielokrotnie krytykował definicję prawdy w modelu jako nieprzydatną dla semantyki języka naturalnego, nie akceptując jednak redukcji semantyki do składni. Do problemu tego powrócimy w dalszej części rozdziału.

4. FILOZOFICZNE PROBLEMY ZWIĄZANE Z SEMANTYCZNĄ DEFINICJĄ PRAWDY

Jednym z często podejmowanych w literaturze zagadnień jest pytanie o to, czy formułując semantyczną definicję prawdy Tarski żywił co do niej jakieś filozoficzne aspiracje. Jak twierdzi Scott Soames w artykule *What is a Theory of Truth?*, wielu filozofów ma poważne wątpliwości czy semantyczna definicja prawdy jest faktycznie tzw. teorią prawdy, tj. teorią, która – mówiąc niebýt ściśle – ma nam wyjaśnić czym jest prawda:

Podstawowy powód tej niepewności wiąże się z trudnością określenia czym powinna być teoria prawdy. Ogólnie rzecz biorąc celem każdej teorii prawdy było:

- i. podanie znaczenia predykatów prawdziwości [*truth predicates*] w języku naturalnym²⁰,
- ii. zastąpienie owych predykatów pewnymi substytutami, często ściśle zdefiniowanymi, aby wykorzystać je w dalszym programie redukcjonistycznym lub
- iii. użycie pewnego uprzednio zrozumianego pojęcia prawdy dla szerszych celów filozoficznych, jak wyjaśnienie pojęcia znaczenia lub obrona jakiegoś poglądu metafizycznego (Soames 1984, s. 411).

Gdyby celem semantycznej definicji prawdy było podanie sensu terminu „prawdziwy”, to – zdaniem Soamesa – formułując swą definicję Tarski musiałby przeanalizować „(...) zawartość paradygmatycznych przykładów, w których to co ma być prawdziwe jest raczej sądem [*proposition*] niż zdaniem lub wypowiedzeniem zdania” (*ibidem*). Jak twierdzi Soames, Tarski nie miał żadnych ambicji filozoficznych, czemu dał wyraz, głosząc neutralność filozoficzną semantycznej definicji prawdy:

Możemy więc przyjąć semantyczną koncepcję prawdy nie porzucając własnego stanowiska epistemologicznego, jakiegokolwiek by ono nie było; możemy pozostać naiwnymi realistami, krytycznymi realistami lub idealistami, empirystami lub metafizykami – kimkolwiek byliśmy przedtem. W stosunku do tych wszystkich spraw semantyczna koncepcja prawdy jest całkowicie neutralna (1944, s. 268).

Oczywiście tak rozumiana neutralność filozoficzna nie musi implikować braku jakichkolwiek filozoficznych ambicji ze strony Tarskiego. Musiałby one jednak wykraczać poza deklarowany cel, dla którego sformułował semantyczną definicję prawdy:

²⁰ Liczba mnoga bierze się stąd, że prawdziwość wyraża się na kilka sposobów, np. „prawdziwy”, „jest prawdą” itd.

Dalsze ważne rezultaty można otrzymać, stosując teorię prawdy do języków sformalizowanych pewnej bardzo szerokiej klasy dyscyplin matematycznych; nie należą do tej klasy jedynie dyscypliny o elementarnym charakterze i bardzo elementarnej strukturze logicznej. Okazuje się, że dla żadnej dyscypliny z tej klasy *pojęcie prawdy nigdy nie pokrywa się z pojęciem dowodliwości* – wszystkie bowiem zdania dowodliwe są prawdziwe, ale są zdania prawdziwe, które nie są dowodliwe (1944, s. 252–253).

Przez język sformalizowany – obecnie używa się raczej terminu „język formalny” – rozumie się język teorii logicznej lub matematycznej. W teorii matematycznej dowodzi się twierdzeń, wyprowadzając je z aksjomatów tej teorii. Okazuje się jednak, że dla niektórych teorii nie da się udowodnić wielu zdań prawdziwych, co stanowi ciekawy problem metamatematyczny. Jak pokazał Tarski, w odpowiednio wzbogaconym w stosunku do języka *L* metajęzyku *M* można te zdania udowodnić, co pozwala na wyznaczenie zbioru wszystkich zdań prawdziwych w *L*. Wyjaśnia to również dlaczego Tarski w większości sformułowań konwencji T zakładał, że metajęzyk obejmuje język przedmiotowy.

Semantyczna definicja prawdy pozwalała na sformułowanie wielu ciekawych rezultatów na temat teorii matematycznych. Ten aspekt semantycznej definicji prawdy nie jest kontrowersyjny i ze względu na cel tych rozważań nie będę się do niego więcej odnosił. Jednak nawet w przypadku zastosowania semantycznej definicji prawdy do języków formalnych, nie likwiduje to jednak wspomnianego wcześniej problem statusu użytych w niej pojęć semantycznych, chociaż likwiduje problem niedopuszczalnego dla zwolennika fizykalizmu statusu znaczeń wyrażeń języka.

Językiem przedmiotowym w oryginalnym sformułowaniu semantycznej definicji prawdy jest język formalnej teorii dedukcyjnej z aksjomatami oraz regułami inferencji. Języki takie mają dodatkową zaletę, którą wykorzystał Tarski. System aksjomatów oraz reguł inferencji określa bowiem za pomocą środków czysto syntaktycznych sensy wyrażeń języka przedmiotowego²¹:

Rezygnując z wyłuszczonych w §1 względów z prób rozwiązania podstawowego zagadnienia w odniesieniu do języka potocznego, w dalszym ciągu rozważań ograniczę się wyłącznie do *języków sformalizowanych*. Mógłbym je scharakteryzować nader ogólnikowo jako tego rodzaju (sztucznie skonstruowane) języki,

²¹ Wprawdzie Davidson adaptował semantyczną definicję prawdy do języków naturalnych, ale jego poglądy nie stoją w całkowitej sprzeczności z poglądami Tarskiego na temat roli aksjomatów w konstytuowaniu sensu wyrażeń. O znaczeniach wyrażeń indywidualnego języka każdego z nas decydują bowiem układy przekonań aktualizowane za pomocą wypowiedzi zdań, w których występują owe wyrażenia. Pogląd ten – nie do końca wyartykułowany – pojawił się w związku z pojęciem interpretacji, które zostanie omówione w rozdziale 4.

w których sens każdego wyrażenia jest jednoznacznie wyznaczony przez jego kształt (1933, s. 31).

O ile zaproponowana przeze mnie interpretacja tego fragmentu jest właściwa, to mówiąc o sensie zdania Tarski miał na myśli coś, co filozofowie Koła wiedeńskiego nazywali znaczeniem analitycznym zdania. Oczywiście w podejściu tym nie da się wskazać żadnego bytu będącego znaczeniem analitycznym, lecz można ogólnie powiedzieć, że układ aksjomatów teorii sprawia, że wyrażenia, które w nich występują, stają się sensowne, zaś dwa wyrażenia języka, które pełnią w nich różną rolę, mają różne sensy²². To ograniczenie pozwoliło w meta-matematycznych zastosowaniach semantycznej definicji prawdy uniknąć Tarskiemu problemów związanych ze znaczeniem i przekładem.

Tarski nie musiał zatem – w odróżnieniu od sytuacji, gdyby definicja prawdy zastosowana została do języków naturalnych lub jedynie sformalizowanych – troszczyć się o znaczenia wyrażen oraz o założony przezeń przekład z języka przedmiotowego na metajęzyk. Jeżeli dodatkowo założyć się, co Tarski często czynił, że metajęzyk obejmuje język przedmiotowy, problem znaczeń można uznać za wyeliminowany. W tym miejscu warto się odwołać do zawartych w (1936a) uwag Tarskiego pod adresem Marii Kokoszyńskiej:

Ze słów referentki wynika m.in., że pojęcie prawdy – w jednej z jego interpretacji – winno być relatywizowane do pojęcia znaczenia. Czy nie byłoby rzeczą prostszą relatywizować pojęcie prawdy do pojęcia języka, które wydaje się pojęciem jaśniejszym i logicznie mniej skomplikowanym od pojęcia znaczenia? (s. 203).

Na pierwszy rzut oka wydaje się, że nie ma między obu podejściami różnicy. Język zinterpretowany można bowiem rozumieć jako układ znaków z przyporządkowanymi im znaczeniami – dla Tarskiego interesującym językiem w tym sensie był system aksjomatyczny, w którym aksjomaty podają sensy wyrażen językowych²³. Wyjaśnia to, dlaczego Tarski uważał, iż zrelatywizowanie do

²² W inny, „semantyczny” sposób interpretuje ten fragment Hartry Field w znanym artykule *Tarski's Theory of Truth*. Cytując ten sam fragment Field pisze: „(...) co znaczy, że dwóch mówiących używa tej samej nazwy (lub jeden mówiący używa jej dwukrotnie) odnoszą się do tego samego obiektu (...)” (1972, s. 348). Przedstawiona przeze mnie interpretacja nawiązuje do tzw. dyrektywnej teorii znaczenia, którą sformułował Kazimierz Ajdukiewicz w (1931, 1934 oraz 1934a).

²³ W tym miejscu konieczna jest uwaga na temat terminów „sens” i „znaczenie”. Nie wchodząc w szczegóły można powiedzieć, że przedteoretyczny termin „znaczenie” można traktować jako sposób rozumienia wyrażenia. Wielu filozofów języka pragnęło uściślić ten termin, budując różne teorie znaczenia. Dla uściśłego pojęcia znaczenia niektórzy z nich używali (za Fregem)

języka (w specyficznym rozumieniu tego terminu jako systemu aksjomatycznego) jest prostsze od zrelatywizowania do znaczenia, którego zdefiniowanie natrafiało w tamtych czasach na nieprzezwyciężalne trudności. W odpowiedzi Kokoszyńska w zasadzie przyznała rację Tarskiemu, stwierdzając, że pojęcie prawdy zależy od „zespołu brzmień składających się na zdania języka” oraz od sposobu przekładu owych brzmień na metajęzyk. A zatem pojęcie prawdy jest zrelatywizowane jedynie do przekładu, który jest wyznaczony przez aksjomaty semantycznej definicji prawdy (s. 204). Nie zmienia to jednak faktu, że aksjomaty semantycznej definicji prawdy zależą od sensów występujących w nich wyrażań.

Zauważmy również, że w semantycznej definicji prawdy język przedmiotowy nie zawiera terminów semantycznych, które pojawiają się dopiero w metajęzyku. W takim przypadku przeniesiony z języka przedmiotowego do metajęzyka układ aksjomatów nie wyznacza sensów terminów semantycznych. Problem fizykalistycznej dopuszczalności pojęć semantycznych nie został zatem rozwiązany przez odwołanie się do sensów i będziemy do niego powracać w dalszych częściach rozdziału.

Nieco wcześniej pisałem, że doraźnym celem semantycznej definicji prawdy było – w intencji jej twórcy – zastosowanie jej do osiągania rezultatów metamatematycznych. Osiągnięcie tego celu wiązało się z eksplikacją pojęcia prawdy, która spełniałaby pewne narzucone kryteria. Innymi słowy, jak twierdził Soames, ostatecznym celem semantycznej definicji prawdy było włączenie jej w charakterystyczny dla filozofii neopozytywizmu program redukcjonistyczny. Nie jest jednak całkowicie bezzasadne przypuszczenie, że ambicje Tarskiego wykraczały poza zastosowania semantycznej definicji prawdy do systemów aksjomatycznych i metajęzyków obejmujących te systemy. Mamy zatem dwie możliwości, które nie muszą się wykluczać. Po pierwsze, semantyczna definicja prawdy może być uznana za użyteczne narzędzie w filozofii matematyki (co jest niewątpliwie faktem) oraz (być może) filozofii języka naturalnego²⁴. Po drugie, może wnieść coś istotnego do filozoficznej dyskusji na temat pojęcia prawdy. Tarski wielokrotnie podkreślał, że jego definicja stanowi eksplikację tzw. klasycznej koncepcji prawdy.

terminu „sens”. Opisowi różnych teorii znaczenia poświęciłem znaczną część mojej monografii (2008).

²⁴ Jak pamiętamy, warunkiem zastosowania semantycznej definicji prawdy do języka naturalnego jest formalizacja fragmentu języka naturalnego, polegająca na zastąpieniu go językiem o ściśle określonej strukturze, tj. bez wieloznaczności, nieostrości i okazjonalności. „Wątpić jednak wolno, czy zracjonalizowany na tej drodze język potoczny zachowałby swą cechę naturalności i czy nie zyskałby wówczas charakterystycznych znamion języków sformalizowanych” (1933, s. 158).

5. SEMANTYCZNA DEFINICJA PRAWDY A KLASYCZNA KONSEPCJA PRAWDY

Ważnym etapem w budowaniu naukowej semantyki jest „doprecyzowanie warunków, przy których sposób operowania pojęciami semantycznymi skłonni jesteśmy uznać za rzeczowo trafny i zgodny z potoczną intuicją” (1936, s. 177)²⁵. Za najbardziej intuicyjną rozumie Tarski klasyczną koncepcję prawdy, którą zazwyczaj interpretuje się jako zgodność zdania myśli lub sądu z rzeczywistością. Innymi słowy bardzo często utożsamia się klasyczną koncepcję prawdy z korespondencyjną definicją prawdy. Ów zwrot „zgodność z rzeczywistością” był – zdaniem Tarskiego – źródłem „poważnych nieporozumień” oraz „mętnych spekulacji”. Jak pokazuje bowiem historia zmagania z pojęciem prawdy, nie udało się wyjaśnić, na czym kryterium zgodności miałoby polegać.

Tarski odróżnił zatem pochodzącą od Arystotelesa klasyczną koncepcję prawdy od późniejszej tzw. korespondencyjnej koncepcji prawdy. Klasyczna koncepcja prawdy w swym oryginalnym sformułowaniu nie posługuje się bowiem pojęciem zgodności z rzeczywistością, lecz stwierdza jedynie, że zdanie jest prawdziwe wtedy i tylko wtedy, gdy jest tak, jak to zdanie głosi, zaś jest fałszywe, gdy tak nie jest. Do takiego rozumienia prawdy odwołał się Tarski *explicite* w wydanym pierwotnie po angielsku artykule *Semantyczna koncepcja prawdy i podstawy semantyki* (1944), uściślając tym samym to, co we wcześniejszych pismach często nie było wyraźnie wyartykułowane²⁶.

Koncepcję klasyczną wyraża znany fragment z *Metafizyki* Arystotelesa [1011b], który Tarski przytacza w postaci następującej: „Jest fałszem powiedzieć o tym co jest, że nie jest, lub o tym, co nie jest, że jest; jest prawdą powiedzieć o tym, co jest, że jest, lub o tym co nie jest, że nie jest” (s. 232). W przytoczonym fragmencie, uznawanym powszechnie za oryginalne sformułowanie klasycznej koncepcji prawdy, nie pojawia się ani termin „korespondencja”, ani „zgodność”, ani „oznaczanie stanu rzeczy”. Do tych właśnie pojęć odwoływały się późniejsze sformułowania klasycznej koncepcji prawdy zwane od czasów św. Tomasza koncepcjami korespondencyjnymi, które można traktować jako próby wyeksplikowania wprowadzonego przez Arystotelesa klasycznego pojęcia prawdy (por. s. 222–223).

W tym miejscu warto zauważyć, że tłumaczenie angielskie, na którym oparł się Tarski, różni się zdecydowanie od znacznie późniejszego tłumaczenia

²⁵ Jest to stawiany procedurze eksplikacji wymóg, aby pojęcie uściślone (*explicatum*) było pod istotnymi względami podobne do pojęcia uściślanego (*explicandum*).

²⁶ Nie twierdzę, że Tarski nie czynił wcześniej tego rozróżnienia, a jedynie że nie akcentował go wyraźnie (por. 1933, s. 15 oraz 1936, s. 174).

tego fragmentu bezpośrednio na język polski: „Twierdzić o Bycie, że nie istnieje, albo o Nie-Bycie, że istnieje, jest fałszem; natomiast twierdzić, że Byt istnieje, a Nie-Byt nie istnieje, jest prawdą” (Arystoteles 1983, s. 99). O ile pierwsze tłumaczenie pozwala na bliską Tarskiemu interpretację nominalistyczną, gdyż „jest” rozumie się tam jako relację orzekania, to drugie prowadzi do rozumienia „jest” jako „istnieje”. Nie jest moim celem rozstrzygnięcie, która z interpretacji klasycznej definicji prawdy jest zgodna z intencjami samego Arystotelesa. Należy jednak wyraźnie podkreślić, że semantyczna definicja prawdy jest do pogodzenia jedynie z interpretacją Rossa.

Odwołanie się do oryginalnego sformułowania pozwala uniknąć Tarskiemu wielu problemów związanych z korespondencyjną koncepcją prawdy, gdyż:

Każde z tych sformułowań²⁷ [korespondencyjnej koncepcji prawdy] może jednak prowadzić do różnych nieporozumień, ponieważ żadne z nich nie jest dostatecznie dokładne i jasne (aczkolwiek odnosi się to w mniejszym stopniu do pierwotnego, Arystotelesowskiego sformułowania niż do każdego z pozostałych): w każdym razie żadnego z nich nie można uważać za zadowalającą definicję prawdy. Poszukianiem dokładniejszego sposobu wyrażenia naszych intuicji zajmujemy się właśnie teraz (1944, s. 233).

W uproszczeniu można powiedzieć, że każde wynikające z semantycznej definicji prawdy T-zdanie stwierdza, że zdanie, które głosi, że jest tak a tak, jest prawdziwe wtedy i tylko wtedy, gdy jest tak a tak. Nie ma tutaj żadnego odwołania do rzeczywistości lub stanów rzeczy. Nie ma również jakiegokolwiek kryterium prawdziwości ani kryterium stwierdzania, że jest tak a tak lub nie jest tak a tak.

Uzasadnienie odwołania się do Arystotelesa można znaleźć w drugiej części *Semantycznej koncepcji prawdy*, w której Tarski bronił swej definicji przed zarzutami oponentów. Oryginalną, pochodzącą od Arystotelesa, wersję klasycznej koncepcji prawdy określił tam jako „sformułowaną w prostych słowach”, w odróżnieniu od sformułowania „filozoficznego” jako zgodności z rzeczywistością lub istniejącym stanem rzeczy. Semantyczną definicję prawdy uznał za eksplikację tego pierwszego ujęcia²⁸:

²⁷ Chodzi tu o pojawiające się w różnych sformułowaniach tej koncepcji określenia „korespondencja”, „zgodność” lub „oznaczanie stanu rzeczy”.

²⁸ Nie twierdzę bynajmniej, że wyeksplikowanie klasycznej koncepcji prawdy w postaci teorii korespondencyjnej jest całkowicie bezpodstawne. Warto natomiast zwrócić uwagę, że każda definicja prawdy, która oparta jest na kryterium prawdziwości, podpada pod zarzut regresu w nieskończoność. Jeśli zdanie jest prawdziwe na mocy pewnego kryterium, np. zgodności z rzeczywistością, to można spytać się o meta-kryterium, które pozwala na zastosowania tego

Mimo wszystko ja akurat jestem przekonany, że koncepcja semantyczna [semantyczna definicja prawdy] zgodna jest w znacznej mierze ze zdroworozsądkowym sposobem posługiwania się pojęciem prawdy (...) (1944, s. 265).

Co więcej, „filozoficzne” sformułowanie klasycznej koncepcji prawdy (czyli koncepcja korespondencyjna) nie jest – zdaniem Tarskiego – intuicyjne:

Jeśli więc zapytamy ucznia szkoły średniej lub nawet dorosłego inteligentnego człowieka bez specjalnego przygotowania filozoficznego, czy uważa, że zdanie jest prawdziwe, jeśli jest zgodne z rzeczywistością, lub jeśli oznacza istniejący stan rzeczy, to może się po prostu okazać, że nie rozumie on pytania (...). Natomiast z punktu widzenia rozważanego problemu bardzo ważna będzie odpowiedź na pytanie, czy zgodziłby się z tym, że „*śnieg pada*” może być prawdziwe, chociaż śnieg nie pada, lub może być fałszywe, chociaż śnieg pada (s. 266).

Przedstawione tu stanowisko Tarskiego ma jeszcze jeden bardzo ważny aspekt. „Filozoficzne” sformułowanie klasycznej koncepcji prawdy zdaje się czynić pewne założenia ontologiczne, których Tarski starał się uniknąć. Podkreślany przez Tarskiego wymóg neutralności filozoficznej nie dopuszcza m.in. założeń ontologicznych dotyczących istnienia obiektów matematycznych, gdyż trudno jest pogodzić neutralność filozoficzną w matematyce z platonizmem. Z kolei w przypadku języka naturalnego pojęcie prawdy odnosimy również do fikcji literackiej. Z problemem tym starało się uporać wielu filozofów. Gottlob Frege twierdził, że zdania na temat fikcji nie są ani prawdziwe, ani fałszywe (por. Frege 1892), zaś Bertrand Russell utrzymywał, że wszystkie takie zdania są fałszami (por. Russell 1905). Zarówno jedno, jak i drugie stanowisko wydaje się skrajnie nieintuicyjne. Z kolei Roman Ingarden stał na stanowisku o wiele bardziej intuicyjnym (dopuszczał zdania prawdziwe, fałszywe oraz pozbawione wartości logicznych), ale uczynił to kosztem wprowadzenia do ontologii obiektów czysto intencjonalnych. Ze względu na swą filozoficzną neutralność semantyczna definicja prawdy może stanowić dogodne narzędzie analizy języka naturalnego i z tego powodu być z powodzeniem wykorzystywana w filozofii języka.

pierwszego. To samo pytanie pojawia się w stosunku do owego meta-kryterium, co prowadzi do regresu w nieskończoność (por. Frege 1978/1918, s. 102). Do sprawy tej jeszcze powrócimy, gdyż Frege – podobnie jak Davidson – uznał, że prawda jest pojęciem pierwotnym i niedefiniowalnym.

6. ROLA T-ZDAŃ W SEMANTYCZNEJ DEFINICJI PRAWDY

Wokół T-zdań narosło wiele nieporozumień, które zazwyczaj mają swe źródło w pomijaniu aksjomatów semantycznej definicji prawdy, z których T-zdania wynikają. Dlatego też warto poświęcić im nieco uwagi, tym bardziej, że Davidson przypisał im odmienną rolę niż Tarski.

Odwołanie się do T-zdań w kontekście prawdy nie było oryginalnym pomysłem Tarskiego. Po raz pierwszy zostały one *explicite* użyte przez zmarłego w 1930 roku Franka Ramseya w wydanym pośmiertnie artykule *The Nature of Truth* (Ramsey 1990). Jednak, jak pisze M. Lynch w wydanym pod tym samym tytułem wyborze prac dotyczących prawdy, „[Tarski] jako pierwszy użył T-zdań do testowania teorii prawdy” (Lynch 2001, s. 323). W tym miejscu mogę jedynie zaznaczyć, że nie do końca zgadzam się z Lynchem w kwestii roli T-zdań w oryginalnym sformułowaniu semantycznej definicji prawdy dla języków formalnych. Jeśli zakłada się, za Tarskim, przekład języka przedmiotowego na metajęzyk (w szczególności metajęzyk może obejmować język przedmiotowy), to – jak twierdził Tarski – T-zdania pełnią funkcję definicji cząstkowych predykatu prawdziwy, gdyż pozwalają zaliczyć dane zdanie do zakresu tego predykatu lub je z tego zakresu wykluczyć. O testowaniu semantycznej definicji prawdy można mówić jedynie wtedy, gdy istnieje możliwość sfalsyfikowania tej definicji, a w konsekwencji sfalsyfikowania przekładu. Z taką sytuacją możemy mieć do czynienia, gdy wynikające z definicji T-zdanie zalicza jakieś zdania do zakresu predykatu prawdziwy, chociaż jego przekład na metajęzyk uznajemy za fałszywy (lub *vice versa*). W takim przypadku możemy zakwestionować przekład, a tym samym zakwestionować samą semantyczną definicję prawdy. Jest to zasadniczo możliwe w przypadku, gdy język przedmiotowy i metajęzyk są językami naturalnymi. Jeżeli zbudujemy nową definicję prawdy, to musimy tym samym założyć nowy przekład, a zatem przypisać wyrażeniom językowym nowe znaczenia. Tą właśnie drogą podążył Donald Davidson, traktując semantyczną definicję prawdy jako teorię podającą znaczenia. Należy tutaj również podkreślić, że żadne T-zdanie nie jest definicją prawdy, gdyż definicję prawdy stanowi ogół jej aksjomatów, przy założeniu przekładu z języka przedmiotowego na metajęzyk.

Mamy zatem do czynienia z dwiema funkcjami T-zdań: funkcją definicji cząstkowej predykatu prawdziwy oraz – przy pewnej modyfikacji założeń czynionych przez Tarskiego – funkcją pozwalającą na testowanie, a w konsekwencji na weryfikację, samej semantycznej definicji prawdy. Dotyczy to nawet sytuacji, gdy język przedmiotowy obejmuje metajęzyk. Nic nie stoi bowiem na przeszkodzie, aby zdanie „Śnieg jest biały” w języku przedmioto-

wym miało inne znaczenie (nawet w potocznym sensie terminu „znaczenie”) niż równokształtny napis w metajęzyku. W takim przypadku, o ile sformułujemy nową teorię prawdy, to metajęzyk nie będzie już obejmował języka przedmiotowego. W tym miejscu natrafiamy jednak na pewną trudność, która nie występowała u Tarskiego. Jeżeli bowiem znany nam metajęzyk, w którym skonstruowaliśmy teorię prawdy, obejmuje nasz własny język przedmiotowy, lub jeżeli znamy przekład z języka przedmiotowego na metajęzyk, to nie ma w ogóle sensu mówić o weryfikacji teorii prawdy, gdyż znamy znaczenia wyrażen języka przedmiotowego. Aby dokonać weryfikacji teorii prawdy potrzebny nam jest drugi człowiek, który włada językiem przedmiotowym i jest w stanie przynajmniej niektóre zdania tego języka zakwalifikować jako prawdziwe.

Oczywiście nie każde T-zdanie pozwala na taką weryfikację. Wyobraźmy sobie bowiem, że ów nasz rozmówca włada językiem przedmiotowym, na który składają się zdania równokształtne ze zdaniami metajęzyka (tj. języka polskiego). Załóżmy, że semantyczna definicja prawdy ma charakter trywialny, tj. niech wynikają z niej T-zdania postaci:

Zdanie „*p*” jest prawdziwe wtw, gdy *p*.

Niech język przedmiotowy różni się od metajęzyka jedynie tym, że słowo „śnieg” znaczy „węgiel” (i *vice versa*) oraz „biały” znaczy „czarny” (i *vice versa*). W takim przypadku wynikające z trywialnej – i jak łatwo zauważyć błędnej – teorii prawdy słynne T-zdanie:

(6.1) Zdanie „Śnieg jest biały” jest prawdziwe wtw, gdy śnieg jest biały

nie pozwoli na jej zweryfikowanie, gdyż obaj jesteśmy zgodni w kwestii, że śnieg (w normalnych okolicznościach) jest biały, a węgiel jest czarny, gdyż to drugie ma na myśli nasz rozmówca, twierdząc, że węgiel jest czarny.

Aby zweryfikować semantyczną definicję prawdy, musimy odwołać się do innych wynikających z niej T-zdań, np. do:

(6.2) Zdanie „Węgiel jest biały” jest prawdziwe wtw, gdy węgiel jest biały.

Metajęzykowe zdanie po prawej stronie (6.2) jest ewidentnie fałszywe, gdyż węgiel rozumiany jako surowiec opałowy (nie zaś diament) w normalnych warunkach biały nie jest. W konsekwencji (6.2) pozwala nam uznać zdanie języka przedmiotowego „Węgiel jest biały” za fałszywe. W tym momencie musimy się odwołać do pewnych argumentów natury pragmatycznej. Jeżeli nasz rozmówca będzie się upierał, że zdanie „Węgiel jest biały” jest prawdziwe, to – przy szeregu zastrzeżeń, o których będzie mowa głównie w rozdziale 4 – możemy podjąć próbę weryfikacji semantycznej definicji

prawdy. Jeśli w konsekwencji ją odrzucimy, a tym samym odrzucimy przykład z języka rozmówcy na nasz własny, to możemy spróbować zastąpić starą definicję definicją nową, o innych aksjomatach²⁹. Oczywiście aby dokonać takiej zmiany nie wystarczy jedynie zakwestionować (6.2), lecz należy się odwołać do wielu innych T-zdań, wynikających ze starej definicji. Gdyby procedura zmiany aksjomatów zakończyła się powodzeniem, z nowej definicji wynikałyby następujące T-zdania:

(6.3) Zdanie „Węgiel jest biały” jest prawdziwe wtw, gdy śnieg jest biały.

(6.4) Zdanie „Śnieg jest czarny” jest prawdziwe wtw, gdy węgiel jest czarny.

W podanym przykładzie dla uproszczenia z góry wskazałem, jak należy przełożyć odpowiednie terminy z języka przedmiotowego na metajęzyk. W praktyce jednak tego nie wiemy, a weryfikacja teorii prawdy w sensie zaproponowanym przez Davidsona jest procesem, który nie zawsze musi zakończyć się całkowitym powodzeniem.

7. PROGRAM NAUKOWEJ SEMANTYKI

Powróćmy obecnie do pytania, czy semantyczna definicja prawdy była semantyczna *sensu stricto*, czy też miała w istocie syntaktyczny charakter, choć występowała w „semantycznym przebraniu”. Jak pamiętamy, definicja prawdy stanowiła – w zamyśle Tarskiego – część szerszego programu redukcjonistycznego. Aby lepiej zrozumieć kierujące Tarskim motywacje oraz kontrowersje narosłe wokół jego teorii, warto odwołać się do artykułu *O ugruntowaniu naukowej semantyki* (1936), w którym autor kreśli podstawy swego programu semantycznego. Przez semantykę Tarski rozumie:

(...) ogół rozważań, dotyczących tego rodzaju pojęć, które – mówiąc ogólnikowo i niezbyt ściśle – wyrażają pewne związki między wyrażeniami języka a przedmiotami, „o których w wyrażeniu mowa”. Jako typowe przykłady pojęć z zakresu semantyki wystarczy przytoczyć pojęcia *oznaczania*, *spełniania*, *definiowania* (...). Zaliczyć tu dalej należy (co dotąd na ogół uchodziło uwagi) pojęcie *prawdy*, przynajmniej przy „klasycznym” rozumieniu tego pojęcia, tzn. przy takim rozumieniu, według którego prawdziwość zdania polega na jego „zgodności z rzeczywistością” (s. 174–175).

²⁹ Przykład ten ma jedynie charakter ilustracyjny. Stosownych aksjomatów podać nie mogę, gdyż terminy „węgiel” i „śnieg” są tzw. terminami masowymi, dla których, jak przyznaje sam Davidson, semantycznej definicji prawdy podać się nie da.

Warto zauważyć, że w omawianym artykule Tarski nie wyróżnia pojęcia prawdy, lecz traktuje ją na równi z innymi pojęciami semantycznymi, które powinny zostać poprawnie i ściśle wyeksplikowane w ramach postulowanej przez niego naukowej semantyki.

Pierwszy etap budowania naukowej semantyki polega na ścisłym opisie morfologii (składni) języka, co jest – zdaniem Tarskiego – w pełni wykonalne jedynie dla języków formalnych. Drugi etap polega na zbudowaniu metajęzyka, który powinien zawierać wyrażenia języka przedmiotowego oraz terminy z zakresu morfologii (składni) tego języka, co stanowiłoby dostateczną podstawę do badań semantycznych:

Sprawę tę przesądza zresztą w pewnej mierze sam charakter pojęć semantycznych: ponieważ pojęcia te wyrażają pewne zależności między przedmiotami, „o których mowa w badanym języku”, a wyrażeniami języka, więc w zdaniach charakteryzujących pojęcia semantyczne w ich właściwej roli [tj. zdaniach metajęzyka], występować muszą zarówno nazwy owych przedmiotów, a zatem wyrażenia badanego języka jak i terminy, którymi posługujemy się przy opisie strukturalnym języka – nazwy konkretnych wyrażen języka, klas tych wyrażen, strukturalnych relacji między wyrażeniami itd. (s. 176).

Fragment ten wyraźnie wskazuje na to, że eksplikacja pojęć semantycznych miała być związana – przynajmniej początkowo – z ich redukcją do pojęć syntaktycznych. Następnie Tarski postuluje stworzenie (jakiejs) semantycznej definicji prawdy rozumianej jako zespół przyjętych w metajęzyku aksjomatów i stwierdza, że jeżeli na gruncie tak rozumianej definicji prawdy da się uzasadnić każde T-zdanie, to oznacza to, że ustaliliśmy w metajęzyku „rzeczowo trafny sposób operowania pojęciem prawdy”. Analogiczną procedurę proponuje Tarski dla pozostałych pojęć semantycznych:

(...) dla każdego z tych pojęć formułujemy układ zdań o postaci równoważności, zdań, które – z intuicyjnego punktu widzenia – noszą charakter „częstkowych definicji” i ustalają sens rozważanego pojęcia w tych wszystkich przypadkach, w których odnosi się ono do konkretnych, strukturalnie opisanych wyrażen badanego języka (s. 178).

A zatem, podobnie jak dla prawdy, Tarski postuluje stworzenie na gruncie metajęzyka definicji pozostałych pojęć semantycznych i przyjęcie odpowiedniej konwencji (odpowiednika konwencji T), która głosi, że operowanie tym pojęciem jest rzeczowo trafne wtedy i tylko wtedy, gdy z owej teorii da się wyprowadzić analogiczne do T-zdań definicje cząstkowe tego pojęcia. Jako jedyny przykład podaje definicję cząstkową pojęcia spełniania (od angielskiego terminu *satisfaction* nazwijmy ją S-zdaniem): „Ludzie Jan i Piotr

spełniają funkcję zdaniową „ X i Y są braćmi” wtedy i tylko wtedy, gdy Jan i Piotr są braćmi” (s. 179).

Pozostaje zatem do ustalenia jaką postać powinna mieć postulowana przez Tarskiego definicja prawdy (lub definicja innego pojęcia semantycznego), z której wynikają T-zdania lub – jak w ostatnim przykładzie – S-zdania. Możliwe są tutaj dwie drogi. Pierwsza to metoda charakterystyczna dla systemów aksjomatycznych, w której – jak w teorii logicznej lub matematycznej – aksjomaty podają sens definiowanych pojęć (np. prawdy i spełniania)³⁰. Metodę tę Tarski odrzuca z powodów następujących:

1. Wybór aksjomatów jest zawsze w pewnej mierze przypadkowy;
2. Pojawia się problem niesprzeczności aksjomatów;
3. Jest nie do przyjęcia z punktu widzenia psychologicznego, gdyż rolę pojęć pierwotnych pełniłyby te pojęcia, które wielokrotnie stanowiły źródło „powiślań i nieporozumień”;
4. Takie ugruntowanie semantyki nie dałoby się pogodzić z „postulatami jedności wiedzy i fizykalizmu, „gdyż pojęcia semantyczne nie byłyby ani pojęciami logicznymi, ani fizycznymi” (s. 180).

Ostatecznie przyjęta przez Tarskiego druga metoda wskazuje raczej na słuszność „syntaktycznej” interpretacji semantycznej definicji prawdy:

Wszystkie trudności odpadłyby, gdyby udało się zastosować tę drugą metodę budowania semantyki, gdybyśmy mianowicie zdołali trafnie zdefiniować pojęcia semantyczne przy pomocy pozostałych pojęć metajęzyka, a więc pojęć ogólnologicznych, pojęć samego języka i pojęć specyficznych morfologii języka. Semantyka języka straciłaby wówczas charakter samodzielnej nauki, stałaby się wyspecjalizowanym rozdziałem dostatecznie szeroko pojmowanej morfologii języka (s. 181).

Tarski zauważa, że w semantycznej definicji prawdy należy „wysunąć na pierwszy plan” pojęcie spełniania, gdyż do tego właśnie pojęcia da się sprowadzić pozostałe pojęcia semantyki. Podobne uwagi można znaleźć w *Semantycznym pojęciu prawdy* (1944). W artykule tym Tarski precyzuje warunki niezbędne do precyzyjnego zrozumienia semantycznej definicji prawdy. Po pierwsze, konieczne jest „uchwycenie rzeczywistego znaczenia pojęcia starego [prawdy]” (s. 230), a po drugie, określenie od czego zależy formalna poprawność definicji. Aby osiągnąć formalną poprawność definicji musimy:

³⁰ Przypomnijmy, że semantyczna definicja prawdy nie jest teorią tego typu, gdyż w jej aksjomatach (poza ostatnim) pojęcie prawdy nie występuje.

(...) wyszczególnić słowa lub pojęcia, których pragniemy użyć, definiując pojęcie prawdy; musimy też podać formalne reguły, do których definicja ta powinna się stosować. Mówiąc ogólniej, musimy opisać formalną strukturę języka, w którym definicja prawdy zostanie podana (s. 230–231).

Jak widać, wyszczególnienie podstawowych pojęć występujących w definicji było – w zamyśle Tarskiego – zabiegiem czysto formalnym i syntaktycznym w ramach metajęzyka. Pojęcia te nie mają zatem mieć charakteru semantycznego w tradycyjnym rozumieniu tego terminu. Redukcja semantyki do składni (morfologii) doskonale się wpisuje w redukcjonistyczny program naukowej semantyki.

Dlaczego jednak, pomimo próby redukcji semantyki do składni, Tarski systematycznie używa terminów „semantyka” i „semantyczny”? Podobnie jak w *O ugruntowaniu naukowej semantyki* Tarski zwraca uwagę, że przedteoretyczne semantyczne pojęcie prawdy różni się od innych przedteoretycznych pojęć semantycznych, tj. oznaczania, spełniania i definiowania, tym, że nie ma charakteru relacyjnego i wyznacza w języku klasę zdań. Semantyczny charakter zdefiniowanej przez Tarskiego prawdy – nawet jeśli jest to pojęcie syntaktyczne jedynie w „semantycznym przebraniu” – zapewniony jest pośrednio dzięki temu, że ścisłą definicję prawdy można uzyskać, korzystając ze ścisłego pojęcia spełniania, które jest rezultatem eksplikacji semantycznego, przedteoretycznego pojęcia denotowania lub oznaczania. Pozwoliło mu to również odróżnić wybrany przez siebie sposób definiowania prawdy od wspomnianego wcześniej, lecz odrzuconego sposobu aksjomatycznego³¹.

8. WSTĘPNA REKONSTRUKCJA SEMANTYCZNEJ DEFINICJI SPEŁNIANIA

Intuicyjnie rzecz biorąc, spełnianie jest „relacją zachodzącą między dowolnymi przedmiotami a pewnymi wyrażeniami zwanymi *funkcjami zdaniowymi*” (1944, s. 249). Takie rozumienie spełniania nie odpowiada jednak „syntaktycznemu” podejściu Tarskiego. Dlatego też nieco dalej proponuje inną definicję:

³¹ Należy podkreślić, że aksjomatyczny sposób definiowania prawdy lub oznaczania charakteryzowałby się tym, że predykat „prawdziwy” i „oznacza” występowałby w aksjomatach, które nadawałyby sens tym pojęciom. W semantycznej definicji prawdy aksjomaty nie nadają sensu ani pojęciu prawdy, ani pojęciu spełniania, lecz pozwalają wyznaczyć zakres tych pojęć w języku przedmiotowym.

(...) dane przedmioty spełniają daną funkcję [zdaniową], jeśli staje się ona zdaniem prawdziwym po zastąpieniu w niej zmiennych wolnych nazwami tych przedmiotów (s. 250).

Definicja ta, chociaż niewątpliwie ma charakter syntaktyczny, odwołuje się *explicite* do pojęcia prawdy, zaś *implicite* do pojęcia spełniania, które miało służyć do zdefiniowania prawdy. Dlatego też Tarski proponuje rekurencyjną procedurę definiowania spełniania:

Wskazujemy jakie przedmioty spełniają najprostsze funkcje zdaniowe, a następnie podajemy warunki określające, kiedy dane przedmioty spełniają złożone funkcje zdaniowe (*ibidem*).

Procedura ta nie rozstrzyga jeszcze czy spełnianie jest pojęciem *sensu stricte* semantycznym, czy też pojęciem syntaktycznym „w semantycznym przeobrażeniu”, gdyż zamiast o nazwie mowa jest o oznaczanym przez ową nazwę przedmiocie. Pewna wzmianka wskazująca na syntaktyczną interpretację pojęcia spełniania znajduje się w *Pojęciu prawdy* z roku 1933:

Powiedzieć, że nazwa x oznacza dany przedmiot a , to to samo co stwierdzić, że przedmiot a (względnie każdy ciąg, którego odpowiednim wyrazem jest a) spełnia funkcję zdaniową pewnego określonego typu; w języku potocznym chodzi tu o funkcje zdaniowe, składające się z trzech kolejnych części: ze zmiennej, ze słowa „jest” i zdanej nazwy x (s. 68, przyp. 45).

Jak pamiętamy, pojęcie spełniania stanowiło w zamyśle Tarskiego *explicatum* potocznego pojęcia oznaczania. Opisana w powyższym cytacie funkcja zdaniowa – nazwijmy ją identycznościową funkcją zdaniową – stanowi bezpośredni odpowiednik (*explicatum*) relacji oznaczania. A zatem zdaniu

Nazwa „Arystoteles” oznacza Arystotelesa

odpowiada

Arystoteles spełnia funkcję „ x_i jest Arystotelesem”.

Na początku zastanówmy się nad tym, kiedy obiekt a spełnia funkcję zdaniową „ x_i jest n ”. Odpowiedź wydaje się bardzo prosta: wtedy i tylko wtedy, gdy a jest n . Ograniczając się do przypadku, gdy metajęzyk obejmuje język przedmiotowy otrzymujemy:

(8.1) a spełnia funkcję zdaniową „ x_i jest n ” wtw, gdy a jest n ³²,

³² „Jest” w tym przypadku stanowi naturalnojęzykowy odpowiednik funktora identyczności.

gdzie x_i jest zmienną, zaś n nazwą własną tego obiektu, którą zastąpiliśmy mylący symbol x występujący w oryginalnym sformułowaniu Tarskiego³³.

Zauważmy, że od razu nasuwa się pewna wątpliwość dotycząca równoważności (8.1). Jej źródłem jest to, że po lewej stronie tej równoważności występuje zmienna (w funkcji zdaniowej ujętej w cudzysłów), zaś po prawej stronie owej zmiennej nie ma. Jeśli zatem a jest n , to a spełnia dowolną funkcję zdaniową postaci „(...) jest n ”, gdzie zamiast (...) występuje zmienna x z dowolnym indeksem. Na przykład Stagiryta spełnia każdą funkcję zdaniową „(...) jest Arystotelesem”:

(8.1ⁱ) Stagiryta spełnia funkcję zdaniową „ x_i jest Arystotelesem” wtw, gdy Stagiryta jest Arystotelesem.

Powodem tych wątpliwości jest to, że indeks przy zmiennej pełni bardzo istotną rolę. Można to intuicyjnie wyjaśnić w sposób następujący. Niech zmienna x_i odpowiada zaimkowi wskazującemu „on”, jeśli używamy go do wskazania osoby i . Oczywiście przy zaimkach nie ma indeksów, ale odpowiednikiem indeksu może być kontekst jego wypowiedzenia lub gest wskazania na dany obiekt. Czy w takim przypadku Stagiryta faktycznie będzie spełniał funkcję zdaniową „ x_i jest Arystotelesem”, gdy x_i na niego nie „wskazuje”, lecz wskazuje na Platona? Nie wypowiadając się w tej chwili na temat trafności oraz charakteru (8.1), zauważmy, że aby nie ignorować indeksu zmiennej należy założyć, że owo „wskazywanie” przez indeks zmiennej powinno gdzieś być zawarte w pojęciu spełniania. W przeciwnym przypadku indeks przy zmiennej byłby zbędny.

Wspomniane „wskazywanie” na obiekt można uściślić, odwołując się do przyporządkowania zmiennej x_i i -tego obiektu w ciągu C . Można zatem zastąpić pojęcie spełniania pojęciem spełniania zrelatywizowanym do ciągu C – oznaczmy je jako „spełnia _{C} ”. W przypadku gdy x_j „wskazuje na Platona”, tj. gdy Platon jest j -tą osobą w ciągu C , Stagiryta nie będzie spełniał _{C} funkcji zdaniowej „ x_j jest Arystotelesem” (choć Stagiryta jest Arystotelesem), gdyż x_j „nie wskazuje” na Stagirytę. Podobnie funkcji tej nie będzie spełniał _{C} Platon, bo chociaż x_j „wskazuje” na Platona, to Platon Arystotelesem nie jest. Zastąpmy zatem (8.1) następującą równoważnością:

(8.2) a spełnia _{C} funkcję zdaniową „ x_i jest n ” wtw, gdy n jest i -tym wyrazem C oraz a jest n .

³³ Symbol x najczęściej rezerwuje się dla zmiennych.

W szczególności dla rozpatrywanych przed chwilą przypadków otrzymamy:

(8.2ⁱ) Stagiryta spełnia_C funkcję zdaniową „ x_i jest Arystotelesem” wtw, gdy Stagiryta jest i -tym wyrazem ciągu C oraz Stagiryta jest Arystotelesem.

Ponieważ Arystoteles oraz Stagiryta to ta sama osoba, to Stagiryta spełnia_C funkcję zdaniową „ x_i jest Arystotelesem”, o ile jest i -tym wyrazem ciągu C . Podobnie:

(8.2ⁱⁱ) Platon spełnia_C funkcję zdaniową „ x_i jest Arystotelesem” wtw, gdy Platon jest i -tym wyrazem ciągu C oraz Platon jest Arystotelesem.

Oczywiście Platon spełniałby tę funkcję zdaniową, gdyby był i -tym wyrazem C oraz okazałoby się, że do tej pory historycy się mylili, bo Platon i Arystoteles to faktycznie ta sama osoba.

Na zakończenie powróćmy do równoważności (8.1). Po lewej stronie tej równoważności nie występuje żadne odwołanie do ciągu C . Jak jednak powiązać wprowadzone przez Tarskiego (zwykłe) pojęcie spełniania pojęciem spełniania_C? Najprostsze rozwiązanie polega na uznaniu, że obiekt a spełnia jednoargumentową funkcję zdaniową, gdy istnieje taki ciąg C , że obiekt a spełnia_C tę funkcję:

(8.3) Obiekt a spełnia funkcję zdaniową „ x_i jest n ” wtw, gdy istnieje ciąg C taki, że a spełnia_C funkcję zdaniową „ x_i jest n ”.

W takim przypadku odpowiednikiem (8.2ⁱ) będzie:

(8.3ⁱ) Stagiryta spełnia funkcję zdaniową „ x_i jest Arystotelesem” wtw, gdy istnieje ciąg C taki, że Stagiryta jest i -tym wyrazem ciągu C oraz Stagiryta jest Arystotelesem,

zaś odpowiednikiem (8.2ⁱⁱ) będzie:

(8.3ⁱⁱ) Platon spełnia funkcję zdaniową „ x_i jest Arystotelesem” wtw, gdy istnieje ciąg C taki, że Platon jest i -tym wyrazem ciągu C oraz Platon jest Arystotelesem.

Rozpatrzmy obecnie przykład, w którym język przedmiotowy jest językiem angielskim, zaś metajęzyk pozostaje językiem polskim. Wówczas (8.2) przybierze postać następującą:

(8.4) a spełnia_C funkcję zdaniową „ x_i is m ” wtw, gdy a jest i -tym wyrazem C oraz a jest n ,

gdzie m jest angielskim odpowiednikiem polskiej nazwy n . W szczególnym przypadku otrzymamy:

(8.4ⁱ) Stagiryta spełnia_C funkcję zdaniową „ x_i is Aristotle” wtw, gdy Stagiryta jest i -tym wyrazem ciągu C oraz Stagiryta jest Arystotelesem

oraz:

(8.4ⁱⁱ) Platon spełnia_C funkcję „ x_i is Aristotle” wtw, gdy Platon jest i -tym wyrazem ciągu C oraz Platon jest Arystotelesem.

Zauważmy, że prawe strony (8.2ⁱⁱ) i (8.4ⁱⁱ) oraz (8.2ⁱ) i (8.4ⁱ) są identyczne. Ponieważ jednak w przypadku (8.4) język przedmiotowy jest językiem angielskim, zaś metajęzyk językiem polskim, widać wyraźnie, że mamy tu do czynienia z zawartym *implicite* przekładem nazwy *Aristotle* na „Arystoteles” (oraz *is* na „jest”).

W jaki sposób dokonuje się przekładu z jednego języka na drugi? Najprostszy sposób polega na przyporządkowaniu wyrażeniom jednego języka ich synonimów w drugim języku, czyli przez odwołanie się do znaczeń. W przypadku nazw własnych sprawa się dodatkowo komplikuje, gdyż nie jest bynajmniej sprawą oczywistą, czy posiadają one w ogóle znaczenia. John Stuart Mill twierdził, że nazwy własne jedynie oznaczają indywidua (czyli denotują), lecz nie konotują³⁴. A zatem znaczeniami nazw własnych są po prostu oznaczane indywidua. Niestety prowadzi to do pewnych komplikacji. Po pierwsze identyczność oznaczanych indywiduów nie wystarczy do dokonania prawidłowego przekładu. *Aristotle* i „Stagiryta” posiadają te same odniesienia, ale trudno uznać, że polskim przekładem nazwy *Aristotle* jest „Stagiryta”. Jeżeli zatem przyznamy rację Millowi, trudno jest wyjaśnić na jakiej podstawie przełożyliśmy *Aristotle* na „Arystoteles”, nie zaś na „Stagiryta”, chociaż te ostatnie nazwy oznaczają ten sam obiekt. Sprawę komplikuje dodatkowo fakt, że do tej pory milcząco zakładałem istnienie indywiduów oznaczanych przez nazwy, np. istnienie Arystotelesa.

W związku z tym nasuwa się pytanie, czy w semantycznej definicji spełniania dla języka naturalnego mogą występować nazwy puste. Wyobraźmy sobie, że Arystoteles ze Stagiry nigdy się nie urodził, jego istnienie jest wymysłem Aleksandra Wielkiego, zaś *Corpus Aristotelicum* zostało na życzenie władcy napisane przez wziętych do niewoli uczonych z ateńskiej Akademii. W takim przypadku równoważności (8.2ⁱ), (8.2ⁱⁱ), (8.4ⁱ) i (8.4ⁱⁱ) byłyby bezsensowne bez dokonania daleko idących założeń ontologicznych na temat statusu przedmiotów nieistniejących, co kłóci się z deklarowaną przez Tarskiego neutralnością filozoficzną semantycznej definicji prawdy. Problem

³⁴ Nazwy konotujące to nazwy pospolite. W przypadku takich nazw ich znaczeniami są zbiory cech oznaczanych obiektów, ze względu na które to cechy, oraz jedynie te cechy, obiekty te są takimi nazwami oznaczane.

ten znika, jeśli zamiast przedmiotów będziemy posługiwać się ich nazwami, a zamiast ciągów przedmiotów – ciągami ich nazw. Innymi słowy, gdy semantyczna definicja spełniania „zrzuci semantyczne przebranie”³⁵.

9. W STRONĘ SEMANTYCZNEJ DEFINICJI PRAWDY W SFORMUŁOWANIU SYNTAKTYCZNYM

Powróćmy do równoważności (8.4ⁱ). W jaki sposób uniezależnić się od założenia o istnieniu *i*-tego wyrazu ciągu *C*? Jedną z możliwości polega na tym, aby uznać, że *i*-tym wyrazem ciągu *C* nie jest obiekt *n*, lecz jego nazwa, która będzie spełniała_C funkcję zdaniową. W takim przypadku równoważność (8.4) mogłaby przybrać postać:

(9.1) „*a*” spełnia_C funkcję zdaniową „*x_i is m*” wtw, gdy „*a*” jest *i*-tym wyrazem *C* oraz *b* jest *n*,

gdzie metajęzykowe nazwy *n* oraz *b* są przekładami nazw *m* i *a* z języka przedmiotowego. W szczególnym przypadku otrzymamy:

(9.1ⁱ) „*Stagirite*” spełnia_C funkcję zdaniową „*x_i is Aristotle*” wtw, gdy „*Stagirite*” jest *i*-tym wyrazem ciągu *C* oraz Stagiryta jest Arystotelesem.

Zauważmy, że przy takim sformułowaniu nie mamy żadnego odwołania do rzeczywistości, zaś zdanie „Stagiryta jest Arystotelesem” możemy zaakceptować lub odrzucić na podstawie dowolnych kryteriów zewnętrznych wobec definicji.

Równoważność (9.1ⁱ) sformułowana jest w kategoriach przekładu rozumianego syntaktycznie, tj. bez odwołania się do znaczeń lub odniesień rozumianych jako pewne byty: *Stagirite* przekładany jest na „Stagiryta”, *Aristotle* na „Arystoteles”, zaś *is* na „jest”. W tym przypadku znika problem istnienia lub sposobu istnienia obiektów, bo wyraz *Stagirite* występuje w języku angielskim. Zatem sama możliwość sformułowania równoważności (9.1ⁱ) gwarantuje istnienie odpowiedniego wyrazu ciągu, co zapewnia spełnienie postulowanego przez Tarskiego wymogu neutralności filozoficznej. Tak rozumiana relacja spełniania stanowiłaby *explicatum* relacji oznaczania, która zachodzi między nazwą obiektu (w języku przedmiotowym) a samym obiektem³⁶. W (9.1ⁱ) relacja spełniania zachodzi bowiem nie między obiektem

³⁵ „Zrzucenie semantycznego przebrania” prowadzi również do sporych problemów. Odpowiedzią na nie jest zaproponowana przez Davidsona semantyka nieodniesieniowa, która będzie przedmiotem dalszych rozdziałów.

³⁶ W tym przypadku między *Stagirite* a Arystotelesem.

a funkcją zdaniową – jak w (8.4ⁱ) – lecz między nazwą obiektu a funkcją zdaniową. Relacja spełniania jest zatem całkowicie określona w języku, co jest akceptowalne z punktu widzenia redukcjonistycznych deklaracji Tarskiego, który pragnął zredukować pojęcia semantyczne do pojęć fizykalistycznie dopuszczalnych³⁷. Przedstawiona tu semantyczna definicja spełniania w sformułowaniu syntaktycznym nawiązuje częściowo do interpretacji, którą sformułował Hartry Field w artykule *Tarski's Theory of Truth* (1972), lecz idzie od niej nieco dalej. Sam Field podaje przykład następujący:

Dla dowolnej nazwy N , obiekt a spełnia funkcję zdaniową „ x_i jest N ” wtw, gdy a jest Francją i N jest „Francja” lub ... lub gdy a jest Niemcami i N jest „Niemcy” (Field 1972, s. 365).

Jeżeli nazwa N będzie identyczna z nazwą „Arystoteles”, zaś a to Stagiryta, otrzymamy:

Obiekt a spełnia funkcję zdaniową „ x_i jest N ” wtw, gdy a jest Arystotelesem i nazwa „ N ” jest „Arystoteles” lub ... lub a jest Platonem i nazwa „ N ” jest „Platon”³⁸.

Ponieważ wyrażenia a oraz N są zmiennymi, to rezygnując z założenia, że metajęzyk obejmuje język przedmiotowy, otrzymujemy:

Stagiryta spełnia funkcję zdaniową „ x_i is *Aristotle*” wtw, gdy „Arystoteles” jest „Arystoteles” i Stagiryta jest Arystotelesem,

zaś po pominięciu oczywistej tożsamości „*Arystoteles* jest *Arystoteles*”:

Stagiryta spełnia funkcję zdaniową „ x_i is *Aristotle*” wtw, gdy Stagiryta jest Arystotelesem.

Jak łatwo zauważyć powyższa równoważność nie jest aksjomatem, lecz wynikającym z definicji S-zdaniem.

Z podanej przeze mnie wcześniej propozycji semantycznej definicji spełniania w sformułowaniu syntaktycznym wynika bardzo podobne S-zdanie:

„*Stagirite*” spełnia funkcję zdaniową „ x_i is *Aristotle*” wtw, gdy Stagiryta jest Arystotelesem,

które można wyprowadzić w sposób następujący:

³⁷ Problem fizykalistycznej redukcji pojęć w semantycznej definicji prawdy jest omawiany dokładnie w (Field 1972) oraz (Soames 1984).

³⁸ Zauważmy, że swej interpretacji semantycznej definicji prawdy Field nie odwołuje się do ciągów.

1. „*Stagirite*” spełnia funkcję zdaniową „ x_i is Aristotle”.
2. Istnieje ciąg C taki, że „*Stagirite*” spełnia _{C} funkcję zdaniową „ x_i is Aristotle” (z definicji spełniania).
3. Istnieje ciąg C taki, że „*Stagirite*” jest i -tym wyrazem C oraz Stagiryta jest Arystotelesem (z 9.1¹).
4. Stagiryta jest Arystotelesem.

Założmy obecnie, że:

1. Stagiryta jest Arystotelesem.
2. Istnieje ciąg C taki, że „*Stagirite*” jest i -tym wyrazem C (bo termin „*Stagirite*” występuje w słowniku języka angielskiego).
3. Istnieje ciąg C taki, że „*Stagirite*” jest i -tym wyrazem C oraz Stagiryta jest Arystotelesem (z 1 oraz 2).
4. Istnieje ciąg C taki, że „*Stagirite*” spełnia _{C} funkcję zdaniową „ x_i is Aristotle” (z 9.1¹).
5. „*Stagirite*” spełnia funkcję zdaniową „ x_i is Aristotle” (z definicji spełniania).

Wystąpienie nazwy *Stagirite* po lewej stronie równoważności (9.1¹) oraz po lewej stronie wyprowadzonego wyżej S-zdania może być zastąpione jakimkolwiek strukturalnym opisem ten nazwy, np. „trzeci wyraz na 10-tej stronie dzieła X ”. A zatem S-zdania w zaproponowanym przeze mnie ujęciu zbudowane są analogicznie do T-zdań.

Dłaczę jednak zestawilem proponowaną przeze mnie semantyczną definicję spełniania w sformułowaniu syntaktycznym z propozycją Fielda, który o przedmiotach pisze *explicite*? Przede wszystkim może się zdarzyć, że nie znamy angielskiego odpowiednika polskiej (czyli metajęzykowej) nazwy „Stagiryta”. W tym przypadku i -tym wyrazem C może być „nazwa, której tłumaczeniem na metajęzyk jest *Stagiryta*”. Ponieważ używanie tego typu opisu może być kłopotliwe, zrozumiałe jest, że zamiast przytoczenia długiego opisu, Field użył odpowiedniej metajęzykowej nazwy bez cudzysłowu. Uproszczenie to nie jest groźne dla semantycznej definicji spełniania w sformułowaniu syntaktycznym, o ile pamięta się wyjściową postać aksjomatów oraz T-zdań.

Zarówno ujęcie semantycznej definicji prawdy prezentowane w tym paragrafie, jak i wspomniana wcześniej interpretacja Fielda, nawiązują bezpośrednio do tych słów Tarskiego, w których deklarował zbudowanie semantycznej definicji prawdy w sformułowaniu syntaktycznym. O ile proponowana

przeze mnie interpretacja poglądów Tarskiego i Fielda jest trafna, to wybrali oni po prostu skrótowy sposób pisania, który polegał na pominięciu cudzysłowu w przypadku nazw metajęzykowych.

S-zdania można potraktować jako definicje cząstkowe pojęcia spełniania, zaś na samą definicję spełniania składają się odpowiednie aksjomaty, np. kandydatem na aksjomat teorii spełniania jest (9.1¹). Definicja ta wyznacza zakres tego pojęcia, którym jest zbiór par, na które składają się nazwy i funkcje zdaniowe, np. *Stagirite* i x_i is *Aristotle*. Same S-zdania (podobnie jak T-zdania w przypadku semantycznej definicji prawdy) pozbawione są wielu „narzędzi semantycznych” występujących w aksjomatach. Dlatego też podanie samego S-zdania, co uczynił Field, sprzyja uproszczeniom³⁹. Przedstawione tu uproszczenie, które polega na pominięciu cudzysłowu, może prowadzić do wielu nieporozumień. Nie wiadomo bowiem czy jest pewna stylizacja definicji budowanej w kategoriach syntaktycznych, czy też faktycznie w definicji posługujemy się ciągami obiektów, zakładając jedynie istnienie przekładu z języka przedmiotowego na metajęzyk.

Na zakończenie rozpatrzmy przykład, który ukaże różnicę we wspomnianych wyżej „narzędziach semantycznych” między definicją spełniania w sformułowaniu syntaktycznym, czyli ciągów nazw obiektów, a definicją spełniania w kategoriach ciągów obiektów, czyli sformułowaniu semantycznym. Jak wiadomo Arystoteles bywał określany za pomocą nazwy „Filozof”, która faktycznie funkcjonowała jako nazwa własna. Wyobraźmy jednak sobie – wbrew powszechnie uznanym i dobrze udokumentowanym faktom historycznym – że nazwa „Filozof” została użyta tylko w jednym zachowanym do dzisiaj średniowiecznym dziele i została odniesiona niejasno do jednego z wielkich filozofów starożytności. Załóżmy, że angielski historyk *A* uznał na podstawie badanego tekstu, że nazwa *Philosopher* oznaczała Arystotelesa. A zatem na podstawie angielskojęzycznej publikacji historyka *A* oraz równoważności będącej wariantem (8.4)⁴⁰ polskojęzyczny (lecz znający język angielski) zwolennik historyka *A* uznaje, że Filozof spełnia_C funkcję zdaniową x_i is *Aristotle*, gdzie *i*-tym wyrazem ciągu *C* jest Filozof.

Wyobraźmy sobie obecnie, że inny angielskojęzyczny historyk *B* dowodzi, że wspomniany Filozof był faktycznie Platonem. W takim przypadku

³⁹ Pomimo poczynionych tutaj uwag uważam, że reprezentowane przez Fielda podejście zarówno do semantycznej definicji spełniania, jak i semantycznej definicji prawdy nie odbiega zasadniczo od mojego ujęcia.

⁴⁰ Zwolennik historyka *A* buduje zatem semantyczną definicję spełniania w kategoriach semantycznych, w której występuje następujący aksjomat spełniania:

(9.4¹) Filozof spełnia_C funkcję zdaniową „ x_i is *Aristotle*” wtw, gdy Filozof jest *i*-tym wyrazem ciągu *C* oraz Filozof jest Arystotelesem.

polskojęzyczny zwolennik historyka *B* stwierdza na podstawie analogicznej równoważności⁴¹, że Filozof spełnia funkcję zdaniową x_i is *Plato*, a w konsekwencji również spełnia_C x_i is *Plato*. Oczywiście dla zwolennika historyka *A* *i*-tym wyrazem ciągu *C* będzie Arystoteles, zaś dla zwolennika *B* *i*-tym wyrazem ciągu *C* będzie Platon.

Nie ulega wątpliwości⁴², że przynajmniej jeden z historyków myli się co do tożsamości Filozofa. Ponadto okazuje się, że każdy z nich posługuje się inną semantyczną definicją spełniania w sformułowaniu semantycznym, gdyż posługuje się różnymi ciągami, oznaczając tym samym mianem „Filozof” różne obiekty będące *i*-tymi wyrazami owych ciągów. Różność owych obiektów suponuje jakiś sposób ich istnienia, co prowadzi – w przypadku sformułowania semantycznego – do wspomnianych już trudności związanych z deklaracją Tarskiego o neutralności filozoficznej. Ostatecznie z konkurencyjnych teorii spełniania w prezentowanej tu wersji semantycznej wynikają odpowiednio następujące S-zdania, w których nie będzie już odniesień do ciągów obiektów:

Filozof spełnia funkcję zdaniową „ x_i is *Aristotle*” wtw, gdy Filozof jest Arystotelesem.

Filozof spełnia funkcję zdaniową „ x_i is *Plato*” wtw, gdy Filozof jest Platonem.

W przypadku definicji spełniania w sformułowaniu syntaktycznym zarówno zwolennik historyka *A*, jak i zwolennik historyka *B* będą posługiwać się dokładnie tym samym ciągiem nazw, co nie rodzi jakichkolwiek problemów związanych z istnieniem lub nieistnieniem obiektów oznaczanych przez owe nazwy:

(9.1^{iv}) Nazwa „*Philosopher*” spełnia_C funkcję zdaniową „ x_i is *Aristotle*” wtw, gdy „*Philosopher*” jest *i*-tym wyrazem ciągu *C* oraz Filozof jest Arystotelesem.

(9.1^v) Nazwa „*Philosopher*” spełnia_C funkcję zdaniową „ x_i is *Plato*” wtw, gdy „*Philosopher*” jest *i*-tym wyrazem ciągu *C* oraz Filozof jest Platonem⁴³.

⁴¹ Odpowiedni aksjomat spełniania miałby postać następującą:

(9.4ⁱⁱ) Filozof spełnia_C funkcję zdaniową „ x_i is *Plato*” wtw, gdy Filozof jest *i*-tym wyrazem *C* oraz Filozof jest Platonem.

⁴² Zakładamy tutaj zgodnie z powszechną i dobrze udokumentowaną wiedzą, że Arystoteles i Platon byli różnymi osobami.

⁴³ Z definicji zwolennika historyka *A* wynika S-zdanie:

„*Philosopher*” spełnia funkcję zdaniową „ x_i is *Aristotle*” wtw, gdy Filozof jest Arystotelesem, zaś z teorii zwolennika historyka *B*:

„*Philosopher*” spełnia funkcję zdaniową „ x_i is *Plato*” wtw, gdy Filozof jest Platonem.

W dalszym ciągu będę odwoływał się do definicji spełniania w sformułowaniu syntaktycznym. Pamiętać jednak należy, że w przypadku prostych aksjomatów definicji prawdy, o których będzie mowa w kolejnym paragrafie, wystarczy pominąć cudzysłów wokół nazwy metajęzykowej, aby uzyskać wersję definicji w sformułowaniu semantycznym. Aby wyjaśnić dlaczego Tarski – wbrew pierwotnym deklaracjom – wybrał ostatecznie drugie sformułowanie, powrócimy w dalszych paragrafach tego rozdziału.

10. UOGÓLNIENIE DEFINICJI SPEŁNIANIA DO DEFINICJI PRAWDY

Do tej pory zajmowaliśmy się – zgodnie z początkowymi sugestiami Tarskiego – jedynie funkcjami zdaniowymi postaci „ x_i jest n ” lub ich angielskimi odpowiednikami, zaś równoważność postaci (8.4) miała stanowić *explicatum* pojęcia denotowania. Jednak pojęcie spełniania poza „identycznościowymi” funkcjami zdaniowymi postaci „ x_i jest n ”, obejmuje również inne funkcje zdaniowe. W takiej też szerszej postaci wykorzystuje się pojęcie spełniania w semantycznej definicji prawdy.

Aby wyjaśnić na czym polega szerzej rozumiane pojęcie spełniania, rozpatrzmy jednoargumentową funkcję zdaniową „ x_i jest człowiekiem”. Funkcja ta nie jest funkcją tożsamościową, zaś „jest” pełni w niej zupełnie inną rolę niż „jest” w przypadku zdań tożsamościowych. Zaczniemy zatem od zmodyfikowanej definicji spełniania w sformułowaniu syntaktycznym dla dowolnej jednoargumentowej funkcji zdaniowej:

(10.1) „ a ” spełnia funkcję zdaniową $\varphi(x_i)$ wtw, gdy istnieje ciąg C taki, że „ a ” spełnia _{C} $\varphi(x_i)$

oraz:

(10.2) „ a ” spełnia _{C} $\varphi(x_i)$ wtw, gdy „ a ” jest i -tym wyrazem ciągu C oraz $\psi(b)$,

gdzie ψ oraz b są odpowiednio przekładami φ oraz a z języka przedmiotowego na metajęzyk. W rezultacie otrzymamy schemat aksjomatu spełniania dla funkcji jednoargumentowej:

(10.3) „ a ” spełnia funkcję zdaniową „ $\varphi(x_i)$ ” wtw, gdy istnieje ciąg C taki, że „ a ” jest i -tym wyrazem ciągu C oraz $\psi(b)$.

Niech $\varphi(x_i)$ będzie funkcją zdaniową x_i is *human*. Otrzymamy wówczas następujący aksjomat semantycznej definicji spełniania dla fragmentu języka angielskiego obejmującego predykat *is human*:

(10.3¹) „*Aristotle*” spełnia funkcję zdaniową „ x_i is human” wtw, gdy istnieje ciąg C taki, że „*Aristotle*” jest i -tym wyrazem C oraz Arystoteles jest człowiekiem.

Wyprowadźmy obecnie S-zdanie:

„*Aristotle*” spełnia „ x_i is human” wtw, gdy Arystoteles jest człowiekiem.

Założmy, że:

1. „*Aristotle*” spełnia „ x_i is human”.
2. Istnieje C taki, że „*Aristotle*” jest i -tym wyrazem C oraz Arystoteles jest człowiekiem (z 10.3¹).
3. Arystoteles jest człowiekiem.

Ponieważ istnienie odpowiedniego ciągu jest zagwarantowane obecnością nazwy *Aristotle*, to wynikanie zachodzi również w drugą stronę.

Wyprowadzone S-zdanie pozwala zaliczyć parę złożoną z nazwy *Aristotle* i funkcji zdaniowej x_i is human do zakresu predykatu „spełnia” dla funkcji jednoargumentowych w języku angielskim jako języku przedmiotowym.

Wyprowadźmy obecnie (z niewielką modyfikacją) S-zdanie dla podanej przez Tarskiego w (1944) dwuargumentowej funkcji zdaniowej x_i is a brother of x_j :

„*John*” i „*Peter*” spełniają funkcję zdaniową „ x_i is a brother of x_j ” wtw, gdy Jan jest bratem Piotra.

Ogólny schemat definicji spełniania dla funkcji dwuargumentowych ma postać następującą:

(10.4) „ a_1 ” i „ a_2 ” spełniają funkcję zdaniową $\varphi(x_i, x_j)$ wtw, gdy istnieje ciąg C taki, że „ a_1 ” jest i -tym wyrazem ciągu C , „ a_2 ” jest j -tym wyrazem ciągu C oraz $\psi(b_1, b_2)$.

Jak poprzednio ψ , b_1 i b_2 są odpowiednio przekładami φ , a_1 oraz a_2 z języka przedmiotowego na metajęzyk. Odpowiedni aksjomat spełniania dla funkcji zdaniowej x_i is a brother of x_j ma postać następującą:

(10.4¹) „*John*” i „*Peter*” spełniają funkcję zdaniową „ x_i is a brother of x_j ” wtw, gdy istnieje ciąg C taki, że „*John*” jest i -tym wyrazem ciągu C , „*Peter*” jest j -tym wyrazem ciągu C oraz Jan jest bratem Piotra.

Wyprowadzenie podanego wyżej S-zdania jest następujące:

1. „*John*” i „*Peter*” spełniają funkcję zdaniową „ x_i is a brother of x_j ”.

2. Istnieje ciąg C , że „*John*” jest i -tym wyrazem C , „*Peter*” jest j -tym wyrazem C oraz Jan jest bratem Piotra (z 10.4ⁱ).

3. Jan jest bratem Piotra.

Podobnie jak w poprzednim przypadku istnienie odpowiedniego ciągu jest zagwarantowane obecnością nazw *John* i *Peter* w języku przedmiotowym, a zatem wynikanie zachodzi również w drugą stronę. Wyprowadzone S-zdanie pozwala zaliczyć trójkę złożoną z nazw *John* i *Peter* oraz dwuargumentowej funkcji zdaniowej x_i is a brother of x_j do zakresu metajęzykowego predykatu „spełnia” dla funkcji dwuargumentowych w języku angielskim jako języku przedmiotowym.

Podany schemat aksjomatu dla funkcji dwuargumentowej łatwo jest przeformułować dla funkcji wieloargumentowych – trójargumentowych, czteroargumentowych itd. Najważniejsze dla naszych celów jest jednak zbudowanie aksjomatu spełniania dla zdań, czyli funkcji zeroargumentowych:

Kiedy już mamy ogólną definicję spełniania, zauważmy, że stosuje się ona automatycznie również do tych szczególnych funkcji zdaniowych, które nie zawierają żadnych zmiennych wolnych, tj. zdań. Okazuje się, że dla zdań możliwe są tylko dwa przypadki: albo zdanie spełnione jest przez wszystkie przedmioty, albo nie jest spełnione przez żaden przedmiot. Dochodzimy więc do definicji prawdy i fałszu mówiąc po prostu, że zdanie jest prawdziwe, gdy jest spełnione przez wszystkie przedmioty, fałszywe zaś – w przeciwnym przypadku (1944, s. 250).

Sformułujmy zatem schemat aksjomatu spełniania dla zdania:

„ a ” spełnia zdanie „ Z ” wtw, gdy istnieje C taki, że Z .

Zauważmy, że nie ma tutaj znaczenia czy nazwa a występuje w zdaniu Z , czy też nie. Ponieważ w konsekwencji nie wymaga się również, aby nazwa a była jakimkolwiek wyrazem ciągu C , można pominąć odwołanie do C :

„ a ” spełnia zdanie „ Z ” wtw, gdy Z .

Jak łatwo zauważyć, taką samą uwagę – jak w przytoczonym wyżej cytacie – można poczynić na temat par nazw, trójek nazw itd. Nie ma bowiem powodu, aby zdanie miało być spełniane jedynie przez nazwy, nie zaś ich pary czy trójki, np.

„ a ”, „ b ” i „ c ” spełniają zdanie „ Z ” wtw, gdy Z .

W tym momencie jesteśmy zatem o krok od wprowadzenia pojęcia prawdy, lecz natrafiamy na pewną trudność techniczną. Konsekwencją jej rozwiązania jest to, że w aksjomatach oryginalnego sformułowania semantycznej

definicji prawdy nie mówi się o spełnianiu funkcji zdaniowych przez obiekty (lub nazwy obiektów), lecz przez ciągi obiektów (lub nazw obiektów).

Na wspomnianą wyżej trudność zwrócił uwagę Tarski w jednym z przypisów w *Semantycznej koncepcji prawdy*, który pozwolę sobie zacytować w całości:

Przy realizacji tego pomysłu [zdanie jest prawdziwe, gdy jest spełnione przez wszystkie przedmioty] wyłaniają się pewne trudności natury technicznej. Funkcja zdaniowa może zawierać dowolną liczbę zmiennych wolnych, a wraz z tą liczbą zmienia się logiczna natura pojęcia spełniania. I tak, pojęcie to zastosowane do funkcji z jedną zmienną jest dwuargumentową relacją między tymi funkcjami a pojedynczymi przedmiotami; zastosowane do funkcji z dwiema zmiennymi staje się trójargumentową relacją między funkcjami a parami przedmiotów – i tak dalej. Dlatego, ściśle mówiąc, mamy do czynienia nie z jednym pojęciem spełniania, ale z nieskończenie wieloma pojęciami; okazuje się, że pojęć tych nie można zdefiniować niezależnie od siebie, lecz trzeba je wszystkie wprowadzić jednocześnie.

Aby tę trudność pokonać, korzystamy z matematycznego pojęcia ciągu nieskończonego [...] Przyjmujemy umowę, że spełnianie nie jest wieloargumentową relacją między funkcjami zdaniowymi i nieokreśloną liczbą przedmiotów, lecz dwuargumentową relacją między funkcjami i ciągami przedmiotów. Przy tym założeniu sformułowanie ogólnej i ścisłej definicji spełniania nie przedstawia już trudności, a zdanie prawdziwe można teraz zdefiniować jako zdanie spełnione przez każdy ciąg (s. 250–251, przyp. 15)⁴⁴.

Trudność przedstawiona przez Tarskiego ma faktycznie jedynie charakter notacyjny i jest w pewnym sensie odziedziczona po próbie zbudowania semantycznej definicji spełniana oddzielnie dla funkcji jednoargumentowych spełnianych przez obiekty (lub nazwy obiektów), dla funkcji dwuargumentowych przez pary obiektów (pary nazw obiektów) itd. W szczególnym przypadku problem polega na tym, że skoro mamy aksjomat spełniania funkcji jednoargumentowej, np. „ x_i jest człowiekiem” przez nazwę „Arystoteles”, to uznajemy, że „Arystoteles” spełnia również zdanie „Arystoteles jest człowiekiem”. Podobnie, jeśli mamy aksjomat spełniania funkcji dwuargumentowej „ x_i jest bratem x_j ” przez Parę nazw „Jan” i „Piotr”, to uznajemy, że nazwy „Jan” i „Piotr” spełniają również zdanie „Jan jest bratem Piotra”. W ogólnym przypadku zdania te spełnione są przez dowolne nazwy i pary nazw.

⁴⁴ Stwierdzenie, że zdanie prawdziwe jest spełniane przez każdy ciąg, nie jest definicją prawdy, lecz jedynie aksjomatem semantycznej definicji prawdy. Aby bowiem określić zakres predykatu „prawdziwy” w języku przedmiotowym, należy podać wszystkie aksjomaty semantycznej definicji prawdy dla tego języka. Jest to jedna z drobnych nieścisłości pojawiających się niekiedy u Tarskiego.

Nic nie stoi również na przeszkodzie, abyśmy przeformułowali aksjomat spełniania przez nazwę funkcji jednoargumentowej, np. „ x_i jest człowiekiem”, do aksjomatu spełniania tej funkcji przez pary nazw:

„Arystoteles” i „Platon” spełniają funkcję „ x_i jest człowiekiem” wtw, gdy istnieje C taki, że „Arystoteles” jest i -tym wyrazem C i Arystoteles jest człowiekiem,

z którego wynika następujące S-zdanie:

„Arystoteles” i „Platon” spełniają funkcję „ x_i jest człowiekiem” wtw, gdy Arystoteles jest człowiekiem.

Oczywiście można podać w naszej teorii spełniania inny aksjomat:

„Arystoteles” i „Platon” spełniają funkcję „ x_i jest człowiekiem” wtw, gdy istnieje C taki, że „Platon” jest i -tym wyrazem C i Platon jest człowiekiem.

Podobne rozumowanie można rozszerzyć na trójki nazw, czwórki nazw, a w końcu na nieskończone ciągi nazw spełniających dowolne funkcje zdaniowe. Na przykład dla funkcji jednoargumentowej „ x_i jest człowiekiem” odpowiedni aksjomat miałby postać następującą:

Ciąg C nazw spełnia funkcję „ x_i jest człowiekiem” wtw, gdy „ n ” jest i -tym wyrazem C i n jest człowiekiem.

To samo rozumowanie można w analogiczny sposób rozszerzyć na funkcje dwuargumentowe, funkcje trójargumentowe itd., oraz na zdania, czyli na funkcje zeroargumentowe. Oczywiście to ostatnie ma charakter trywialny:

Ciąg nazw C spełnia zdanie „ Z ” wtw, gdy Z .

Aby rozszerzyć semantyczną definicję spełniania do semantycznej definicji prawdy, wystarczy wprowadzić aksjomat prawdziwości, skąd natychmiast otrzymujemy schemat T-zdań wynikających z teorii prawdy.

11. PROBLEMY SEMANTYCZNEJ DEFINICJI PRAWDY W SFORMUŁOWANIU SYNTAKTYCZNYM

Przejsie od spełniania funkcji zdaniowej przez nazwy do spełniania przez nieskończone ciągi nazw pozwala nie tylko na sformułowanie aksjomatu prawdziwości, co prowadzi do nadbudowania semantycznej definicji prawdy nad semantyczną definicją spełniania, ale również pozwala na ograniczenie ilości aksjomatów. Jest to możliwe, gdyż spełnianie formuł (tj. funkcji zdaniowych lub zdań) przez ciągi umożliwia jednolite potraktowanie stałych oraz zmiennych, czyli termów. W tym celu należy wprowadzić aksjomaty odnie-

sienia, które określają, jakie nazwy ciągu C przyporządkowuje zmiennym oraz stałym jednostkowym:

(11.1) Dla każdego ciągu C , C przyporządkowuje zmiennej x_i swój i -ty wyraz.

(11.2) Dla każdego ciągu C , C przyporządkowuje nazwie „ a ” tę samą nazwę, tj. nazwę „ a ”.

Jak widać (11.2) jest aksjomatem trywialnym. Z kolei aksjomat spełniania formuły powstałej z funkcji zdaniowej k -argumentowej oraz ciągu k termów można przedstawić w następującej postaci:

(11.3) Ciąg C spełnia funkcję zdaniową „ $\varphi(t_i, t_j, \dots, t_n)$ ” wtw, gdy ciąg C przyporządkowuje odpowiednio termowi „ t_i ” nazwę „ a_i ”, termowi „ t_j ” nazwę „ a_j ”, ..., termowi „ t_n ” nazwę „ a_n ” oraz $\psi(b_i, b_j, \dots, b_n)$,

gdzie b_i, b_j, \dots, b_n są tłumaczeniami na metajęzyk nazw a_i, a_j, \dots, a_n , zaś ψ jest tłumaczeniem na metajęzyk funkcji zdaniowej φ , obejmującym przekład wszystkich występujących w niej predykatów oraz nazw jednostkowych.

Zauważmy, że (11.2) oraz (11.3) faktycznie nie są aksjomatami, lecz schematami aksjomatów, pod które (przy założeniu, że słownik języka przedmiotowego jest skończony) podpada skończenie wiele aksjomatów. Zatem przykładowymi aksjomatami byłyby:

(11.2ⁱ) Dla każdego ciągu C , C przyporządkowuje nazwie „*John*” tę samą nazwę, tj. „*John*”.

(11.2ⁱⁱ) Dla każdego ciągu C , C przyporządkowuje „*Peter*” tę samą nazwę, tj. „*Peter*”.

(11.3ⁱ) Ciąg C spełnia formułę „ t_i is a brother of t_j ” wtw, gdy C przyporządkowuje termowi „ t_i ” nazwę „*John*”, termowi „ t_j ” nazwę „*Peter*” oraz Jan jest bratem Piotra.

Dodatkowo można wprowadzić aksjomaty spełniania dla podstawowych stałych logicznych:

(11.4) Ciąg C spełnia formułę „ $\sim\varphi$ ” wtw, gdy C nie spełnia formuły „ φ ”.

(11.5) Ciąg C spełnia formułę „ $\varphi \wedge \psi$ ” wtw, gdy C spełnia formułę „ φ ” oraz C spełnia formułę „ ψ ”.

(11.6) Ciąg C spełnia formułę „ $\varphi \vee \psi$ ” wtw, gdy C spełnia formułę „ φ ” lub C spełnia formułę „ ψ ”.

(11.7) Ciąg C spełnia formułę „ $\exists x_i \varphi$ ” wtw, gdy dla pewnego ciągu C' , różniącego się od C co najwyżej i -tym wyrazem, C' spełnia formułę „ φ ”.

(11.8) Ciąg C spełnia formułę „ $\forall x_i \varphi$ ” wtw, gdy dla każdego ciągu C' różniącego się od C co najwyżej i -tym wyrazem C' spełnia formułę „ φ ”.

Zauważmy, że w aksjomatach (11.4)–(11.8) symbole logiczne języka przedmiotowego można zastąpić odpowiednio ich naturalnojęzykowymi odpowiednikami *it is not the case that*⁴⁵, *and*, *or*, *some ... are* oraz *all ... are*.

Aby uzupełnić semantyczną definicję spełniania do semantycznej definicji prawdy, wystarczy dołączyć do powyższych aksjomatów aksjomat prawdziwości (10.5). Wyprowadźmy obecnie T-zdanie postaci:

Zdanie „*John is a brother of Peter*” jest prawdziwe wtw, gdy Jan jest bratem Piotra.

Dowód:

1. Zdanie „*John is a brother of Peter*” jest prawdziwe.
2. Dla każdego ciągu C , C spełnia zdanie „*John is a brother of Peter*” (z 10.5).
3. Dowolny wybrany ciąg C spełnia zdanie „*John is a brother of Peter*”.
4. C przyporządkowuje nazwie „*John*” nazwę „*John*”, nazwie „*Peter*” nazwę „*Peter*” oraz Jan jest bratem Piotra (z 11.2ⁱ, 11.2ⁱⁱ i 11.3ⁱ).
5. Jan jest bratem Piotra.

Założmy obecnie, że:

1. Jan jest bratem Piotra.
2. Dla każdego C , C przyporządkowuje nazwie „*John*” nazwę „*John*”, nazwie „*Peter*” nazwę „*Peter*” (z 11.2ⁱ oraz 11.2ⁱⁱ).
3. Dla każdego C , C przyporządkowuje nazwie „*John*” nazwę „*John*”, nazwie „*Peter*” nazwę „*Peter*” oraz Jan jest bratem Piotra (z 1 i 2).
4. Dla każdego ciągu C , C spełnia zdanie „*John is a brother of Peter*” (z 11.3ⁱ).
5. Zdanie „*John is a brother of Peter*” jest prawdziwe (z 10.5).

Dla podanych do tej pory przykładów nie było żadnego problemu z przeformułowaniem aksjomatów oraz dowodów T-zdań lub S-zdań w sformułowaniu syntaktycznym na ich odpowiedniki w sformułowaniu semantycznym. Na zakończenie wyprowadźmy jednak T-zdanie, w którym występuje kwantyfikator:

⁴⁵ Negacja zdaniowa może być realizowana w języku naturalnym na wiele sposobów, dlatego też nie można po prostu powiedzieć, że jej realizacją jest partykuła *not*.

Zdanie „ $\exists x_i(John\ is\ a\ brother\ of\ x_i)$ ” jest prawdziwe wtw, gdy Jan jest czyimś bratem.

Dowód:

1. Zdanie „ $\exists x_i(John\ is\ a\ brother\ of\ x_i)$ ” jest prawdziwe.
2. Dla każdego C , C spełnia „ $\exists x_i(John\ is\ a\ brother\ of\ x_i)$ ” (z 10.5).
3. C spełnia „ $\exists x_i(John\ is\ a\ brother\ of\ x_i)$ ” (gdzie C jest dowolnym ciągiem).
4. Istnieje C' różniące się od C co najwyżej i -tym wyrazem taki, że C' spełnia „ $(John\ is\ a\ brother\ of\ x_i)$ ” (z 11.7).
5. i -tym wyrazem C' jest „ a ” oraz Jan jest bratem b , gdzie „ b ” jest przekładem „ a ” na metajęzyk (z 11.3¹).
6. Dla pewnej nazwy „ a ”, Jan jest bratem b , gdzie „ b ” jest przekładem „ a ” na metajęzyk (z 5).
7. Jan jest czyimś bratem (parafraza wiersza 6).

Założmy obecnie, że:

1. Jan jest czyimś bratem.
2. Dla pewnej nazwy „ a ”, Jan jest bratem b , gdzie „ b ” jest przekładem „ a ” na metajęzyk (parafraza wiersza 1).
3. Dla pewnego C' , którego i -tym wyrazem jest „ a ” Jan jest bratem b , gdzie „ b ” jest przekładem „ a ” na metajęzyk.
4. Dla każdego C istnieje C' różniące się od C co najwyżej i -tym wyrazem taki, że i -tym wyrazem C' jest „ a ” oraz Jan jest bratem b , gdzie „ b ” jest przekładem „ a ” na metajęzyk.
5. Dla każdego C istnieje C' różniące się od C co najwyżej i -tym wyrazem taki, że C' spełnia „ $(John\ is\ a\ brother\ of\ x_i)$ ” (z 11.3¹).
6. Dla każdego C , C spełnia „ $\exists x_i(John\ is\ a\ brother\ of\ x_i)$ ” (z 11.7).
7. Zdanie „ $\exists x_i(John\ is\ a\ brother\ of\ x_i)$ ” jest prawdziwe.

Do przedstawionego wyżej dowodu należy zachować dużą rezerwę. Czyni się w nim bowiem nieuprawnione założenie, że zarówno język przedmiotowy, jak i metajęzyk zawierają wszystkie nazwy własne, które mógłby

nosić brat Jana⁴⁶. Stanowi to niepożądaną, lecz nieuniknioną konsekwencję aksjomatów dla formuł z kwantyfikatorami. Wyjaśnia to jednocześnie, dlaczego semantyczna definicja prawdy (lub definicja spełniania) nie może być wyrażana w sformułowaniu syntaktycznym. Zagadnieniu temu poświęcony jest kolejny paragraf.

12. SEMANTYCZNA DEFINICJA PRAWDY W SFORMUŁOWANIU SEMANTYCZNYM

Przedstawiona w ostatnich paragrafach syntaktyczna wersja semantycznej definicji prawdy (oraz semantycznej definicji spełniania) miała spełnić dwa zadania. Po pierwsze miała podkreślić syntaktyczny charakter intuicji Tarskiego, a po drugie zachować filozoficzną neutralność w kwestii istnienia lub nieistnienia pewnych obiektów. Niestety okazuje się, że podejście to jest trudne do utrzymania z innych względów. Problemy pojawiają się w ostatnim dowodzie T-zdania, w którym zakłada się w sposób całkowicie nieuprawniony istnienie pewnych nazw w metajęzyku. Przedstawiona wersja semantycznej definicji prawdy spełniałaby swoje zadanie, gdyby język przedmiotowy zawierał tyle nazw własnych, ile nam potrzeba, tj. gdyby każdy obiekt (realnie istniejący lub jedynie pomyślany) miał swą nazwę w metajęzyku.

Na analogiczny problem zwrócił uwagę Tarski w pracy *O pojęciu wynikania logicznego* (1936a). Przed wprowadzeniem pojęcia wynikania logicznego Tarski odwołuje się do wspomnianej wcześniej idei „naukowej semantyki” (s. 194), licząc na to, że dostarczy ona środków niezbędnych do zdefiniowania tego pojęcia, rozumianego jako konieczna relacja dziedziczenia prawdy. Charakteryzując intuicje związane z relacją wynikania logicznego Tarski formułuje warunek konieczny na to, aby zdanie X wynikało z klasy zdań K :

(W) Jeśli terminy stałe, nie będące terminami ogólnologicznymi, zastąpimy odpowiednio w zdaniach klasy K i w zdaniu X przez dowolne inne terminy stałe (przy czym terminy równokształtne zastępować wszędzie będziemy stałymi równokształtnymi) i w ten sposób uzyskamy nową klasę zdań K' i nowe zdanie X' , to zdanie X' musi być prawdziwe, jeśli tylko wszystkie zdania klasy K' będą prawdziwe (s. 195).

⁴⁶ W przypadku języka przedmiotowego nieistniejącą nazwę własną ewentualnie można zastąpić przez „nazwa, której przekładem na metajęzyk jest...”. W przypadku metajęzyka musimy dysponować nazwą, której użyjemy w aksjomatach oraz po prawej stronie T-zdania.

gdzie przez terminy ogólnologiczne rozumie się po prostu stałe logiczne: koniunkcję, alternatywę, negację, kwantyfikatory oraz predykat równości.

Przedstawione wyżej podejście do relacji wynikania ma tę samą wadę, na którą wskazałem poprzednio dla semantycznej definicji prawdy w sformułowaniu syntaktycznym. Polegała one na tym, że w języku przedmiotowym może nie być dostatecznie wielu nazw, które występowałyby jako wyrazy ciągów:

Nie trudno pokazać (...), że mimo spełnienia warunku (W) zdanie *X* nie musi bynajmniej wynikać – z punktu widzenia potocznych intuicji – ze zdań klasy *K*; warunek ten może być bowiem spełniony li tylko wskutek niedostatecznego zasobu terminów stałych w języku, którego dotyczą nasze rozważania. Wtedy tylko można by uznać warunek (W) za wystarczający na to, by zdanie *X* wynikało formalnie ze zdań klasy *K*, gdyby założyć, że rozważany język zawiera w sobie nazwy wszystkich możliwych przedmiotów (s. 196).

Aby temu zaradzić w przypadku konsekwencji logicznej, Tarski odwołuje się do znanego nam już pojęcia spełniania funkcji zdaniowej przez ciąg przedmiotów. Następnie proponuje, aby każde zdanie pewnej klasy zdań *L* zastąpić funkcją zdaniową powstałą przez zastąpienie stałych pozalogicznych występujących w zdaniu symbolami zmiennych, otrzymując w ten sposób klasę funkcji zdaniowych *L'*. Ostatecznie modelem klasy zdań *L* będzie dowolny ciąg przedmiotów, który spełnia każdą funkcję zdaniową klasy *L'* (s. 197), zaś wniosek wynika z przesłanek, gdy każdy model zbioru przesłanek jest jednocześnie modelem wniosku. Pojęcie modelu ewoluowało u Tarskiego i doprowadziło go ostatecznie do przeformułowania semantycznej definicji prawdy jako prawdy w języku na definicję prawdy w modelu, rozumianym jednak inaczej niż w *O pojęciu wynikania logicznego*. W nowym pojęciu modelu wyróżnia tzw. dziedzinę interpretacji, składającą się z obiektów, co do których nie czyni się żadnych założeń ontologicznych co do ich istnienia lub sposobu istnienia.

Do sprawy teoriomodelowej interpretacji semantycznej definicji prawdy powrócimy nieco dalej. W tym miejscu warto zauważyć, że problem istnienia lub nieistnienia obiektów może być rozwiązany w inny sposób, bez konieczności przeformułowania semantycznej definicji prawdy do postaci syntaktycznej. Jak się zdaje, uczynił to już Tarski, lecz *explicite* tego nie wyartykułował. Można bowiem – jak ujął to Donald Davidson – potraktować obiekty tworzące ciągi jako byty postulowane przez semantyczną definicję prawdy. Potraktowanie obiektów jako bytów postulowanych przez semantyczną definicję prawdy nie narusza postulowanej przez Tarskiego neutralności filozoficznej. Aby zbudować semantyczną definicję prawdy nie musimy bowiem faktycznie

zakładać fizycznego (czy jakiegokolwiek innego) sposobu istnienia obiektów wchodzących w skład ciągów. Z tego punktu widzenia, dziedzinę interpretacji wchodzącą w skład modelu można potraktować po prostu jako wyraźną realizację tego założenia.

Neutralność filozoficzna, którą postulował Tarski, przejawiała się m.in. odrzuceniem korespondencyjnej koncepcji prawdy na rzecz koncepcji klasycznej. Klasyczna koncepcja prawdy w rozumieniu Tarskiego nie zakłada bowiem – w odróżnieniu od koncepcji korespondencyjnej – istnienia obiektów, lecz istnienie obiektów jedynie postuluje⁴⁷. Nieco wcześniej była mowa o nominalistycznej interpretacji klasycznej koncepcji prawdy, która możliwa była na gruncie przekładu Rossa, w którym „jest” oznacza orzekanie. Z kolei polski przekład Kazimierza Leśniaka, który operuje pojęciem Bytu oraz Niebytu, prowadzi do realistycznego traktowania obiektów i jest z semantyczną definicją prawdy nie do pogodzenia. Dlatego też semantyczna definicja prawdy w oryginalnym ujęciu Tarskiego może być stosowana zarówno do języka opisującego fizycznie istniejące rzeczy, języka matematyki odnoszącego się do bytów abstrakcyjnych (cokolwiek miałyby to znaczyć) oraz języka fikcji literackiej. Aby zastosować wprost korespondencyjną teorię prawdy do zdań matematyki, należałoby uznać istnienie obiektów matematycznych jako obiektów platońskich, zaś jej zastosowanie do zdań o bohaterach fikcyjnych wymagałoby uznania pewnego sposobu istnienia tych obiektów, np. jako obiektów czysto intencjonalnych⁴⁸.

Aby zilustrować neutralność filozoficzną oryginalnego sformułowania semantycznej definicji prawdy, przyjrzymy się przykładowi wzorowanemu na *Prawdzie i znaczeniu* (Davidson 1967), w którym przedstawia się teorię nazwy jednostkowej *the father of x* jako teorię w pewnym sensie analogiczną do semantycznej definicji prawdy. Dlatego też zamiast dla nazwy jednostkowej *the father of x*, poniżej formułuję analogiczną teorię dla nazwy jednostkowej (deskrypcji określonej) *the oldest son of x*⁴⁹. Teoria ta posiada dwa aksjomaty:

(12.1) Wyrażenie „*x*” odnosi się do *x-a* (odniesieniem wyrażenia „*x*” jest *x*).

(12.2) Wyrażenie „*the oldest son of x*” odnosi się do najstarszego syna odniesienia *x-a*.

⁴⁷ Tak interpretuje semantyczną definicję prawdy Davidson. Pisząc o bytach postulowanych przez teorię Davidson nawiązuje bezpośrednio do wprowadzonego przez Quine’a pojęcia zobowiązań ontologicznych.

⁴⁸ Obiekty takie pojawiają się w tradycji fenomenologicznej, np. wprowadza je w swojej ontologii Roman Ingarden.

⁴⁹ Powód jest bardzo prosty, gdyż o ile każdy człowiek ma lub miał ojca, to nie każdy ma syna.

Zauważmy, że aksjomat (12.2) ma charakter rekurencyjny, gdyż można z niego wyprowadzić

Wyrażenie „*the oldest son of the oldest son of x*” odnosi się do najstarszego syna odniesienia wyrażenia „*the oldest son of x*”,

czyli:

Wyrażenie „*the oldest son of the oldest son of x*” odnosi się do najstarszego syna najstarszego syna *x-a*.

Zdania takie jak powyższe – nazwijmy je O-zdaniami – stanowią, podobnie jak T-zdania w teorii prawdy, warunki adekwatności teorii, tj. pozwalają na jej testowanie. Wynikające z powyższej „teorii” O-zdania mają postać:

„*t*” odnosi się do *y*,

gdzie *t* zastąpione jest opisem strukturalnym terminu jednostkowego, zaś *y* przez sam ten termin (por. 1992/1967, s. 5). Przykładami O-zdań byłyby:

„*Leopold Mozart*” odnosi się do Leopolda Mozarta.

„*The oldest son of Leopold Mozart*” odnosi się do najstarszego syna Leopolda Mozarta.

Zauważmy, że przedstawiona namiastka teorii odniesienia faktycznie rezygnuje z wymogu istnienia obiektów, chociaż jak wiadomo Leopold Mozart miał najstarszego syna, podobnie jak jego najstarszy syn, czyli Wolfgang Amadeusz (był nim Raimund Leopold Mozart, który zmarł we wczesnym dzieciństwie). Nie zmienia to jednak faktu, że na gruncie tej teorii można również mówić o najstarszym synu przedwcześnie zmarłego Reimunda Leopolda Mozarta. Jest tak, gdyż – jak zauważa Davidson – w teorii tej używa się wprawdzie pojęcia odniesienia, ale nie jest to pojęcie semantyczne *sensu stricto*, gdyż pojęcie odniesienia i opisana procedura wyznaczania odniesień traktowane są tutaj jedynie instrumentalnie. Owo instrumentalnie traktowane odniesienie służy – jak twierdzi Davidson, i o czym będzie jeszcze mowa – podaniu znaczenia wyrażenia „najstarszy syn”. Przedstawiona namiastka teorii odniesienia oczywiście nie wystarczy, aby podać znaczenie predykatu „najstarszy syn”. Rolę taką może jednak pełnić cała semantyczna definicja prawdy dla języka, do którego należy to wyrażenie. W konsekwencji semantyczna definicja prawdy nie będzie podawała znaczenia wybranego wyrażenia, lecz całego języka, co związane jest z jej holistycznym aspektem.

13. HOLISTYCZNY ASPEKT SEMANTYCZNEJ DEFINICJI PRAWDY

Zarówno w standardowej interpretacji semantycznej definicji prawdy, jak i odrzuconej przez nas jej wersji w sformułowaniu syntaktycznym, zakłada się, że istnieje przekład z języka przedmiotowego na metajęzyk, a w szczególnym przypadku metajęzyk obejmuje język przedmiotowy. Tarski nigdy nie kwestionował poprawności przekładu, co jest zrozumiałe w przypadku metamatematycznych zastosowań semantycznej definicji prawdy. Ponieważ naszym celem jest adaptacja semantycznej definicji prawdy do języka naturalnego, odejźmy od założenia o poprawności przekładu i uznajmy, że nie jest z góry przesądzone czy mamy do czynienia z przekładem „właściwym”, a – co za tym idzie – z „właściwą” definicją prawdy. Aby się przekonać, czy mamy do czynienia z właściwym przekładem, można – o czym była mowa nieco wcześniej – wykorzystać T-zdania do testowania definicji. Owo testowanie ujawnia holistyczny aspekt semantycznej definicji prawdy.

Wyobraźmy sobie, że zwiedziony podobieństwem nazw ktoś błędnie przekłada angielską nazwę *Plotinus* na „Platon” oraz angielską nazwę *Ariston* na „Arystoteles”⁵⁰. Dla odpowiadającej tej sytuacji definicji prawdy mielibyśmy następujące aksjomaty:

Każdy ciąg *C* przyporządkowuje nazwie „*Plotinus*” Platona.

Każdy ciąg *C* przyporządkowuje nazwie „*Ariston*” Arystotelesa.

Ciąg *C* spełnia funkcję t_1 is t_2 wtw, gdy obiekt przyporządkowany termowi t_1 przez *C* jest obiektem przyporządkowany termowi t_2 przez *C*.

Jak łatwo pokazać, z podanych aksjomatów wynika następujące T-zdanie:

Zdanie „*Ariston is Plotinus*” jest prawdziwe wtw, gdy Arystoteles jest Platonem.

Ponieważ zarówno zdanie po lewej stronie, jak i prawa strona równoważności są fałszywe⁵¹, powyższe T-zdanie nie pozwala nam zakwestionować przekładu z języka przedmiotowego na metajęzyk. Aby to uczynić musi się pojawić niezgodność między przekonaniem wyrażonym w metajęzyku (po prawej stronie), a przekonaniem naszego rozmówcy wyrażonym po lewej stronie. W szczególnym przypadku owym rozmówcą możemy być my sami. W takim przypadku sytuacja jest bardzo prosta. Jeśli zdanie wyrażone w języku obcym akceptujemy, zaś jego przekład na nasz własny język

⁵⁰ „Aryston” (ang. *Ariston*) to właściwe imię Platona.

⁵¹ A dokładnie, rozmówca odrzuca zdanie *Ariston is Plotinus*, podobnie jak my odrzucamy „Arystoteles jest Platonem”.

odrzucaamy, to uznamy, że przekład jest po prostu błędny. Oznacza to, że jeżeli językiem przedmiotowym jest język obcy, to pewnym jego wyrażeniom przypisujemy niewłaściwe znaczenia. W przypadku, gdy interpretowany i interpretator nie są tą samą osobą, wówczas nie każda niezgodność przekonań wystarczy, aby przekład zakwestionować, gdyż obaj rozmówcy mogą się różnić w kwestii faktów. W takim przypadku należy się odwołać do pewnych warunków pragmatycznych, o których będzie mowa w dalszych rozdziałach.

Aby znaleźć „różnicę zdań”, a tym samym zakwestionować przekład, a zarazem sfalsyfikować semantyczną definicję prawdy, należy odwołać się do innych aksjomatów (wzorem Davidson założmy, że dysponujemy definicją prawdy dla całego języka angielskiego jako języka przedmiotowego). Oczywiście nie każdy nowy aksjomat pozwoli na zweryfikowanie definicji, np. nie pozwoli na to aksjomat:

Ciąg C spełnia funkcję „ t_1 is *philosopher*” wtw, gdy obiekt przyporządkowany termowi t_1 przez C jest filozofem,

gdyż osoby oznaczane przez wszystkie z wymienionych nazw własnych (zarówno w języku przedmiotowym, jak i metajęzyku) były filozofami.

Na sfalsyfikowanie przekładu pozwoli natomiast aksjomat następujący:

Ciąg C spełnia funkcję „ t_1 is a *pupil of* t_2 ” wtw, gdy obiekt przyporządkowany termowi t_1 przez ciąg C jest uczniem obiektu przyporządkowanego termowi t_2 przez C .

Z aksjomatu tego wynika bowiem T-zdanie:

Zdanie „*Ariston is a pupil of Plotinus*” jest prawdziwe wtw, gdy Arystoteles jest uczniem Platona,

w którym prawa strona jest prawdziwa, lecz zdanie po lewej stronie równoważności jest ewidentnie fałszywe (bo Aryston, czyli Platon nie był uczniem Plotyna).

To, że sfalsyfikowaliśmy definicję prawdy, nie daje nam żadnej wskazówki, jak powinna wyglądać „właściwa” definicja prawdy. Wiadomo jedynie, że należy zmienić jej aksjomaty jako wyraz zmiany przekładu z metajęzyka na język przedmiotowy. Nie powinna nas tu mylić sugestia na temat błędnego przekładu nazw własnych. Ten sam rezultat moglibyśmy uzyskać, gdyby błędnie przełożono jakiś predykat (np. *pupil* na „nauczyciel”). Oczywiście w naszym przypadku najprostsze rozwiązanie polegałoby na zmianie aksjomatów odniesienia jako konsekwencji zmiany przekładu nazw *Plotinus* i *Ariston*. Tak otrzymana definicja prawdy może być akceptowana do momentu, gdy wyprowadzimy z niej T-zdanie, które ponownie pozwoli ją sfalsyfikować.

Ponieważ jednak możliwych T-zdań jest nieskończenie wiele, definicja nie może zostać nigdy ostatecznie potwierdzona.

Gdy podobnie jak Tarski akceptujemy przekład, to T-zdania służą do wyznaczenia zakresu predykatu „prawdziwy”. Jeśli jednak dopuścimy możliwość weryfikacji przekładu, T-zdania służą do testowania definicji. Jest tak, gdyż przynajmniej w pewnym zakresie możemy odgadnąć, które zdania języka przedmiotowego nasz rozmówca uznaje za prawdziwe. W takim przypadku każde wynikające z semantycznej definicji prawdy T-zdanie zestawia przekonanie rozmówcy (wyrażone w języku przedmiotowym) z naszym własnym przekonaniem (wyrażonym w metajęzyku). Mamy zatem do rozwikłania zagadkę z wieloma niewiadomymi. Problematyce interpretacji słów drugiego człowieka – bo o niej tu mowa – poświęcony zostanie ostatni rozdział monografii.

14. HARTRY FIELDA INTERPRETACJA SEMANTYCZNEJ DEFINICJI PRAWDY

Na temat semantycznej definicji prawdy istnieje bardzo bogata literatura. Na koniec rozdziału odwołam się do dwóch „kanonicznych” tekstów. Są to wspomniane wcześniej artykuły Hartry Fielda (1972) oraz Scotta Soamesa (1984). Obaj podkreślają, że motywem Tarskiego do zbudowania semantycznej definicji prawdy była chęć redukcji pojęć semantycznych akceptowalnej z punktu widzenia postulatów jedności wiedzy i fizykalizmu.

Hartry Field zdecydowanie akcentuje syntaktyczny (morfologiczny) charakter oryginalnego sformułowania semantycznej definicji prawdy, wyrażony przez samego Tarskiego *implicite* w omówionej wcześniej pracy *O ugruntowaniu naukowej semantyki* (1936). Zwracając uwagę przede wszystkim na filozoficzne, a nawet na ideologiczne motywacje związane z fizykalizmem, które legły u podstaw semantycznej definicji prawdy, zadaje pytanie, czy Tarski zdołał osiągnąć zamierzony cel. Jak pamiętamy, cel ten polegał na redukcji pojęć semantycznych do innych pojęć, które byłyby fizykalistycznie dopuszczalne. Aby odpowiedzieć na to pytanie, Field dokonał eksplikacji semantycznej definicji prawdy w kategoriach syntaktycznych, czyli zgodnie z intencjami samego Tarskiego (teoria T_2), i porównuje ją z własną propozycją semantycznej definicji prawdy (teoria T_1), która – zdaniem Fielda – jest faktycznie semantyczna i jest taka, jaką Tarski „powinien był sformułować”.

Teoria T_2 stanowi jedynie notacyjny wariant Tarskiego definicji prawdy w sformułowaniu syntaktycznym, dlatego też nie będę jej tutaj przytaczał, gdyż wspominałem o niej już wcześniej. Zdaniem Fielda definicja ta częściowo spełnia wymóg redukcji fizykalistycznej, gdyż oparta jest na przekładzie

z języka przedmiotowego na język angielski (metajęzyk), co umożliwia eliminację takich pojęć semantycznych jak „odnosi się do” lub „oznacza”. Oczywiście samo pojęcie przekładu nie jest fizykalistycznie dopuszczalne, zaś próba zdefiniowania go musi się odwołać do pojęcia „współznaczania” wyrażień, które samo jest pojęciem semantycznym, które nie zostało zredukowane do pojęć niesemantycznych (*op. cit.*, s. 355)⁵².

Okazuje się jednak, że wbrew deklaracjom, rekonstruując teorię Tarskiego, Field nie jest zbyt konsekwentny. W przeciwieństwie do przedstawionej przeze mnie wcześniej semantycznej definicji prawdy w sformułowaniu syntaktycznym, Field nie formułuje aksjomatów T_2 , odwołując się *explicite* do ciągów nazw i spełniania funkcji zdaniowej przez nazwy, lecz do ciągów obiektów oraz spełniania funkcji zdaniowych przez obiekty⁵³, traktując *implicite* owe obiekty – podobnie jak Tarski i Davidson – jako byty postulowane przez teorię. W tym miejscu pojawia się problem następujący – teoria może postulować istnienie tylko tych obiektów, których nazwy są sensowne. W przypadku semantycznej definicji prawdy dla języków teorii matematycznych, czyli języków, dla których była ona pierwotnie przeznaczona, problem ten się nie pojawia, gdyż sensy nazw przedmiotów matematycznych wyznaczone są przez aksjomaty teorii. Posiadają one – w myśl weryfikacjonistycznej teorii znaczenia uznawanej przez filozofów Koła wiedeńskiego, z którymi Tarski utrzymywał rozliczne kontakty naukowe – tzw. znaczenia analityczne. Używanie nazw tych bytów jest zatem fizykalistycznie dopuszczalne. W przypadku zdań syntetycznych, a takie występują głównie w językach naturalnych, można mówić jedynie o znaczeniach empirycznych, gdyż zdaniem weryfikacjonistów znaczenia empiryczne muszą być ugruntowane w rzeczywistości. Problem związku wyrażień z rzeczywistością podjął Field, analizując teorię T_1 .

Jak pamiętamy, Tarski budował semantyczną definicję prawdy dla języków, w których „sens każdego wyrażenia jest jednoznacznie wyznaczony przez jego kształt”. Field interpretuje postulat Tarskiego – o czym już była mowa – w sposób nieco inny niż ja. Utrzymuje mianowicie, że Tarskiemu chodziło o języki, w których każdemu wyrażeniu jednoznacznie odpowiada

⁵² W przypadku, gdy język przedmiotowy jest teorią aksjomatyczną, zaś metajęzyk obejmuje język przedmiotowy, zarzut Fielda jest bezpodstawny. Ma on natomiast rację, gdy mamy do czynienia z językiem naturalnym, gdyż nawet równokształtne wyrażenia mogą mieć różne znaczenia.

⁵³ Do teorii T_2 odwołałem się nieco wcześniej, zauważając, że różni się od wersji rozpatrywanej przeze mnie opuszczeniem cudzysłówów.

określone odniesienie⁵⁴. W językach takich nie mogłyby występować wyrażenia niejednoznaczne, których sens i odniesienie dookreślane są przez kontekst użycia, a nie byłyby jednoznacznie zdeterminowane przez ich kształt:

(...) tj. gdy dwóch użytkowników wypowiada tę samą nazwę (lub jeden mówiący wypowiada ją w dwóch różnych okolicznościach), odnoszą się do tego samego obiektu oraz gdy dwóch użytkowników wypowiada to samo zdanie, to albo obaj mówią prawdę, albo żaden z nich nie mówi prawdy (Field 1972, s. 348).

Oznaczmy to sformułowanie jako warunek (*).

Aby przedstawić teorię T_I rozważmy za Fieldem język, w którym występują jedynie terminy jednostkowe, predykaty jednoargumentowe oraz podstawowe stałe logiczne⁵⁵. Niech s będzie ciągiem obiektów $\langle s_1, s_1, s_3, \dots \rangle$. Podstawowymi pojęciami semantycznymi, które występują w teorii T_I są denotowanie_s obiektów przez terminy jednostkowe (dla ciągu s), oraz odnoszenie się_s predykatu do obiektu, co znajduje wyraz w następujących aksjomatach:

- (A) 1. zmienna ' x_k ' denotuje_s s_k ,
2. stała ' c_k ' denotuje_s to, co denotuje.
- (B) 1. formuła ' $P_k(e)$ ' jest prawdziwa_s wtw, gdy
 - (i) istnieje obiekt a denotowany_s przez e oraz
 - (ii) ' P_k ' odnosi się_s (*applies to*_s) do a .
- 2. formuła ' $\sim\varphi$ ' jest prawdziwa_s wtw, gdy ' φ ' nie jest prawdziwa_s.
- 3. formuła ' $\varphi \wedge \psi$ ' jest prawdziwa_s wtw, gdy ' φ ' oraz ' ψ ' są prawdziwe_s.
- 4. formuła ' $\forall x_k \varphi$ ' jest prawdziwa_s wtw, gdy dla każdego ciągu s^k , który różni się od ciągu s co najwyżej na k -tym miejscu, ' φ ' jest prawdziwa_s^k.
- (C) Zdanie jest prawdziwe wtw, gdy jest spełnione przez wszystkie ciągi (*op. cit.* s. 350)⁵⁶.

W przypadku wyrażeń syntaktycznie prostych Field proponuje zastąpić terminy „denotuje_s” oraz „odnosi się_s” przez „pierwotnie denotuje_s”. Oznacza to, że predykat jednoargumentowy, podobnie jak nazwa własna, pierwotnie denotuje obiekty. Oczywiście pierwotne denotowanie w obu przypadkach oznacza coś innego – nazwa własna po prostu oznacza obiekt i tylko ten

⁵⁴ Faktycznie nasze podejścia się nie wykluczają, o ile przez odniesienie rozumiemy byt postulowany przez teorię.

⁵⁵ Field wprowadza również do języka jednoargumentowe wyrażenia funkcyjne, które dla uproszczenia pomijam.

⁵⁶ Podaję z niewielkimi zmianami notacyjnymi, które – moim zdaniem – sprawiają, że podane przez Fielda aksjomaty są bardziej czytelne.

obiekt, zaś w przypadku predykatu jednoargumentowego oznacza to, że obiekt należy do zakresu predykatu.

Chociaż w teorii T_1 prawdziwość została zredukowana do pierwotnego denotowania, nie jest to jednak – zdaniem Fielda – redukcja pełna, czyli fizykalistycznie dopuszczalna, gdyż samo pojęcie pierwotnego denotowania nie jest fizykalistycznie dopuszczalne. Zaletę tego sformułowania Field upatruje jednak w tym, że pozwala ono na łatwe rozszerzenie na języki, które nie spełniają warunku (*), czyli języki dopuszczające wieloznaczność. Oryginalny przykład podany przez Fielda to *John takes grass*, co może znaczyć w pewnym kontekście, że Jan zażywa marihuanę lub (w innym kontekście), że Jan kradnie nasiona traw. Odpowiednia modyfikacja T_1 polegałaby – zdaniem Fielda – na uznaniu, że podane wyżej aksjomaty nie mówią nic o pierwotnym denotowaniu lub byciu prawdziwym dla typów wyrażeń, lecz o pierwotnym denotowaniu lub byciu prawdziwym dla wypowiedzi lub okazów wyrażeń (*tokens*). W takim przypadku możliwe jest podanie różnych aksjomatów dla okazów tego samego typu wyrażenia. Z alternatywnych aksjomatów korzystamy w zależności od kontekstu wypowiedzenia. Z tak zmodyfikowanej definicji powinny wynikać następujące dwa T-zdania:

Wypowiedzenie zdania „*John takes grass*” jest prawdziwe wtw, gdy Jan pali marihuanę.

Wypowiedzenie zdania „*John takes Grass*” jest prawdziwe wtw, gdy Jan kradnie nasiona traw.

Druga zaleta T_1 polega na tym, że wbrew deklaracjom Tarskiego nie jest konieczne założenie o ustalonym słowniku języka przedmiotowego. Zdaniem Fielda wystarczy założyć, że język ma ustalone kategorie syntaktyczne wyrażeń i w razie potrzeby uzupełniać język przedmiotowy o nowe wyrażenia. W oryginalnym ujęciu Tarskiego (a przynajmniej w rekonstruującej jego ujęcie teorii T_2) nie da się wprowadzić nowego wyrażenia do języka przedmiotowego bez konieczności wzbogacenia metajęzyka. Takie wzbogacenie metajęzyka powoduje jednak, że faktycznie mamy do czynienia z nową definicją prawdy. W przypadku T_1 nic nie stoi na przeszkodzie, aby – o ile w języku występuje odpowiednia kategoria wyrażeń – wprowadzać nowe wyrażenia bez zmiany teorii (por. *op. cit.*, s. 352–353).

W tym miejscu warto pokusić się o pewną dygresję, dotyczącą podanej przez Fielda interpretacji programu semantycznego Davidsona. Field – moim zdaniem słusznie – przypisuje Davidsonowi zaadaptowanie do potrzeb swego programu semantycznej definicji prawdy w postaci T_2 , traktując ją – tym

razem niesłusznie – jako definicję prawdy w sformułowaniu syntaktycznym. Następnie wyciąga z tego konsekwencję, że połączenie programu Davidsona z teorią T_1 , która daje możliwość wzbogacania języka oraz analizy wypowiedzi zdań, zapewnia przedsięwzięciu Davidsona pewną szansę powodzenia⁵⁷. Stwierdzenie to oparte jest jednak na oczywistym nieporozumieniu, gdyż T_2 jest faktycznie teorią, w której postuluje się istnienie obiektów. Ponadto definicja prawdy spełnia swą funkcję teorii znaczenia nawet – a może nawet przede wszystkim – gdy rozszerzenie języka wiąże się ze zmianą aksjomatów semantycznej definicji prawdy. Zamyśl Davidsona przewiduje bowiem ciągłe przechodzenie od teorii do teorii. Ponadto semantyczny program Davidsona, o czym będzie jeszcze wielokrotnie mowa, faktycznie uwzględnia wieloznaczności i okazjonalność, dla których nie ma miejsca w oryginalnym sformułowaniu Tarskiego. Jak wielokrotnie uzasadniał Davidson, semantyczną definicję prawdy da się bez trudu adaptować dla wypowiedzi wyrażen języka naturalnego, co pozwala na objęcie nią zarówno okazjonalności, jak i różnego rodzaju wieloznaczności. W tej sytuacji semantyczna definicja prawdy stosuje się faktycznie nie do zdań, lecz do wypowiedzi zdań⁵⁸. Można zatem powiedzieć, że Field wyważa otwarte drzwi.

Dalsze rozważania Fielda również dają się pogodzić z Davidsonowskim rozumieniem semantycznej definicji prawdy, chociaż pozornie wskazują na to, że Field nie traktował obiektów jako byty postulowane, lecz raczej na sposób realistyczny. To co może – zdaniem Fielda – niepokoić w aksjomatach teorii T_1 , to pojawiające się w nich *explicite* założenia egzystencjalne. Założenia te są jednak przynajmniej do pewnego stopnia do uzgodnienia z postulatem redukcji pojęć semantycznych. Jak deklarował Field, redukcja pojęcia prawdy nie została przeprowadzona w T_1 w sposób całkowity. Aby zakończyć to przedsięwzięcie należy zredukować również pojęcie pierwotnego denotowania do pojęć fizykalistycznie dopuszczalnych. Sprawa nie okazuje się jednak prosta, zaś Field wskazuje na dwa możliwe kierunki takiej redukcji. Pierwszy z nich odwołuje się do wprowadzonego przez Russella pojęcia „prawdziwych

⁵⁷ „Przewaga takich przykładów [tj. wieloznaczności] w języku naturalnym pozwala zadać pytanie, czy teoria semantyczna typu Tarskiego daje się zastosować do języków, w których sens *nie* jest zdeterminowany przez formę [kształt wyrażenia]; jeśli odpowiedź jest negatywna, wówczas interesujący projekt Davidsona, polegający na podaniu warunków prawdziwości dla zdań języka naturalnego, jest od początku skazany na niepowodzenie” (*op. cit.*, s. 351–352).

⁵⁸ Przedstawione tu uwagi nie powinny być traktowane jako zarzut pod adresem Fielda, że ten źle zrozumiał Davidsona. Należy bowiem wziąć pod uwagę, że artykuł *Tarski's Theory of Truth* został opublikowany przed powstaniem większości prac Davidsona poświęconych teorii interpretacji. W szczególności przed 1972 rokiem Davidson nie ujawnił się jeszcze jako przeciwnik pojęcia języka jako abstrakcyjnego tworu opartego na konwencjach.

nazw własnych”, które odpowiadają bezpośrednio indywiduom istniejącym w świecie. Owymi prawdziwymi nazwami własnymi są po prostu zaimki wskazujące na realnie istniejące obiekty. Relacja między takimi wyrażeniami a światem nie da się już w żaden sposób zredukować do innych pojęć i jest fizykalistycznie dopuszczalna⁵⁹. Ponieważ Russell opisał procedurę redukcji „zwykłych nazw” do „nazw prawdziwych” (*op. cit.*, s. 366), pojęcie pierwotnego denotowania dałoby się w ten sposób zredukować do pojęć fizykalistycznie dopuszczalnych.

Druga możliwa redukcja odwołuje się do tzw. historyczno-przyczynowej teorii oznaczania nazw własnych oraz nazw gatunków naturalnych rozwiniętych przez Saula Kripkego w *Naming and Necessity* (1980) oraz Hilarego Putnama w *The Meaning of Meaning* (1975). W teorii tej relacja łącząca nazwy z rzeczywistością jest ustanowiona aktem „pierwszego chrztu”, podczas którego dane indywiduum zostaje opatrzone nazwą własną, np. niemowlę imieniem „Arystoteles”, lub próbka substancji naturalnej nazwą gatunkową, np. próbka cieczy nazwą „woda” lub zwierzę nazwą „tygrys”. Od tej chwili nazwę własną odnosimy do nazwanego przez nią indywiduum (lub nazwę gatunkową do przedstawicieli lub próbek rodzaju naturalnego), gdyż użycie tej nazwy łączy się z aktem pierwszego chrztu poprzez zachowujący odniesienie historyczny łańcuch przyczynowo-skutkowy oznaczania tego indywiduum (lub oznaczania przedstawiciela rodzaju naturalnego).

Próba fizykalistycznej redukcji relacji pierwotnego denotowania za pomocą teorii Kripkego-Putnama generuje pewien rodzaj esencjalizmu, który jest nieuprawniony z punktu widzenia postulowanej przez Tarskiego neutralności filozoficznej. Redukcjonistyczny program Koła wiedeńskiego, którego celem było m.in. wyeliminowanie z nauki wszelkich naleciałości metafizycznych, został zarzucony między innymi ze względu na to, że sam *implicite* czynił założenia metafizyczne, nie spełniając tym samym zakładanych przez siebie kryteriów naukowości. Pomijając powyższy zarzut, sugestia Fielda, aby semantyczna definicja prawdy dla języka naturalnego była w pewien sposób zakotwiczona w rzeczywistości, jest do pogodzenia z poglądami Davidsona, który traktował semantyczną definicję prawdy jako teorię empiryczną,

⁵⁹ Koncepcja Russella dla prawdziwych nazw własnych przypomina analogiczne pojęcie zdań protokolarnych, popularne w początkowym okresie działalności Koła wiedeńskiego, które niewątpliwie propagowało fizykalizm. Zdanie protokolarne miało być – podobnie jak prawdziwa nazwa własna – bezpośrednio związane z rzeczywistością. Oznaczało to, że zdania protokolarne uznawane miały być bezpośrednio i niezależnie od siebie na podstawie danych doświadczenia, co było niewątpliwie fizykalistycznie dopuszczalne. Koncepcja ta została z czasem zarzucona, a ostatecznie przekonanie o możliwości sformułowania zdań protokolarnych zostało zdyskredytowane przez Quine’a w *Dwóch dogmatach empiryzmu* (1952).

która może być testowana za pomocą wynikających z niej T-zdań. Szczególną rolę pełnią w niej T-zdania, w których po obu stronach równoważności występują zdania okazjonalne o treści empirycznej, np. „To jest pies”. Zdania okazjonalne tego typu uznajemy najczęściej jako skutki pewnych zdarzeń zachodzących w świecie zewnętrznym. Oczywiście nie każde zdanie tej postaci uznawane jest w taki sposób – za przykład posłużyć nam może zdanie „Oto centaur”, wypowiedziane przez bohatera powieści fantastycznej. Jak jednak twierdzi Davidson, warunkiem istnienia języka jest to, że bardzo wiele zdań tego typu uznajemy właśnie w ten sposób. Język jest zatem zakotwiczony w rzeczywistości jako pewna organiczna całość. Jak się zdaje Field nie docenił holistycznego aspektu semantycznej definicji prawdy. Do sprawy języka jako „organicznej całości” powrócę w kolejnym rozdziale.

15. SCOTTA SOAMESA INTERPRETACJA SEMANTYCZNEJ DEFINICJI PRAWDY

W artykule *What Is a Theory of Truth* Scott Soames (1984) podejmuje trzy zadania. Po pierwsze, dokonuje krytycznej analizy Hartry Fielda interpretacji semantycznej definicji prawdy, po drugie, argumentuje przeciwko wykorzystaniu semantycznej definicji prawdy do definiowania pojęcia znaczenia, po trzecie, wskazuje jaka miała być rola semantycznej definicji prawdy w zamierzeniach samego Tarskiego⁶⁰.

Jak pamiętamy, teoria T_I posłużyła Fieldowi do pokazania, że Tarski nie zdołał przeprowadzić fizykalistycznej redukcji pojęcia prawdy, gdyż pojęcie to zostało zredukowane do pojęcia pierwotnego denotowania, które to pojęcie

⁶⁰ Ogólnie rzecz biorąc, celem Soamesa było wykazanie, że dyskusja nad filozoficzną doniosłością semantycznej definicji prawdy oparta jest na nieporozumieniu. Źródłem tego nieporozumienia mogła być – zdaniem Soamesa – rola predykatu „prawdziwy” w formułowaniu pewnych twierdzeń logicznych i filozoficznych w metajęzyku. Zjawisko to nazywa (za Quinem) semantycznym wznoszeniem na wyższy poziom języka (*semantic ascent*). Na przykład zamiast wymieniać wszystkie zdania prawdziwe postaci $p \Rightarrow (q \Rightarrow p)$, można napisać „dla wszystkich zdań p i q , (p jest prawdziwe \Rightarrow (q jest prawdziwe $\Rightarrow p$ jest prawdziwe))”. Ponadto jedną z wersji tezy realizmu metafizycznego można, zamiast (nieskończonej) alternatywy zdań w języku przedmiotowym, wyrazić w sposób następujący: „istnieje przynajmniej jedno prawdziwe zdanie s , do którego uznania nigdy nie znajdziemy wystarczających podstaw”. Zdaniem Soamesa, użyteczność predykatu „prawdziwy” skłoniła niektórych filozofów do uznania pojęcia prawdy za „pojęcie głęboko metafizyczne”. Ponieważ jednak zdania metajęzykowe z predykatem „prawdziwy” posiadają tę samą zawartość informacyjną, co ich odpowiedniki z języka przedmiotowego, to jedyna rola tego predykatu polega na ułatwieniu formułowania tez filozoficznych. Zdaniem Soamesa, pogląd ten podzielał *implicite* sam Tarski, twierdząc, że predykat „prawdziwy” jest filozoficznie neutralny (s. 414).

samo nie było fizykalistycznie dopuszczalne. Soames interpretuje intencje Fielda w sposób następujący:

(...) autentyczna redukcja musi pokazać, że semantyczne własności wyrażeń superwenują na fizycznych faktach dotyczących ich użytkowników oraz środowisku, w którym zostały użyte. Definicja Tarskiego tego nie zapewnia (*op. cit.*, s. 418).

W tym miejscu konieczne jest pewne wyjaśnienie. Pojęcie superwencji pochodzi od Moore'a, który głosił, że jeżeli zachodzi jakakolwiek różnica w ocenie etycznej jakiegoś faktu, to istnieje również różnica w jego opisie fizykalnym, co oznacza, że etyczna ocena działania superwenuje na jego opisie fizykalnym. A zatem jeżeli dwa działania różnie oceniamy pod względem etycznym, to opisy tych działań w języku fizykalnym nie mogą być identyczne. U Fielda, który odwołał się do historyczno-przyczynowej teorii oznaczania nazw własnych, prowadzi to do konsekwencji, że jeżeli dwie nazwy własne pierwotnie denotują dwa różne obiekty, to muszą istnieć dwa różne łańcuchy przyczynowo-skutkowe użycia tych nazw⁶¹.

Pozornie mogłoby się wydawać, że pomysł Fielda z zastosowaniem historyczno-przyczynowej teorii oznaczania nazw własnych lub Bertranda Russella redukcji do prawdziwych nazw własnych, rozwiązuje problem fizykalistycznej redukcji pojęcia pierwotnego denotowania. Soames jest jednak innego zdania:

Jeżeli rzekoma zależność własności semantycznych od faktów fizycznych dotyczących ich użytkowników wskazuje na to, że Tarski nie zdołał zredukować pierwotnej denotacji do faktów fizycznych, to ten sam argument może być użyty do wykazania, że nie zdołał zredukować [pojęcia] prawdy do [pojęcia] pierwotnej denotacji (s. 419).

Oznacza to, że propozycja Fielda zawieszona jest w próżni, gdyż wprowadzenie pojęcia pierwotnego denotowania nie zapewnia fizykalistycznej redukcji predykatu „prawdziwy”.

Jak pamiętamy z paragrafu poświęconego pojęciu wynikania logicznego, Tarski zrelatywizował to pojęcie do wyboru stałych logicznych. Oczywiście był świadomy faktu, że wybór stałych logicznych ma charakter arbitralny, czemu dał wyraz na zakończenie *O pojęciu wynikania logicznego*:

⁶¹ Zauważmy, że odwrotna zależność nie zachodzi. Mogą bowiem istnieć dwie nazwy o dwóch różnych łańcuchach przyczynowych ich użycia (np. „Tuliusz” i „Cyceron”), które pierwotnie denotują ten sam obiekt.

U podstaw całej konstrukcji leży podział wszystkich wyrazów języka na logiczne i pozalogiczne. Podział ten nie jest całkowicie dowolny (...). Z drugiej jednak strony, nie znam żadnych obiektywnych względów, które pozwalałyby przeprowadzić dokładną granicę między obiema kategoriami terminów (s. 200).

Ten sam arbitralny podział założony jest w semantycznej definicji prawdy, gdyż dla każdej stałej logicznej zarezerwowany jest stosowny aksjomat. Upraszczając nieco argumentację Soamesa, aby dokonać fizykalistycznej redukcji pojęcia prawdy należałoby również pokazać, w jaki sposób semantyczne własności stałych logicznych superwenują na fizycznych faktach ich użycia przez użytkowników języka, czego Field nie zdołał przedstawić. Definicja prawdy nie podaje ponadto kryteriów, w jaki sposób pewne symbole stałych języka przedmiotowego odpowiadają negacji, koniunkcji, alternatywie czy kwantyfikatorom (s. 421).

Argumentacja Soamesa odwołuje się do faktu, że semantyczna definicja prawdy (w wersji T_2 *explicite*, zaś w wersji T_1 *implicite*) zakłada pojęcie przekładu, np. symbolu \wedge na odpowiednik koniunkcji w metajęzyku. Jak twierdzi Soames, ani przekład nie jest pojęciem fizykalistycznie dopuszczalnym, ani nie podano jego fizykalistycznie dopuszczalnej redukcji. Jest to stwierdzenie oczywiste, jednak – jak starałem się pokazać – nie dotyczy to przypadku, gdy metajęzyk obejmuje język przedmiotowy teorii aksjomatycznej, który to przypadek najbardziej Tarskiego interesował. W przypadku języka naturalnego argumentacja wymierzona w Fielda, a zarazem w możliwość dokonania fizykalistycznie dopuszczalnej redukcji pojęcia prawdy, okazuje się być zbieżna z pewnymi poglądami Quine’a i Davidsona. Donald Davidson wykazał, że aparat semantycznej definicji prawdy pozwala jedynie na częściowe zredukowanie pojęcia przekładu do pewnych faktów fizycznych użycia wyrażen. Przekład w sensie Davidsona (z pewnymi zastrzeżeniami, o których będzie mowa w rozdziale 4) podaje w sposób systematyczny znaczenia wyrażen języka przedmiotowego osobie, która zna metajęzyk, zaś częściowa możliwość jego redukcji wyjaśnia dlaczego nie jest on całkowicie zdeterminowany przez ogół faktów dotyczących użycia wyrażen przez mówiącego i interpretatora. Stwierdzenie to jest w zasadzie pewnym sformulowaniem Quine’a tezy o niezeterminowaniu przekładu oraz tezy Davidsona o niezeterminowaniu interpretacji, o których będzie mowa w rozdziale 4.

Na zakończenie przyjrzyjmy się stosunkowi Soamesa do kluczowego w tej monografii problemu, czy semantyczna definicja prawdy może pełnić funkcję teorii znaczenia. W trzecim rozdziale omawianego artykułu Soames podaje kilka argumentów na rzecz tezy, że znajomość warunków prawdziwości zdania nie stanowi warunku koniecznego znajomości jego znaczenia.

Przyjrzyjmy się jednemu z tych argumentów. Zwolennik krytykowanej przez Soamesa tezy zaakceptowałby – jego zdaniem – następujący schemat zdania:

(16) Gdyby „*S*” znaczyło w języku *L*, że *p*, to „*S*” byłoby prawdziwe w *L* wtw, gdy *p* (s. 422).

Niech *Ws* będzie zdaniem języka *L*, które znaczy, że śnieg jest biały. Jak twierdzi Soames użycie semantycznej definicji prawdy prowadzi do absurdałnej tezy:

(16_T) Gdyby „*Ws*” znaczyło w *L*, że śnieg jest czarny, to śnieg byłby biały wtw, gdy śnieg jest czarny (s. 423).

Przejdźcie od (16) do (16_T) odbywa się w następujących krokach:

(16ⁱ) Gdyby „*Ws*” znaczyło w języku *L*, że *śnieg jest czarny*, to „*Ws*” byłoby prawdziwe w *L* wtw, gdy *śnieg jest czarny* (*ibidem*).

Niech semantyczna definicja prawdy będzie taka, że wynika z niej T-zdanie:

„*Ws*” jest prawdziwe w *L* wtw, gdy śnieg jest biały (*ibidem*).

Po zastąpieniu w (16ⁱ) zdania „*Ws* jest prawdziwe” przez „śnieg jest biały” otrzymamy (16_T). Można mieć wątpliwości, czy zastąpienie takie jest uprawnione, gdyż okres warunkowy (16) tworzy kontekst intensjonalny, ale Soames uzasadnia jego zasadność (s. 423, przyp. 21). Gdyby jednak uzasadnienie było niesłuszne, to podobny rezultat można byłoby otrzymać, zamieniając (16) na jego ekstensjonalną wersję:

(16*) Jeżeli „*S*” znaczy w języku *L*, że *p*, to „*S*” jest prawdziwe w *L* wtw, gdy *p* (*ibidem*).

z którego w ten sam sposób otrzymujemy absurdałne zdanie:

Jeżeli „*Ws*” znaczy w języku *L*, że *śnieg jest czarny*, to śnieg jest biały wtw, gdy *śnieg jest czarny* (*ibidem*).

Czy przedstawiony argument zagraża stanowisku Davidsona? Absurdałność przytoczonego wyżej zdania bierze się stąd, że pojęcie znaczenia zostało tu użyte w dwóch sensach. W pierwszym sensie pojęcie znaczenia występuje w (16*), a – co więcej – (16*) miało nadać znaczenie temu

pojęciu⁶². W innym sensie rozumie się pojęcie znaczenia, gdy zakłada się, że *Ws* znaczy, że śnieg jest biały, co w konsekwencji pozwala na sformułowanie T-zdania:

„*Ws*” jest prawdziwe wtw, gdy śnieg jest biały.

Oznacza to, że znaczenie zdania *Ws* w drugim sensie jest podane przez przekład założony w teorii prawdy. Mamy zatem w argumentacji Soamesa do czynienia z pewną formą ekwiwokacji.

Pomimo tego wydaje się, że przykład ten, jak i inne podane przez Soamesa, ilustrują pewną ważną tezę. Wbrew Soamesowi znaczenie zdania nie podaje żaden pojedynczy postulat znaczeniowy jak (16), ani nawet pojedyncze T-zdanie jako warunek jego prawdziwości, lecz cała semantyczna definicja prawdy. Davidson stał bowiem na stanowisku holizmu znaczeniowego. Co więcej, w teorii znaczenia sformułowanej w kategoriach semantycznej definicji prawdy termin „znaczenie” nie musi w ogóle *explicite* występować. Soames do pewnego stopnia akceptował tę tezę, gdyż pisał:

(...) prawdopodobnie najlepiej należy rozumieć [teorię znaczenia] Davidsona nie jako próbę analizowania znaczenia w kategoriach prawdy, lecz jako eliminację pojęcia znaczenia na rzecz pojęcia prawdy (s. 424, przyp. 23).

Czy Davidson faktycznie wyeliminował pojęcie znaczenia na rzecz pojęcia prawdy? On sam pisał, że teoria prawdy w pewnym sensie podaje znaczenia wyrażeń języka. Oczywiście sformułowanie to można rozumieć na wiele sposobów i w kolejnych publikacjach starał się będzie dookreślić rolę semantycznej definicji prawdy w teorii znaczenia.

DODATEK POJĘCIE PRAWDY W MODELU

Nawiązując do słynnego tekstu Tarskiego, Adam Nowaczyk – w artykule *Co naprawdę powiedział Tarski o prawdzie w roku 1933?* (2003) – odróżnia omawianą w tym rozdziale semantyczną definicję prawdy, która sformułowana została przez Tarskiego w pracach z lat 30. i 40. XX w., od późniejszej definicji prawdy w modelu. O ile pierwsza relatywizuje pojęcie prawdy do przekładu z języka przedmiotowego na metajęzyk, to ta późniejsza relatywizuje je do modelu języka. Celem tego paragrafu jest pokazanie, że zdefiniowanie prawdy w modelu pozwala uniknąć wielu problemów, o których była

⁶² Ma zatem stanowić postulat znaczeniowy terminu „znaczenie”.

mowa w tym rozdziale, ale – jak pokazuje Nowaczyk – powoduje, że propozycja przestaje być filozoficznie interesująca.

Nieco wcześniej stwierdziłem, że chociaż semantyczna definicja prawdy początkowo była zaprojektowana w sformułowaniu syntaktycznym, to bez żadnej szkody można ją przedstawić w sformułowaniu semantycznym. Podejście to w pełni zaakceptował Davidson, który w artykule *Reality without Reference* (1977) twierdzi, że podobnie jak teorie fizyczne postulują istnienie pewnych obiektów, tak semantyczna definicja prawdy jedynie postuluje istnienie indywiduów, a zatem faktycznie ma charakter odniesieniowy, gdyż nie wymagamy od żadnego wyrażenia językowego, aby oznaczało jakikolwiek byt. We wczesnych pracach Tarski podzielał milcząco ten pogląd, gdyż w przeciwnym przypadku semantyczna definicja prawdy nie byłaby neutralna filozoficznie, np. implikowałaby konieczność istnienia obiektów matematycznych. A zatem semantyczna definicja prawdy zarówno w sformułowaniu syntaktycznym, jak i semantycznym dokonuje – zgodnie z postulatem fizykalizmu – redukcji pojęć *sensu stricte* semantycznych do języka bądź do pojęć, które faktycznie semantyczne nie są. Przez semantykę rozumiał bowiem Tarski badanie relacji między znakami językowymi, a tym co owe znaki denotują, co oznacza, że semantyka była dla niego z definicji odniesieniowa.

W późniejszym okresie semantyczna definicja prawdy została jednak sformułowana teoriomodelowo, przez co stała się typową semantyką odniesieniową. Nie jest to jednak wersja semantyki odniesieniowej nie do zaakceptowania z punktu widzenia fizykalistycznych założeń Tarskiego, gdyż obiekty w modelu są, podobnie jak w ciągach, jedynie obiektami postulowanymi.

Zdefiniowane przez Tarskiego w *O pojęciu wynikania logicznego* (Tarski 1936a) pojęcie modelu nie jest dokładnie tym pojęciem, które wykształciło się w ciągu lat w związku z teoriomodelową interpretacją semantycznej definicji prawdy. Ponieważ ścisła definicja modelu nie jest nam potrzebna, poniżej postaram się w przybliżeniu określić jak rozumie się model w zastosowaniach związanych z filozofią języka. Aby określić model dla języka przedmiotowego (sam model opisany jest w metajęzyku) wprowadza się dziedzinę interpretacji, którą można rozumieć jako zbiór obiektów postulowanych przez semantyczną definicję prawdy. O obiektach tych nie rozstrzygamy czy faktycznie istnieją, czy też nie istnieją i traktujemy je jako odniesienia terminów jednostkowych (nazw własnych oraz deskrypcji określonych) języka przedmiotowego. Dodatkowo wprowadza się (za Fregem) dwa abstrakcyjne obiekty Prawda (oznaczana zazwyczaj jako *I*) oraz Fałsz (oznaczany zazwy-

czaj jako 0), które pełnią funkcję odniesień zdań⁶³. Odniesieniami predykatów jednoargumentowych są funkcje, których argumentami są obiekty z dziedziny interpretacji, zaś wartościami są 1 oraz 0. Odniesieniami predykatów dwuarargumentowych są funkcje, których argumentami są obiekty z dziedziny interpretacji, zaś wartościami funkcje, których argumentami są obiekty z dziedziny interpretacji, zaś wartościami są 1 oraz 0 itd. Oczywiście możliwe są tutaj alternatywne ujęcia, np. odniesieniami predykatów jednoargumentowych mogą być podzbiory dziedziny interpretacji, odniesieniami predykatów dwuarargumentowych zbiory par obiektów dziedziny interpretacji itd.

Jak pamiętamy, semantyczna definicja prawdy z roku 1933 relatywizowała zakres predykatu „prawdziwy” do języka, a w ogólnym przypadku – do przekładu z języka przedmiotowego na metajęzyk. W przypadku definicji prawdy w modelu rolę przekładu bierze na siebie funkcja interpretacyjna, która przypisuje wyrażeniom języka przedmiotowego ich odniesienia. Dlatego przez model rozumie się tutaj parę złożoną z dziedziny interpretacji oraz funkcji interpretacyjnej. Poza dziedziną interpretacji oraz funkcją interpretacyjną wprowadza się również wartościowania zmiennych, czyli przyporządkowania zmiennym obiektów z dziedziny interpretacji. Wartościowanie zmiennej x_i , która przyporządkowuje jej obiekt a , jest odpowiednikiem przyporządkowania tej zmiennej swojego i -tego wyrazu przez ciąg C . Zdanie jest prawdziwe w modelu, jeśli funkcja interpretacyjna przyporządkowuje mu wartość 1. Na funkcje interpretacyjne nakłada się pewne ograniczenia dotyczące interpretacji znaków logicznych. Jako przykład może nam posłużyć spójnik zdaniowy \wedge , któremu (o ile w danym modelu interpretowany jest standardowo, tj. jako koniunkcja), każda funkcja interpretacyjna przyporządkowuje jako odniesienie funkcję przyporządkowującą parze $(1, 1)$ wartość 1, zaś pozostałym parom wartości logicznych wartość 0. Ograniczenia nakładane na funkcje interpretacyjne pozwalają na wprowadzenie pojęcia zdania logicznie prawdziwego, tj. zdania prawdziwego w każdym modelu. Warto na to zwrócić uwagę, gdyż semantyczna definicja prawdy w interesującym nas, pierwotnym sformułowaniu nie pozwala na takie rozróżnienie.

⁶³ Chodzi tu oczywiście o zdania oznajmujące. Różnica w stosunku do Fregego polega na tym, że Frege utrzymywał, że niektóre zdania oznajmujące (np. o fikcji literackiej) nie mają odniesienia, tj. wartości logicznej, gdyż nazwy przedmiotów fikcyjnych, które w nich występują, same nie mają odniesienia. W semantykach teoriomodelowych języka naturalnego każda nazwa ma odniesienie, którym jest jakiś element dziedziny interpretacji. W konsekwencji każde zdanie języka przedmiotowego jest albo prawdziwe, albo fałszywe (w danym modelu), o ile nie zmodyfikujemy pojęcia modelu i nie wprowadzimy dla zdań o fikcji trzeciej wartości logicznej.

Poza matematyką definicja prawdy w modelu znalazła również zastosowania w filozofii języka. Stanowi ona podstawę teoriomodelowych semantyk budowanych przez Richarda Montague dla języka naturalnego⁶⁴. Warto, choćby pobieżnie, przyjrzeć się semantyce teoriomodelowej dla ograniczonego fragmentu języka naturalnego, gdyż semantyki tego typu były bardzo zdecydowanie krytykowane przez Donalda Davidsona⁶⁵.

Namiastkę gramatyki Montague można przedstawić rozpatrując następujący przykład. Niech język przedmiotowy L składa się z trzech nazw własnych „Jaś”, „Małgosia” i „Piotr” oraz jednoargumentowego predykatu „pali”, zaś dziedzina interpretacji jest zbiorem złożonym z trzech obiektów: j , m oraz p . Niech funkcja interpretacyjna I będzie taka, że $I(„Jaś”) = j$, $I(„Małgosia”) = m$, $I(„Piotr”) = p$ oraz $I(„pali”) = p$ oraz $I(„pali”) = p$ będzie funkcją, która obiektom j oraz p przyporządkowuje 1 , zaś obiektowi m przyporządkowuje 0 (co oznacza ni mniej ni więcej, że obaj chłopcy palą, zaś Małgosia nie)⁶⁶. W języku tym można zbudować trzy zdania „Jaś pali”, „Małgosia pali” oraz „Piotr pali”. Jaki będzie zakres predykatu „prawdziwy” dla zdań języka L ? Ponieważ $I(„Jaś”) = j$, zaś $I(„pali”) = 1$ jest funkcją, która obiektowi j przyporządkowuje 1 , $I(„Jaś pali”) = 1$. Analogicznie otrzymuje my $I(„Małgosia pali”) = 0$ oraz $I(„Piotr pali”) = 1$. W konkwencji zakresem predykatu „prawdziwy” dla języka L w powyższym modelu jest zbiór zdań „Jaś pali” i „Piotr pali”.

Przydatność definicji prawdy w modelu nie ulega wątpliwości w przypadku zastosowań metamatematycznych. Spore kontrowersje sprawia natomiast oparta na niej semantyka teoriomodelowa języka naturalnego w stylu Montague. Jest to niewątpliwie teoria odniesienia zrelatywizowana do modelu, który jest arbitralny zarówno ze względu na dziedzinę interpretacji, której nie można utożsamiać ze zbiorem realnie istniejących rzeczy, jak i ze względu na funkcję interpretacyjną (nie jesteśmy w dziedzinie złożonej ze wszystkich ludzi wskazać odniesienia predykatu „pali”, gdyż wymagałoby to wiedzy o każdym człowieku, który żyje na świecie). Zaleta takiej semantyki polega jednak na czymś innym. Pozwala ona bowiem dobrać typ odniesienia wyrażień określonych kategorii syntaktycznych tak, aby odniesieniami zdań, w których

⁶⁴ Gramatyki Richarda Montague (zwane również semantykami Montague) były sformułowane w szeregu jego artykułów zebranych w (Thomason 1976).

⁶⁵ Problem ten omówię dokładnie w rozdziale 3.

⁶⁶ Odniesienie funktora musi być funkcją, której typ musi być zgodny z pochodzącą również od Fregego zasadą składalności odniesienia. Zasada ta głosi, że odniesienie wyrażenia złożonego z funktora i jego argumentu jest wartością funkcji, która jest odniesieniem funktora, dla odniesienia jej argumentu, co w postaci symbolicznej da się wyrazić za pomocą równości $I(F(x)) = I(F)(I(x))$. W naszym przypadku $I(„pali”) = 1$ musi być funkcją, która obiektom z dziedziny interpretacji przyporządkowuje wartości logiczne.

te wyrażenia występują, były wartości logiczne. W naszym przykładzie odniesienie predykatu „pali” musiało być tak dobrane, aby odniesieniem zdania postaci „ X pali” była wartość logiczna. A zatem semantyki Montague pozwalają się czegoś dowiedzieć na temat zależności między możliwymi odniesieniami wyrażen różnych kategorii syntaktycznych oraz wyjaśnić niektóre zjawiska występujące w języku naturalnym.

Należy również zauważyć, że semantyka teoriomodelowa nie może być traktowana jako teoria znaczenia wyrażen języka przedmiotowego. W przypadku tzw. stałych logicznych znajomość znaczeń się zakłada, co wpływa na wybór funkcji interpretacyjnej. Na przykład przypisanie symbolowi \wedge podanej wcześniej funkcji jako odniesienia oznacza, że przyporządkowaliśmy mu znaczenie takie samo, jakie w sposób intuicyjny przypisujemy koniunkcji. W przypadku wyrażen pozalogicznych, np. predykatów, znaczenia przypisane im w języku przedmiotowym nie mają wpływu na wybór funkcji interpretacyjnej, a w konsekwencji na ich odniesienie w modelu, np. odniesienie predykatu „pali” zostało wybrane arbitralnie i nie pozostaje w jakimkolwiek związku ze znaczeniem tego wyrażenia. W szczególności odniesieniem tego predykatu może być – w danej dziedzinie – funkcja pełna (wszystkie obiekty dziedziny palą) lub pusta (żaden obiekt nie pali).

ROZDZIAŁ 2

TEORIA ZNACZENIA A WYUCZALNOŚĆ

(...) wielu współczesnych teorii znaczenia nie da się, nawet w zasadzie, zastosować do języków naturalnych, gdyż języki, do których się te teorie stosuje, nie są wyuczane

(Davidson 1984/1965, s. 9).

Theories of Meaning and Learnable Languages (1965) jest jedną z pierwszych publikacji Davidsona, a jednocześnie jego pierwszą znaną publikacją z zakresu filozofii języka, w której przygotowuje grunt pod swój program semantyczny. Nie jest to tekst łatwy, gdyż Davidson był wymagający wobec swoich czytelników, a jego wywody mają – co podkreśla wielu komentatorów – charakter eliptyczny. Jednak przy uważnej lekturze można odnaleźć w nim bardzo ciekawe i inspirujące wątki filozoficzne oraz wnikliwą krytykę zastanych stanowisk w kwestii teorii znaczenia wyrażeń językowych. Rozważania wokół wyuczalności języka pozwoliły Davidsonowi sformułować tezę, że teoria znaczenia języka naturalnego powinna być rekurencyjna oraz – co zasugerował w tym tekście, a dokładnie uzasadnił później w *Prawdzie i znaczeniu* (1967) – powinna przybrać postać Tarskiego semantycznej definicji prawdy.

Zasadniczym celem *Theories of Meaning* nie było przedstawienie własnej teorii znaczenia, lecz wyciągnięcie wniosków na temat kształtu teorii znaczenia języka naturalnego ze skądinąd oczywistego stwierdzenia, że posługiwanie się językiem naturalnym jest nieskończoną dyspozycją nabytą w skończonym procesie uczenia się języka¹. Cechę języka, która umożliwia jego

¹ Jest to dyspozycja nieskończona, gdyż pozwala na rozumienie i wypowiedzanie (ze zrozumieniem) potencjalnie nieskończenie wielu zdań. Proces uczenia się języka jest skończony, gdyż odbywa się w skończonym czasie, a co za tym idzie, opiera się na skończonym materiale empirycznym. Każdy użytkownik języka zna bowiem skończoną ilość słów i zetknął się ze skończoną ilością zdań.

użytkownikowi posiadanie owej dyspozycji nazywa Davidson wyuczalnością, cechą – dodajmy – oczywistą i niekontrowersyjną. Okazuje się jednak, że nie wszyscy filozofowie języka wyciągnęli z faktu wyuczalności języka właściwe konsekwencje co do teorii języka naturalnego.

W pierwszej części *Theories of Meaning* Davidson krytykuje niewłaściwe i zbyt daleko idące – jego zdaniem – wnioski z wyuczalności języka wyprowadzane przez empirystów. W drugiej, zdecydowanie dłuższej części, krytykuje kilka wybranych teorii znaczenia (nie tylko empirystycznych), które z wyuczalnością języka naturalnego się nie liczą. Jak to się często zdarza w polemikach zawartych w pismach Davidsona, bardziej niż głoszona teza interesujący jest sposób argumentacji, który poza bardzo drobiazgową analizą, wydobywa z krytykowanych poglądów to, co – jego zdaniem – jest w nich cenne. Ponieważ wywody Davidsona są niezwykle skompresowane, rozwijam jego argumenty w tych miejscach, w których mogą być niezrozumiałe oraz dodaję własne przykłady i wyjaśnienia. Ma to również stanowić okazję do krytycznego przyjrzenia się kilku dobrze znanym teoriom znaczenia.

Davidson, podobnie jak jego nauczyciel Quine, był empirystą i podobnie jak Quine zdawał sobie doskonale sprawę ze słabych punktów tego stanowiska. W pierwszej części *Theories of Meaning* Davidson rozprawia się z empirystycznymi teorią języka, które – najogólniej rzecz ujmując – utożsamiają porządek epistemologiczny związany z uzasadnianiem zdań (tez, twierdzeń itp.) z porządkiem uczenia się języka. Utożsamienie to prowadzi – zdaniem Davidsona – do całkowicie błędnych teorii języka. Zarzuty te nie powinny być jednak w żadnym wypadku rozumiane jako odrzucenie przez Davidsona empiryzmu jako postawy filozoficznej. Davidson podąża tu drogą wyznaczoną przez Quine'a w *Dwóch dogmatach empiryzmu*, który – nie odrzucając empiryzmu – wskazywał na nieuprawnione, a przy tym często nieuświadomione i przyjmowane bezkrytycznie, dogmaty związane z tym stanowiskiem (por. Quine 1951).

Ponieważ krytyka przeprowadzona przez Davidsona w pierwszej części artykułu dotyczy dość zaawansowanego pojęciowo poglądu prezentowanego przez Petera Strawsona w artykule *Singular Terms, Ontology and Identity* (Strawson 1956) bez nawiązania do innych, o wiele lepiej znanych teorii empirystycznych, warto wcześniej przyjrzeć się pokrótce poglądom empirystów na język. Na początku przyjrzymy się Arystotelesowi teorii znaku językowego, która była ściśle związana z jego empirystyczną epistemologią, oraz teoriom języka, które powstały pod wpływem Arystotelesa. Krytyka przeprowadzona przez Davidsona w *Theories of Meaning* bezpośrednio do tradycyjnych empirystycznych teorii znaczenia się nie odnosi, ale problemy, na które zwraca uwagę można już dostrzec u samego Arystotelesa, a jeszcze wyraźniej

w ideacyjnej teorii znaczenia (zwanej niekiedy asocjacionistyczną teorią znaczenia) pochodzącej od empirystów angielskich (J. Locke i D. Hume) oraz w konotacyjnej teorii znaczenia J. S. Milla. Można je również dostrzec u Quine'a, którego Davidson niezwykle cenił i prawie nigdy nie krytykował wprost.

1. EMPIRYSTYCZNE TEORIE ZNACZENIA I ICH KRYTYKA

Przegląd empirystycznych teorii znaczenia rozpoczynamy od teorii znaku Arystotelesa, którą można zrekonstruować na podstawie krótkich fragmentów rozproszonych w różnych pismach². Najbardziej znany fragment, w którym Arystoteles naszkicował teorię, którą nazwalibyśmy obecnie teorią znaczenia, pochodzi z *Hermeneutyki*:

Słowa są symbolicznymi znakami wrażeń doznawanych w duszy, a dźwięki pisane są znakami dźwięków mówionych. Jako pisane a także jako mówione nie są dla wszystkich ludzi te same. Ale to, czego przede wszystkim są znakami, wrażeń doznawanych w duszy, jest takie samo dla wszystkich; a więc i to, do czego są te wrażenia podobne, mianowicie rzeczy są również takie same (Arystoteles 1990, s. 69).

Owe wrażenia doznawane w duszy pełnią u Arystotelesa funkcję znaczenia.

Aby wyjaśnić, dlaczego rozumiemy nazwę ogólną „pies”, „kot”, „człowiek” itp. (dla uproszczenia ograniczmy się do nazw gatunków biologicznych), korzystamy – zdaniem Arystotelesa – z umiejętności abstrahowania, która pozwala uchwycić w poszczególnych indywiduach³ (psach, kotach, ludziach), to co mają ogólnego. Abstrahowanie jest operacją umysłu, która pełni rolę podwójną. Z jednej strony pozwala tworzyć lub chwytąć pojęcia ogólne, co stanowi integralny element uczenia się języka i zdobywania wiedzy, z drugiej zaś strony pozwala osobie znającej już język orzekać powszechniki o indywiduach, czyli rozpoznawać indywidua jako przynależne do określonych gatunków. Przez powszechnik (gatunek⁴) rozumiemy odniesienie nazwy ogólnej, czyli – wedle Platona oraz Arystotelesa – realnie istniejący byt idealny, zaś przez pojęcie – odpowiadającą mu treść nazwy

² Rekonstrukcję Arystotelesa teorii znaku językowego można znaleźć w (Modrak 2001 i 2006).

³ W ontologii Arystotelesa indywidua noszą nazwę substancji pierwszych.

⁴ Oczywiście nie każdy powszechnik (uniwersale) jest gatunkiem. Jednak sam Arystoteles zajmował się przede wszystkim gatunkami, co poniekąd usprawiedliwia ograniczenie się do gatunków (w terminologii Arystotelesa substancji wtórych).

ogólnej, czyli – stosując terminologię z *Hermeneutyki* – „wrażenie [rzeczy] doznawane w duszy”.

Rozumiane na oba wymienione sposoby abstrahowanie pozwala nie tylko orzekać gatunki o indywiduach, lecz również rodzaje o gatunkach. Mamy wówczas do czynienia ze zdaniami ogólnymi, np. „Każdy pies jest zwierzęciem”, które za pomocą indukcji wyprowadzamy ze zdań jednostkowych, tj. zdań o indywiduach (np. o poszczególnych psach, kotach czy ludziach).

To, co może niepokoić w ostatnim sformułowaniu, to pewna niejednoznaczność terminu „wyprowadzić”. Może to bowiem oznaczać, że zdania ogólne uzasadniamy w oparciu o zdania jednostkowe, odwołując się do indukcji rozumianej jako pewna forma rozumowania, dająca się wyrazić w języku, lub że owe zdania odkrywamy, tj. dochodzimy do przekonania o ich prawdziwości poprzez abstrahowanie, które rozumiemy jako operację umysłu. Pierwszy przypadek zakłada uprzednią znajomość języka, w drugim przypadku mamy do czynienia z elementem procesu uczenia się języka. Przez uczenie się języka rozumiem tutaj nie tylko proces nabywania znajomości języka naturalnego, ale również proces tworzenia lub uczenia się teorii naukowej, który polega również na tworzeniu lub uczeniu się nowych pojęć w języku tej teorii. W obydwu przypadkach uczenie się języka wiąże się z nabywaniem przekonań (w języku teorii naukowej byłyby to twierdzenia), które da się w tym języku wyartykułować jako asercje⁵. Oczywiście nie chodzi tu o wszystkie przekonania, ale o te, które pełnią funkcje tzw. postulatów znaczeniowych. O kimś, ktoś nie wie, że psy są zwierzętami, nie powiemy, że jest słaby z biologii, lecz raczej, że nie rozumie nazw „pies” lub „zwierzę”. Podobnie kogoś, kto nie wie, że dwa i dwa to cztery, nie uznamy za słabego matematyka, lecz powiemy, że nie rozumie liczebników języka polskiego⁶. A zatem uczenie się języka związane jest nieodłącznie z czymś, co zazwyczaj nazywa się porządkiem odkrycia⁷.

Typowy błąd empirystów polega – zdaniem Davidsona – na utożsamieniu tych dwóch zupełnie różnych porządków. Z jednej strony mamy do czynienia z apriorycznym, logicznym czy też epistemologicznym porządkiem

⁵ Wszystkie teorie znaczenia wyrażen językowych wiążą znaczenia z jakąś wiedzą o świecie, którą da się w języku wyartykułować. Dlatego też – jak twierdził Frege – znaczenie jest nośnikiem wiedzy.

⁶ Związek znajomości języka z wiedzą podkreślał w dyrektywnej teorii znaczenia Kazimierz Ajdukiewicz (por. 1931).

⁷ Wykorzystanie terminów „porządek odkrycia” oraz „porządek uzasadniania” pochodzi ode mnie. Jestem jednak przekonany, że pojęcia te doskonale współgrają z poglądami Davidsona wyrażonymi w *Theories of Meaning*.

uzasadniania⁸, a z drugiej, z przyczynowym porządkiem uczenia się języka i zdobywania wiedzy o świecie, która może być wyartykułowana jedynie w języku⁹:

Często się twierdzi lub wnosi, że czysto aprioryczne rozważania wystarczą do określenia mechanizmu nauki języka lub wyróżnienia jej etapów; takie stwierdzenia wydają się jednak wątpliwe. W pierwszej części tego artykułu rozważam typowy przykład takiego stanowiska i próbuję oddzielić to, co jest w nim do zaakceptowania, a co nie (1984/1965, s. 3).

Opisywany przez empirystów sposób weryfikowania lub uzasadniania zdań, wyrażający porządek epistemologiczny, nie powinien być zatem rzutowany na proces nauki języka: „(...) nie ma żadnego oczywistego powodu, dla którego należy sądzić, że proces nauki [języka] jest związany z pierwszeństwem epistemologicznym” (s. 4).

Na początku paragrafu odwoływałem się do teorii znaku Arystotelesa (nieco dalej przedstawię dokładniej pewne jej konsekwencje), która jest chronologicznie najwcześniejszą teorią znaczenia. Już w tej teorii można dostrzec to, co stanowiło – wedle Davidsona – główną wadę empirystycznych teorii znaczenia. Jednak w teoriach empirystycznych Davidson dostrzega również pewne zalety. Podstawowa polega na próbie uwzględnienia faktu, że znajomość języka jest dyspozycją, którą nabywamy w procesie uczenia się.

Aby nie popaść w błędy empirystów, przy budowaniu teorii znaczenia języka naturalnego należy starannie odróżnić dwa zagadnienia związane z uczeniem się języka. Po pierwsze, teoria znaczenia musi uwzględniać to, że języka można się w ogóle nauczyć. W tej kwestii Davidson jest całkowicie z empirystami zgodny. Wedle Davidsona, podobnie zresztą jak w tym samym czasie dla Noama Chomskiego, istotną cechą języka naturalnego jest jego wyuczalność. Cecha ta musi być uwzględniana przez każdą poprawną teorię znaczenia języka naturalnego:

W przeciwieństwie do mętnych intuicji na temat sposobów uczenia się języka, proponuję coś, co jest konieczną cechą charakterystyczną języka wyuczanego: musi być możliwe podanie konstruktywnego opisu (*constructive account*) znaczenia zdań tego języka. Taki opis nazywam teorią znaczenia (s. 3).

Fakt wyuczalności powinien zatem być uwzględniony w „konstruktywnym opisie” języka, tj. takim opisie, który wyjaśnia w jaki sposób ze skończonej ilości słów możemy wyprowadzić nieskończoną ilość zdań.

⁸ Terminy te używane były przez Davidsona wymiennie.

⁹ Davidson utrzymywał, że bez języka nie może istnieć żadna wiedza, ani żadne przekonania.

W ten sposób dochodzimy do zagadnienia, na temat którego Davidson z empirystami polemizuje. Zdaniem Davidsona proces nauki języka należy do porządku przyczynowego – pozwoliłem sobie go nazwać nieco wcześniej porządkiem odkrycia – którego dokładny opis nie musi, a nawet nie powinien, być w teorii znaczenia uwzględniany, gdyż opis tego procesu jest raczej sprawą psychologii niż semantyki. Należy tutaj podkreślić, że w całej swej twórczości Davidson bardzo starannie odróżniał wszelkie operacje na języku (w tym operacje logiczne w wąskim tego słowa znaczeniu, np. przechodzenie od przesłanek do wynikającego z nich wniosku) od związków przyczynowych zachodzących między zdarzeniami. Uczenie się języka należy, przynajmniej częściowo, do porządku przyczynowego.

Theories of Meaning Davidson rozpoczyna od stwierdzenia, że filozofowie bardzo często wskazują na te właściwości języka, które zapewniają mu wyuczalność. Własności te miały zazwyczaj „wspierać lub obalać jakąś doktrynę filozoficzną, epistemologiczną, metafizyczną, ontologiczną lub etyczną” (s. 3). Ponieważ są to charakterystyczne dla filozofii rozważania natury apriorycznej, ma to niewątpliwie przełożenie na teorię języka:

Jeżeli argumenty [na rzecz doktryny filozoficznej, odwołujące się do własności języka zapewniające mu wyuczalność] są mocne, muszą pociągać za sobą konsekwencje dotyczące empirycznej nauki budowania pojęć (*empirical science of concept formation*) – być może są to konsekwencje polegające jedynie na określeniu granic tego co empiryczne (*ibidem*).

Termin „empiryczna nauka budowania pojęć” jest tutaj rozumiany jako empirystyczna teoria znaczenia, w której opis procesu tworzenia pojęć złożonych z pojęć prostych przebiega wedle ustalonego *a priori* porządku epistemologicznego. Porządek taki zakładał zarówno Arystoteles, jak i empiryści angielscy. Jak pokazał Davidson opis ten jest całkowicie błędny, zaś ewentualny opis tego procesu formułowany przez psychologię, nie powinien być w teorii znaczenia uwzględniany.

Z niewłaściwych założeń na temat wyuczalności, a dokładniej z przekonania, że wyuczalność da się uwzględnić w teorii znaczenia jedynie poprzez włączenie do teorii znaczenia opisu procesu uczenia się języka utożsamianego z porządkiem uzasadniania, bierze się charakterystyczne dla empiryzmu założenie, że „rozważania *a priori*” wystarczą do określenia mechanizmu tego procesu¹⁰. Krytyka tego stanowiska nie oznacza jednak, że na poprawną teorię

¹⁰ Wprawdzie wyuczalność języka wydaje się faktem oczywistym, ale w tradycji empirystycznej pojawiały się teorie, które ten wymóg ignorowały. W słynnej recenzji książki Skinnera *Verbal behavior* z roku 1957, Chomsky zwrócił uwagę na to, że behawiorystyczny (czyli w istocie

znaczenia nie powinno się nakładać ograniczeń formułowanych *a priori*. W tej sprawie Davidson podążył śladami Noama Chomsky'ego, którego do empirystów zaliczyć się nie da. Nie ma w tym jednak nic dziwnego, gdyż uwzględnienie wyuczalności języka w teorii znaczenia przejawia się – zdaniem Davidsona – narzuceniem na nią apriorycznych ograniczeń. Odwołując się do terminologii wprowadzonej przez Chomsky'ego, Davidson twierdzi, że znajomość języka oparta jest na dwóch rodzajach kompetencji językowej: syntaktycznej i semantycznej, które można *a priori* opisać w ramach teorii języka¹¹.

Zdaniem Davidsona jakakolwiek akceptowalna teoria znaczenia nie może być teorią całkowicie aprioryczną, gdyż musi podlegać weryfikacji empirycznej wobec dostępnych świadectw. Całkowicie aprioryczne są natomiast warunki narzucone na klasę adekwatnych teorii znaczenia. Tego samego zdania był Michael Dummett, odróżniając od teorii znaczenia tzw. teorię znaczeniową:

Używam zwrotu „teoria znaczeniowa” jako odpowiednika „teorii poznania”, dla określonej gałęzi filozofii znanej też pod mniej znaną nazwą „filozofia języka”. Aby odróżnić ją od tego, o czym Davidson i inni mówią jako o „*pewnej* teorii znaczenia”, a więc od pełnego wykazu znaczeń wszystkich słów i wyrażen jednego szczególnego języka, dla określenia tej drugiej użyję wyrażenia „teoria znaczenia”. Zgadzam się z Davidsonem, że właściwa metoda badawcza [prowadząca do] teorii znaczeniowej polega na poszukiwaniu ogólnych zasad, stanowiących podstawę konstrukcji teorii znaczenia (Dummett 1998/1976, s. 40)¹².

Adekwatna teoria znaczenia języka naturalnego, teoria, która jest weryfikowalna empirycznie, musi spełniać ogólne, aprioryczne warunki narzucone przez teorię znaczeniową. Takim apriorycznym warunkiem ma być właśnie

skrajnie empirystyczny) opis posługiwania się językiem, który odwołuje się do pojęć uwarunkowania i wzmocnienia, nie jest w stanie wyjaśnić ludzkich zachowań językowych (Chomsky 1959). Podstawowy zarzut wobec teorii Skinnera polega na tym, że teoria ta ignoruje produktywność lub – trzymając się terminologii Davidsona – wyuczalność języka. Zgodnie z krytykowaną teorią behawioralną, aby zrozumieć zdanie, należy skorelować je z odpowiednimi bodźcami, co nie pozwala wyjaśnić, w jaki sposób rozumiemy te zdania znanego nam języka, z którymi stykamy się po raz pierwszy. Chomsky poszukiwał cech języka warunkujących jego wyuczalność na gruncie syntaktycznym, zaś Davidson również na gruncie semantycznym. Obaj natomiast byli zgodni w krytyce prób opisu „mechanizmu” uczenia się języka, który był charakterystyczny dla behawioryzmu, będącego jednym z przejawów empiryzmu w filozofii języka.

¹¹ Semantyczna definicja prawdy opisuje oba rodzaje kompetencji.

¹² Rozważając w *Theories of Meaning* aprioryczne warunki nakładane na teorię znaczenia, Davidson formułuje w istocie teorię znaczeniową w sensie Dummetta.

możliwość „konstruktywnego opisu”, będąca konsekwencją wyuczalności języka. Możliwość konstruktywnego opisu zapewniona jest w przypadku semantycznej definicji prawdy jej rekurencyjnością. Wedle Davidsona empiryści wyciągnęli jednak błędne wnioski z faktu wyuczalności języka, uznając, że teoria znaczenia uwzględnia wyuczalność tylko wtedy, gdy jednocześnie podaje opis procesu nauki języka. Opis ten – zdaniem Davidsona – nie tylko był niepotrzebny, ale był również całkowicie błędny, gdyż miał charakter właściwy dla apriorycznego porządku uzasadniania. W konsekwencji empiryści nie byli w stanie podać ani poprawnej teorii znaczenia, ani poprawnej teorii uczenia się języka. Przyjrzyjmy się zatem, jakie są tego konsekwencje.

Wstępny zarzut formułowany przez Davidsona przeciw empirystycznym teoriom znaczenia dotyczy tego, że opis procesu nauki języka – od nazw realnie istniejących obiektów do pojęć ogólnych – nie zawsze ma miejsce. Każdy użytkownik języka naturalnego jest bowiem w stanie używać fikcyjnych pojęć ogólnych, np. „centaur” lub „złoty górą”, chociaż odpowiadające im indywidua po prostu nie istnieją. Dzieci często uczą się pojęć ogólnych bez bezpośredniego kontaktu z odpowiednimi obiektami (lwy, wilki, dinozaury, mastodonty), przeglądając książeczki dla dzieci, słuchając bajek lub oglądając filmy. Oczywiście zarzut ten może być przez empirystów bardzo łatwo odparty i Davidson z pewnością zdawał sobie z tego sprawę¹³. W tym celu przywołajmy ponownie teorię znaku Arystotelesa, wedle którego predykaty puste (tj. nieoznaczające realnie istniejących indywiduów), np. „centaur”, którym w ramach jego ontologii nie odpowiadają byty ogólne (powszechniki), nie mają tzw. definicji realnych, lecz jedynie podające ich znaczenia definicje nominalne (por. Modrak 2006). W ramach ontologii Arystotelesa odniesieniami predykatów, które można orzekać o realnie istniejących indywiduach, są bowiem powszechniki, zaś predykaty fikcyjne, których nie da się orzec o żadnych realnie istniejących indywiduach, tak rozumianych odnie-sień po prostu nie mają. Oba rodzaje predykatów posiadają natomiast znaczenia, które Arystoteles utożsamiał z pewnymi bytami mentalnymi, czyli pojęciami, wyobrażeniami lub „wrażeniami w duszy”.

Wprowadzenie psychologistycznie rozumianego znaczenia pozwoliło Arystotelesowi na podanie warunku sensowności predykatów fikcyjnych, gdyż definicja nominalna terminów „centaur” lub „jednorożec” jest możliwa dzięki dokonywanym w umyśle operacjom na znaczeniach innych predykatów, które można orzekać o realnie istniejących indywiduach. Widać to jeszcze wyraźniej

¹³ Krytykowane przez Davidsona poglądy Strawsona uwzględniają już rozwiązanie tej pozornej trudności.

u empirystów angielskich. Locke i Hume odrzucili wprawdzie Arystotelesowską ontologię (szczególnie powszechniki), lecz nie odrzucili pomysłu, że znaczeniem predykatu jest byt mentalny, zwany przez nich ideą. Arystotelesowskim definicjom nominalnym, które przynależą do epistemologicznego porządku uzasadnienia, odpowiadają u nich operacje na ideach, które pozwalają na budowanie takich pojęć fikcyjnych. Błądność tego rozwiązania polegałaby – zdaniem Davidsona – na ścisłym skorelowaniu definicji nominalnej, podającej znaczenie predykatu, z pewną operacją na bytach mentalnych. Jeśli bowiem definiujemy centaury jako konia z ludzkim tułowiem, to w pewien sposób łączymy w umyśle wyobrażenie człowieka z wyobrażeniem konia. Ponieważ predykaty „człowiek” i „koń” fikcyjne nie są, odpowiednie znaczenia tych predykatów „chwytamy” przez abstrakcję, obserwując odpowiednie obiekty¹⁴. W konsekwencji każde pojęcie jakim operujemy, nawet pojęcie, które nie odpowiada bytom istniejącym, jest zakotwiczone w rzeczywistości w opisany sposób. Davidson krytykuje ten pogląd, gdyż epistemologiczny porządek uzasadniania, na którym jest on oparty, nie opisuje faktycznego procesu tworzenia pojęć, a jedynie podaje uzasadnienie ich istnienia w naszych umysłach.

W dalszej części tym staram się pokazać, że poglądy Davidsona na temat empirystycznych teorii znaczenia są doskonale egzemplifikowane w tradycyjnych poglądach empirystów na język. Podsumowując warto zauważyć, że rozwiązania charakterystyczne dla empiryzmu prowadzą w kwestii znaczenia zawsze do pewnej wersji redukcjonizmu, która polega na wyróżnieniu znaczeń prostych lub atomowych, np. u Locke’a byłyby to idee proste¹⁵. Tendencja redukcjonistyczna była szczególnie widoczna w weryfikacjonistycznej teorii znaczenia, dominującej przez pewien czas wśród filozofów tzw. Koła wiedeńskiego. Wprawdzie weryfikacjoniści mówili o znaczeniach zdań, nie zaś o znaczeniach nazw, ale za warunek sensowności zdania uznali jego redukowalność do tzw. zdań protokolarnych, które odpowiadają bezpośrednio pewnym danym doświadczenia. Przekonanie o możliwości odkrycia zdań protokolarnych było jednym z dwóch dogmatów empiryzmu skrytykowanych przez Quine’a w (1951). Krytykując empirystyczny dogmat o istnieniu zdań protokolarnych (oraz powiązany z nim dogmat o ścisłym

¹⁴ Empiryzm nie jest poglądem jednolitym. Wedle Locke’a, a jeszcze bardziej Hume’a, nie ma podstaw do mówienia o indywiduach, lecz o impresjach (które współcześnie odpowiadają bodźcom). Uwzględnienie tych różnic nie jest jednak istotne z punktu widzenia omawianej tu krytyki empiryzmu przedstawionej przez Davidsona.

¹⁵ W przypadku teorii konotacyjnej Milla funkcje atomów znaczeniowych (dla nazw) pełnią cechy. Warto jednak zauważyć, że w przypadku zdań, ich znaczenia podawane są przez warunki prawdziwości (por. Mill 1843, rozdz. 2 „O znaczeniu zdań”).

rozróżnieniu między zdaniem analitycznymi i syntetycznymi) Quine wprowadził tezę holizmu confirmacyjnego, która jest znana również jako teza Duhema-Quine'a. W myśl owej tezy, te same dane doświadczenia mogą posłużyć do potwierdzania niewspółmiernych teorii naukowych, zaś każde nowe doświadczenie może być w pewien sposób uzgodnione ze wszystkimi konkurencyjnymi teoriami. Teoria naukowa została metaforycznie opisana przez Quine'a jako pole sił dostosowujące się do tego, co dzieje się na jego brzegu (por. Quine 2000/1951, s. 71–75).

Uwzględniając wspomniane wcześniej powiązanie języka z wiedzą, można uznać, że język naturalny jest pewną popularną (ludową) teorią rzeczywistości. Z tego punktu widzenia odpowiednikiem tezy holizmu confirmacyjnego jest hipoteza Sapira-Whorfa, która głosi, że język naturalny determinuje obraz świata jego użytkowników. W konsekwencji obraz świata związany z wybranym językiem może być niedostępny dla użytkowników innych języków. Dlatego też niektóre odległe geograficznie, środowiskowo oraz kulturowo języki (np. język angielski i język Indian Hopi) mogą być, zdaniem autorów hipotezy, całkowicie nieprzetłumaczalne. Bardziej umiarkowane stanowisko głosił Quine w tezie o niezdeterminowaniu przekładu, twierdząc, że przekład jest wprawdzie możliwy, ale nie jest zdeterminowany przez ogół bodźców¹⁶. Zarówno hipoteza Sapira-Whorfa, jak i teza o niezdeterminowaniu przekładu zakładają istnienie tzw. schematu pojęciowego. Sensowność tego pojęcia zakwestionował Davidson w artykule *O pojęciu schematu pojęciowego* (1974), akceptując jednak za Quinem holistyczny aspekt języka, czego konsekwencją jest pojawiające się w *Theories of Meaning* stwierdzenie, że język stanowi organiczną całość. Stwierdzenie jest konsekwencją jego krytyki poglądów Strawsona¹⁷.

2. KRYTYKA POGLĄDÓW STRAWSONA

Po wstępnym zarysowaniu problemów związanych z empirystycznymi teoriami języka możemy przejść do przeprowadzonej przez Davidsona krytyki bardzo szczegółowej argumentacji Strawsona, która oparta jest na charakterystycznych dla empiryzmu, a zdaniem Davidsona błędnych, założeniach na temat

¹⁶ Teoria znaczenia w kategoriach podręcznika przekładu, dla której Quine sformułował słynną tezę o niezdeterminowaniu, została przedstawiona w *Słowie i rzeczy* (1960). Była to, podobnie jak teoria Davidsona, teoria holistyczna. Do teorii Quine'a powrócimy w ostatnim rozdziale.

¹⁷ Drugą konsekwencją klasycznej wersji empirystycznej teorii znaczenia było reifikowanie znaczeń nazw jako bytów mentalnych. Davidson nie odniósł się bezpośrednio do tej kwestii w *Theories of Meaning*, przeprowadzając gruntowną krytykę teorii znaczenia utożsamiających znaczenia z pewnymi bytami w *Prawdzie i znaczeniu*.

nauki języka. Krytykowana argumentacja skierowana jest przeciwko głoszonej przez Quine'a tezie o eliminowalności terminów jednostkowych (tj. nazw własnych i deskrypcji określonych) z języka. Nie wnikając w słuszność tezy głoszonej przez Quine'a, gdyż jest to kwestia, którą Davidson się nie zajmuje, wystarczy powiedzieć, że wedle Strawsona terminy jednostkowe są wprawdzie eliminowane z danego języka, ale nikt nie jest w stanie takiego języka (tj. języka pozbawionego terminów jednostkowych) się nauczyć. Przedmiotem krytyki Davidsona nie jest przy tym powyższe stwierdzenie, lecz argumenty, które zostały użyte przez Strawsona do jego uzasadnienia. Argumenty te Davidson rekonstruuje w sposób następujący:

Dwa podstawowe założenia na rzecz tego argumentu są następujące:

- (1) aby zrozumieć jakikolwiek predykat, niektóre predykaty muszą być nauczone przez ostensję lub „bezpośrednią konfrontację”;
- (2) aby taka nauka mogła mieć miejsce, nauczanie przez ostensję musi być „wyartykułowane w języku” za pomocą elementu demonstratywnego, który wybiera [*picks out*] lub identyfikuje indywidua typu, do którego stosuje się predykat (s. 5)¹⁸.

Zauważmy, że założenie (1) wyraża typowy pogląd charakterystyczny dla omawianych wcześniej empirystycznych teorii znaczenia. Aby zrozumieć nazwę „centaur” należy nauczyć się przez ostensję znaczenia nazwy „koń” oraz znaczenia nazwy „człowiek”. Aby jednak nauczyć się tych predykatów przez ostensję – założenie (2) – należy dysponować środkami językowymi służącymi do nazywania indywiduów, czyli poszczególnych koni i ludzi. Mogą to być nazwy własne, np. „Bucefal” lub „Sokrates”, które arbitralnie wybierają (*pick out*) indywidua, lub deskrypcje określone, np. „rumak bojowy Aleksandra Wielkiego” lub „nauczyciel i mentor Platona”, które służą do identyfikacji owych indywiduów przez odwołanie się do posiadanych przez nie cech¹⁹.

Davidsona interesuje przede wszystkim wynikająca z założeń (1) i (2) teza, że „istotne problemy dotyczące nauki języka mogą być rozwiązane *a priori*” (*ibidem*). Byłoby tak – a twierdzi tak Strawson – gdyby założenia (1) oraz (2) były nie tylko prawdami, ale prawdami koniecznymi na temat nauki

¹⁸ Na tej samej stronie można znaleźć odniesienia bibliograficzne dotyczące sporu między Strawsonem a Quinem.

¹⁹ W tym miejscu można rzecz jasna podnieść zarzut, że pierwszego języka uczymy się zaczynając od takich terminów jak „mama”, „tata” czy „miś”. Jednak zgodnie z intencją Strawsona słowa te pełnią w początkowym okresie nauki języka funkcję terminów jednostkowych, czyli nazw konkretnych indywiduów. Na tym etapie nauki języka nie ma sensu odróżnianie nazw własnych od deskrypcji określonych.

języka (*ibidem*, przyp. 3). Na konieczny charakter tych założeń jednoznacznie wskazuje czasownik „musi”, występujący w obu założeniach. W konsekwencji z (1) i (2) wynikałoby, że każdy kto nabył pewną zdolność do posługiwania się językiem, musiał przejść opisany w nich proces nauki. Tymczasem, argumentuje Davidson, zdolność do posługiwania się językiem jest dyspozycją nabytą w wyniku pewnych doświadczeń, co jest niewątpliwie kwestią empiryczną (s. 6). Davidson uzasadnia, że (1) i (2) nie są prawdami koniecznymi, wskazując na przypadki, oba bardzo mało prawdopodobne lecz możliwe, w których nauka przez ostensję oraz użycie w procesie nauki terminów jednostkowych nie muszą zachodzić równolegle. Pierwszy przypadek polegałby na celowym zabiegu nauczyciela języka, aby w nauczaniu predykatów języka używać obiektów (tj. pokazywać uczniowi obiekty), lecz nie nazywać ich za pomocą terminów jednostkowych. Nie można również wykluczyć możliwości opanowania języka bez ostensji w sprytnie zaaranżowanym, „fałszywym”, czy – jak raczej powiedzielibyśmy dzisiaj – wirtualnym środowisku. Ostensja wymaga bowiem obecności właściwego obiektu, zaś poprawne użycie terminu jednostkowego nie wymaga jego obecności. W konsekwencji: „(..) nie ma powodu, aby uczenie się predykatów przez ostensję musiało być artykułowane w języku w ten, a nie inny sposób (...)” (*ibidem*).

Jak pokazuje Davidson uznawana przez Strawsona konieczność tez (1) oraz (2), a w konsekwencji wniosek o konieczności uczenia się języka w opisany w nich sposób, stanowią błędną konkluzję z prawdziwych założeń co do natury języka naturalnego. Krytykując Quine’a tezę o eliminowalności terminów jednostkowych, Strawson odwołuje się do holizmu znaczeniowego i faktycznie akceptuje to stanowisko. Wedle Strawsona część języka nie może istnieć i funkcjonować w izolacji od całego języka w taki sam sposób, w jaki funkcjonuje w ramach całości, jaką stanowi język. Z tego punktu widzenia język – zdaniem Strawsona – jest całością organiczną, z której nie da się eliminować żadnych jego części²⁰.

Sugeruję, że z argumentów na rzecz pojęciowej współzależności różnych podstawowych części języka, Strawson wyprowadził w sposób nieuprawniony wnioski dotyczące mechanizmu i kolejności uczenia się języka (s. 7).

A zatem – zdaniem Davidsona – Strawson słusznie rozpoznał holistyczny aspekt języka, lecz zamiast stworzyć holistyczną teorię znaczenia uwzględniającą współzależność znaczeń, starał się wyjaśnić zależność między znaczeniami, odwołując się do procesu nabywania kompetencji językowych. Skoro

²⁰ Jest to zarazem ilustracja typowej dla Davidsona strategii, który w krytykowanych stanowiskach odnajdywał zawsze elementy zgodne z jego własnymi poglądami.

język jest całością organiczną, to – jak twierdził Strawson – żadna jego część (np. określona część mowy) nie może być pominięta w procesie jego nauki.

Oczywiście określenie „całość organiczna” jest bardzo nieściśle i ma niewątpliwie metaforyczny charakter. Niemniej jednak w pewnym słabym sensie tego terminu wszyscy empiryści intuicyjnie uznawali tezę, że język stanowi całość organiczną, gdyż w sposób oczywisty między przynajmniej niektórymi wyrażeniami zachodzą związki znaczeniowe. To stwierdzenie jest raczej niekontrowersyjne i nie zależy od stanowiska w kwestii znaczenia, gdyż znaczenia pewnych terminów, np. „zwierzę” i „pies”, są niewątpliwie powiązane. W tak rozumianym słabym sensie tezę o tym, że język stanowi organiczną całość, można niewątpliwie empirystom przypisać, gdyż formułowane przez nich teorie znaczenia niektóre zależności między znaczeniami faktycznie uwzględniają. Błąd empirystów polegał jednak na tym, że aby owe zależności uchwycić, tworzyli opis nauki języka jako podlegającej ścisłej i *a priori* wyznaczonej kolejności oraz procedurze, gdyż zależności między znaczeniami były konsekwencją zależności między bytami uznawanymi za znaczenia. Wiązało się to niewątpliwie z błędnymi poglądami na temat charakteru znaczeń. Wprawdzie Davidson nie pisze tego *explicite* w *Theories of Meaning*, ale przeprowadzona tam krytyka empirystycznych teorii znaczenia przyczyniła się do odrzucenia w *Prawdzie i znaczeniu* (1967) znaczeń jako bytów.

Powróćmy jednak do argumentacji Davidsona z *Theories of Meaning*. Jeżeli – jak twierdzi – faktycznie uznamy tezę o współzależności znaczeń wyrażen językowych, to nie jesteśmy się w stanie nauczyć języka bez terminów jednostkowych. Nie jest jednak tak dlatego, że – jak uważa w krytykowanym tekście Strawson – brak terminów jednostkowych uniemożliwia sam proces uczenia się języka, lecz dlatego, że „(...) nie mógłby to być język” (s. 7). Wedle Strawsona brak terminów jednostkowych „zaburza” związki znaczeniowe, a uczący się tak okrojonego języka nie jest w stanie poznać znaczenia pozostałych wyrażen, np. predykatów. Wedle Davidsona błąd, który popełnił Strawson polegał, o czym już była mowa wcześniej, na pomieszanu porządku przyczynowego uczenia się języka (który jest pewną formą porządku odkrycia) z porządkiem logicznego lub epistemologicznego uzasadniania przekonań. Odwołując się do tez (1) oraz (2) zrekonstruowanego przez siebie argumentu Strawsona, Davidson pisze:

Jest to z pewnością kwestia empiryczna, czy dana osoba w wyniku pewnych doświadczeń zdobywa umiejętności, których wcześniej nie posiadała. Jednak stwierdzenia (1) i (2), głoszące, że każdy kto nabył pewną kompetencję językową musiał przejść opisana drogę, mają charakter ściśle „logiczny” (s. 6).

Ostateczna konkluzja polemiki ze Strawsonem jest następująca:

Wnioski dla różnych teorii nauki języka są całkowicie negatywne, ale nie bez znaczenia: jak długo poważnie traktujemy „organiczny” charakter języka, nie możemy precyzyjnie opisać pierwszych kroków prowadzących do jego opanowania jako nauki części języka; jest to raczej częściowe uczenie się języka (*ibidem*)²¹.

Nie można się zatem nauczyć części języka jako tworu powstałego z języka w wyniku wyeliminowania terminów jednostkowych, gdyż czegoś takiego jak część języka po prostu nie ma, czy też raczej – część języka nie jest dla Davidsona językiem. Język jest całością, którą możemy częściowo opanować, zaś opis procesu jego nauki nie należy do filozofii języka, lecz do psychologii.

Chociaż stanowisko Davidsona wydaje się dziwne i zaskakujące, nie jest ono całkowicie odosobnione. Podobne poglądy na temat języka naturalnego głosił Kazimierz Ajdukiewicz w artykule *O znaczeniu wyrażeń* (1931). Dla Ajdukiewicza, który jest autorem holistycznej teorii znaczenia zwanej teorią dyrektywalną, język również jest całością organiczną, zaś nabywanie kompetencji językowych (Ajdukiewicz pisze o nabywaniu dyspozycji do używania wyrażeń należących do zasobu wyrażeń danego języka jako wyrażeń danego języka) wymyka się opisowi. Ajdukiewicz wyraźnie deklaruje, że nie będzie łączył go do teorii znaczenia, ani podejmował próby opisu tego procesu.

3. WYUCZALNOŚĆ JĘZYKA A TEORIA ZNACZENIA

W rezultacie polemiki z argumentem Strawsona, Davidson stwierdza – co sygnalizowałem już w paragrafie 1 – że wprawdzie nie możemy na podstawie apriorycznych „rozważań logicznych” opisać mechanizmu nauki języka, ale możemy:

(...) rozważać, zanim zajmiemy się empirycznymi badaniami znajomości języka, w jaki sposób należy opisać umiejętność lub zdolność osoby, która nauczyła się mówić danym językiem (s. 7–8).

Rozważania natury logicznej lub apriorycznej pozwalają bowiem narzucić pewne ograniczenia na opis dyspozycji zwanej powszechnie znajomością języka. Ponieważ znajomość języka polega *de facto* na znajomości znaczeń jego wyrażeń, opis znajomości języka można utożsamić z teorią znaczenia.

²¹ Podkreślenia moje.

Określenie „język wyuczalny”, które pełni kluczową rolę w *Theories of Meaning*, może prowadzić do pewnych nieporozumień. Na początku, dla uproszczenia, określiłem wyuczalność jako własność języka. Wyuczalność we właściwym tego słowa znaczeniu nie jest jednak własnością języka, podobnie jak „rozpuszczalność” nie jest własnością jakiejś substancji. Określenia te urobione są od nazw rezultatów zdarzeń bądź procesów, np. termin „rozpuszczalność” urobiony jest od „rozpuszczania się”.

W teorii znaczenia języka wyuczalnego nie pojawi się *explicite* predykat „wyuczalny”, podobnie jak w poważnej teorii fizycznej lub chemicznej nie pojawi się predykat „rozpuszczalny”, lecz predykaty odnoszące się do budowy molekularnej. W szczególności stwierdzenie, że przedmiot rozpuścił się w wodzie bo jest rozpuszczalny, jest trywialne, gdyż predykat „rozpuszczalny” definiuje się za pomocą czasownika „rozpuścić się”²². „Rozpuszczalność”, podobnie jak „wyuczalność” są bowiem predykatami dyspozycyjnymi, które powinno się definiować za pomocą tzw. definicji redukcyjnej w sposób podany przez Carnapa (por. Pawłowski 1977). W przypadku języka jego wyuczalność definiowana byłaby mniej więcej w następujący sposób: jeżeli grupa osób o normalnych zdolnościach została poddana procesowi uczenia się języka, to język jest wyuczalny wtedy i tylko wtedy, gdy w skończonym czasie osoby te nauczyły się tego języka, tj. są w stanie używać (tj. wypowiadać i rozumieć) potencjalnie nieskończenie wiele jego zdań. Podana definicja redukcyjna jest po prostu definicją sprawozdawczą terminu „wyuczalny”, gdyż tak termin „wyuczalny” rozumiemy.

Przywołajmy ponownie analogię do terminu „rozpuszczalny”. Nietrywialna teoria fizyczna rozpuszczalności powinna umożliwić zbudowanie definicji równościowej predykatu „rozpuszczalny”, która odwołuje się do pewnych własności budowy molekularnej przedmiotów rozpuszczalnych. Należy również zwrócić uwagę, że sama znajomość „teorii rozpuszczalności” nie pozwala nam jeszcze na to, aby o konkretnym przedmiocie orzec, czy faktycznie w wodzie się rozpuści. Musimy bowiem wiedzieć o tym przedmiocie jaka jest jego budowa molekularna lub przynajmniej empirycznie stwierdzić, że mamy do czynienia z próbką substancji o znanej nam budowie, która zapewnia jej rozpuszczalność, nie zaś z substancją ludząco do niej podobną. W obu przypadkach należy odwołać się do doświadczenia. Podobnie jest z adekwatną teorią znaczenia. Sama teoria znaczenia dla pewnego języka wyuczalnego nie pozwoli nam stwierdzić, co mówiący rozumie przez okreśło-

²² Davidson w kilku innych artykułach odwołuje się do predykatów dyspozycyjnych, wskazując na to, że ich nietrywialne definicje muszą być konstruowane w ramach jakiejś teorii. Inaczej odpowiedzi na pytania „dlaczego?” będą trywialne.

ne zdania, czyli przypisać znaczenia jego słowom. Musimy empirycznie stwierdzić, czy mówiący faktycznie posługuje się językiem opisywanym przez teorię, czy też językiem, w którym występują wyrażenia o tym samym kształcie fizycznym, co wyrażenia naszego języka, lecz o innych znaczeniach. Aby było to możliwe, musimy coś wiedzieć o stanach psychicznych użytkownika, a przede wszystkim przypisać mu przekonania i inne postawy propozycyjalne. Mówiący może bowiem używać zdań o kształcie fizycznym nieodróżnialnym od zdań języka opisanego przez daną teorię znaczenia, lecz o zupełnie innych znaczeniach. Aby teorię stosować, tj. przypisać za jej pomocą znaczenia czymś słowom, należy poza samą teorią dysponować odpowiednimi empirycznymi świadectwami. Innymi słowy należy ją poddawać weryfikacji empirycznej i w miarę potrzeby modyfikować. Jednak sama teoria znaczenia powinna być, wedle Davidsona, formalną teorią języka wyuczalnego, w której nie używa się żadnych pojęć psychologicznych²³. Przyjrzyjmy się zatem, jakie warunki należy – zgodnie z poglądami Davidsona – narzucić na teorię znaczenia języka wyuczanego.

Zdaniem Davidsona pierwszym z warunków, które należy narzucić na teorię znaczenia, jest możliwość zdefiniowania predykatu w metajęzyku tego języka, który to predykat stosuje się do klasy wyrażen sensownych (*meaningful*) tego języka, czyli do zdań. Predykat taki można orzekać o wszystkich zdaniach i jedynie o zdaniach tego języka. Predykat ten podaje również gramatykę tego języka (s. 8), gdyż stanowi jego konstruktywny opis. Stwierdzenie to jest pewnym skrótem myślowym – Davidsonowi chodzi o to, że predykat ten powinien być zdefiniowany za pomocą wszystkich reguł gramatycznych tego języka i tylko za pomocą tych reguł. Jak się okaże, predykatem tym będzie predykat „prawdziwy”. Analogia do rozpuszczalności (w określonej cieczy) jest tu oczywista. Zbudowanie teorii rozpuszczalności polega na zdefiniowaniu takiego predykatu, który byłby prawdziwy dla wszystkich przedmiotów rozpuszczalnych i jedynie dla nich. Musiałby to być prawdopodobnie predykat bardzo złożony. Należy zauważyć, że predykat „prawdziwy” jest w semantycznej definicji prawdy również predykatem bardzo złożonym, gdyż definiuje się go za pomocą złożonych aksjomatów.

Drugi warunek narzucony na teorię znaczenia wymaga, aby pozwalała ona określić jedynie na podstawie formalnych własności języka, co znaczy dowolne jego zdanie. Wyposażeni w taką teorię moglibyśmy rozstrzygać, co mówiący miał na myśli, wypowiadając dane zdanie:

²³ Stąd Davidson używał wymiennie terminów „teoria języka” i „teoria znaczenia”.

Stosując odpowiednie wybiegi psychologiczne (*psychological trappings*) nasza teoria powinna pozwolić nam odkryć, w przypadku dowolnego zdania, co mówiący ma na myśli (*ibidem*).

Sprzeczność, która zdaje się tutaj pojawiać, jest jedynie pozorna. Zdaniem Davidsona teoria języka pozwala *a priori*, tj. jedynie na podstawie formalnych – czy jak pisał Tarski – morfologicznych własności zdań przypisywać im znaczenia. Jednak samo formułowanie teorii znaczenia oraz jej weryfikacja, oparte są na przesłankach empirycznych, np. na empirycznych obserwacjach zachowań językowych i pozajęzykowych mówiącego, które niekiedy są przez nas prowokowane (wybiegi psychologiczne): „Mając na względzie adekwatną teorię widzimy jak działania i dyspozycje mówiącego odciskają na zdaniach języka strukturę semantyczną (*ibidem*)”. Uzbrojeni w taką teorię możemy interpretować wypowiedzenia mówiącego w oparciu o obserwowane zachowania i przypisywane mu dyspozycje, czyli przekonania oraz inne postawy propozycyjalne.

W ostatnim cytacie pojawia się zapowiedź traktowania teorii znaczenia jako narzędzia do interpretowania nie tyle zdań rozumianych jako pewne twory abstrakcyjne rozumiane jako typy wypowiedzeń, co samych wypowiedzeń zdań w konkretnym kontekście. Pozwala to wyjaśnić skuteczność interpretacji w przypadku pomyłek, przejęzyczeń, niedopowiedzeń itp. Jednak w swych wczesnych pracach Davidson nie dokonuje jeszcze *explicite* tego rozróżnienia. W późniejszych pracach staje wyraźnie na stanowisku antykonwencjonalizmu, które głosi, że teoria znaczenia nie służy do interpretowania języka rozumianego jako abstrakcyjny twór oparty na konwencjach, lecz do interpretowania idiolektu mówiącego, zaś konwencje językowe (ustawiczne łamane przez pomyłki i przejęzyczenia) stanowią jedynie ułatwienie procesu interpretacji.

Podane wyżej kryteria spełnia, wedle Davidsona, zmodyfikowana przezeń Tarskiego semantyczna definicja prawdy:

Choć niewątpliwie zrelatywizowana do czasów, miejsc i okoliczności, rodzaj wymaganej struktury [gramatycznej] wydaje się być albo identyczny, albo ściśle związany z rodzajem [struktury] podanej przez definicję prawdy zbieżnej z wprowadzoną po raz pierwszy przez Tarskiego. Taka definicja zapewnia efektywną metodę podawania znaczenia każdego zdania (tj. podaje warunki, dla których jest ono prawdziwe) (s. 8).

W *Theories of Meaning* mamy do czynienia z zapowiedzią semantyki warunków prawdziwości²⁴, która zdominowała poglądy semantyczne Davidsona. W tym miejscu warto jednak przypomnieć kilka szczegółów z poprzedniego rozdziału. Ponieważ Tarski starał się zdefiniować prawdę dla języków formalnych, zastosowanie tej definicji do języka naturalnego wymaga wielu jej modyfikacji, która pozwoliłaby interpretować zdania okazjonalne oraz inne konstrukcje, np. zdania wyrażające postawy propozycyjne, zdania performatywne i wyrażenia cudzysłowowe, które w językach formalnych nie występują. Celem Tarskiego było zdefiniowanie prawdy w języku, co było możliwe przy założeniu znajomości przekładu z języka przedmiotowego na metajęzyk. Wedle Davidsona, co *explicite* stwierdza w wielu późniejszych artykułach, prawda jest pojęciem pierwotnym i niedefiniowalnym, co pozwoliło mu niejako odwrócić rolę definicji prawdy, która po tej modyfikacji mogła pełnić rolę teorii znaczenia.

Dlaczego jednak definicja prawdy Tarskiego spełnia wymogi narzucone przez Davidsona na teorie znaczenia języka wyuczalnego? Jak pamiętamy, opis „zdolności osoby, która nauczyła się języka” wymaga zdefiniowania predykatu, który pozwala wyróżnić klasę zdań sensownych (*meaningful*), odwołując się jedynie do ich własności formalnych (gramatycznych, syntaktycznych lub morfologicznych). Warunek ten spełnia niewątpliwie Tarskiego semantyczna definicja prawdy, zaś predykatem, o którym mowa jest predykat „prawdziwy”.

Pojawiający się tutaj wymóg zdefiniowania predykatu „prawdziwy” zdaje się pozostawać w sprzeczności z przytoczoną przed chwilą, a deklarowaną w późniejszych pracach, tezą o niedefiniowalności prawdy. Wspomniana sprzeczność nie jest jednak groźna dla koncepcji Davidsona. Przede wszystkim niedefiniowalne jest – jego zdaniem – przedteoretyczne i absolutne pojęcie prawdy, a dokładniej treść tego pojęcia. W swym oryginalnym sformułowaniu definicji prawdy Tarski zakłada znajomość znaczeń wyrażeń języka, co pozwala mu zdefiniować zakres predykatu „prawdziwy” w języku. Kiedy Davidson pisze o definiowaniu predykatu „prawdziwy”, to ma na myśli definiowanie zakresu predykatu „prawdziwy” w określonym języku, nie zaś treści tego pojęcia rozumianego w sensie absolutnym. A zatem Tarskiego semantyczna definicja prawdy może pełnić podwójną rolę: definiować predykat „prawdziwy” w języku przedmiotowym (jeśli znamy znaczenia wyrażeń tego języka) oraz podawać, w sposób, który będzie omówiony dalej, znaczenia wyrażeń tego języka (przy założeniu, że znamy treść pojęcia

²⁴ Sam termin „warunki prawdziwości” nie jest jednoznaczny. W przypadku teorii Tarskiego warunkami prawdziwości są tzw. T-zdania. Do sprawy tej powrócę w kolejnych rozdziałach.

prawdy). Aby uniknąć niezręczności związanej w podwójną rolę semantycznej definicji prawdy, w swych późniejszych pracach Davidson zastępował termin „semantyczna definicja prawdy” zwrotem „teoria prawdy”.

Teoria prawdy Tarskiego jest teorią rekurencyjną. Sama rekurencyjność tej teorii nie stanowi jednak warunku wystarczającego, aby podawała znaczenia nieskończenie wielu zdań języka wyuczanego. Aby mogła pełnić to zadanie, w jej (rekurencyjnych) aksjomatach musi pojawiać się skończona ilość wyrażen semantycznie prostych (*semantical primitives*):

Nazwijmy wyrażenie semantycznie prostym o ile reguły podające znaczenie zdań, w których to wyrażenie nie występuje, nie wystarczą do określenia znaczenia zdań, w których to wyrażenie występuje (s. 9).

Zauważmy, że wyrażenia semantycznie proste nie są wyrażeniami o znaczeniach atomowych, w sensie jakim asocjacyoniści mówili o ideach prostych, z których konstruowane są idee złożone. Wyrażeniem semantycznie prostym będzie nie tylko „koń” i „człowiek”, ale również „centaur”. Do wyrażen semantycznie prostych należą nie tylko wyrazy, ale również idiomy, gdyż znaczenie idiomu nie jest wyznaczone przez znaczenie wyrazów, które w nich występują. Ponieważ jest ich w każdym języku naturalnym skończenie wiele, można je traktować na równi z wyrazami. Nie stanowi również problemu wieloznaczność wyrazów, gdyż różnych znaczeń danego wyrazu jest zawsze skończenie wiele. Każdemu ze skończonej ilości wyrażen semantycznie prostych odpowiada na gruncie teorii prawdy odpowiedni aksjomat. Po wprowadzeniu tych ograniczeń, Tarskiego semantyczna definicja prawdy może zatem pełnić – wedle określenia Davidsona – funkcję adekwatnej teorii znaczenia języka naturalnego.

Należy zauważyć, że faktyczny problem z potraktowaniem teorii prawdy jako teorii znaczenia języka naturalnego, jest zupełnie innej natury. Rekurencyjność teorii Tarskiego zapewniona jest nie przez reguły gramatyczne języka naturalnego podawane przez filologów, lecz przez reguły języka rachunku predykatów, gdyż faktycznie dla takich reguł podaje się aksjomaty teorii prawdy. Aby zatem teoria Tarskiego mogła pełnić funkcję teorii znaczenia języka naturalnego należałoby wykazać, że język naturalny ma formę logiczną lub strukturę głęboką charakterystyczną dla języka rachunku predykatów. Oznacza to zazwyczaj konieczność wprowadzenia do języka naturalnego nieskończonej ilości zmiennych indywiduowych. Jest to konieczne ze względu na budowanie aksjomatów dla kwantyfikatorów wiążących owe zmienne. Zauważmy, że w języku naturalnym zmienne wiązane przez kwanty-

fikatory *explicite* nie występują²⁵. Zaproponowany przez Davidsona i realizowany przez niego program semantyczny, miał na celu budowanie teorii prawdy dla kolejnych fragmentów języka naturalnego.

Do programu semantycznego, rozwijanego w następnych publikacjach Davidsona, powrócimy jeszcze wielokrotnie. W *Theories of Meaning* Davidson ograniczył się jedynie do pokazania, że kilka znanych teorii języka²⁶ nie może być uznanych za adekwatne teorie znaczenia języka naturalnego, gdyż systemy znaków, które owe teorie opisują, nie są wyuczane, a zatem nie zasługują *de facto* na miano języka naturalnego, chociaż występujące w nich wyrażenia są równokształtne z wyrażeniami takiego języka. Takie systemy znaków Davidson nazywa językami niewyuczalnymi. Niewyuczalność nie oznacza tutaj braku rekurencyjnych reguł składni lub reguł semantyki. Jako przykład można tu podać język sztuczny rachunku predykatów z nieskończoną ilością stałych indywiduowych a_1, a_2, a_3, \dots oraz z nieskończoną ilością predykatów jednoargumentowych A_1, A_2, A_3, \dots . Nie jest to język wyuczany w sensie Davidsona, co przejawia się tym, że formuł tego języka nie rozumiemy w takim samym sensie, w jakim rozumiemy zdania języka polskiego czy angielskiego²⁷. Dla takiego języka można wprowadzić zbudować Tarskiego definicję prawdy, lecz nieskończonej ilości wyrażen semantycznie prostych musi odpowiadać nieskończona ilość aksjomatów. Ponieważ jednak nieskończonej ilości aksjomatów podać się nie da, zawsze będą istniały stałe, a co za tym idzie formuły, dla których teoria prawdy Tarskiego nie poda znaczenia. Możemy sobie bowiem wyobrazić teorię prawdy dla tego języka, która stałej a_1 przyporządkuje (na mocy odpowiedniego aksjomatu spełniania) Piotra, stałej a_2 (na mocy kolejnego aksjomatu) Jana, zaś stałej a_3 (na mocy jeszcze innego aksjomatu) Roberta. Jednak stałych jest nieskończenie wiele, co sprawia, że nie możemy dla każdej z nich podać aksjomatu spełniania. W szczególności nie potrafimy na podstawie aksjomatu dla stałej a_n podać aksjomatu dla stałej a_{n+1} . W konsekwencji język taki jest niewyuczalny w sensie podanym przez Davidsona²⁸. Znajomość języka wyuczanego nie

²⁵ Jeśli za odpowiedniki zmiennych potraktujemy zaimki, to i tak nie możemy potraktować formuły z zaimkiem jako formuły otwartej w rachunku predykatów, lecz jako zdanie okazjonalne. Do sprawy zdań okazjonalnych powrócimy nieco dalej.

²⁶ Faktycznie są to teorie opisujące jedynie pewne fragmenty języka naturalnego.

²⁷ Wyjątek stanowią tu tzw. stałe logiczne, które rozumiemy na mocy odpowiednich aksjomatów. Warto również zwrócić uwagę, że na wyuczalność języka przedmiotowego nie ma wpływu fakt, że włączamy doń nieskończenie wiele zmiennych, gdyż wyuczalność związana jest ze zdaniami, czyli formułami, w których nie występują zmienne wolne.

²⁸ Przykład ten (pochodzący ode mnie) ma pokazać, że brak wyuczalności nie dyskwalifikuje przydatności języków sztucznych jako przedmiotu badań logików czy matematyków. Dla takich języków nie ma bowiem potrzeby budowania teorii znaczenia i rozważania Davidsona się do

oznacza bowiem jedynie umiejętności operowania regułami składni, lecz rozumienie tego języka, zaś teoria, która to zapewnia, powinna być tzw. teorią interpretacyjną. Do sprawy interpretacyjności teorii powrócę pod koniec rozdziału 3. W tym miejscu warto jedynie zauważyć, że teoria znaczeniowa (w sensie podanym przez Dummetta) rozwinięta przez Davidsona w *Theories of Meaning* ma przede wszystkim na celu budowanie interpretacyjnej teorii znaczenia wyjaśniającej możliwość rozumienia języka wyuczalnego²⁹. Rozumienie to jest możliwe, gdy teoria znaczenia przyporządkuje znaczenia potencjalnie nieskończenie wielu zdaniom, korzystając ze skończonej ilości aksjomatów. W kolejnych paragrafach przedstawiam – za Davidsonem – analizę kilku wybranych teorii znaczenia języka naturalnego, które jego zdaniem ignorują fakt wyuczalności języka.

4. KRYTYKA TEORII WYRAŻEŃ CUDZYSŁOWOWYCH

Przez wyrażenia lub nazwy cudzysłowowe (*quotations*) rozumiemy zazwyczaj przytoczenia pewnych wypowiedzi. Oczywiście funkcje cudzysłowu wykraczają poza cytowania lub przytoczenia, nas jednak będzie tu interesowała jedynie ta funkcja. Przytaczać wypowiedzi można w języku mówionym, ale jest o wiele łatwiej analizować je w języku pisanym, ze względu na użycie cudzysłowu (*quotation marks*), wewnątrz którego znajdują się przytaczane wyrażenia. Jako element występujący powszechnie w języku, wyrażenia cudzysłowowe są przez nas doskonale rozumiane. Ponieważ wyrażeń cudzysłowowych jest nieskończenie wiele, owo rozumienie musi być oparte na

nich nie odnoszą. W oryginalnym sformułowaniu pochodzącym od Tarskiego warunek wyuczalności nie musi być spełniony. W szczególności jeśli metajęzyk obejmuje język przedmiotowy, to zamiast nieskończonej ilości aksjomatów dla stałych i predykatów możemy podać skończoną ilość schematów aksjomatów, np.

Dla każdego ciągu C , C przyporządkowuje stałą „ a_i ” obiekt a_i .

Dla każdego ciągu C , C spełnia formułę „ $A_i(t)$ ” wtw, gdy obiekt przyporządkowany termowi „ t ” przez ciąg C jest A_i .

Dla języków naturalnych, jeśli metajęzyk obejmuje język przedmiotowy, to można zastosować ten sam zabieg.

²⁹ Możemy sobie wyobrazić ekscentryka, który zamiast normalnych słów języka polskiego używa w sposób konsekwentny stałych indywidualnych oraz predykatów. Taki język może być wyuczany pod warunkiem użycia jedynie skończonej ilości takich wyrażeń w pewien regularny sposób (np. odnosząc przynajmniej niektóre wyrażenia do rzeczywistości), gdyż wówczas można zbudować teorię prawdy dla tego języka ze skończoną ilością aksjomatów. Co więcej, budując taką teorię, druga osoba może zinterpretować ten język, a w konsekwencji nauczyć się go tak, jak każdego obcego języka.

jakiejs regule, jednak – jak twierdzi Davidson – „zagadka się pojawia, gdy chcemy użyć tej reguły jako elementu teorii znaczenia” (*ibidem*).

Zarówno potoczna intuicja, jak i elementarne podręczniki podpowiadają nam, że wyrażenie cudzysłowowe lub nazwa cudzysłowowa jest po prostu nazwą wyrażenia znajdującego się między cudzysłowami. Davidson twierdzi jednak, że takie podejście do wyrażen cudzysłowowych nie może nawet stanowić namiastki teorii wyrażen cudzysłowowych jako fragmentu teorii języka naturalnego. Przytoczony wyżej intuicyjny pogląd na naturę wyrażen cudzysłowowych przedstawiają Quine w *Logice matematycznej* (1974/1940, s. 29) oraz Tarski w artykule *Pojęcie prawdy w językach nauk dedukcyjnych* (s. 17–31)³⁰. Wymienieni autorzy wskazują, że powyższe podejście stanowi źródło sporych kłopotów, które napotykamy, analizując język zawierający wyrażenia cudzysłowowe, lecz nie są w stanie zaproponować zadowalającego rozwiązania tego problemu³¹.

Na czym polega główna trudność teorii nazw cudzysłowowych? Wyrażen cudzysłowowych jest w każdym języku naturalnym potencjalnie nieskończenie wiele, gdyż wewnątrz cudzysłowów mogą występować nie tylko wyrażenia semantycznie proste, lecz całe zdania, które są wyrażeniami semantycznie złożonymi. Jednak w przypadku omawianego poglądu wszystkie wyrażenia cudzysłowowe muszą być traktowane jako wyrażenia semantycznie proste, czemu dał wyraz Tarski:

Nazwy cudzysłowowe można traktować tak jak pojedyncze wyrazy języka, a zatem jak wyrażenia syntaktycznie niezłożone; poszczególne części składowe tych nazw – cudzysłowy i wyrażenia zawarte między cudzysłowami – pełnią tę samą funkcję, co litery lub zespoły kolejnych liter w pojedynczych wyrazach, nie posiadają zatem w tym kontekście żadnego samodzielnego znaczenia (s. 23)³².

Skoro nie jesteśmy w stanie zrozumieć nieznanego nam słowa na podstawie znajomości liter, za pomocą których zostało zapisane, tak samo – zgodnie

³⁰ Tego samego zdania był Frege: „Zdanie w mowie niezależnej oznacza jakieś zdanie (...)” (por. 1975/1892, s. 73).

³¹ Tarski wskazuje, że wyrażenia cudzysłowowe mogą być źródłem paradoksów. Jak pamiętamy, w semantycznej definicji prawdy paradoksy są eliminowane przez wyłączenie wyrażen semantycznych z języka przedmiotowego.

³² W pierwszym rozdziale *Pojęcia prawdy* Tarski próbował uściślić tzw. klasyczną definicję prawdy. Trudności, na które natrafił, wiązały się z wyrażeniami cudzysłowowymi, które pojawiają się w zwrocie „Zdanie *p* jest prawdziwe”: „(...) z tą chwilą, gdy zaczynamy bliżej analizować znaczenie występujących w (5) i (6) wyrażen cudzysłowowych, dostrzegamy szereg trudności i niebezpieczeństw” (*ibidem*). Definicja prawdy, którą buduje w dalszych częściach artykułu, pozwala uniknąć mu tych trudności, chociaż sformułowanego przez Davidsona problemu nazw cudzysłowowych Tarski nie rozwiązał.

z tą teorią – nie możemy rozumieć wyrażenia cudzysłowowego na podstawie rozumienia wyrazów wewnątrz cudzysłowu. W konsekwencji teoria nazw cudzysłowowych Tarskiego-Quine’a traktuje – zdaniem Davidsona – język zawierający wyrażenia cudzysłowowe za niewyuczalny (s. 10)³³.

Ma zatem miejsce swoisty paradoks polegający na tym, że jeżeli rozumiemy wyrażenie semantycznie złożone wewnątrz cudzysłowu, to krytykowana teoria nie pozwala na wyjaśnienie, w jaki sposób rozumiemy (a faktycznie je rozumiemy) to samo wyrażenie wraz z otaczającym je cudzysłowem. Nie ulega przecież najmniejszej wątpliwości, że nazwy cudzysłowowe stanowią integralną część języka naturalnego, zaś kompetentny użytkownik tego języka rozumie każdą nazwę cudzysłowową, o ile rozumie wyrażenia znajdujące się między cudzysłowami.

Alfred Tarski sugerował również inne możliwe podejścia do wyrażen cudzysłowowych, które jednak nie były satysfakcjonujące:

Chcąc uratować sens zdań (5) i (6) trzeba uciec się do innej zupełnie interpretacji nazw cudzysłowowych. Nazwy te musimy już traktować jako wyrażenia syntaktycznie złożone, których składnikami syntaktycznymi są zarówno cudzysłowy, jaki i wyrażenia zawarte między cudzysłowami. (...) cudzysłowy stają się wtedy samodzielnymi wyrazami z zakresu semantyki, zbliżonymi znaczeniowo do wyrazu „nazwa”, a pod względem syntaktycznym grającymi rolę funktorów. Powstają również nowe powikłania. Sens intuicyjny nazwy cudzysłowowej i samych cudzysłowów nie jest dostatecznie jasny. W każdym razie nie są to funktory ekstensjonalne (...) (1933, s. 25).

Cudzysłowy nie są funktorami ekstensjonalnymi, gdyż zastąpienia wyrażenia *a* w zdaniu:

(4.1) *X* powiedział „*a*”

jakimkolwiek różnym od niego co do fizycznego kształtu wyrażeniem *b* o tym samym odniesieniu, a nawet tym samym znaczeniu, może zmienić wartość logiczną całego zdania z prawdy na fałsz, o ile *X* nie powiedział również *b*. Chcąc być bardziej precyzyjnym warto zauważyć, że analizowane zdanie jest eliptyczne i w związku z tym możliwe są dwie jego interpretacje. Można je bowiem rozumieć jako zdanie okazjonalne odnoszące się do konkretnej wypowiedzi w określonym momencie czasu. W takim przypadku nie ma znaczenia czy *X* kiedykolwiek indziej powiedział *b* czy też nie, a zamiana *a* na *b* zmienia wartość logiczną zdania. W drugim przypadku zdanie (4.1) głosi, że

³³ Jak zauważa Davidson, Tarski oraz Quine zdawali sobie sprawę z problemów, które niesie za sobą takie podejście do cudzysłowów. Żaden z nich nie zdołał jednak zbudować satysfakcjonującej teorii wyrażen cudzysłowowych.

X przynajmniej raz powiedział *a*. Wówczas do zmiany wartości logicznej nie dojdzie, o ile *X* przynajmniej raz powiedział *b*. Uwaga ta ilustruje co następuje. Po pierwsze przed budowaniem teorii prawdy język naturalny musi być ujednoznaczony, zaś jednym z możliwych narzędzi ujednoznacznienia języka jest sprowadzenie jego wyrażeń do postaci wyrażeń języka rachunku predykatów. Po drugie, nie istnieje jakikolwiek dający się wyrazić za pomocą reguł związek logiczny między wyrażeniem wewnątrz cudzysłowu, a całym zdaniem. Jak pokazał nieco dalej Davidson, nieekstensjonalność funktorów (w naszym przypadku byłby to cudzysłów) wyklucza wyuczalność języka, odrzucając tym samym sugestię Tarskiego³⁴.

Na zakończenie przyjrzymy się dokładniej propozycji Tarskiego. Podanej niżej analizy Davidson nie przedstawia, ale pozostaje ona całkowicie w duchu Davidsona, który bardzo często wydobywał z krytykowanych poglądów te ich aspekty, które zapowiadają jego własne rozwiązanie. Rozważmy pojawiające się w przytoczonym powyżej cytacie, dokonane przez Tarskiego, porównanie cudzysłowu do funktora „nazwa” oraz dokonanego przezeń nieco wcześniej przyrównania wyrażeń cudzysłowowych do imion własnych ludzi:

Każde wyrażenie cudzysłowowe jest wówczas stałą nazwą jednostkową pewnego określonego wyrażenia (tego mianowicie, które jest ujęte w cudzysłowy) i to nazwą o tym samym charakterze co imiona własne ludzi (s. 23–24).

Oczywiście funktor „nazwa” przyporządkowujący obiektom (ludziom) ich imiona własne jest odwrotnością semantycznej relacji oznaczania. Tak rozumiany funktor w języku przedmiotowym nie występuje, gdyż aby funktor występował w języku, musiałyby również występować w tymże języku argumenty tego funktora. Okazuje się jednak, że ten sam zarzut można podnieść, jeśli funktor „nazwa”, tym razem rozumiany jako cudzysłów, potraktujemy jako wyrażenie metajęzykowe – jego argumentami nie byłyby wyrażenia metajęzyka lecz języka przedmiotowego. Mamy zatem do czynienia z sytuacją następującą: cudzysłów nie tylko jest nieekstensjonalny, ale nie może być przyporządkowany do żadnego języka, czyli *de facto* żadnym funktorem nie jest. Jak zatem można analizować zdania mieszane, takie jak (4.1)? Oczywiście bez możliwości przyporządkowania cudzysłowowi odpowiedniego aksjomatu, zdania te w ogóle nie mogą być brane w semantycznej

³⁴ Do sprawy tej wrócimy w następnym paragrafie. W tym miejscu warto jedynie zaznaczyć, że teoria prawdy jest teorią ekstensjonalną i jej stosowanie do języka naturalnego pociąga za sobą założenie, że intensjonalność w języku naturalnym jest zjawiskiem jedynie pozornym.

definicji prawdy pod uwagę. Nie da się bowiem zbudować takiej definicji prawdy, z której wynikałoby T-zdanie postaci:

Zdanie „ X powiedział p ” jest prawdziwe wtw, gdy X powiedział p .

Oznaczałoby to, że semantyczna definicja prawdy nie obejmowałaby bardzo ważnego fragmentu języka. Sprawa jest istotna również z innego powodu. W semantycznej definicji prawdy cudzysłów występuje zarówno w aksjomatach, jak i w T-zdaniach. Brak możliwości analizy cudzysłówów mógłby stanowić wobec niej bardzo poważny zarzut.

Okazuje się jednak, że wskazana przez Tarskiego analogia cudzysłówów do hipotetycznego funktora „nazwa” (na obu poziomach: relacja rzeczywistość – język oraz relacja język przedmiotowy – metajęzyk) niesie ze sobą bardzo owocną intuicję. Chodzi tu o element wskazujący, który występuje w tym funktorze. Na poziomie relacji rzeczywistość – język, funktor ten wskazuje na pewien obiekt, zaś na poziomie język przedmiotowy – metajęzyk, na to wyrażenie języka przedmiotowego, które znajduje się wewnątrz cudzysłowu. Co więcej, owo wyrażenie zawsze występuje jako część w wypowiedzi, w której jest przytaczane, czyli pozostaje wewnątrz cudzysłowu. W konsekwencji zdanie zawierające wyrażenie cudzysłowowe, np.

„Jan” składa się z trzech liter,

można przekształcić do postaci:

To słowo składa się z trzech liter. Jan,

gdzie zaimek „to” odnosi się do wypowiedzenia słowa „Jan”. Z kolei zdanie

Galileusz powiedział „*Eppur si muove*”

można rozumieć jako:

Galileusz wypowiedział zdanie o tym kształcie: *Eppur si muove*.

W ten sposób cudzysłów traci charakter funktora oznaczającego relację między językiem przedmiotowym a metajęzykiem, a w konsekwencji zdania z wyrażeniami cudzysłowowymi można analizować po sprowadzeniu ich do postaci kanonicznej, jako dwa niezależne od siebie wyrażenia języka przedmiotowego.

Wskazująca funkcja cudzysłówów stanowi istotę rozwiązania tego problemu zwanego analizą parataktyczną (równoległą), które Davidson przedstawił w artykule *Quotation* (1979). Niewątpliwą zaletą ten analizy jest wyeliminowanie ze zdania z wyrażeniem cudzysłowowym nieekstensjonalnego cudzysłowu, który w dodatku nie może być nawet traktowany jako funktor.

W konsekwencji otrzymujemy dwa niezależne logicznie wyrażenia: zdanie wskazujące na wyrażenie przytoczone wewnątrz cudzysłowu oraz samo to wyrażenie³⁵. W przypadku „zwykłych” T-zdań cudzysłów również można wyeliminować:

Zdanie (języka angielskiego) o kształcie fizycznym identycznym z następnym zdaniem jest prawdziwe wtw, gdy Arystoteles był filozofem. *Aristotle was a philosopher.*

Podobny zabieg można zastosować do aksjomatów semantycznej definicji prawdy.

5. KRYTYKA WYBRANYCH TEORII MOWY ZALEŻNEJ

Podobnie jak wyrażenia cudzysłowowe, mowa zależna stanowi integralną część języka naturalnego i, podobnie jak cudzysłów, mowa zależna tworzy konteksty nieekstensjonalne, które stanowią spore wyzwanie dla semantyki języka naturalnego. Wyzwanie to dostrzegł jako pierwszy Frege i zaproponował jego rozwiązanie. W mowie zależnej występują bowiem zdania, które wyrażają treści myśli, przekonań, obaw, wątpliń itd. Z punktu widzenia składni stanowią one zdania podrzędne w tzw. zdaniach psychologicznych. Problem, który się z tym wiąże, polega na tym, że w zdaniu psychologicznym zdaje się nie zachodzić tzw. zasada zastępowania ekstensjonalnego (por. Frege 1977/1892, s. 72). Zasada ta głosi, że jeżeli w wyrażeniu złożonym zastąpimy jego część innym wyrażeniem o tym samym odniesieniu, to odniesienie całego wyrażenia złożonego nie ulegnie zmianie. W szczególności, jeśli wyrażenie złożone jest zdaniem, to opisana wyżej substytucja nie zmieni jego wartości logicznej³⁶. Zasada ta zdaje się wyrażać oczywistą intuicję, że rzeczywistość nie zależy od tego, w jaki sposób została opisana. Jeżeli zatem prawdą jest, że Jan pali, to prawdą jest również, że mąż Marii pali, o ile „Jan” oraz „mąż Marii” oznaczają tę samą osobę.

³⁵ Szczegółowe omówienie analizy parataktycznej nazw cudzysłowowych wymaga oczywiście osobnego opracowania.

³⁶ Dla Fregego to właśnie wartości logiczne, tj. Prawda lub Fałsz, są odniesieniami tych zdań, które posiadają wartość logiczną. Pozbawione wartości logicznej są zdania opisujące fikcję, np. fikcję literacką, gdyż zawierają wyrażenia pozbawione odniesień, np. nazwy bohaterów fikcyjnych.

W tym miejscu pojawia się problem, który nie występuje u Fregego³⁷. Davidson znany jest jako autor semantyki odniesieniowej i nie operuje pojęciem odniesienia rozumianym jako korelat semantyczny wyrażenia językowego. Z impasu tego można wybrnąć, definiując pojęcie równoważności wyrażen jako odpowiednika posiadania przez nie tego samego odniesienia. W przypadku zdań równoważność sprowadza się do posiadania tej samej wartości logicznej, w przypadku nazw jednostkowych a i b , są one równoważne, gdy dla każdego predykatu P , $P(a)$ ma tę samą wartość logiczną, co $P(b)$, zaś predykaty P i Q są równoważne, gdy dla każdego terminu jednostkowego a , $P(a)$ ma tę samą wartość logiczną, co $Q(a)$.

Zasada zastępowania ekstensjonalnego (zarówno w wersji „odniesieniowej”, jak i w wersji „równoważnościowej”) nie zachodzi jednak w przypadku zdań lub ich części w mowie zależnej (tj. zdań podrzędnych w zdaniach psychologicznych), np. zastąpienie nazwy „Jan” nazwą „mąż Marii” może bowiem spowodować zmianę wartości logicznej zdania „Piotr powiedział, że Jan pali” lub „Piotr przypuszcza, że Jan pali”.

Jak pamiętamy z poprzedniego paragrafu, w wyrażeniach cudzysłowowych zasada zastępowania ekstensjonalnego również nie zachodzi, np. zdanie „*Jan* składa się z trzech liter” jest prawdziwe, lecz zdanie „*Mąż Marii* składa się z trzech liter” jest fałszem, chociaż Jan jest faktycznie mężem Marii³⁸. Skrytykowana przez Davidsona intuicyjna teoria wyrażen cudzysłowowych autorstwa Tarskiego omija wprawdzie tę trudność³⁹, lecz przez niego odrzucona z powodu konfliktu z wymogiem wyuczalności.

Podobnie mają się sprawy w przypadku kilku wybranych teorii mowy zależnej. Davidson pokazuje, że trzy skrytykowane przez niego rozwiązania problemu mowy zależnej powstały w tym celu, aby uporać się z niezachodzeniem zasady zastępowania oraz miały – w opinii ich autorów – stanowić ważny element analizy języka naturalnego. Jak argumentuje Davidson, błąd

³⁷ W semantyce Fregego wszystkie wyrażenia mogą mieć – lecz nie muszą – odniesienia. Pozwala to formułować zasadę zastępowania ekstensjonalnego odwołując się bezpośrednio do odniesień wyrażen (o ile istnieją).

³⁸ Z punktu widzenia analizy parataktycznej faktycznie nie są to zdania, lecz pary złożone ze zdania oraz wyrażenia znajdującego się w cudzysłowie.

³⁹ Problem niezachodzenia zasady zastępowania ekstensjonalnego nam nie grozi, gdyż zgodnie ze stanowiskiem Tarskiego wyrażenie cudzysłowowe tworzy całość, w której nie można zastąpić zawartości cudzysłowu wyrażeniem o tym samym odniesieniu. Ponieważ całe wyrażenie cudzysłowowe stanowi nazwę wyrażenia wewnątrz cudzysłowu, nie da się go zastąpić innym wyrażeniem cudzysłowowym, lecz – jak proponował Tarski – opisem strukturalnym tego wyrażenia, np. „imię składające się z liter J , a oraz n ” lub „imię męża Marii”. Przy takiej zamianie zdania „imię składające się z liter J , a oraz n składa się z trzech liter” oraz „Imię męża Marii składa się z trzech liter” są prawdziwe.

polegał na zignorowaniu wymogu, aby teorie te były teoriami języka wyuczanego, a w konsekwencji teorie te nie były i nie mogły być adekwatnymi teoriami języka naturalnego.

Pierwsza z krytkowanych przez Davidsona teorii, to tzw. inskrypcyjna teoria mowy zależnej, zaproponowana przez Izraela Schefflera w *An Inscriptional Approach to Indirect Quotation* (1954). Zgodnie z propozycją Schefflera zdanie „Tonkin powiedział, że śnieg jest biały” (*Tonkin said that snow is white*)⁴⁰ powinno być analizowane jako „Tonkin wymówił wypowiedzenie że-śnieg-jest-biały” (*Tonkin spoke a that-snow-is-white utterance*). Wedle Schefflera sztucznie brzmiące wyrażenie „wypowiedzenie, że-śnieg-jest-biały” jest terminem ogólnym (predykatem jednoargumentowym), którego zakres stanowią wypowiedzenia lub napisy (*inscriptions*) zdania *Snow is white* lub jego tłumaczenia na inne języki. Wówczas zdanie *Tonkin spoke a that-snow-is-white utterance* jest prawdziwe wtedy i tylko wtedy, gdy Tonkin wymówił jedno z tych wypowiedzeń. Predykat ten – jak argumentuje Davidson – nie może być już dalej analizowany na gruncie teorii Schefflera, stanowiąc tym samym wyrażenie semantycznie proste. Ponieważ predykatów takich jest nieskończenie wiele (gdyż repertuar Tonkina obejmuje nieskończenie wiele wypowiedzeń zdań), teoria Schefflera nie może być teorią języka wyuczanego (s. 12).

Podobnie jak w przypadku wyrażeń cudzysłowowych, odejdźmy na chwilę od rozważań Davidsona z *Theories of Meaning* i odwołajmy się do znacznie późniejszej pracy Schefflera, w której autor deklaruje skrajny nominalizm. W przypadku wyrażeń językowych nominalizm ten przejawia się właśnie inskrypcjonizmem:

(...) moje podejście zawsze było nominalistyczne, unikające obiektów abstrakcyjnych i intensjonalnych oraz zakładające jedynie indywidualne obiekty oznaczające i indywidualne obiekty oznaczane. W szczególności dla języka podejście to, zwane inskrypcjonizmem, zakłada jedynie indywidualne okazy znaków językowych (wypowiedzenia i inskrypcje) oraz indywidualne obiekty, do których się one odnoszą (Scheffler 1997, s. 6).

Stanowisko Schefflera do złudzenia przypomina reizm Tadeusza Kotarbińskiego, który również wystrzegał się mówienia o wszelkich bytach abstrakcyjnych oraz intensjonalnych⁴¹. W szczególności obaj unikali mówienia o typach wypowiedzeń, chociaż nie byli w tej sprawie konsekwentni.

⁴⁰ Oryginalny przykład cytowany przez Davidsona.

⁴¹ Należy zauważyć również, że stanowisko Kotarbińskiego antycypowało Davidsona analizę parataktyczną zdań psychologicznych.

W konstrukcji Schefflera używa się *de facto* typów wypowiedzeń, gdyż zdanie „Tonkin wymówił wypowiedzenie że-śnieg-jest-biały” można sparafrazować jako „Tonkin wymówił wypowiedzenie typu, do którego należy to wypowiedzenie. Śnieg jest biały”. Ponieważ do typu, o który tu chodzi, należą wszystkie wypowiedzenia synonimiczne z wypowiedzeniem „Śnieg jest biały”, ostateczna wersja może przybrać postać następującą: „Tonkin wypowiedział zdanie synonimiczne ze zdaniem: Śnieg jest biały”.

Okazuje się zatem, że pomimo podniesionego tu zarzutu wprowadzania tylnymi drzwiami niedozwolonych pojęć typów wypowiedzeń opartych na synonimiczności⁴², rozwiązanie Schefflera sugeruje rozwiązanie problemu mowy zależnej, które Davidson znacznie wcześniej zaproponował w *On Saying That* (1968). W myśl tego rozwiązania w poprawnej analizie zdania postaci „*X* powiedział, że *p*” musi pojawić się zaimek wskazujący na wypowiedzenie zdania *p* dokonane przez osobę, która wypowiedziała lub napisała zdanie „*X* powiedział, że *p*”. Wypowiedzenie to jest synonimiczne z wypowiedzeniem zdania *p* przez *X*-a. Podobnie jak w przypadku analizy wyrażen cudzysłowowych, zdanie wskazujące oraz zdanie wskazywane są niezależne logicznie. Co więcej, podejście to zakłada eliminowalność kontekstów nieekstensjonalnych, co pozwala na podanie dla tych zdań adekwatnej, tj. uwzględniającej wyuczalność, semantyki⁴³.

Podobnie jak w przypadku teorii Schefflera zarzut ignorowania wyuczalności stawia Davidson również zaproponowanemu przez Quine’a w *Słowie i rzeczy* (1960) rozwiązaniu problemu zdań o wierzeniach (*belief sentences*). Wedle Quine’a zdanie postaci „Tom wierzy, że Cyceron wydał Katylinę” powinno być analizowane w postaci *F(a)*, gdzie *a* zastępuje imię „Tom”, zaś *F* jest złożonym terminem ogólnym powstałym z operatora „wierzy, że” i zdania będącego jego argumentem. W podejściu tym – jak zauważa Davidson – nawet główny czasownik „wierzy” jest niedostępny jakiegokolwiek logicznej analizie⁴⁴: „Mówienie o składnikach i operatorach musi, rzecz jasna, być traktowane jako czysto syntaktyczne, bez oparcia na jakiegokolwiek teorii semantycznej” (s. 13), gdyż ta musi zakładać zawsze jakąś ontologię. Semantyka bez ontologii jest nieciekawa i faktycznie sprowadza się do syntaktyki. W konsekwencji język, dla którego nie jesteśmy w stanie sformułować teorii semantycznej opartej na jakiejś ontologii postulującej istnienie pewnych bytów, nie może być – zdaniem Davidsona – wyuczany (s. 14).

⁴² Te same zarzuty często stawiane są również Kotarbińskiemu.

⁴³ Podobnie jak w przypadku analizy zdań cudzysłowowych zagadnienie analizy parataktycznej mowy zależnej wymaga dokładniejszego omówienia w oddzielnej monografii.

⁴⁴ Zauważmy, że w proponowanym przez Quine’a podejściu zasady zastępowania ekstensjonalnego nawet się nie da zastosować.

W tym miejscu konieczne są dwie uwagi. Po pierwsze, mówiąc o ontologii Davidson ma na myśli ontologię postulowaną przez teorię semantyczną. Semantyka warunków prawdziwości oparta na semantycznej definicji prawdy postuluje bowiem istnienie obiektów oraz istnienie zdarzeń. Po drugie, fakt niewyuczalności języka opisywanego przez teorię Quine'a jest oczywisty, gdyż predykaty *F* są w myśl tej teorii syntaktycznie złożone, lecz semantycznie proste, zaś syntaktyczna złożoność sprawia, że mamy ich nieskończenie wiele.

O ile teorie Schefflera oraz Quine'a nie są teoriami języka wyuczanego, gdyż cierpią, zdaniem Davidsona, na niedostatek postulowanych przez nie bytów, to ostatnia z omawianych teorii, oparta na Fregego analizie mowy zależnej Alonzo Churcha teoria sensu i odniesienia jest pod tym względem niezwykle hojna. Niestety – jak pokazuje Davidson – wprowadzenie do semantyki nadmiaru bytów nie sprawia bynajmniej, że mamy do czynienia z semantyką języka wyuczanego.

Analizując mowę zależną w artykule *Sens i znaczenie* (1892) Frege przypisał predykatom psychologicznym wyrażającym postawy propozycjonalne, np. „wierzyć, że”, podwójną rolę. Po pierwsze, predykaty te tworzą kontekst, w którym odniesieniem występującego w nich zdania podrzędnego nie jest jego zwykłe odniesienie, czyli wartość logiczna, lecz zwykły sens tego zdania, czyli myśl, którą owo zdanie wyraża. Ponieważ dla zdań podrzędnych obowiązują zasady składalności sensu oraz składalności odniesienia⁴⁵, ta sama uwaga odnosi się do wszystkich składników owych zdań – odniesieniami wyrażen w kontekście intensjonalnym stają się ich zwykłe sensy. Zabieg ten pozwolił Fregemu ocalić zasadę zastępowania ekstensjonalnego, gdyż w zdaniu podrzędnym nie można wymieniać wyrażen o tym samym odniesieniu, lecz jedynie wyrażenia o tym samym sensie (czyli synonimy). W ten sposób w zdaniu „Piotr myśli, że Jan pali”, terminu „Jan” nie można zastąpić terminem „mąż Marii”, gdyż wprowadzie terminy te oznaczają ten sam obiekt, ale ich odniesieniami w analizowanym kontekście nie są obiekty, lecz sensy owych wyrażen. Ponieważ sensy są różne, terminy te są w podanym kontekście niewymienialne.

Po drugie, predykaty psychologiczne, np. „myśli”, są funktorami, które parom złożonym z obiektu (osoby) oraz zwykłego sensu zdania podrzędnego, czyli specjalnego odniesienia, przypisują wartości logiczne (Frege 1975, s. 60–88). Ponieważ zwykłe sensy zdań oraz sensy ich składników stają się

⁴⁵ W najogólniejszej postaci zasada składalności odniesienia (sensu) głosi, że odniesienie (sens) wyrażenia złożonego zależy od odniesień (sensów) wyrażen składowych i sposobu, w jaki tworzą wyrażenie złożone. Do sprawy tej powrócę w ostatnim paragrafie.

w tych kontekstach ich odniesieniami, to zarówno zdania podrzędne, jak i ich składniki muszą posiadać w tych kontekstach również inne, specjalne sensory⁴⁶. Niestety teoria Fregego oraz oparta na niej teoria Churcha, do której bezpośrednio odwołuje się Davidson, nie podają reguł, które pozwalają generować zwykłe sensory ze zwykłych odniesień, a specjalne sensory z sensów zwykłych. Sprawę dodatkowo pogarsza fakt, że zdania w mowie zależnej mogą być zdaniami podrzędnymi w kolejnych zdaniach z czasownikami wyrażającymi kolejne postawy propozycyjne, co prowadzi do konieczności wprowadzania sensów coraz to wyższych rzędów. W rezultacie z każdym zdaniem języka należy wiązać nieskończenie wiele sensów, a uwzględniając zasadę składalności, nieskończenie wiele sensów jesteśmy zmuszeni wiązać również z każdym słowem.

Wielokrotnie wspominałem, że teoria języka wyuczanego musi być rekurencyjna, zaś teorie Churcha oraz Fregego są rekurencyjne. Jednak język opisywany przez owe teorie jest niewyuczalny, ponieważ obie teorie zakładają nieskończoną ilość wyrażen semantycznie prostych⁴⁷, gdyż każde słowo lub idiom może pełnić w nieskończenie wielu możliwych kontekstach użycia funkcję innego wyrażenia semantycznie prostego.

6. WYUCZALNOŚĆ A SKŁADALNOŚĆ

Każdy, kto zna semantykę Fregego, w sposób oczywisty skojarzy omówiony wcześniej wymóg wyuczalności z przypisywaną Fregemu zasadą składalności odniesień oraz zasadą składalności sensów. Ponieważ Frege traktował sensory jako szczególnego rodzaju odniesienia, najczęściej mówi się po prostu o zasadzie składalności (*compositionality*). W swym najogólniejszym sformułowaniu zasada składalności głosi, że znaczenie wyrażenia złożonego – bez względu jak rozumiemy termin „znaczenie” – zależy od znaczeń jego części oraz sposobu, w jaki są one połączone w tym wyrażeniu. W paragrafie tym przyjrzymy się związkom zachodzącym między wyuczalnością a składalnością.

Podane wyżej sformułowanie zasady składalności – jedno z wielu spotykanych w literaturze – odwołuje się bezpośrednio do wprowadzonego przez

⁴⁶ Wedle Fregego każde wyrażenie językowe, w każdym możliwym kontekście musi posiadać sens, gdyż inaczej nie byłoby zrozumiałe.

⁴⁷ Już w latach 50. Davidson krytykował analizę zdań psychologicznych zaproponowaną przez Fregego w *Sensie i znaczeniu*. Napisany w tym czasie artykuł *The Method of Extension and Intension*, który opublikowany został dopiero w roku 1963, wskazywał na wady zaproponowanej przez Rudolfa Carnapa w *Meaning and Necessity* (1947) korekty klasycznego rozwiązania autorstwa Fregego.

Fregego podziału wyrażeń językowych na nasycone i nienasycone. Wyrażeniami nienasyconymi są funktory, zaś nasyconymi wyrażenia, które funktorami nie są, a mogą być jedynie argumentami funktorów. Wyrażeniami nasyconymi są zatem wyrażenia jednostkowe (deskrypcje, nazwy własne) oraz zdania. Ich odniesieniami są odpowiednio indywidua oraz wartości logiczne, a sensami tzw. pojęcia jednostkowe oraz sensy zdań, które w przypadku zdań oznajmujących Frege nazywa myślami. Odniesieniami i sensami funktorów są funkcje różnego typu.

Semantyka odniesieniowa Fregego zapoczątkowała rozwój tzw. semantyki teoriomodelowej języka naturalnego, której wybitnym przedstawicielem był wspomniany już Richard Montague. To właśnie semantyka Montague stanowiła punkt odniesienia dla przeglądowego artykułu Barbary H. Partee *Compositionality* (1984), w którym autorka wskazała na wiele problemów związanych z zachodzeniem zasady składalności. Nie wchodząc w szczegóły analizy przedstawionej w tym artykule, wiele konstrukcji składowych opisanych w ramach paradygmatu semantyki Montague zdaje się ignorować zasadę składalności.

Według Davidsona wyuczalność jest pewną cechą języka⁴⁸, zaś adekwatna teoria znaczenia powinna ową dyspozycję uwzględniać i wyjaśniać. Aby tak było, w myśl podanej poprzednio analogii do rozpuszczalności oraz teorii rozpuszczalności, adekwatna teoria znaczenia (teoria semantyczna) powinna posiadać pewne określone cechy. Problem polega na identyfikacji owych cech. Pierwsza nasuwająca się propozycja to zachodzenie zasady składalności. W tym miejscu natrafiamy jednak na spory problem, gdyż w podanym wyżej sformułowaniu zasada ta ma zastosowanie do teorii odniesieniowych, czyli takich, które przyporządkowują wyrażeniom językowym pewne byty jako odniesienia i znaczenia. Wprawdzie oparta na semantycznej definicji prawdy semantyka Davidsona postuluje pewne byty jako odniesienia, ale jest ona w istocie teorią nieodniesieniową. W szczególności teoria ta nie przyporządkowuje wyrażeniom żadnych bytów jako znaczeń i nie daje się do niej bezpośrednio zastosować zasady składalności w wersji przedstawionej przez Fregego dla jego własnej semantyki odniesieniowej. Pewne zarzuty wobec

⁴⁸ Jest to pewien skrót myślowy, gdyż predykat „wyuczalny” jest predykatem dyspozycyjnym. Na początku tego rozdziału naszkicowałem analogię między wyuczalnością a rozpuszczalnością, gdzie predykat „rozpuszczalny” jest predykatem dyspozycyjnym. Aby przedmiot był rozpuszczalny musi posiadać pewne cechy dające się opisać w języku fizykalnym. Podobnie, aby język był wyuczalny, musi posiadać cechy, które dają się opisać za pomocą odpowiedniej teorii rekurencyjnej (np. semantycznej definicji prawdy). Nazywanie cechy za pomocą predykatu dyspozycyjnego jest całkowicie zgodne z poglądami Davidsona.

semantyk odniesieniowych zostaną przedstawione na początku następnego rozdziału.

Można oczywiście przeformułować wymóg składalności do postaci następującej: znaczenie wyrażenia złożonego w jakiś sposób zależy od znaczeń jego części oraz sposobu, w jaki są połączone. W takiej postaci zasada składalności jest raczej niekontrowersyjna, a jej stosowanie nie ogranicza się to teorii odniesieniowych. Przy takim sformułowaniu nie operujemy bowiem *explicite* funkcjami, będącymi odniesieniami czy sensami funktorów. Należy jednak zauważyć, że zachodzenie tak rozumianego wymogu składalności nie zapewnia jeszcze, aby teoria semantyczna była teorią języka wyuczalnego⁴⁹. Dlatego też Davidson pisał w *Theories of Meaning*, że teoria języka powinna dodatkowo podawać konstruktywny opis (*constructive account*) znaczeń jego wyrażeń. Konstruktywny opis to nic innego niż opis dany za pomocą skończonej ilości reguł (lub aksjomatów), które pozwalają w pewien, wymagający jeszcze doprecyzowania sposób, podać (dla danego języka) znaczenie dowolnego wyrażenia. Tak rozumiany wymóg konstruktywności opisu znaczeń zawiera już w sobie *implicite* wymóg skończonej ilości wyrażeń semantycznie prostych, którym odpowiada skończona ilość reguł (aksjomatów) semantycznej definicji prawdy.

W przypadku „odniesieniowej” wersji zasady składalności Partee wskazała na szereg konstrukcji gramatycznych, dla których owa zasada zdaje się nie obowiązywać, gdyż znaczenie wyrażenia złożonego nie jest funkcją znaczeń wyrażeń prostych. Przenosząc to na ogólniejsze sformułowanie zasady składalności wzbogacone o wymóg skończonej ilości wyrażeń semantycznie prostych zauważmy, że wyuczalność języka nie jest w tym przypadku zagrożona, o ile znaczenia wyrażeń złożonych są opisane za pomocą skończonej ilości reguł (aksjomatów). Pochodząca od Fregego zasada składalności w wersji „odniesieniowej” nie była motywowana wyuczalnością, gdyż nie była pierwotnie przeznaczona do języków naturalnych, lecz formalnych teorii logicznych i matematycznych⁵⁰. Nie należy się zatem dziwić, że w przypadku języka naturalnego sens (znaczenie) lub odniesienie wyrażenia złożonego w niektórych przypadkach nie jest funkcją sensów (znaczeń) lub odniesień wyrażeń składowych. Tak jest w przypadku choćby idiomów, których istnienie w języku nie zagraża jego wyuczalności ze względu na ich skończoną ilość. Gdyby zachodzenie zasady składalności miało gwarantować, aby teoria

⁴⁹ Tak jest, gdy w języku mamy nieskończenie wiele wyrazów.

⁵⁰ W językach tych mamy zazwyczaj nieskończenie wiele wyrażeń semantycznie prostych, np. zmiennych, stałych indywidualnych czy predykatów. Nie ma jednak powodu, aby do takich języków stosować wymóg wyuczalności. Języków takich nie uczymy się i nie rozumiemy w taki sposób, w jaki uczymy się i rozumiemy języki naturalne.

odniesieniowa była semantyką języka wyuczanego, należałoby nałożyć na ową zasadę pewne dodatkowe warunki. Takim warunkiem mógłby być wymóg, aby sensy wszystkich wyrażeń języka podawane były przez skończoną ilość funkcji. Możliwości tej nie będę tutaj analizował, gdyż Davidson z powodów omówionych w kolejnym rozdziale odrzucił semantyki odniesieniowe i za teorię znaczenia uznał semantyczną definicję prawdy, w której termin „znaczenie” w ogóle nie występuje.

Jako zwolennik semantyk nieodniesieniowych Davidson prawdopodobnie zdawał sobie sprawę, że zasada składalności w tradycyjnym sformułowaniu nie daje się bezpośrednio zastosować do jego teorii, tj. zapewnić, aby teoria ta była teorią języka wyuczanego. Dlatego też w *Theories of Meaning*, podobnie jak w innych publikacjach, nie używa w ogóle tego terminu. Wymóg składalności znaczenia został u niego wyrażony – o czym już była mowa – wymogiem podawania konstruktywnego opisu znaczeń, co miało gwarantować, aby teoria semantyczna była teorią języka wyuczanego.

Przedstawiona powyżej interpretacja wymogu wyuczalności jako zmodyfikowanej wersji warunku składalności znajduje również oparcie w literaturze poświęconej filozofii Davidsona. Na przykład drugi rozdział monografii Lepore’a i Ludwiga *Donald Davidson. Meaning, Truth, Language, and Reality* nosi tytuł „Języki wyuczane a wymóg składalności” (por. Lepore i Ludwig 2005, s. 25–37). Już na początku tego rozdziału autorzy utożsamiają wyuczalność ze składalnością, pisząc wprost, że wymóg składalności został wprowadzony przez Davidsona w *Theories of Meaning* (s. 25). Nieco dalej eksplikują wymóg konstruktywnego opisu znaczeń podawanych przez teorię zwaną przez nich składaną teorią znaczenia:

[CM] Składalna teoria znaczenia języka *L* jest to teoria formalna, która pozwala każdemu, kto zna język, w którym ta teoria jest sformułowana, rozumieć wyrażenia proste języka *L* oraz na tej podstawie rozumieć wyrażenia złożone tego języka (*op. cit.* s. 26–27).

W sformułowaniu tym „pozwała rozumieć wyrażenie” znaczy tyle, co „podaje znaczenie wyrażenia”. Podawanie znaczenia zdania przez zbiór reguł teorii polega, jak twierdzą Lepore i Ludwig, na tym, że każdy, kto zna owe reguły, może zrozumieć odpowiednie zdanie (s. 27). Oczywiście reguły te (bądź aksjomaty) mogą mieć postać funkcji, lecz argumentami owych funkcji nie są w żadnym razie znaczenia wyrażeń wchodzących w skład zdania. Warunek ten spełnia semantyczna definicja prawdy, która jest teorią rekurencyjną. To właśnie rekurencyjność teorii wraz z wymogiem skończoności słownika wyrażeń semantycznie prostych gwarantuje, że znaczenie wyrażenia złożonego (czyli zdania) zależy od znaczeń wyrażeń składowych oraz od sposobu,

w jaki zdanie to jest zbudowane z wyrażen składowych. Do sprawy tej powrócimy w kolejnym rozdziale.

Na zakończenie warto wspomnieć, że wymóg wyuczalności spotkał się z wieloma obiekcjami. Zostały one przedstawione m.in. w artykule Matthews'a *Learnability of Semantic Theory* (1986) oraz w cytowanym powyżej rozdziale drugim monografii Lepore'a i Ludwiga, w którym autorzy krytycznie odnoszą się również do pewnych propozycji Matthews'a.

Theories of Meaning jest tekstem wstępnym, w którym Davidson nakłada ograniczenia na właściwą teorię znaczenia języka naturalnego. Jak pokazuje w krytyce teorii Fregego i Churcha oraz analizie wyrażen cudzysłowowych teoria znaczenia powinna być ekstensjonalna, gdyż konteksty intensjonalne generują dla każdego nieskończoną ilość sensów. Oczywiście ten ostatni wymóg oznacza, że język naturalny musi być *de facto* uznany za ekstensjonalny, zaś konteksty intensjonalne powinno dać się z niego eliminować. Teoria prawdy jest teorią ekstensjonalną, a zatem jest dobrym kandydatem na adekwatną teorię znaczenia. Davidson stwierdza wprawdzie, że Tarskiego teoria prawdy traktowana jako teoria znaczenia spełnia wymóg wyuczalności (lub składalności w podanym tu sformułowaniu), ale nie twierdzi, że jakaś zupełnie inna teoria tego wymogu nie może spełniać. Wymóg wyuczalności zapewniony jest nie tylko przez ekstensjonalność teorii prawdy, ale również przez jej rekurencyjność. Oczywiście sama rekurencyjność teorii nie zapewnia wyuczalności opisywanego języka, gdyż teorie rekurencyjne, w tym teoria prawdy, mogą również opisywać języki niewyuczalne, np. języki pewnych teorii logicznych lub matematycznych. Jednak języki naturalne zawierają skończoną ilość słów, idiomów i konstrukcji gramatycznych. Dlatego też rekurencyjna teoria prawdy może służyć jako teoria znaczenia owych języków. Poczynione do tej pory uwagi nie wykluczają możliwości, że adekwatna teoria języka może być odpowiednio zaadaptowaną do tego celu semantyką odniesieniową. W słynnym artykule *Prawda i znaczenie* (1967), Davidson przeprowadza gruntowną krytykę semantyk odniesieniowych, ograniczając tym samym zbiór potencjalnych teorii znaczenia języka do nieodniesieniowych teorii w stylu teorii prawdy Tarskiego, która spełnia wszystkie wymogi narzucone na adekwatną teorię znaczenia języka naturalnego. O ile zatem w *Theories of Meaning* Davidson wskazał jakie warunki powinna spełniać teoria znaczenia języka naturalnego, to w *Prawdzie i znaczeniu* pokazał przede wszystkim jaką teorią znaczenia być nie powinna. W tym celu przeprowadził gruntowną krytykę semantyk odniesieniowych, podając szereg argumentów na rzecz ich bezużyteczności jako teorii języka naturalnego.

ROZDZIAŁ 3

SEMANTYKA NIEODNIESIENIOWA

Mój zarzut wobec znaczeń w teorii znaczenia dotyczy nie tego, że są one abstrakcyjne, czy też, że ich kryteria identyczności są niejasne, ale tego, że nie mają one jasnego zastosowania (Davidson 1992/1967, s. 8).

Termin „semantyka nieodniesieniowa” jest niewątpliwie bardzo wieloznaczny – w skrajnym przypadku może oznaczać nawet skrajny antyrealizm. Nie taka była jednak intencja Davidsona, który nie negował istnienia zewnętrznej wobec nas i niezależnej od nas rzeczywistości. Wedle Davidsona nie ma powodu, aby wątpić w realne istnienie mnogości obiektów, lecz w pewnych sytuacjach można się mylić co do istnienia tego, czy tamtego obiektu oraz można mówić o obiektach nieistniejących (fikcyjnych) lub abstrakcyjnych. Jak pamiętamy semantyczna definicja prawdy była traktowana przez Davidsona jako teoria, która postulowała istnienie obiektów. To czy postulowane przez teorię obiekty faktycznie istnieją nie da się rozstrzygnąć na jej gruncie. Nie ma jednak powodu, abyśmy nie mogli mówić w ten sam sposób o obiektach, o których wierzymy, że istnieją fizycznie oraz o obiektach, o których wiemy, że są wytworem fantazji literackiej lub operacji umysłu matematyka. Podejście to przypomina bardzo poglądy Quine’a, który twierdził, że istnieć, to być wartością zmiennej. Jednak waga, jaką w swych poglądach przykładał do relacji przyczynowej, sprawiała, że musiał uznać realne istnienie zdarzeń oraz obiektów, które w owych zdarzeniach uczestniczą. Problem nie dotyczy zatem ontologicznego pytania o istnienie rzeczywistości, lecz raczej tego, że możemy się co do owej rzeczywistości mylić, przypisując istnienie rzeczom nieistniejącym i *vice versa*.

Jak pamiętamy, wadą empirystycznych teorii języka było to, że konsekwencją ich przyjęcia była konieczność uznania, że istnieją pewne wyróż-

nione obiekty – te mianowicie, bez których nie byłoby możliwe nauczenie się języka. Innymi słowy teorie te zakładały warunkującą możliwość nauczenia się języka relacji oznaczania, łączącej niektóre wyrażenia językowe z ich realnie istniejącymi odniesieniami. Relacja ta gwarantowała sensowność nie tylko tych wyrażen, które posiadały odniesienia, ale również innych (np. nazw obiektów fikcyjnych), o których było wiadomo, że odniesień nie posiadały. Aby zapewnić sensowność takim wyrażeniom można było postąpić na jeden z dwóch sposobów. Po pierwsze, można było dokonać pewnej operacji na znaczeniach wyrażen oznaczających, co prowadziło do utożsamienia sensowności wyrażenia językowego z posiadaniem przezeń znaczenia rozumianego jako pewnego rodzaju byt. Znaczenie stawało się zatem szczególnym odniesieniem, które każde sensowne wyrażenie języka musiało posiadać¹. O takich teoriach powiemy, że reifikują znaczenia. Przykładem takiego podejścia była Fregego teoria sensu. Po drugie, można dokonać pewnych operacji na języku, które pozwalały sprowadzić wszystkie wyrażenia sensowne do wyrażen oznaczających. Przykładem zastosowania tej strategii jest teoria deskrypcji Russella w wersji z artykułu *On denoting* (1905). W artykule tym Russell proponuje rozwiązanie paradoksów (zagadek) – związanych m.in. z deskrypcjami nieoznaczającymi – na płaszczyźnie syntaktycznej, bez konieczności wprowadzania specjalnych sposobów istnienia odniesień oraz reifikowania znaczeń². Zanim przyjrzymy się podanym przez Davidsona argumentom wymierzonym przeciw semantykom odniesieniowym, warto przypomnieć kilka takich semantyk, pamiętając jednak, że nie każdą semantykę odniesieniową można zaliczyć do teorii empirystycznych.

1. PRZEGLĄD WYBRANYCH SEMANTYK ODNIESIENIOWYCH

Wśród teorii semantycznych, które określiłem etykietą „odniesieniowe” można wyróżnić dwa zasadnicze typy. Pierwszy typ to teorie oparte na relacji odniesienia, która musi zachodzić między określonymi wyrażeniami języka

¹ Relację między wyrażeniem językowym a jego znaczeniem określano tradycyjnie mianem relacji wyrażania. A zatem wyrażenie językowe oznaczało swój desygnat (zakres, odniesienie itp.), lecz wyrażało swoje znaczenie.

² Zagadki (*puzzles*), które Russell próbuje rozwiązać, to pozorne odnoszenie się deskrypcji do obiektów nieistniejących, np. w zdaniu „Obecny król Francji jest łysy”, zagadka negatywnych zdań egzystencjalnych z deskrypcją nieoznaczającą w funkcji podmiotu, np. „Obecny król Francji nie istnieje”, zagadka Fregego (polegająca różnicy wartości poznawczej zdań tożsamościowych postaci $a = a$ oraz $a = b$) oraz niezachodzenie zasady zastępowania ekstensjonalnego w mowie zależnej. Dwie ostatnie zagadki rozwiązał Frege reifikując sensy. Dokładne omówienie zagadek Fregego można znaleźć w (Lycan 2000).

(zazwyczaj nazwami lub zdaniami) a jakimiś fragmentami rzeczywistości pozajęzykowej. Jak już podkreślałem, w tak rozumianych semantykach odniesieniowych nie wymaga się, aby relacja nazwa – rzecz zachodziła dla wszystkich wyrażen, np. w teorii znaku Arystotelesa oraz w teorii asocjacionistów za sensowne uznawano również nazwy nieoznaczające, tj. nazwy puste. Sensowność tym nazwom nadają znaczenia, które Locke utożsamiał z ideami, rozumianymi jako pewne byty mentalne zwane przez niego ideami³. W przypadku nazw pustych ich znaczeniami są idee złożone, będące pewnymi konstrukcjami umysłu, które budują je z idei prostych. Te z kolei muszą bezpośrednio odpowiadać pewnym danym doświadczenia, a dokładniej realnie istniejącym rzeczom oznaczanym przez odpowiednie nazwy. A zatem relacja słowo – rzecz stanowi podstawowe pojęcie teorii semantycznej asocjacionizmu, która umożliwia zbudowanie całej teorii. Dlatego też semantyka ta w pełni zasługuje na miano semantyki odniesieniowej pierwszego typu.

Jak pamiętamy, Davidson nie neguje faktu zachodzenia relacji między słowem a rzeczą. Krytykuje natomiast charakterystyczny dla empiryzmu pogląd, że można wyróżnić pierwotne atomy znaczeniowe (np. idee proste), z których rekonstruuje się znaczenia pozostałych wyrażen. Idee proste są wyróżnione spośród pozostałych idei (znanen) ze względu na fakt pozostawania w ścisłej relacji z rzeczami⁴. Wedle empirystów proces konstruowania pozostałych znaczen, czyli tworzenie idei złożonych, pokrywa się – jak pamiętamy – z procesem uczenia się pierwszego języka. Początkowo dziecko poznaje jedynie część języka, tj. atomy znaczeniowe, później stopniowo zaczyna rozumieć znaczenia złożone. Jak starał się pokazać Davidson, ten z gruntu fałszywy pogląd na uczenie się języka jest ściśle związany z błędną teorią znaczenia. Poprawna, czy też adekwatna teoria znaczenia nie może opierać się na atomach znaczeniowych, lecz musi dotyczyć języka jako całości⁵.

³ W przypadku teorii znaku Arystotelesa sprawa nie jest bynajmniej tak prosta. Znaczeniem ogólnej nazwy nieoznaczającej, np. „centaur”, jest wyobrażenie (pojęcie) centaury, zaś nazwa ta posiada jedynie definicję nominalną. W przypadku nazw oznaczających znaczenie ma jakby „dwie strony”, które odpowiadają definicji nominalnej i realnej. Są to wyobrażenie oraz forma gatunkowa. Odniesieniem w tym przypadku byłby gatunek rozumiany jako uniwersale.

⁴ Ujął to doskonale David Hume, pisząc, że idee to po prostu mniej wyraźne impresje, czyli mniej wyraźne doznania zmysłowe.

⁵ W późniejszych pismach epistemologicznych, zebranych w tomie *Objective, Subjective, Intersubjective* (2001a), Davidson wielokrotnie deklaruje, że uczenie się języka nie polega na rozpoznawaniu semantycznej relacji znak językowy – fragment rzeczywistości, lecz ma charakter przyczynowy. Podobną myśl znajdziemy już u Arystotelesa, który w *Kategoriach* deklaruje, że „istnienie rzeczy wydaje się w pewien sposób przyczyną istnienia prawdziwego twierdzenia” (por. Arystoteles 1990b, 14b). Różnica między omawianymi do tej pory teoriami

Z kolei John S. Mill uważał, że każda nazwa musi być zawsze nazwą czegoś rzeczywistego lub wyobrażonego. Nazwy rzeczy wyobrażonych nie mają realnie istniejących desygnatów, lecz – podobnie jak nazwy posiadające takie desygnaty – mają znaczenia, czyli konotacje. Konotacje nazw rzeczy jedynie wyobrażonych składają się z cech, które występują wprawdzie w obiektach rzeczywistych, ale nie występują łącznie w żadnym obiekcie rzeczywistym, będąc jedynie połączone w obiektach wyobrażonych. Zatem warunkiem sensowności nazw pustych, np. „centaur”, byłoby to, że posiadają te same cechy, które posiadają obiekty istniejące realnie. Pozwala to zaliczyć teorię Milla do pierwszego typu teorii odniesieniowych, gdyż zakłada się w niej istnienie obiektów jako nośników cech, których pewne kombinacje są konotacjami nazw pustych.

Wśród teorii odniesieniowych pierwszego typu warto również wymienić weryfikacjonistyczną teorię znaczenia uznawaną przez pewien czas przez przedstawicieli Koła wiedeńskiego. Teoria ta uznawana jest niekiedy za współczesną wersję asocjacionizmu Locke’a i Hume’a, z zastrzeżeniem, że dla weryfikacjonistów nośnikiem znaczenia nie była nazwa, lecz zdanie. Weryfikacjoniści uważali, że warunkiem sensowności empirycznej zdania jest istnienie sposobu jego empirycznej weryfikacji, czego konsekwencją była możliwość redukcji tego zdania do tzw. zdań protokolarnych, odpowiadających bezpośrednio danym doświadczenia. W tym sensie można powiedzieć, że zdania protokolarne posiadały odniesienie i stanowiły człony konstytuujące relację język – rzeczywistość. Inne zdania sensowne powinny się do zdań protokolarnych zredukować, zaś redukcja ta była możliwa dzięki występowaniu w nich wyrażen o znaczeniu analitycznym, czyli wyrażen logicznych oraz matematycznych. Sensowność analityczna tych wyrażen była z kolei zapewniona przez ich występowanie w postulatach znaczeniowych czyli aksjomatach teorii logicznych i matematycznych. Uczenie się języka postępowałoby oczywiście odwrotnie do owego procesu redukcji, czyli od zdań protokolarnych, odpowiadających prostym danym wrażeniowym, do zdań złożonych⁶. Należy również zauważyć, że wprawdzie weryfikacjoniści używali często terminu „znaczenie”, ale nie utożsamiali znaczeń z żadnymi

odniesieniowymi a teorią Davidsona ma w tym kontekście charakter zasadniczy. W omawianych do tej pory teoriach zakłada się bowiem *a priori* zachodzenie relacji odniesienia dla określonych wyrażen, np. tych wyrażen, których znaczeniami są idee proste, zaś relacja język – rzeczywistość jest konstruowana z takich par. Charakterystyczny dla poglądów Davidsona przyczynowy charakter związku na poziomie cały język – rzeczywistość nie pociąga za sobą konieczności istnienia par „określone wyrażenie” – „odniesienie tego wyrażenia”.

⁶ Jak wiadomo program szukania zdań protokolarnych się nie powiódł, zaś klęskę weryfikacjonizmu podsumował Quine w *Dwóch dogmatach empiryzmu*.

bytami. Zwrot „posiadanie znaczenia przez wyrażenie” rozumiany był przez nich jako synonim jego sensowności.

Jak widać, tradycyjny empiryzm włączał do teorii znaczenia relację oznaczania, utrzymując, że jej zachodzenie stanowi bezpośredni warunek sensowności niektórych wyrażen języka (nazw odpowiadających ideom prostym, zdaniom protokolarnym itp.) oraz pośredni warunek sensowności pozostałych wyrażen języka. Nawet Quine, który zwalczał „dogmaty empiryzmu”, nie do końca wyzwolił się spod jego wpływu. W *Słowie i rzeczy* (1960) przedstawił teorię znaczenia w postaci podręcznika przekładu⁷. Jest to teoria holistyczna, w której przyjmuje się, że znaczenia wyrażen nieznanego nam języka są podawane przez podręcznik przekładu (zdanie na zdanie) owego języka na znany nam język. Teoria ta, podobnie jak weryfikacjonizm, a w przeciwieństwie do asocjacionizmu i millianizmu (teorii Milla), nie reifikuje znaczeń. Aby skonstruować podręcznik przekładu Quine wprowadził pojęcie znaczenia bodźcowego, które – podobnie jak typowe koncepcje empirystyczne – operuje semantyczną relacją między językiem a rzeczywistością (a dokładniej między zdaniem a bodźcem). Znaczenie bodźcowe zdania stanowi zatem parę, której pierwszy człon jest to klasa bodźców, przy których zdanie jest akceptowane, a drugi człon – klasa bodźców, przy których zdanie jest odrzucane przez kompetentnego użytkownika języka. Zbudowanie podręcznika przekładu jest możliwe dzięki zdaniom okazjonalnym (a dokładniej obserwacyjnym), np. „Oto królik”, których tłumaczenia na język obcy są możliwe dzięki identyczności znaczeń bodźcowych w obu językach. Innymi słowy to właśnie semantyczna relacja język – rzeczywistość (a dokładniej zdania – bodźce) nadaje sens nie tylko zdaniom okazjonalnym, ale również wszystkim pozostałym zdaniom języka, co pozwala zaliczyć tę teorię do grupy teorii odniesieniowych pierwszego typu. Krytyka empiryzmu przedstawiona w *Theories of Meaning* dotyczy również Quine’a. Podany przez Quine’a opis tzw. przekładu radykalnego, który stanowi zarazem odpowiednik procesu uczenia się pierwszego języka, stanowi – z punktu widzenia *Theories of Meaning* – przykład rzutowania błędnej teorii uczenia się języka na teorię znaczenia. Nie oznacza to jednak, że Davidson odrzuca ideę przekładu radykalnego. Wręcz przeciwnie – wykorzystuje ją w tzw. interpretacji radykalnej, lecz – jak się przekonamy w ostatnim rozdziale – poważnie modyfikuje empiryczne warunki przekładu.

Do teorii odniesieniowych pierwszego typu można także zaliczyć – przy pewnych zastrzeżeniach oraz wbrew intencjom jej autora – Kazimierza Ajdukiewicza dyrektywalną teorię znaczenia. Zastrzeżenia co do zakwalifiko-

⁷ Dokładniejszy opis teorii Quine’a podaję w ostatnim rozdziale.

wania jej do teorii odniesieniowych pierwszego typu biorą się stąd, że podobnie jak u Davidsona, teoria ta dotyczyła języka jako całości, będąc tym samym teorią holistyczną⁸. Pomimo deklaracji, że teoria dyrektywalna nie jest teorią semantyczną, Ajdukiewicz nie zrezygnował w niej do końca z semantycznej relacji język – rzeczywistość. W konstrukcji tzw. macierzy języka pojawia się tzw. część empiryczna, na którą składają się pary złożone z danych doświadczenia i zdań. Ponieważ znaczenie wyrażenia jest reprezentowane przez jej miejsce w tzw. macierzy języka, zachodzenie tak rozumianej relacji między zdaniami empirycznymi a rzeczywistością stanowi warunek sensowności przynajmniej niektórych wyrażen. Dyrektywy empiryczne stanowią, jak starałem się pokazać w (2013) nie tylko relikty weryfikacjonizmu, ale rodzaje trudności polegające na tym, że – wbrew konwencjonalistycznym deklaracjom samego Ajdukiewicza – dane doświadczenia są w tej konstrukcji jednoznacznie nazywane⁹.

Teorią odniesieniową pierwszego typu jest również Saula Kripkego historyczno-przyczynowa teoria nazw własnych oraz Saula Kripkego i Hilarego Putnama historyczno-przyczynowa teoria nazw gatunków naturalnych. Ograniczając się do nazw własnych, w teorii tej akt „pierwszego chrztu” danego obiektu nadaje sensowność jego nazwie własnej, która od tej pory oznacza zawsze to samo indywiduum, chociaż różni użytkownicy tej nazwy mogą przypisywać temu indywiduum różne własności. Pomimo tych różnic odnoszą ową nazwę do tego samego indywiduum, co zapewnione jest przez istnienie historyczno-przyczynowego łańcucha użycia danej nazwy własnej, który zachowuje relację odniesienia. W teorii tej nazwa własna nie może być traktowana jako synonim jakiegokolwiek deskrypcji. Na przykład dwaj historycy filozofii mogą się spierać, czy Arystoteles był faktycznym autorem tego czy innego dzieła, ale używają nazwy „Arystoteles” w sposób sensowny i w tym samym znaczeniu¹⁰, gdyż odnosząc ją do tego samego indywiduum. A zatem podobnie jak w przypadku omawianych do tej pory teorii znaczenia, sensow-

⁸ W oryginalnym sformułowaniu teoria Ajdukiewicza nie spełniała podstawowego wymogu stawianego przez Davidsona teoriom znaczenia, gdyż nie jest to teoria rekurencyjna, a zatem nie jest teorią języka wyuczanego (por. Maciaszek 2015a).

⁹ Oczywiście dyrektywy empiryczne pozwalają uznawać różne zdania w różnych językach przedmiotowych, co jest zgodne z konwencjonalistycznymi poglądami Ajdukiewicza z tego okresu. Mój zarzut polega na tym, że dane doświadczenia są w sposób jednoznaczny indeksowane, co prowadzi do konstatacji, że istnieje wyróżniony język, dla którego zachodzi – jak w weryfikacjonizmie – jednoznaczna relacja między zdaniami a danymi doświadczenia.

¹⁰ Kripke nie próbował definiować znaczenia natomiast Hilary Putnam zaproponował pewną definicję znaczenia nazw gatunków naturalnych w (1975).

ność nazw własnych oraz nazw gatunków naturalnych oparta jest na relacji nazwa – obiekt¹¹.

Na zakończenie tego krótkiego przeglądu warto wspomnieć o semantykach proponowanych przez fenomenologów, np. Edmunda Husserla oraz Romana Ingardena. O ile w teoriach przedstawionych do tej pory sensowność wszystkich wyrażen języka zapewniona była przez posiadanie odniesień przez niektóre wyrażenia języka, to w teoriach fenomenologicznych wszystkie sensowne wyrażenia języka posiadały odniesienia. Było to możliwe dzięki temu, że fenomenolodzy dopuszczali różne sposoby istnienia. W szczególności odniesieniami nazw fikcyjnych były obiekty intencjonalne, powołane do istnienia i utrzymywane w istnieniu przez akty psychiczne użytkowników języka, np. czytelników powieści (por. Ingarden 1961, rozdz. IX „Forma I przedmiotu czysto intencjonalnego”).

Do tej pory była mowa o teoriach odniesieniowych pierwszego typu, tj. o teoriach, w których zakładano istnienie obiektów lub fragmentów rzeczywistości (np. zdarzeń), będących odniesieniami wyróżnionych wyrażen języka. W niektórych teoriach reifikowano znaczenia jako byty pewnego rodzaju, w innych zadowalano się jedynie sensownością wyrażen. Można zatem mówić o teoriach odniesieniowych również w innym niż w podanym powyżej sensie. Chodzi tu o teorie, które reifikują znaczenia, tj. przypisują wyrażeniom języka byty jako ich znaczenia, które to byty można uznać jako specjalne odniesienia. Dlatego też każda teoria drugiego typu jest zarazem teorią pierwszego typu, chociaż nie każda teoria pierwszego typu reifikuje znaczenia.

Jak łatwo zauważyć teoriami drugiego typu są niektóre z wymienionych teorii odniesieniowych typu pierwszego, tj. teoria znaku Arystotelesa, asocjacionizm i konotacyjna teoria Milla, jak również zasygnalizowana nieco wcześniej teoria Edmunda Husserla¹² oraz teoria Romana Ingardena¹³. Nie należy do

¹¹ W przypadku nazw gatunków naturalnych ich sensowność opiera się na akcie „pierwszego chrztu”, podczas którego pewna próbka substancji (wody, złota) lub pewien egzemplarz gatunku biologicznego został opatrzony daną nazwą. W kolejnych użyciach tej nazwy, nazywana próbka lub obiekt musi pozostawać w relacji należenia do tego samego gatunku naturalnego, co nazwana po raz pierwszy próbka lub obiekt. Ponieważ moim celem nie jest dokładna analiza historyczno-przyczynowej teorii nazywania, zainteresowanego czytelnika odsyłam do licznych opracowań tego tematu, np. (Lycan 2000).

¹² Teoria znaku Husserla została omówiona bardzo klarownie w monografii Bella *Husserl* (1990). Nie wchodząc w szczegóły, znaczenie w sensie Husserla jest bytem idealnym, co upodabnia teorię Husserla do Fregowskiej teorii sensu.

¹³ Wprawdzie Ingarden był uczniem Husserla, ale jego teoria znaczenia różni się znacznie od teorii nauczyciela. Dokładne przedstawienie ingardenowskiej teorii znaczenia można znaleźć w *O dziele literackim* (1960, rozdział V „Warstwa tworców znaczeniowych”).

nich natomiast ani teoria dyrektywalna, ani teoria weryfikacjonistyczna¹⁴. W przypadku historyczno-przyczynowej teorii Kripkego i Putnama sprawa nie jest jednoznaczna, gdyż Kripke znaczeń nie reifikuje, zaś w *Znaczenie wyrazu „znaczenie”* Hilary Putnam utożsamia znaczenie nazwy rodzaju naturalnego z wektorem, którego jedną ze składowych jest odniesienie (por. 1998/1975, s. 179–184)¹⁵.

Paradygmatycznym przykładem teorii odniesieniowej drugiego typu była dla Davidsona teoria sensu i odniesienia Fregego. To właśnie do niej stosuje się przede wszystkim krytyka semantyk odniesieniowych, którą przeprowadził w *Prawdzie i znaczeniu* (1966). Zwykłym odniesieniem wyrażenia językowego (o ile owo wyrażenie w ogóle je posiada)¹⁶ było dla Fregego indywiduum (dla nazw jednostkowych), wartość logiczna (dla zdań) oraz funkcja (dla wyrażen pozostałych kategorii syntaktycznych). Niezależnie od posiadania odniesienia, każde wyrażenie językowe posiadało znaczenie (sens), gdyż „Sens nazwy chwytają każdy, kto zna wystarczająco język lub symbolikę, do których ona należy” (Frege 1977, s. 62)¹⁷. Sens u Fregego jest specjalnym rodzajem odniesienia, gdyż – o czym już była mowa – pełni tę funkcję w kontekstach intensjonalnych, np. w zdaniach wyrażających postawy propozycyjne, czyli w mowie zależnej (s. 73 i nast.).

W *Prawdzie i znaczeniu* Davidson podał argumenty przeciwko przyporządkowywaniu wyrażeniom językowym jakichkolwiek odniesień – zarówno „zwykłych”, jak i „specjalnych”, czyli znaczeń. Argumentacja Davidsona skierowana była głównie przeciw Fregemu. Ponieważ w teorii Fregego znaczenie (sens) staje się w pewnych kontekstach odniesieniem, zaciera się rozróżnienie między tymi pojęciami. Warto o tym pamiętać, gdyż wyjaśnia to, dlaczego w *Prawdzie i znaczeniu* Davidson nie czyni rozróżnienia między

¹⁴ Charakterystyczne dla weryfikacjonistów stwierdzenie, że znaczeniem empirycznym zdania jest sposób jego empirycznej weryfikacji nie jest w istocie reifikacją znaczeń, lecz wskazaniem kryterium sensowności zdania.

¹⁵ Wydaje się, że podjęta przez Putnama próba rekonstrukcji pojęcia znaczenia jako złożonego bytu jest zabiegiem zasadniczo zbędnym w semantyce języka naturalnego. Posiadanie odniesienia jest bowiem w tej teorii wystarczającym warunkiem sensowności wyrażenia. Putnam starał się po prostu zbudować definicję sprawozdawczą terminu „znaczenie” dla nazw gatunków naturalnych.

¹⁶ Wedle Fregego każde wyrażenie języka (czyli wyrażenie sensowne) posiada zreifikowany sens, lecz nie musi posiadać odniesienia. Dzięki temu rozumiemy fikcję literacką.

¹⁷ Wprawdzie Frege wspomina tu o nazwach, ale uwagę tę można odnieść do wszystkich wyrażen języka. Zauważmy, że w przeciwieństwie do krytykowanych przez Davidsona teorii empirystycznych sensowność wyrażen językowych nie zależy u Fregego od posiadania odniesień przez określone wyrażenia.

odniesieniem a zreifikowanym znaczeniem¹⁸. Brak tego rozróżnienia ma jednak tę zaletę, że wysuwane przed Davidsona argumenty uderzają w teorie odniesieniowe obydwu typów.

2. W STRONĘ SEMANTYKI NIEODNIESIENIOWEJ

Prawda i znaczenie był jednym z najbardziej wpływowych i inspirujących tekstów napisanych przez Davidsona. W artykule tym twierdzi, że adekwatna teoria znaczenia nie powinna operować pojęciem znaczenia oznaczającym jakiś byt, a zatem powinna być – w pewnym sensie – teorią znaczenia „bez znaczeń”. A zatem adekwatna teoria znaczenia nie powinna znaczeń reifikować, jak to było w przypadku kilku wymienionych wcześniej teorii odniesieniowych drugiego typu, a nawet uderza w teorie odniesieniowe typu pierwszego. Za przewrotnym określeniem „teoria znaczenia bez znaczeń” kryje się jednak coś więcej. Zdaniem Davidsona w adekwatnej teorii znaczenia nie powinien występować w ogóle termin „znaczenie”. Teza ta zdaje się na pierwszy rzut oka niezgodna z deklaracjami Davidsona z *Theories of Meaning*. Jednym z zadań teorii znaczenia jest bowiem wyjaśnienie, w jaki sposób znaczenia zdań zależą od znaczeń słów, oraz w jaki sposób użytkownik języka może „wyznaczyć” znaczenie zdania na podstawie znajomości znaczeń słów. Jednak w sformułowaniu tym termin „znaczenie” występuje na poziomie przedteoretycznym i nie wiąże się ani z koniecznością użycia terminu „znaczenie” w samej teorii, ani – tym bardziej – z koniecznością reifikowania znaczeń.

Argumentacja Davidsona przedstawiona w *Prawdzie i znaczeniu* sprowadza się do wykazania, że w teorii znaczenia odwołującej się do znaczeń jako pewnych bytów, nie da się spełnić wymogu, aby znaczenie wyrażenia złożonego można było wyznaczyć na podstawie znaczeń wyrażen prostych, nawet gdyby to była – jak w semantyce Fregego zastosowanej do języka naturalnego – teoria rekurencyjna, operująca skończoną ilością znaczeń wyrażen semantycznie prostych¹⁹. Co więcej, Davidson argumentuje, że w adekwatnej teorii

¹⁸ Przy pierwszej lekturze może to stanowić dla czytelnika spory kłopot. Być może zamiast terminów „znaczenie” lub „zreifikowane znaczenie” zrzęcniej byłoby używać pojęcie „korelat semantyczny”.

¹⁹ Warto w tym miejscu ponownie podkreślić różnicę między wyrażeniami semantycznie prostymi (*semantic primitives*), a wyrażeniami, których znaczeniami są atomy znaczeniowe. Pojęcie wyrażenia semantycznie prostego, w przeciwieństwie do wyrażenia atomicznego, nie implikuje tezy, że znaczenie jest jakimkolwiek bytem. Przez atom znaczeniowy rozumiem zreifikowane znaczenie, które służy do budowania znaczeń złożonych. Przykładem atomu znaczeniowego jest idea prosta u asocjacionistów.

znaczenia termin „znaczenie” nie powinien *explicite* występować, gdyż termin ten wyznacza kontekst intensjonalny, zaś teoria znaczenia powinna być ekstensjonalna.

Przez to, że teoria przyporządkowuje wyrażeniom zreifikowane znaczenia Davidson rozumie, że wyrażenia w pewien sposób odnoszą się do znaczeń lub oznaczają znaczenia. W określeniu tym widoczna jest wspomniana już niezręczność, a zarazem dwuznaczność, gdyż utarło się, że wyrażenia oznaczają swoje odniesienia lub odnoszą się do swych odniesień, zaś wyrażają swe znaczenia. Jeżeli jednak przyjmiemy na wzór Fregego istnienie zreifikowanych znaczeń (sensów), które dodatkowo pełnią w pewnych kontekstach (np. w mowie zależnej) funkcję odniesień, niezręczność ta okazuje się jedynie pozorna, gdyż w semantykach odniesieniowych wyrażenie językowe oznacza swoje odniesienie (o ile w ogóle je posiada)²⁰, jak i – w pewien specjalny sposób – oznacza swoje znaczenie.

Wprowadzenie zreifikowanych znaczeń rodziło wielokrotnie problemy wyrażone pytaniem „Czego znakiem jest słowo – przedmiotu czy pojęcia?”. Problem ten widoczny już był w *Hermeneutyce*, gdzie Arystoteles pisze, że słowo jest znakiem rzeczy, lecz przede wszystkim znakiem wrażenia doznawanego w duszy (Arystoteles 1990, s. 69). Aby uniknąć tego typu problemów Mill pisał, że nazwa (konotacyjna) denotuje przedmiot (denotację), a współoznacza (konotuje) swą konotację (Mill 1962/1843, s. 51). Z kolei Carnap pisze w *Meaning and Necessity*, że wyrażenia posiadają ekstensję i intensję (1947, s. 23 i nast.). Jednak zdaniem Davidsona we wszystkich tych podejściach chodzi o zamaskowanie faktu, że w owych teoriach wyrażenie językowe jest znakiem dwóch bytów: odniesienia i znaczenia, rozumianego jako specjalne odniesienie. Usprawiedliwia to sposób, w jaki Davidson pisze w *Meaning and Necessity* o zreifikowanych znaczeniach, nie odróżniając ich od odniesień.

Pierwszym argumentem wymierzonym w teorie reifikujące znaczenia jest pewna wersja tzw. Argumentu trzeciego człowieka. Argument, który wziął nazwę od słynnego argumentu przytoczonego przez Platona przeciw istnieniu idei²¹, dotyczy sposobu konstytuowania znaczeń wyrażen złożonych,

²⁰ Jak łatwo zauważyć, wszystkie teorie znaczenia w jakiś sposób muszą uwzględniać fakt, że nie wszystkie wyrażenia języka posiadają odniesienia.

²¹ Argument trzeciego człowieka został przedstawiony przez Platona w późnym dialogu *Parmenides*. W sporym uproszczeniu argument ten przedstawia się następująco. Weźmy pod uwagę dwóch ludzi, którzy jako ludzie są do siebie podobni pod względem bycia człowiekiem. Jeśli się spytamy się o kryterium tego podobieństwa, platoński Sokrates odpowiedziałby, że podpadają pod wspólną ideę człowieka pełniącą funkcję wzorca, do której każdy z nich musi być podobny. Jeśli jednak spytamy się dlaczego poszczególny człowiek jest podobny do idei

np. zdań, przez znaczenia ich części, np. wyrazów. Argument ten przedstawię w dwóch wersjach: oryginalnej, będącej rozwinięciem niezwykle eliptycznej uwagi pochodzącej z *Prawdy i znaczenia* oraz w zmodyfikowanej przeze mnie, uwzględniającej wymóg wyuczalności języka. W obu wersjach wykazuje się, że przypisanie częściom wyrażenia złożonego zreifikowanych znaczeń nie pozwala na przypisanie (w jakimkolwiek tego słowa znaczeniu) zreifikowanego znaczenia wyrażeniu złożonemu, gdyż prowadzi to do regresu w nieskończoność.

Rozpatrzmy proste zdanie złożone jedynie z nazwy własnej w funkcji podmiotu oraz czasownika nieprzechodniego w funkcji orzeczenia – może to być przytoczone przez Davidsona zdanie „Teajtet lata”²². Z syntaktycznego punktu widzenia zdanie to stanowi konkatenację wyrazów „Teajtet” oraz „lata”. Jeżeli przypiszemy zreifikowane znaczenia zarówno zdaniu, jak i obu wyrażeniom składowym²³, to pojawia się problem, w jaki sposób byty będące zreifikowanymi znaczeniami wyrazów „Teajtet” i „lata” konstytuują byt będący zreifikowanym znaczeniem zdania „Teajtet lata”. Zreifikowanego znaczenia tego zdania nie można po prostu utożsamić z uporządkowaną parą znaczeń wyrazów składowych, gdyż nie dałoby się odróżnić znaczenia zdania „Teajtet lata” od pary znaczeń dwóch kolejno wymienionych, lecz niepołączonych w zdanie słów „Teajtet” i „lata”.

Owo połączenie wyrazów w zdanie zapewnia właśnie ich konkatenacja. Konkatenacja jest jednak pojęciem syntaktycznym, które pozwala, zgodnie z regułami składni języka, łączyć wyrażenia składowe w wyrażenie złożone, nie jest natomiast pojęciem semantycznym, które pozwalałoby „łączyć” zreifikowane znaczenia wyrazów składowych w jakiś byt będący zreifikowanym znaczeniem wyrażenia złożonego. Aby rozwiązać ten problem na poziomie semantycznym należy uznać konkatenację za ukryty w zdaniu funktor i przypisać mu również zreifikowane znaczenie. Zdanie przybiera wówczas postać „Konkatenacja Teajtet lata”. Z punktu widzenia syntaktycznego struktura zdania uległa skomplikowaniu i mamy obecnie do czynienia z konkatenacją trzech wyrazów „Konkatenacja” „Teajtet” i „lata”, a zatem

człowieka jako swego wzorca, musimy wprowadzić ideę drugiego stopnia (trzeciego człowieka), która pełni funkcję wzorca dla poszczególnego człowieka oraz idei człowieka. Na tym jednak sprawa się nie kończy, gdyż możemy pytać się o kryterium podobieństwa idei człowieka i idei drugiego stopnia, wprowadzając ideę trzeciego stopnia. Ponieważ rozumowanie to można prowadzić *ad infinitum* mamy do czynienia z regresem w nieskończoność.

²² Bardziej skomplikowane przykłady nie zmieniają istoty tego argumentu.

²³ Dla rozumowania tego nie jest istotne, jakie byty utożsamimy ze znaczeniami. Sam Davidson mówi po prostu o Teajciecie jako o znaczeniu wyrażenia „Teajtet” oraz o własności latania jako znaczeniu wyrażenia „lata”.

konkatenacja nie została wyeliminowana. Wprowadzenie funktora konkatenacji i przypisanie mu zreifikowanego znaczenia nie rozwiązuje problemu, gdyż mamy obecnie zamiast dwóch, trzy zreifikowane znaczenia i nadal nie wiadomo, w jaki sposób konstytuują one znaczenie zdania „Teajtet lata”. Kolejny krok na drodze do usunięcia problemu na płaszczyźnie semantycznej polega na wprowadzeniu kolejnego egzemplarza terminu „konkatenacja” i przypisaniu mu zreifikowanego znaczenia – otrzymamy wówczas zdanie „Konkatenacja konkatenacja Teajtet lata”. Kontynuując to rozumowanie dochodzimy do regresu w nieskończoność. W konsekwencji przypisanie zreifikowanych znaczeń składnikom zdania nie pozwala – bez groźby popadnięcia w regres w nieskończoność – na przypisanie zreifikowanego znaczenia zdaniu.

Argumentacja, którą Davidson przedstawia w niezwykle eliptycznej postaci (1992, s. 3–4), zdaje się pozostawać w sprzeczności z wprowadzonym przez Fregego podziałem wyrażeń językowych na funktory i ich argumenty. W standardowym podejściu funkcję funktora pełni czasownik „lata”, zaś funkcję jego argumentu nazwa „Teajtet”. Funktory określił Frege jako wyrażenia nienasycone i przyporządkował im jako odniesienia funkcje, zaś argumenty funktorów, które same nie są funktorami określił jako wyrażenia nasycone. Wyrażeniami nasyconymi są u Fregego nazwy jednostkowe i zdania, którym jako odniesienia odpowiadają odpowiednio indywidua oraz wartości logiczne. W tym miejscu wykroczę poza zwięzły tekst Davidsona, pokazując, że Argument trzeciego człowieka nie daje się unieważnić przez wprowadzenie rozróżnienia dokonanego przez Fregego. Sugeruje to sam Davidson twierdząc, że opisany wyżej manewr Fregego „raczej wskazuje na trudności niż je rozwiązuje” (s. 4), chociaż dalej nie rozwija tej myśli.

Przyjrzyjmy się zatem temu, jakie trudności niesie ze sobą odniesieniowa semantyka Fregego. Ponieważ czasownik „lata” jest funktorem, zaś nazwa jednostkowa „Teajtet” jest jego argumentem, forma logiczna²⁴ tego zdania przybierze postać „lata(Teajtet)”²⁵. Zgodnie z rozwiązaniem Fregego zreifikowane znaczenie (lub odniesienie) znaczenie czasownika „latać” jest funkcją, która zreifikowanemu znaczeniu (odniesieniu) każdej nazwy własnej przypo-

²⁴ Forma logiczna zdania z języka naturalnego jest wprawdzie pojęciem spornym, ale zawsze wiąże się z tłumaczeniem zdania na jakiś język sformalizowany. Zdanie „Teajtet lata” tłumacząc na język rachunku predykatów, w którym „latać” jest predykatem jednoargumentowym (czyli funktorem), zaś „Teajtet” stałą jednostkową (czyli argumentem tego funktora). Takie wyrażenie formy logicznej jest całkowicie zgodne z poglądami Fregego.

²⁵ Można oczywiście potraktować, jak w gramatykach Montague, nazwę własną „Teajtet” za tzw. frazę nominalną. W takim przypadku będziemy mieli do czynienia z sytuacją odwrotną, „Teajtet” będzie funktorem, zaś „lata” argumentem. Nie zmienia to w niczym przedstawionej tu argumentacji.

rzędkowuje zreifikowane znaczenie (lub odniesienie) odpowiedniego zdania powstałego z tej nazwy oraz czasownika „latać”. Rozwiązanie takie okazuje się jednak nie do przyjęcia w świetle dotychczasowych ustaleń. Widać to wyraźnie na przykładzie Fregowskiej teorii odniesienia i sensu. Można bowiem znać odniesienie zdania „Jaś pali” – niech będzie to Prawda – oraz uchwycić sens tego zdania, nie znając ani funkcji będącej odniesieniem funktora „palić”, ani funkcji będącej sensem tego funktora. Taka semantyka nie byłaby z pewnością semantyką języka wyuczanego, gdyż w języku wyuczonym znaczenie wyrażenia złożonego (zdania) jest nam znane na podstawie znajomości znaczeń wyrażań składowych. Należy zatem odrzucić możliwość, że znaczenie funktora jest funkcją, która przyporządkowuje znaczeniom jego argumentów znaczenia zdań (powstałych z konkatencji tego funktora oraz owych argumentów).

Trudność ta jest widoczna we wspomnianej w pierwszym rozdziale teoriomodelowej semantyce języka naturalnego, która nawiązywała do semantyki Fregego. W semantyce teoriomodelowej znaczenie (odniesienie) zdania należy znać z góry, i do niego właśnie należy dobrać znaczenia (odniesienia) wyrażań prostych, np. funktorów. Aby wyznaczyć odniesienie funktora „pali” należy znać odniesienia (tj. wartości logiczne) wszystkich zdań „Jaś pali”, „Małgosia pali” itd. Wówczas odniesieniem funktora „pali” będzie funkcja, która indywiduom przyporządkowuje wartości logiczne zdań (np. Małgosi przyporządkowuje wartość logiczną zdania „Małgosia pali”).

A zatem wspomniany przez Davidsona „manewr” Fregego polega w istocie na tym, aby tak dobrać odniesienie funktora (funkcję), by ukryć fakt, że zreifikowane znaczenie (lub odniesienie) wyrażenia złożonego języka naturalnego nie może być wyznaczone przez zreifikowane znaczenie (odniesienie) funktora oraz jego argumentu. Problem polega na tym, że trudno jest uznać, że ktokolwiek zna tak określone zreifikowane znaczenie (lub odniesienie) funktora. Oczywiście w ramach konkretnego modelu o skończonej dziedzinie interpretacji można znać zreifikowane znaczenie (odniesienie) funktora²⁶. Jednak w przypadku języka naturalnego nie da się utożsamić znajomości zreifikowanego znaczenia (lub odniesienia) funktora ze znajomością jakiegokolwiek funkcji w modelu o ustalonej dziedzinie.

Można próbować zaradzić opisanej wyżej trudności, pozostając zarazem w ramach semantyki Fregego, traktując zarówno „lata”, jak i „Teajtet” jako argumenty jakiegoś funktora, a zatem wyrażenia nasycone. Na przykład

²⁶ Na przykład dla dziedziny, która składa się z trzech obiektów, odniesieniem funktora (czasownika nieprzechodniego) będzie funkcja, która każdemu z tych obiektów przypisuje wartość logiczną.

występujący w zdaniu „lata(Teajtet)” nawias można potraktować jako ukryty funktor dwuargumentowy – nazwijmy go „przysługuje”. Otrzymamy wówczas zdanie o sztucznej formie „Przysługuje (lata, Teajtet)”, co w języku naturalnym odpowiadałoby mniej więcej zdaniu „Teajtetowi przysługuje latanie”. W takim przypadku należałoby przypisać zreifikowane znaczenie (lub odniesienie) wyrażeniu nienasyconemu „przysługuje”, co zgodnie z Argumentem trzeciego człowieka prowadzi do regresu w nieskończoność. Jest tak, gdyż zreifikowane znaczenie (lub odniesienie) tego funktora jest funkcją, która parom zreifikowanych znaczeń (lub odniesień) jego argumentów przyporządkowuje zreifikowane znaczenia (lub odniesienia) odpowiednich zdań. Wyjaśnia to dlaczego dokonany przez Fregego podział wyrażen na nasycone i nienasycone nie chroni przed regresem w nieskończoność.

Argument trzeciego człowieka, zarówno w wersji oryginalnej, jak i wersji rozszerzonej o krytykę semantyki Fregego, prowadzi w konsekwencji do wniosku, że w semantyce języka wyuczalnego częściom zdań nie powinno się przypisywać znaczeń jako bytów. W przypadku języka wyuczanego jakim jest język naturalny, znajomość znaczeń wyrażen składowych (czyli ich rozumienie) jest zarówno warunkiem koniecznym, jak i – przy założeniu znajomości reguł gramatycznych – warunkiem wystarczającym dla rozumienia wyrażenia złożonego (zdania). Jednak uchwycenie zreifikowanych znaczeń (lub odniesień) wyrażen składowych nie wystarcza – nawet zakładając znajomość reguł gramatycznych – aby uchwycić zreifikowane znaczenie (lub odniesienie) wyrażenia złożonego²⁷. Jak pokazuje Argument trzeciego człowieka zawsze brakuje możliwości uchwycenia zreifikowanego znaczenia (lub odniesienia) jakiegoś składnika obecnego w tym zdaniu, które jednak w owym zdaniu *explicite* nie występuje (np. funktora konkatencji). Oczywiście sama wiedza, że „Teajtet” oznacza Teajteta, a „lata” oznacza zbiór rzeczy latających (lub odpowiednią funkcję charakterystyczną tego zbioru), nie wystarcza do rozstrzygnięcia, czy zdanie „Teajtet lata” jest prawdziwe czy fałszywe (jeżeli uznamy wartość logiczną za odniesienie zdania)²⁸. Aby to stwierdzić należy bowiem utożsamić obiekt oznaczany przez słowo „Teajtet” z odpowiednim obiektem znajdującym się w zbiorze rzeczy latających, tj. w zakresie predykatu „lata”, co w ogólnym przypadku nie jest możliwe.

Znajomość zreifikowanych znaczeń (lub odniesień) słów nie stanowi również warunku koniecznego znajomości – czyli uchwycenia – zreifikowanego

²⁷ Stosuję tutaj terminologię Fregego, który pisał o chwytności sensów, czyli zreifikowanych znaczeń słów.

²⁸ W przypadku zreifikowanego znaczenia zdania (np. Fregeowskiego sensu) można argumentować w ten sam sposób.

znaczenia (lub odniesienia) zdania, bo nawet jeśli uznamy, że zrozumienie zdania polega na uchwyceniu jego zreifikowanego znaczenia, to nie pociąga to za sobą konieczności przypisania zreifikowanych znaczeń wszystkim występującym w nim słowom lub innym elementom składni. Podobnie jest w przypadku odniesień. Wiadomo bowiem, że zdanie „Ojcem Wolfganga Amadeusza Mozarta był Leopold Mozart” jest prawdziwe (załóżmy, że odniesieniem tego zdania jest Prawda), mimo że nikt z nas nie zna funkcji, która jest odniesieniem funktora „ojciec x -a”.

Zdaniem Davidsona semantyczna definicja prawdy nie jest narażona na Argument trzeciego człowieka, gdyż wprawdzie używa pojęcia odniesienia, ale – w przeciwieństwie do semantyki Fregego oraz opartej nań semantyki teoriomodelowej – nie wskazuje na żadne konkretne byty jako odniesienia. Zagadnieniu temu poświęcony jest kolejny paragraf.

3. TEORIA ODNIESIENIA BEZ ODNIESIEŃ

Rezygnacja z semantyki odniesieniowej nie oznacza, że w semantyce nie można posługiwać się pojęciem odniesienia. Odniesienie staje się jednak w takim przypadku jedynie bytem postulowanym przez teorię semantyczną. W takim przypadku znajomość zreifikowanych znaczeń lub odniesień słów – o czym była mowa na końcu poprzedniego paragrafu – nie jest ani warunkiem koniecznym, ani warunkiem wystarczającym znajomości zreifikowanych znaczeń ani odniesień zdań²⁹.

Argumentując na rzecz uznania semantycznej definicji prawdy za teorię znaczenia, Davidson odwołuje się do przytoczonego w zmienionej postaci w rozdziale 1 przykładu funktora „ojciec”, który tworzy nazwę jednostkową, poprzedzając (dowolną ilość razy) inną nazwę jednostkową, np. „ojciec ojca ojca x -a”. Podając ten przykład Davidson pokazuje w jaki sposób można używając pojęcia odniesienia jako bytu postulowanego podać znaczenie funktora „ojciec”, nie czyniąc żadnych założeń na temat istnienia zreifikowanych odniesień i znaczeń.

Niech naszą nazwą jednostkową będzie –jak w rozdziale 1 – „Wolfgang Amadeusz Mozart”. Teoria, którą szkicuje Davidson w *Prawdzie i znaczeniu*, można uznać za namiastkę teorii terminu „odniesienie”. W teorii tej występuje bowiem reguła, która przyporządkowuje nazwie własnej jej odniesienie, np. nazwie własnej „Wolfgang Amadeusz Mozart” przyporządkowuje Wolfganga Amadeusza Mozarta, zaś wyrażeniu jednostkowemu „ojciec x -a”

²⁹ W tym momencie nie rozstrzygamy jeszcze jakim bytem postulowanym przez teorię będzie odniesienie zdania.

przyporządkowuje ojca osoby, do której odnosi się termin x , np. wyrażeniu „ojciec Wolfganga Amadeusza Mozarta”, przyporządkowuje ojca Wolfganga Amadeusza Mozarta. Teoria ta nie odwołuje się jednak do żadnego zreifikowanego znaczenia funktora „ojciec”, a nawet – wbrew pozorom – nie odwołuje się do odniesienia tego funktora, lecz posługuje się nim w sposób instrumentalny. W przeciwnym przypadku odniesieniem funktora „ojciec” byłyby bowiem funkcja, która przyporządkowywałaby indywiduom inne indywidua, które są ich ojcami. Gdyby teoria ta korzystała w ten sposób z odniesień funktora „ojciec”, to musiałaby przyporządkowywać konkretne indywidua tym indywiduom, które są ojcami indywiduów będących odniesieniami argumentów funktora „ojciec”. Oczywiście nie jesteśmy w stanie podać tak rozumianego odniesienia wyrażenia „ojciec”, gdyż wymagałoby to opanowania niedostępnej nam wiedzy.

Nie oznacza to jednak, że nieznajomość odniesienia wyrażenia „ojciec” uniemożliwia nam rozumienie znaczenia tego wyrażenia – oznacza to jedynie, że nie potrafimy w każdym przypadku identyfikować ojców znanych nam indywiduów. Problem, na który wskazuję, polega na tym, że semantyka odniesieniowa wymagałaby takiej wiedzy, chociaż ojcowie bywają niekiedy nieznani lub możemy – co niekiedy się zdarza – mylić się na temat tego, kto faktycznie jest czymś ojcem³⁰. Okazuje się jednak, że fakt istnienia odniesienia (nawet nieznanego) nie jest warunkiem koniecznym sensowności wyrażenia, czyli posiadania przez owo wyrażenie znaczenia. Można bowiem przedstawić teorię (jak w rozdziale 1), w której termin „ojciec” jest zastąpiony terminem „syn”. Teoria taka również podaje znaczenie terminu „syn x -a”, chociaż nie każdy ma lub miał syna. Nawet jeśli x jest bezdzietny, wyrażenie „syn x -a” posiada znaczenie i poprawna teoria znaczenia powinna owo znaczenie w pewien sposób podać.

Jak pokazuje Davidson zarówno identyfikacja odniesień, jak i ich istnienie, nie jest nam potrzebne do rozumienia języka³¹. Teoria deskrypcji „ojciec x -a”, którą przedstawił Davidson, nie wyróżnia żadnego indywiduum, czy innego obiektu, jako ojca Wolfganga Amadeusza Mozarta. Nie wskazuje

³⁰ Chociaż każdy niewątpliwie ma lub miał ojca, czyli jeżeli x ma odniesienie, to „ojciec x -a” również musi mieć odniesienie. Jest to jednak fakt przygodny, o czym świadczą przypadki dzieworódtwa wśród niektórych organizmów. W takim przypadku doskonale rozumiemy, co to znaczy być ojcem jakiegoś indywiduum, chociaż wiemy on nim, że ojca nie posiada.

³¹ W tym miejscu należy poczynić uwagę następującą: zdaniem Davidsona warunkiem znajomości języka jest istnienie zewnętrznego wobec nas świata, w którym znajdują się odniesienia bardzo wielu nazw. Nie ma natomiast powodu utrzymywać, że znajomość języka zależy od istnienia określonych nazw. Do sprawy tej powrócę w kolejnym rozdziale omawiając warunki interpretacji.

zatem, do czego się wyrażenie „ojciec Wolfganga Amadeusza Mozarta” odnosi, ale mimo to w swych regułach (aksjomatach) używa pojęcia odniesienia (por. Lepore i Ludwig 2005, s. 47). Teoria ta pozwala jednak – zdaniem Davidsona – częściowo zrozumieć wyrażenie „ojciec x -a”, gdyż w skład tej teorii wchodzi rekurencyjne reguły postaci³²:

Wyrażenie „ x ” odnosi się do x -a (odniesieniem wyrażenia „ x ” jest x).

Wyrażenie „ojciec x -a” odnosi się do ojca odniesienia x -a.

W tym momencie ktoś może zaprotestować, gdyż teoria taka może nie tylko wydawać się całkowicie trywialna, ale wręcz cyrkularna. Wszak zadaniem tej teorii jest wyjaśnienie znaczenia terminu „ojciec”, co teoria czyni, używając tego samego terminu. W podanym przypadku tak właśnie jest, gdyż metajęzyk, czyli język, w którym sformułowana jest teoria, obejmuje fragment języka przedmiotowego, dla którego została zbudowana ta teoria. Teorię tę można jednak budować w innym języku niż język polski, chociaż nadal będzie to teoria dla polskiego funktora „ojciec x -a”:

The expression “ x ” refers to x (the referent of the expression “ x ” is x).

The expression “ojciec x -a” refers to the father of the referent of “ x ”.

Z teorii tej wynikają wszystkie zdania³³ postaci *t refers to y* , gdzie *t* zastępuje jest opisem strukturalnym terminu jednostkowego, a *y* przez sam ten termin, np.

The expression “Wolfgang Amadeusz Mozart” refers to Wolfgang Amadeusz Mozart.

The expression “ojciec Wolfganga Amadeusza Mozarta” refers to the father of Wolfgang Amadeus Mozart.

The expression “ojciec ojca Wolfganga Amadeusza Mozarta” refers to the father of the father of Wolfgang Amadeus Mozart.

W przypadku podanego wyżej wariantu namiastki teorii odniesienia, rozumienie wyrażenia „ojciec”, czyli przypisanie znaczenia temu wyrażeniu, zapewnione jest dzięki regule (aksjomatowi), która używa przekładu wyrażenia „ojciec” na metajęzyk, czyli na wyrażenie *the father of*. Jednak nawet w przypadku, gdy metajęzyk obejmuje język przedmiotowy, teoria nie musi być trywialna. W analogicznej regule może bowiem zamiast samego

³² Faktycznie tak uboga teoria nie może pełnić funkcji teorii znaczenia. Stanowi ona jednak ważny element argumentacji Davidsona.

³³ Można je określić jako R-zdania.

wyrażenia „ojciec” używać jakiegoś synonimu tego wyrażenia, np. „najbliższy przodek biologiczny płci męskiej”.

Przytoczony tu za *Prawdą i znaczeniem* przykład stanowi namiastkę teorii znaczenia proponowanej przez Davidsona i została przez niego przedstawiona jedynie w celu wstępnego zarysowania stanowiska w kwestii podawania znaczenia przez teorię odniesienia, która – paradoksalnie – nie jest teorią odniesieniową. W szczególności nie należy tego rozumieć w ten sposób, że teoria deskrypcji „ojciec x -a” stanowi uproszczoną wersję teorii znaczenia, ograniczoną do części języka składającego się z terminów jednostkowych i deskrypcji określonej „ojciec”. Takie rozwiązanie jest w świetle ustaleń *Theories of Meaning* nie do zaakceptowania, gdyż – jak pamiętamy – Davidson stał na gruncie holizmu znaczeniowego i uważał, że teoria znaczenia musi być teorią całego języka, rozumianego jako „organiczna całość”. Aby teoria nazwy jednostkowej (deskrypcji) „ojciec x -a” faktycznie podawała znaczenie wyrażenia „ojciec”, powinna stanowić fragment teorii odniesienia dla całego języka³⁴.

W świetle powyższych ustaleń jest oczywiste, że teoria ta nie ma za zadanie ustalenie kto faktycznie był ojcem x -a, np. kto był ojcem Wolfganga Amadeusza Mozarta, a kto ojcem ojca Wolfganga Amadeusza Mozarta itd., lecz wskazanie efektywnej procedury ustalania odniesień, która to procedura faktycznie służy do podawania znaczeń: „(...) teoria ściśle określa efektywną procedurę ustalania – dla każdego terminu jednostkowego w ramach jego uniwersum – do czego ów termin się odnosi” (s. 5). Procedura ta polega na tym, że odniesieniem terminu „ojciec x -a” jest ojciec odniesienia terminu x , tj. ojciec x -a, odniesieniem terminu „ojciec ojca x -a” jest ojciec odniesienia terminu „ojciec x -a” itd. To, że teoria określa w ten właśnie sposób procedurę ustalania odniesień terminów jednostkowych, nie oznacza, że sama teoria wskazuje byty, które miałyby być ich odniesieniami³⁵.

Kolejnym krokiem w stronę zbudowania nieodniesieniowej teorii używającej odniesień, która pełniłaby jednocześnie funkcję teorii znaczenia, była przedstawiona przez Davidson namiastka teorii dla predykatów jednoargumentowych. W poprzedniej „teorii” znaczenie funktora „ojciec” jest podane (częściowo) dzięki użyciu odniesień nazw jednostkowych: x , „ojciec x -a” itd. Predykaty jednoargumentowe są również funktorami, lecz funktorami zdaniotwórczymi od jednego argumentu nazwowego. Budując teorię dla tych

³⁴ W języku angielskim *the father of x* . Dlatego też wyrażenie „ojciec x -a” traktuję nie jako predykat, lecz jako deskrypcję określoną.

³⁵ Mówiąc nieco żartobliwie, lecz zgodnie z intencjami Davidsona, taka teoria nie pozwala na ustalanie ojcostwa, lecz jedynie (częściowe) wyjaśnienie co słowo „ojciec” znaczy. Aby zbudować teorię odniesieniową, należałoby móc ustalić, kto faktycznie był czyim ojcem.

predykatów Davidson odwołuje się do wskazanej przez Fregego analogii między nazwami jednostkowymi a zdaniami. Frege traktował bowiem zdania oznajmujące jako szczególne nazwy jednostkowe, których odniesieniami w kontekstach ekstensjonalnych mogą być dwa abstrakcyjne obiekty: Prawda i Fałsz, zaś w mowie zależnej – sensy lub myśli. Rozważania Davidsona były na tym etapie bardzo ogólne, w szczególności nie precyzuje on na tym etapie, co miałyby być odniesieniem zdania (równie dobrze mógł się odwołać do tradycji fenomenologicznej, w której odniesieniami zdań były tzw. stany rzeczy). Jednak bez względu na to, na jakie odniesienia się zdecydujemy, zdania oznajmujące można potraktować jako jednostkowe nazwy czegoś, i to coś, jako odniesienie zdania, będzie używane w tej teorii, podobnie jak w teorii funktora „ojciec” były używane odniesienia terminów jednostkowych.

Niech przykładem takiej teorii będzie teoria predykatu jednoargumentowego „jest kompozytorem” (nadbudowana nad teorią funktora „ojciec”), który z nazw jednostkowych może tworzyć zdania, np. „Wolfgang Amadeusz Mozart jest kompozytorem”, „Ojciec Wolfganga Amadeusza Mozarta jest kompozytorem” itd. Aksjomatami tej teorii byłyby m.in. zdania:

Dla dowolnej nazwy jednostkowej „x”, „x” odnosi się do x-a (odniesieniem nazwy jednostkowej „x” jest x).

Dla dowolnej nazwy jednostkowej „x”, zdanie „x jest kompozytorem” odnosi się do (odniesienie „x” jest kompozytorem).

Z teorii tej wynika między innymi sztucznie brzmiące zdanie³⁶:

Zdanie „Wolfgang Amadeusz Mozart jest kompozytorem” odnosi się do (odniesienie terminu jednostkowego „Wolfgang Amadeusz Mozart” jest kompozytorem).

Uwzględniając zapis funktor-argument można to zdanie przeformułować do następującej postaci:

Zdanie „Wolfgang Amadeusz Mozart jest kompozytorem” odnosi się do (jest-kompozytorem odniesienie terminu jednostkowego „Wolfgang Amadeusz Mozart”),

czyli:

(3.1) Zdanie „Wolfgang Amadeusz Mozart jest kompozytorem” odnosi się do (Wolfgang Amadeusz Mozart jest kompozytorem).

³⁶ Nie jest to ani S-zdanie, ani T-zdanie, lecz raczej coś pośredniego między nimi.

Niewątpliwa sztuczność przytoczonych przykładów wynika z tego, że w sformułowaniu teorii odwołujemy się do wspomnianej wyżej analogii między nazwami jednostkowymi a zdaniami, czyli używamy odniesień zdań (czykolwiek by one nie były), podobnie jak w teorii funktora „ojciec” używaliśmy odniesień nazw jednostkowych. Przykład ten może sugerować jednak, że zdania mogą się odnosić do pewnych bytów (sytuacji, faktów, stanów rzeczy). Gdyby tak było, semantyczna definicja prawdy nie mogłaby pełnić funkcji teorii znaczenia, a funkcję taką mogłaby pełnić jedynie jakaś teoria sytuacji. Przeciwno takiej możliwości skierowany jest kolejny, niezwykle kontrowersyjny argument Davidsona.

4. ARGUMENT PROCY

W przypadku nazw jednostkowych nie mamy większego problemu z odpowiedzią na pytanie, czym są (czy też raczej, czym mogą być) ich odniesienia. Odniesieniami tymi są (mogą być) obiekty lub indywidua. Podobnie jest w przypadku nazw ogólnych (predykatów), dla których również nie musimy troszczyć się o odniesienia, gdyż dysponując teorią nieodniesieniową, odniesienia predykatów są bytami jedynie postulowanymi, wykorzystywanymi w teorii jedynie instrumentalnie. Wprawdzie można by w nieco sztuczny i niezgodny z intencjami Tarskiego sposób sformułować aksjomaty definicji prawdy w taki sposób, aby wynikało z niej następujące T-zdanie:

Zdanie „Wolfgang Amadeusz Mozart jest kompozytorem” jest prawdziwe wtw, gdy odniesienie terminu „Wolfgang Amadeusz Mozart” jest elementem zbioru będącego odniesieniem predykatu „jest kompozytorem”,

ale zdanie po prawej stronie jest równoważne temu, że:

Odniesienie terminu „Wolfgang Amadeusz Mozart” należy do elementu zbioru kompozytorów,

czyli po prostu:

Wolfgang Amadeusz Mozart jest kompozytorem.

W semantycznej teorii języka naturalnego (semantycznej definicji prawdy) wystarczy postulować istnienie indywiduów (elementów ciągów), zaś postulowanie istnienia zbiorów, funkcji itp. okazuje się być całkowicie zbędne, gdyż nazwy odniesień mogą być w metajęzyku eliminowane (funkcje,

zbiory)³⁷. Aby jednak semantyczna definicja prawdy mogła być teorią języka naturalnego, należy rozstrzygnąć problem postulowanych odniesień zdań. Czym może być zatem odniesienie zdania?

W *Über Sinn und Bedeutung* Frege argumentuje na rzecz tezy, że odniesieniem zdania oznajmującego jest jego wartość logiczna (por. Frege 1975, s. 70). Ponieważ argumentacja Fregego nie jest według niektórych komentatorów konkluzywna, próbowano uzasadnić tezę Fregego w inny sposób. Najbardziej znany jest tzw. argument Churcha-Gödla, opisany i skrytykowany w artykule Barwise'a i Perry'ego *Semantic Innocence and Uncompromising Situations* (1985). Od autorów pochodzi również popularna nazwa „Argument procy” (*Slingshot Argument*)³⁸.

Argument procy został wykorzystany przez Davidsona wielokrotnie i w kilku wersjach – po raz pierwszy w *Prawdzie i znaczeniu*. Wprawdzie argument ten został oryginalnie wykorzystany przez Churcha na rzecz tezy Fregego, że (zwykłym) odniesieniem zdania musi być jego wartość logiczna, Davidson wykorzystywał pewne jego wersje przeciw poglądom Fregego. Chodziło tu przede wszystkim o dokonaną przez Fregego reifikację sensów zdań, które mogą pełnić w pewnych kontekstach funkcję „specjalnych” odniesień. Pamiętając, że według Fregego zdania mogą być traktowane jako szczególny przypadek terminów jednostkowych, Davidson zauważa, iż:

Teraz jednak, o ile w dalszym ciągu będziemy identyfikować (*implicite*) znaczenie terminu jednostkowego z jego odniesieniem, wkradnie się do naszej teorii pewna trudność (s. 6).

³⁷ Davidson był nominalistą i stosował wprowadzony przez Ockhama postulat ograniczania przyjętych bytów do niezbędnego minimum. Poza postulatem istnienia indywiduów Davidson postulował również istnienie zdarzeń (por. 1967a). Z kolei Quine postulował poza indywiduami istnienie klas jako bytów nieredukowalnych w semantyce teorii matematycznych. Oczywiście postulowana przez Davidsona ontologia indywiduów oraz zdarzeń może być uzupełniona przez inne byty, gdyby odpowiednie fragmenty języka naturalnego zostały sprowadzone do postaci, która pozwoliłaby zastosować do nich teorię prawdy. Wówczas byty te musiałyby móc występować jako wyrazy ciągów.

³⁸ Argument ten został określony tym mianem (niekiedy bywa określanym jako „argument z procy”), gdyż jak piszą autorzy „(...) zajmuje tak niewiele miejsca, rzadko ponad pół strony, i używa tak niewielkiej ilości amunicji – teorii deskrypcji i popularnego pojęcia równoważności logicznej – że nazwiemy go *procq*” (Barwise i Perry 1981, s. 398). Za nazwą kryje się oczywiście aluzja do broni, którą biblijny Dawid pokonał olbrzyma Goliata, podobnie jak Davidson za pomocą swego argumentu „pokonał” Fregego. Faktycznie argument procy może być wykorzystany zarówno na rzecz tezy Fregego, że odniesieniami zdań mogą być jedynie wartości logiczne, jak i przeciw tezie Fregego, że specjalnymi odniesieniami zdań mogą być ich sensy, czyli myśli. Prostota tego argumentu wzbudza – zdaniem Barwise'a i Perry'ego – tyle wątpliwości, że przyrównali go do rysunków Eschera.

Ową trudność stanowi konkluzja Argumentu procy:

Widać stąd, że dwa zdania o tej samej wartości logicznej mają to samo odniesienie. Jeśli jednak znaczenie zdania jest jego odniesieniem, to wszystkie zdania o tej samej wartości logicznej są synonimiczne, a taki wniosek trudno zaakceptować (*ibidem*).

Poniżej przedstawiam w wersji rozwiniętej Argument procy w oryginalnym sformułowaniu z *Prawdy i znaczenia*. Czynię to głównie dlatego, że oryginalne sformułowanie ma charakter niezwykle eliptyczny.

4.1 Zdanie jest złożonym terminem jednostkowym.

4.2. Logicznie równoważne terminy jednostkowe mają to samo odniesienie.

4.3. Termin jednostkowy nie zmienia swego odniesienia, gdy jakiś jego termin zastąpimy innym terminem, o tym samym odniesieniu.

4.4. R i S są zdaniami o tej samej wartości logicznej.

Davidson pokazuje, że następujące cztery zdania, traktowane jako złożone terminy jednostkowe, mają te same odniesienia:

(1) R

(2) $\{x: (x = x) \wedge R\} = \{x: x = x\}$.

(3) $\{x: (x = x) \wedge S\} = \{x: x = x\}$.

(4) S

Rozwijając przedstawioną argumentację otrzymamy:

1. R jest logicznie równoważne zdaniu $\{x: (x = x) \wedge R\} = \{x: x = x\}$.

Nazwę jednostkową $\{x: x = x\}$ rozumiemy jako zbiór tych obiektów, które są identyczne same ze sobą. Jest to zatem zbiór wszystkich obiektów. Jeżeli R jest prawdziwe, to obiekt opisany za pomocą nazwy „zbiór takich obiektów x , że $x = x$ ” jest identyczny z obiektem opisanym za pomocą nazwy „zbiór takich obiektów x , że $x = x$ i R ”, co oznacza, że $\{x: (x = x) \wedge R\} = \{x: x = x\}$ jest zdaniem prawdziwym. Jeżeli R jest fałszywe, to $\{x: (x = x) \wedge R\}$ jest zbiorem pustym, a zatem $\{x: (x = x) \wedge R\} = \{x: x = x\}$ jest zdaniem fałszywym, co dowodzi (1).

2. Zdania R oraz $\{x: (x = x) \wedge R\} = \{x: x = x\}$ posiadają te same odniesienia.

Zgodnie z (4.1) zdania posiadają odniesienia. Skoro oba zdania są równoważne logicznie (1), to zgodnie z (4.2) mają te same odniesienia.

3. Deskrypcje $\{x: (x = x) \wedge R\}$ i $\{x: (x = x) \wedge Z\}$ posiadają te same odniesienia.

Jeżeli założymy prawdziwość R i Z (4.4), to odniesieniami obu deskrypcji jest zbiór pełny. W przypadku fałszywości tych zdań – zbiór pusty.

4. Zdania $\{x: (x = x) \wedge R\} = \{x: x = x\}$ i $\{x: (x = x) \wedge Z\} = \{x: x = x\}$ posiadają te same odniesienia.

Wykorzystujemy zasadę zastępowania ekstensjonalnego (4.3) oraz (3), gdyż drugie zdanie powstaje z pierwszego przez zastąpienie nazwy jednostkowej $\{x: (x = x) \wedge R\}$ przez nazwę jednostkową o tym samym odniesieniu czyli $\{x: (x = x) \wedge Z\}$.

5. Z jest logicznie równoważne zdaniu $\{x: (x = x) \wedge Z\} = \{x: x = x\}$.

Powtarzamy rozumowanie z kroku (1) dowodu.

6. Z i $\{x: (x = x) \wedge Z\} = \{x: x = x\}$ posiadają te same odniesienia.

Powtarzamy rozumowanie z punktu (2).

7. R i Z posiadają te same odniesienia.

Wykorzystujemy wiersz (4).

Konkluzja z argumentu procy jest następująca: dowolne zdania o tej samej wartości logicznej muszą mieć te same odniesienia, a zatem istnieje wzajemnie jednoznaczna zależność między odniesieniami zdań a ich wartościami logicznymi. Odniesieniami zdań nie mogą być zatem ani sytuacje, ani stany rzeczy, ani fakty. Chyba że – jak twierdził Davidson – zdanie prawdziwe oznacza po prostu jeden Wielki Fakt. Utożsamienie prawdy z Wielkim Faktem prowadzi wprost do pojęcia prawdy w sensie ontologicznym. Prawda to po prostu Rzeczywistość, zaś pojedyncze zdania prawdziwe są po prostu jej różnymi nazwami. Rozumienie prawdy jako Wielkiego Faktu jest również zgodne z duchem poglądów Fregego³⁹.

Na temat ewentualnej błędności Argumentu procy istnieje bogata literatura. Dlatego też w tym miejscu ograniczę się jedynie do dwóch uwag. Pierwsza dotyczy komentarza, którym Davidson opatrzył Argument procy w *Prawdzie i znaczeniu*: „Warto chyba zauważyć, że powyższe uzasadnienie nie zależy od żadnej szczególnej identyfikacji bytów, którym zdania te miałyby odpowiadać” (s. 6, przyp.) Okazuje się jednak, że stwierdzenie to wymaga pewnej ostrożności.

³⁹ W podobny sposób prawdę u Fregego interpretuje Jan Łukasiewicz (por. Frege 1975, s. 70–71, przypis tłumacza).

Jako potencjalną kandydaturę na byt, który mógłby być zreifikowanym znaczeniem (czyli specjalnym odniesieniem) zdania, proponuję rozważyć intensję wprowadzoną przez Carnapa w *Meaning and Necessity* jako odpowiednik przedteoretycznego terminu „znaczenie”. Przyjrzyjmy się zatem, czy korzystając ze zmodyfikowanej wersji Argumentu procy da się wykazać, że zdania R i S o tej samej wartości logicznej mają te same intensje.

W takim przypadku na samym wstępie należy zmodyfikować dwa założenia Argumentu procy:

4.2'. Logicznie równoważne terminy jednostkowe mają te same intensje⁴⁰.

4.3'. Termin jednostkowy nie zmienia swej intensji, gdy jakiś jego termin składowy zastąpimy innym terminem, o tej samej intensji⁴¹.

Problem pojawia się, gdy chcemy skorzystać z założenia o współprawdziwości zdań R i S (odpowiednik 4.4). Okazuje się bowiem, że podając intensję zdania nie można już po prostu mówić o prawdziwości lub fałszywości zdań. Intensje zdań rozumie się bowiem jako funkcję z tzw. opisów stanu (Carnap) lub możliwych światów (Kripke oraz Montague) w wartości logiczne. W obu przypadkach prawdziwość lub fałszywość zdania (w możliwym świecie lub dla opisu stanu) zależy od jego intensji. Oczywiście jeśli założymy, że R i S są prawdziwe w tych samych światach lub dla tych samych opisów stanu, to tym samym już w punkcie (4.4) zakładamy, że mają te same intensje i Argument procy jest zbędny. Załóżmy zatem, że:

4.4'. R i S są zdaniami o tej samej wartości logicznej dla określonego (np. rzeczywistego) opisu stanu (lub w określonym możliwym świecie).

Dwa pierwsze wiersze dowodu nie budzą zastrzeżeń:

1'. R jest logicznie równoważne zdaniu $\{x: (x = x) \wedge R\} = \{x: x = x\}$.

Analogicznie do (1) w poprzednim dowodzie.

2'. Zdania R oraz $\{x: (x = x) \wedge R\} = \{x: x = x\}$ posiadają te same intensje.

Zgodnie z (4.3'). Problem pojawia się jednak w wierszu trzecim:

3'. Deskrypcje $\{x: (x = x) \wedge R\}$ i $\{x: (x = x) \wedge Z\}$ posiadają te same odniesienia.

⁴⁰ Zauważmy, że w tym przypadku zachodzi również zależność odwrotna. Terminy jednostkowe o tych samych intensjach są logicznie równoważne.

⁴¹ Ponieważ synonimy posiadają te same intensje, to ich wzajemna zamiana w jakimś wyrażeniu nie zmieni jego intensji. Ponieważ w języku polskim synonimami są „ziemniak” i „kartofel”, to wyrażenia „Jaś rzucił ziemniakiem” oraz „Jaś rzucił kartoflem” mają te same intensje.

Jeśli odniesieniami tych deskrypcji są zbiory obiektów dla wybranego wyżej opisu stanu (lub świata możliwego), to kolejny wiersz dowodu jest ewidentnie fałszywy:

4'. Zdania $\{x: (x = x) \wedge R\} = \{x: x = x\}$ i $\{x: (x = x) \wedge Z\} = \{x: x = x\}$ posiadają te same intensje.

Gdybyśmy uznali, że odniesieniami deskrypcji są intensje rozumiane jako funkcje z możliwych światów lub opisów stanu w zbiory, to fałszywy byłby już wiersz 3:

3". Deskrypcje $\{x: (x = x) \wedge R\}$ i $\{x: (x = x) \wedge Z\}$ mają te same odniesienia (tj. intensje).

A zatem Argument procy spełnia swe zadanie niezależnie od wyboru bytów na zreifikowane znaczenia zdań pod warunkiem, że byty te nie zależą od wartości logicznych. Intensje zdań są natomiast funkcjami, których wartościami są prawda lub fałsz.

Druga uwaga ma charakter ogólniejszy. Dotyczy ona tzw. niekonserwatywności rozszerzenia języka. Przez rozszerzenie niekonserwatywne rozumiemy wzbogacenie języka L – w naszym przypadku języka, w którym występują zdania R i S – o takie wyrażenie – w naszym przypadku o operator deskrypcji $\{x: (...)\}$ – które pozwala na udowodnienie czegoś na temat zdań języka L nie zawierających owego wyrażenia (w naszym przypadku udowodnienie, że zdania R i S mają te same odniesienia), przy czym nie dałoby się zbudować dowodu bez owego rozszerzenia. W przypadku prostych i niesformalizowanych dowodów – jak ma to miejsce u Davidsona – rozszerzenie niekonserwatywne pozwala często na dowiedzenie wszystkiego, o ile nie narzuci się na dowód określonych ograniczeń, co jest możliwe jedynie w przypadku przyjęcia pewnej formalnej teorii dowodu. Na temat niekonserwatywności rozszerzeń istnieje bardzo bogata literatura zainicjowana polemiką między Priorem (1960) a Belnapem (1962). Polemika ta wykracza jednak zdecydowanie poza cele tego opracowania, gdyż w tym miejscu nie interesują nas formalne teorie dowodu, a jedynie dowody przeprowadzane w języku naturalnym, ewentualnie wzbogaconym o pewne wyrażenia pochodzące z języków sformalizowanych⁴².

Argument procy posiada zarówno żarliwych obrońców, jak i krytyków. Sam Davidson używa go w sposób niesformalizowany i intuicyjny, zaś liczne kontrowersje wokół tego argumentu dotyczą przede wszystkim nieprecyzyjnej

⁴² Pewne uwagi na ten temat można znaleźć w mojej monografii *Znaki logiczne. Granice logiki oraz logiczność teorii i relacji wynikania* (2003, s. 94–97).

wersji argumentu⁴³. Argument procy wykorzystywany jest również w logice formalnej, najczęściej przeciw tzw. semantynom sytuacyjnym⁴⁴. Sformalizowane wersje tego argumentu budowane są w bardzo precyzyjnych językach logiki dla ściśle zdefiniowanych pojęć. Celem Davidsona było jednak formułowanie tego argumentu dla języka naturalnego i dla intuicyjnych pojęć, które stosujemy do języka naturalnego. Dlatego też nie jest możliwe zajęcie wobec niego jednoznacznego stanowiska.

5. W STRONĘ ALFREDA TARSKIEGO SEMANTYCZNEJ DEFINICJI PRAWDY

Jak twierdzi Davidson odrzucenie zreifikowanych znaczeń nie oznacza konieczności rezygnacji z posługiwania się terminami „znaczenie” oraz „teoria znaczenia”. Oznacza to jedynie, że nie będziemy przypisywać wyrażeniom języka jakichkolwiek bytów zwanych znaczeniami, co prowadzi do pozornie paradoksalnego określenia jego teorii jako „teorii znaczenia bez znaczeń”, którego to określenia Davidson użył w *Prawdzie i znaczeniu*. Pomimo oksymoronicznej nazwy teoria taka doskonale spełnia swą funkcję – wystarczy, że przez znaczenia słów będziemy rozumieć systematyczny wpływ, jaki te słowa wywierają na znaczenia zdań.

Wypada zatem zmodyfikować warunki, których spełnienia oczekujemy od zadowalającej teorii znaczenia w taki sposób, by nie sugerować już, że pojedyncze słowa mają jakiegokolwiek znaczenia rozumiane szerzej niż wywieranie systematycznego wpływu na znaczenia zdań, w których te słowa się pojawiają (s. 5).

Powyższy cytat można rozumieć w sposób następujący – teoria znaczenia powinna wyjaśnić, w jaki sposób rozumienie słów wpływa na rozumienie zdań, przy czym przez rozumienia nie należy utożsamiać z „chwytaniem” znaczeń jako bytów. Sposobem na to, aby wyrazić ów wpływ, jest – zdaniem Davidsona – podanie dla danego języka warunków prawdziwości dla wszystkich jego zdań. Owymi warunkami prawdziwości są oczywiście wynikające z Tarskiego teorii prawdy T-zdania. Rozumienie zdania danego języka można zatem utożsamiać ze znajomością warunków prawdziwości wszystkich zdań tego języka lub – równoważnie – ze znajomością teorii prawdy, z której (zgodnie z konwencją T) wynikają wszystkie T-zdania.

⁴³ Przegląd podstawowych zarzutów przeciwko argumentowi procy można znaleźć w (Lepore i Ludwig 2005).

⁴⁴ Przykład formalnej wersji Argumentu procy można znaleźć w (Shramko i Wansing 2009).

W świetle dotychczasowych ustaleń Tarskiego teorię prawdy dla języka naturalnego można określić jako skończoną (w sensie ilości aksjomatów) i rekurencyjną teorię, w której używa się pojęcia odniesienia. Teoria taka spełnia wówczas wymóg teorii języka wyuczanego. Jak pamiętamy, teoria prawdy posługuje się wprowadzając pojęciem odniesienia – a dokładniej pojęciem spełniania, które jest rezultatem eksplikacji pojęcia odniesienia – ale nie wyróżnia konkretnych odniesień. Jeżeli za odniesienie zdania oznajmującego uznamy jego wartość logiczną, teoria Tarskiego podaje w jaki sposób odniesienia zdania zależą od odniesień ich składników. Wystarczy to, aby – wedle Davidsona – podać znaczenie zdania, gdyż rozumieć zdanie, to wiedzieć nie tylko kiedy jest ono prawdziwe a kiedy fałszywe, ale kiedy są prawdziwe pozostałe zdania języka. W konsekwencji T-zdanie dla określonego zdania danego języka podaje jego warunek prawdziwości o ile stoi za nim cała teoria prawdy. W rozdziale 2 „O znaczeniu zdań” *Zarysu logiki* John S. Mill również podaje znaczenia zdań w postaci ich warunków prawdziwości. Różnica między obydwojema podejściami jest jednak zasadnicza. O ile dla Milla znaczenie zdania „Jan jest człowiekiem” jest podane przez jego warunek prawdziwości, np. wyrażony zdaniem „Jan należy do zbioru ludzi”, to dla Davidsona pojedyncze T-zdanie wprowadzając podaje warunek prawdziwości zdania po jego lewej stronie, ale nie podaje jego znaczenia. Znaczenie to podawane jest bowiem przez warunki prawdziwości (T-zdania) dla wszystkich zdań rozpatrywanego języka.

W przypadku utożsamienia znaczenia zdania z jego warunkami prawdziwości bardzo łatwo jest o nieporozumienie. Aby się o tym przekonać, przeanalizujmy przykład najbardziej znanego T-zdania podanego przez samego Alfreda Tarskiego:

Zdanie „*Śnieg jest biały*” jest prawdziwe wtw, gdy śnieg jest biały.

Wadą przytoczonego T-zdania jest przede wszystkim to, że nie istnieje prosta teoria prawdy, z której miałyby ono wynikać⁴⁵, niemniej jednak może nam ono posłużyć – podobnie jak Tarskiemu – jako przykład warunku prawdziwości. Załóżmy zatem, że istnieje odpowiednia teoria prawdy i na początek rozpatrzmy przypadek, gdy metajęzyk obejmuje język przedmiotowy. Jeśli w takim przypadku wypowiem zdanie „*Śnieg jest biały*”, to odpowiednie T-zdanie wynikające z teorii prawdy, którą sam buduję we własnym języku (idiolekcie) dla tego samego języka (idiolektu), ma postać następującą:

5.1 (Moje zdanie) „*Śnieg jest biały*” jest prawdziwe wtw, gdy śnieg jest biały.

⁴⁵ Teoria taka – na co zwraca uwagę Davidson – wymagałaby uporania się z tzw. terminami masowymi jak śnieg, woda itp.

Powyższe T-zdanie jest banalne i oczywiste. Jest tak jednak dlatego, że zdanie „Śnieg jest biały” po prostu rozumiem, czyli znam jego znaczenie. A zatem to nie teoria prawdy, z której wynika (5.1), podaje znaczenie zdania „Śnieg jest biały”, lecz znajomość znaczenia tego zdania (oraz innych zdań mojego własnego idiolektu) pozwala uznać (5.1) za jego warunek prawdziwości. W konsekwencji teoria prawdy dla mojego własnego idiolektu, zbudowana w moim własnym idiolektcie, okazuje się trywialna. Można się o tym łatwo przekonać rozpatrując zdanie o tym samym kształcie fizycznym, tj. „Śnieg jest biały”, lecz wypowiedziane przez inną osobę. W takim przypadku wynikające ze zbudowanej przeze mnie teorii prawdy dla idiolektu⁴⁶ tej osoby odpowiednie T-zdanie dla wypowiedzianego przez nią zdania „Śnieg jest biały” ma postać:

5.2 (Zdanie X-a) „Śnieg jest biały” jest prawdziwe wtw, gdy śnieg jest biały.

Jak widać T-zdanie (5.2) nie ma już tak trywialnego charakteru jak (5.1). Czy można je jednak traktować jako warunek prawdziwości zdania „Śnieg jest biały”? W przypadku oryginalnej wersji Tarskiego definicji prawdy nie stanowiło to żadnego problemu, gdyż Tarski zakładał, że znamy przekład z języka przedmiotowego na metajęzyk, zaś każde T-zdanie wynikające z teorii prawdy pełniło funkcję definicji częściowej predykatu „prawdziwy”. W prezentowanym przez Davidsona ujęciu teorii prawdy założenia o istnieniu przekładu się jednak nie czyni. Wręcz przeciwnie, to teoria prawdy ma nam podać przekład zdań języka przedmiotowego na metajęzyk. A zatem (5.2) może być warunkiem prawdziwości wypowiedzianego przez X-a zdania „Śnieg jest biały” dopiero wtedy, gdy zostanie ono jakoś zinterpretowane. Trudno bowiem odnosić się do prawdziwości czy fałszywości niezinterpretowanych ciągów dźwięków lub napisów. Kiedy jednak zdanie to zostanie zinterpretowane oznacza to, że musimy dysponować gotową teorią prawdy, z której – zgodnie z konwencją T – wynikają wszystkie T-zdania, również (5.2). Zdanie (5.2) stanowi zatem warunek prawdziwości zdania „Śnieg jest biały” jedynie w ramach teorii prawdy, z której wynika. Co więcej, może się okazać, że jakieś inne T-zdanie wynikające z tej teorii skłoni nas do jej weryfikacji. W takim przypadku można zasadnie zakwestionować stwierdzenie, że (5.2) jest warunkiem prawdziwości zdania „Śnieg jest biały”.

Aby się o tym przekonać wyobraźmy sobie, że wyrażenie „śnieg” w moim idiolektcie znaczy tyle, co „lód” w idiolektcie X-a i na odwrót, zaś „biały” w moim idiolektcie znaczy tyle, co „zimny” w idiolektcie X-a i na odwrót.

⁴⁶ Przez idiolekt będziemy (za Davidsonem) rozumieć indywidualny język danej osoby. Zagadnienie to zostanie rozwinięte w rozdziale 4.

Znajduje to odbicie w aksjomatach „właściwej”, lecz nieznaney nam w tym momencie teorii prawdy⁴⁷:

Ciąg C spełnia funkcję zdaniową „śnieg(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi *t* jest lodem.

Ciąg C spełnia funkcję zdaniową „lód(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi *t* jest śniegiem⁴⁸.

Ciąg C spełnia funkcję zdaniową „biały(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi *t* jest zimny.

Ciąg C spełnia funkcję zdaniową „zimny(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi *t* jest biały.

Założmy jednak, że nie wiedząc o tym, dysponujemy teorią prawdy, z której wynika (5.2) oraz szereg analogicznych T-zdań, odpowiadających zdaniom „Lód jest biały”, „Śnieg jest zimny” oraz „Lód jest zimny”⁴⁹. Jaki użytek można – przy tych założeniach – poczynić z (5.2) i pozostałych T-zdań odpowiadających wymienionym wyżej zdaniom, skoro w prezentowanym ujęciu Davidsona nie pełnią one funkcji częściowych definicji prawdy przy założeniu znajomości przekładu? Użytek ten mógłby polegać, na przykład, na poprawnym zinterpretowaniu zdania *X*-a „Śnieg jest biały” jako synonimu „Lód jest zimny” w naszym własnym języku (idiolekcie). Jeśli jednak ograniczymy się do samego (5.2) oraz wspomnianych T-zdań wynikających z podanych w przypisie aksjomatów, to nie uda się nam tego osiągnąć, gdyż równokształtne zdania „Śnieg jest biały” w obu idiolektach znaczą wprowadzić coś innego, ale są prawdziwe⁵⁰. A zatem nic nie stoi na przeszkodzie, aby uznać T-zdanie (5.2), podobnie jak T-zdanie:

5.3 (Zdanie *X*-a) „Lód jest zimny” jest prawdziwe wtw, gdy lód jest zimny.

⁴⁷ W przykładzie tym zakładam istnienie wszechwiedzącego obserwatora.

⁴⁸ Aksjomaty te pełnią jedynie funkcję ilustracyjną. Nie należy ich traktować jako aksjomatów poprawnej teorii prawdy dla terminów masowych „śnieg” i „lód”.

⁴⁹ Na teorię tę składają się m.in. trywialne aksjomaty:

Ciąg C spełnia funkcję zdaniową „śnieg(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi „*t*” jest śniegiem.

Ciąg C spełnia funkcję zdaniową „lód(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi „*t*” jest lodem.

Ciąg C spełnia funkcję zdaniową „biały(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi „*t*” jest biały.

Ciąg C spełnia funkcję zdaniową „zimny(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg C termowi „*t*” jest zimny.

⁵⁰ Chodzi tu o prawdziwość w normalnych warunkach.

W opisanym przypadku przyjęliśmy wstępnie najprostszą (trywialną) teorię prawdy, w której zakłada się wzajemną synonimiczność wszystkich równokształtnych wyrażeń języka (idiolektu) *X*-a oraz języka (idiolektu) interpretatora⁵¹. Założenie o synonimiczności jest jednak w świetle naszej umowy ewidentnie fałszywe⁵². Z problemem można się uporać, testując teorię dla innych zdań języka *X*-a, np. „Łód jest miękki”, „Śnieg jest twardy” itp. Z naszej trywialnej teorii prawdy wynikają bowiem następujące T-zdania⁵³:

5.4 (Zdanie *X*-a) „Łód jest miękki” jest prawdziwe wtw, gdy lód jest miękki

oraz

5.5 (Zdanie *X*-a) „Śnieg jest twardy” jest prawdziwe wtw, gdy śnieg jest twardy.

Ponieważ zdania naszego idiolektu „Łód jest miękki” oraz „Śnieg jest twardy” uznajemy – zgodnie z naszymi przekonaniem – w normalnych warunkach za fałszywe, to na gruncie naszej trywialnej teorii prawdy zdania *X*-a: „Śnieg jest twardy” i „Łód jest miękki” uznajemy również za fałszywe. Otwartym problemem jest jednak to, czy nasz rozmówca jest tego samego zdania. Aby się o tym przekonać należy – odwołując się do odpowiednich świadectw – przypisać lub nie przypisywać przekonania, które są aktualizowane przez asertywne wypowiedzenia powyższych zdań. W tym momencie ujawnia się pewien szczególny pożytek z warunków prawdziwości rozumianych jako T-zdania wynikające z teorii prawdy. Jeśli bowiem stwierdzimy, że *X* szczerze obstaje przy twierdzeniu, że wypowiedziane przez niego zdania „Łód jest miękki” oraz „Śnieg jest twardy” są prawdziwe, to przy założeniu, że *X* nie żartuje, nie kłamie i jest przy zdrowych zmysłach, musimy odrzucić T-zdania (5.4) oraz (5.5), gdyż w przypadku w przypadku uogólnień typu

⁵¹ Interpretując słowa innych użytkowników naszego własnego języka zazwyczaj z góry zakładamy taką najprostszą, a zarazem trywialną teorię prawdy. Jednak nawet w takim przypadku mogą się ujawnić pewne, zazwyczaj drobne, różnice znaczenia równokształtnych wyrazów obu użytkowników. Jest to jeden z powodów, dla których Davidson odrzucił pojęcie języka, jako teoretycznie bezużyteczne. Będzie o tym mowa w kolejnym rozdziale.

⁵² Oczywiście umowa, że „śnieg” znaczy dla *X*-a tyle co „lód” w moim własnym idiolektcie, wymaga spojrzenia z perspektywy zewnętrznego obserwatora, który dodatkowo dysponuje pełną wiedzą na temat znaczeń w językach obu interlokutorów. Ponieważ założenie to nigdy nie jest spełnione, ma ono jedynie ułatwić zrozumienie rozpatrywanego przykładu.

⁵³ Przy założeniu, że składają się na nią również aksjomaty:

Ciąg *C* spełnia funkcję zdaniową „miękki(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg *C* termowi „*t*” jest miękki.

Ciąg *C* spełnia funkcję zdaniową „twardy(*t*)” wtw, gdy obiekt przyporządkowany przez ciąg *C* termowi „*t*” jest twardy.

„Lód jest (nie jest) twardy” oraz „Śnieg jest (nie jest) miękki” można rozsądnie założyć, że rozmówcy powinni być zgodni⁵⁴.

Ponieważ T-zdania wynikające z teorii prawdy muszą być prawdziwe, gdyż pełnią funkcje twierdzeń, należy odrzucić dotychczasową teorię prawdy i zastąpić ją inną. Nowa teoria prawdy, której aksjomaty zostały nieco wcześniej podane, może się różnić od poprzedniej tym, że termin „lód” w idiolekcie *X*-a znaczy w naszym języku „śnieg” i na odwrót, zaś termin „biały” znaczy „zimny” i na odwrót. Z nowej teorii będą wynikały m.in. następujące T-zdania:

5.6 (Zdanie *X*-a) „Śnieg jest biały” jest prawdziwe wtw, gdy lód jest twardy

oraz

5.7 (Zdanie *X*-a) „Lód jest twardy” jest prawdziwe wtw, gdy śnieg jest biały.

Przykład ten pokazuje jaka jest rola warunków prawdziwości zdań w zaproponowanej przez Davidsona teorii znaczenia. Warunki prawdziwości rozumiane jako T-zdania wynikające z teorii prawdy, służą do jej weryfikacji. Należy jednak pamiętać, że żadne T-zdanie wzięte w izolacji (tj. bez podania odpowiedniej teorii prawdy) nie podaje znaczenia zdania po lewej stronie równoważności. Jeśli w wyniku owej weryfikacji starą teorię sfalsyfikujemy, a następnie zastąpimy nową, to nowa teoria różni się od poprzedniej tym, że podaje inne znaczenia wyrażeń języka przedmiotowego:

Teoria znaczenia (w moim, nieco przewrotnym, rozumieniu) jest teorią empiryczną, z ambicjami wyjaśnienia działania językowego. Jak każda teoria, może być testowana przez porównanie jej konsekwencji [T-zdań] z faktami (s. 14).

Davidson nazywa teorię prawdy mianem teorii empirycznej, gdyż weryfikacja odbywa się w oparciu o świadectwa empiryczne obejmujące nie tylko dane doświadczenia, ale również obserwację zachowań językowych oraz pozajęzykowych rozmówcy⁵⁵.

Jak można było zauważyć, określone T-zdania nie zawsze pozwalają na falsyfikowanie teorii, z której wynikają. Teoria prawdy budowana jest jednak dla całego języka i ma charakter holistyczny, który ujawnia się poprzez

⁵⁴ Jest to jedna z wersji tzw. Zasady życzliwości, o której będzie mowa w ostatnim rozdziale. Oczywiście zgoda nie zawsze jest wymagana. Jeśli przypiszemy rozmówcy fałszywe przekonanie, to odpowiednie T-zdanie pozostanie prawdziwe i nie ma powodu, aby teorię prawdy falsyfikować.

⁵⁵ Zachowania pozajęzykowe pomagają nam stwierdzić kłamstwo lub zaburzenia psychiczne rozmówcy.

związki między warunkami prawdziwości. Na poziomie T-zdań związki te polegają na współwystępowaniu pewnych wyrażeń języka przedmiotowego w zdaniach po lewej stronie równoważności oraz na współwystępowaniu pewnych wyrażeń metajęzyka w zdaniach po prawej stronie równoważności. Na poziomie teorii prawdy związki te ujawniają się na poziomie aksjomatów. Pozwala to, jak w powyższym przykładzie, sfalsyfikować teorię za pomocą innych, lecz powiązanych znaczeniowo z nimi, T-zdań:

Działanie teorii polega na ustalaniu związków pomiędzy znanymi warunkami prawdziwości każdego zdania, a tymi jego aspektami („słowami”), które pojawiają się w innych zdaniach i pełnią w nich takie same funkcje (s. 15).

Powróćmy jednak do problemu, w jaki sposób teoria prawdy podaje znaczenia zdań. Aby sformułować warunek prawdziwości zdania *X*-a „Śnieg jest biały” w postaci (5.6) musimy zbudować, jak to zasugerowałem w przykładzie, teorię prawdy dla całego języka (idiolektu) *X*-a. Z teorii tej będą wynikać wszystkie odpowiednie *T*-zdania, w tym również (5.6), oraz szereg innych, w tym również równokształtnych, z niektórymi *T*-zdaniami wynikającymi z trywialnej teorii prawdy. Stanowi to kolejny argument na rzecz przytaczanej już tezy, że pojedyncze *T*-zdania, chociaż są warunkami prawdziwości poszczególnych zdań języka *X*-a, nie podają ich znaczeń. Znaczenia zdań podaje cała teoria prawdy:

(...) definicja [prawdy Tarskiego] działa poprzez podanie koniecznych i wystarczających warunków prawdziwości każdego zdania, a to oznacza podanie w pewien sposób znaczenia zdania. Jeśli się zna semantyczne pojęcie prawdy dla danego języka, to wie się, co znaczy – dla dowolnego zdania – być prawdziwym, a to z kolei jest – w pewnym uzasadnionym sensie – równoważne z rozumieniem tego języka (s. 13-14).

Z kolei znaczenia słów ograniczają się do ich wkładu w znaczenia wszystkich zdań, w których owe słowa występują, czyli – innymi słowy – do wkładu w warunki prawdziwości wszystkich zdań.

Skąd się jednak bierze pozór, że pojedynczy warunek prawdziwości (tj. *T*-zdanie) podaje znaczenie zdania, które znajduje się po lewej stronie równoważności? Jest tak dlatego, że gdy zaakceptujemy teorię prawdy dla języka (idiolektu) *X*-a, możemy zastąpić predykat „jest prawdziwy” predykatem „znaczy, że” we wszystkich wynikających z tej teorii *T*-zdaniami. W naszym przypadku jedno z tak przekształconych *T*-zdań (od angielskiego *meaning* możemy je nazywać *M*-zdaniami) miałoby postać:

(Zdanie *X*-a) „Śnieg jest biały” znaczy, że lód jest zimny.

Oczywiście na gruncie tej samej teorii będziemy mieli M-zdanie:

(Zdanie X-a) „Łód jest zimny” znaczy, że śnieg jest biały

oraz nieskończenie wiele innych M-zdań.

Do sprawy zastępowania predykatu „jest prawdziwy” przez predykat „znaczy, że” powrócę nieco dalej. W *Prawdzie i znaczeniu* Davidson pokazuje, że nie da się zbudować teorii znaczenia, w której występuje *explicite* predykat „znaczy, że” i z której wynikają M-zdania. Podstawowy argument przeciwko takiej możliwości polega na tym, że predykat „znaczy, że”, łączący opis (nazwę) zdania z jego metajęzykowym synonimem, nie jest ekstensjonalny (s. 11) ponieważ jak twierdzi Davidson:

Teoria [znaczenia] spełni swe zadanie, jeśli każdemu zdaniu *s* badanego języka przyporządkuje takie zdanie (zastępujące „*p*”), które, w sposób wymagający jeszcze wyjaśnienia, „poda znaczenie” *s* (*ibidem*).

Aby ustanowić to przyporządkowanie nieco dalej postuluje, aby zrezygnować z niejasnego i nieekstensjonalnego zwrotu „znaczy, że”, opatrzyć zdanie *p* ekstensjonalnym spójnikiem równoważności oraz, aby po lewej stronie otrzymać zdanie, dołączyć do opisu zastępującego *s* (czyli strukturalnoopisowej nazwy zdania *s*) odnoszący⁵⁶ się doń predykat (s. 12). Wówczas zamiast „*s* znaczy, że *p*” otrzymamy: „*s* jest *T* wtedy i tylko wtedy, gdy *p*”.

W ten, jak sam stwierdza, nieco okrężny – i dodajmy, że raczej perswazyjny sposób – Davidson dochodzi do teorii prawdy jako teorii znaczenia, gdyż predykat *T*, to po prostu predykat „jest prawdziwy”. To właśnie teoria prawdy, za pomocą wynikających z niej T-zdań, przyporządkowuje zdaniom języka przedmiotowego zdania metajęzyka, które (przy założeniu znajomości całej teorii prawdy) „podają” ich znaczenia. Kiedy jednak uznamy, że określona teoria prawdy jest (wobec wszelkich dostępnych nam świadectw) poprawną teorią prawdy dla języka (idiolektu) naszego rozmówcy, można – bez jakiegokolwiek szkody – zastąpić we wszystkich wynikających z niej T-zdaniach predykat „prawdziwy” zwrotem „znaczy, że”. Zaznaczyć jednak należy, że jest to jedynie powierzchowny zabieg mający na celu nawiązanie do naszych elementarnych intuicji i dotyczy jedynie T-zdań, nie zaś teorii prawdy, z której te zdania wynikają.

⁵⁶ W oryginale czytamy: *supply the description that replaces “s” with its own predicate*. W polskim tłumaczeniu mamy: „dołączymy do opisu zastępującego *s* jego własny predykat” (s. 12). Fragment nie jest do końca jasny, ale warto przypomnieć, że w *Theories of Meaning* Davidson pisze o tym predykanie jako odnoszącym się do wszystkich zdań języka.

6. OKAZJONALNOŚĆ

W podawanych do tej pory przykładach teorii prawdy nie występowały zdania okazjonalne, tj. zdania, których wartość logiczna zależy od okoliczności (kontekstu) ich wypowiedzenia: osoby mówiącej, miejsca oraz momentu wypowiedzenia. W przykładzie podanym w poprzednim paragrafie do weryfikacji teorii prawdy służyły nam zdania postaci „Śnieg jest biały”, które można rozumieć jako uogólnienia wielu zdań okazjonalnych postaci „Próbka śniegu znajdująca się w obecności X -a w momencie czasu t jest biała”. Uogólnienia te miałyby faktycznie postać „Każda próbka śniegu w obecności każdego X -a, w każdym momencie czasu jest biała”.

W oryginalnym sformułowaniu Tarskiego semantyczna definicja prawdy z oczywistych względów nie uwzględniała zdań okazjonalnych. Dokonane przez Davidson rozszerzenie semantycznej definicji prawdy na zdania okazjonalne ma zatem dwa powody. Jeżeli teoria prawdy ma być teorią znaczenia języka naturalnego, to tzw. zdania okazjonalne – tj. takie, których wypowiedzenia mogą zmieniać wartość logiczną w zależności od tego, kto, kiedy i gdzie ich użył – muszą niewątpliwie być również brane w niej pod uwagę. Jest tak nie tylko dlatego, że nasze codzienne rozmowy obfitują w zdania okazjonalne, ale także dlatego, że warunki prawdziwości wypowiedzeń takich zdań są szczególnie użyteczne w weryfikacji empirycznej teorii prawdy. Zamiast bowiem używać do tego celu uogólnień indukcyjnych (jak w podanym przeze mnie przykładzie), o wiele łatwiej jest wykorzystać wypowiedzenia zdań okazjonalnych, których prawdziwość lub fałszywość można bezpośrednio w danych warunkach stwierdzić. Zamiast odwoływać się do uogólnień i liczyć na to, że nasz rozmówca dokonał tych samych uogólnień, o wiele łatwiej uzyskać zgodę między rozmówcami w przypadku wypowiedzeń zdań: „To jest śnieg”, „To jest lód”, „To jest białe”, czy „To jest zimne”.

Aby to zilustrować powróćmy do podanego wcześniej przykładu teorii prawdy. Teorię tę byłoby o wiele łatwiej weryfikować, gdyby wynikało z niej między innymi T-zdanie:

Zdanie wypowiedziane przez X -a w momencie t „To jest śnieg” jest prawdziwe wtw, gdy próbka, na którą wskazuje X w momencie t , jest śniegiem.

Gdyby okazało się, że X (po stwierdzeniu, że nie kłamie, nie żartuje i we właściwy sposób odbiera bodźce) odrzuca zdanie „To jest śnieg”, wskazując jednocześnie na typową próbkę śniegu, mielibyśmy podstawę, aby sfalsyfikować naszą teorię. Konieczna jest zatem pewna modyfikacja teorii prawdy, o czym *explicit*e pisze Davidson:

Potraktujmy prawdę jako własność nie zdań, ale wypowiedzi⁵⁷, aktów mowy, czy też uporządkowanych trójek złożonych ze zdań, czasu i osób, albo prościej: jako relacje pomiędzy zdaniem, osobą i czasem. Przy takim ujęciu obowiązują nadal zwykłe relacje logiczne między zdaniami, ale tylko zrelatywizowanymi do tego samego czasu i wypowiadającej je osoby; relacje między pozostałymi zdaniami można opisać za pomocą nowych aksjomatów (czyż nie będę się tu zajmował). Teoria ulegnie systematycznej, choć nie zagadkowej, zmianie; każdemu wyrażeniu z elementem okazjonalnym odpowiadać będzie w teorii teza relatywizująca warunki prawdziwości zdań, w których to wyrażenie się pojawia, do różnego czasu i wypowiadających je osób (s. 29).

Rozszerzenie teorii na konstrukcje okazjonalne ma dla Davidsona jeszcze jeden niezwykle ważny aspekt. Jak pamiętamy z poprzedniego rozdziału, Davidson starał się wyeliminować z języka naturalnego wszechobecne w nim konteksty nieekstensjonalne. Sposobem osiągnięcia tego celu było dla Davidsona wskazanie w kontekście nieekstensjonalnym ukrytego elementu okazjonalnego, który pozwala na taką eliminację. Jest to rezultat równie doniosły, jak zaadaptowanie teorii prawdy do typowych wypowiedzi okazjonalnych, np. „To jest królik”⁵⁸.

Fakt, że wyrażenia okazjonalne poddają się obróbce formalnej, zwiększa znacznie nadzieję na poważną semantykę języka naturalnego. Jest bowiem prawdopodobne, iż wiele istotnych problemów, takich jak analiza cytatów czy zdań o postawach propozycjonalnych może być rozwiązanych, jeśli odnajdziemy tam ukrytą konstrukcję okazjonalną (s. 30).

Do sprawy formułowania aksjomatów teorii prawdy dla zdań okazjonalnych powrócę pod koniec następnego paragrafu, przy okazji omawiania odpowiedzi na zarzuty Fostera. W dalszej części tego paragrafu ograniczę się do przedstawienia pewnych zastosowań teorii prawdy rozszerzonej na zdania okazjonalne.

Jak już wspominałem Davidson stosował teorię prawdy nie do zdań rozumianych jako pewne twory abstrakcyjne, lecz do wypowiedzi zdań, przy czym przez wypowiedzenia zdań rozumie się również napisy. Podejście to umożliwiło analizę nowego rodzaju elementów okazjonalnych⁵⁹. Zazwyczaj przez elementy okazjonalne rozumie się wyrażenia (najczęściej zaimki), które zmieniają odniesienie w zależności od kontekstu i okoliczności wypowiedzenia, np. obecnych obiektów, osób, czasu oraz miejsca. Zdaniem Davidsona

⁵⁷ Być może lepszym tłumaczeniem terminu *utterance* byłoby „wypowiedzenie”, które konsekwentnie stosuję w tej pracy.

⁵⁸ Jest to oczywiście aluzja do znanego przykładu podanego przez Quine’a w *Słowie i rzeczy*.

⁵⁹ Używam tu terminu „element okazjonalny”, gdyż nie musi to być wyrażenie występujące w zdaniu *explicite*.

w wielu wypowiedzeniach można znaleźć, zazwyczaj ukryte, elementy okazjonalne odnoszące się do części owych wypowiedzeń⁶⁰. Analiza, która wydobyla na powierzchnię okazjonalność względem (części) wypowiedzeń oraz pozwala przedstawić wypowiedzenia w ich formie kanonicznej nosi nazwę analizy parataktycznej.

O analizie parataktycznej była już mowa w poprzednim rozdziale. Pozwoliła ona Davidsonowi zastosować ekstensjonalną teorię prawdy do wypowiedzeń nieekstensjonalnych. Jak pamiętamy, nieekstensjonalność wypowiedzeń była wedle Davidsona pozorna, gdyż po wydobyciu na powierzchnię elementów okazjonalnych, forma kanoniczna całego wypowiedzenia okazywała się ekstensjonalna. Analizę parataktyczną Davidson zastosował do wyrażen cudzysłowowych, trybów gramatycznych oraz do zdań zawierających zwrot „powiedział, że”. Widział również jej potencjalne zastosowanie do zdań performatywnych oraz zdań wyrażających postawy propozycyjne. Analiza parataktyczna z pewnością zasługuje na oddzielne opracowanie i dlatego w tym miejscu zarysuję jedynie podstawową ideę tej analizy, unikając podejmowania oraz analizowania kontrowersji, które owa analiza wzbudza.

W artykule *Quotation* (1979) zdania, w których występują wyrażenia cudzysłowowe, analizowane są jako pary złożone z wypowiedzenia zdania oraz wypowiedzenia wyrażenia wewnątrz cudzysłowu, przy czym wspomniane zdanie ma charakter okazjonalny, gdyż zawiera ukryty element wskazujący na wyrażenia wewnątrz cudzysłowu. Na przykład zdanie:

Lecha Wałęsa powiedział „Zbudujemy drugą Japonię”

można analizować jako:

Lecha Wałęsa powiedział zdanie o kształcie fizycznym identycznym z następnym zdaniem. Zbudujemy drugą Japonię.

Oczywiście w analizie wypowiedzeń cudzysłowowych należy rozważyć wiele okoliczności wiążących się z różnymi sposobami użycia cudzysłowu, które w tym miejscu pomijam.

Analizę trybów gramatycznych oraz zapowiedź wypowiedzeń performatywnych przedstawił Davidson w *Moods and Performances* (1979a). O zastosowaniu analizy parataktycznej do wypowiedzeń performatywnych wspomina

⁶⁰ W gramatyce odróżnia się zaimki wskazujące (tzw. deiktyczne) od tzw. zaimków anaforycznych, które wskazują na pewne wypowiedzenia. Rozpatrzmy zdania „Spotkałem wczoraj Jan a. Widziałem go na ulicy”. Zaimek „on” w dopełniaczu („go”) pełni tutaj rolę zaimka anaforycznego, gdyż nie odnosi się bezpośrednio do osoby, lecz czyni to za pośrednictwem wypowiedzianej w poprzednim zdaniu nazwy „Jan”.

również pod koniec *On Saying That* (1979). Rozważmy przykład zdania z czasownikiem performatywnym „rozkazać”:

Oficer X rozkazał żołnierzowi *Y* biec do najbliższego drzewa⁶¹.

Zgodnie z analizą parataktyczną zdanie to należy interpretować w następujący sposób:

Oficer X wypowiedział rozkaz o znaczeniu identycznym ze znaczeniem następnego zdania. *Y* biegnie do najbliższego drzewa,

przy czym pierwsze zdanie jest zdaniem wprowadzającym, w którym podaje się typ spełnionego aktu mowy (w tym przypadku jest to rozkaz), zaś drugie zdanie podaje treść (znaczenie) owego aktu.

Co daje analiza parataktyczna zdań performatywnych? Przede wszystkim potraktowanie wypowiedzenia zdania z czasownikiem performatywnym jako pary wypowiedzeń pozwala zastosować teorię prawdy niezależnie do obu wypowiedzeń, co umożliwia ich zinterpretowanie za pomocą środków ekstensjonalnych. Po drugie propozycja ta pozwala ocenić dwa, logicznie niezależne od siebie wypowiedzenia, w kategoriach prawdy i fałszu. Pierwsze z tych wypowiedzeń głosi, że został spełniony odpowiedni akt mowy. A zatem w pewnych okolicznościach zdanie wprowadzające może okazać się fałszywe. Może tak być, gdy *X* nie jest osobą uprawnioną do wydawania *Y*-owi rozkazów, czyli nie jest oficerem, lecz jedynie osobą pod oficera się podszywającą. Niezależnie od tego rozkaz lub rzekomy rozkaz może być wykonany lub nie (np. może zostać zbojkotowany). W pierwszym przypadku zdanie podające treść rozkazu (lub rzekomego rozkazu) będzie prawdziwe, a drugim przypadku fałszywe, gdyż *Y* do drzewa się nie czołga⁶². Ponieważ mamy do czynienia ze zdaniami niezależnymi od siebie logicznie, możemy wskazać na cztery możliwe przypadki odpowiadające prawdziwości lub fałszywości obu zdań.

W przypadku zdań performatywnych wypowiedzianych w I osobie czasu teraźniejszego strony czynnej, np. „Biegnij do tamtego drzewa. To jest rozkaz” lub „Rozkazuję, biegnij do tamtego drzewa”, analiza parataktyczna wyjaśnia pewnego rodzaju sprzężenie zwrotne, które zauważył już Austina w *Czynnościach mowy* (Austin 1962). Zgodnie z analizą parataktyczną zdanie to miałyby następującą postać:

⁶¹ W zdaniu tym występuje element okazjonalny „najbliższe drzewo”, którego analizę w tym przykładzie pomijam.

⁶² Wszechstronna analiza tego przykładu wymagałaby uwzględnienie „zwykłej” okazjonalności, np. momentu czasu wypowiedzenia rozkazu (lub rzekomego rozkazu) oraz momentu jego wykonania (lub zbojkotowania).

To co teraz mówię jest rozkazem o treści podanej przez moje następne wypowiedzenie. Y biegnie do tamtego drzewa.

Ów efekt polega na tym, że samo wypowiedzenie zdania wprowadzającego stanowi jeden z warunków sukcesu (spełnienia aktu mowy), czyli prawdziwości zdania wprowadzającego.

W tym miejscu konieczne jest pewne wyjaśnienie, które dotyczy dwójakiego rozumienia terminu „warunek prawdziwości”. Można bowiem powiedzieć, że samo wypowiedzenie zdania performatywnego w I osobie czasu teraźniejszego strony czynnej jest jednym z warunków prawdziwości zdania wprowadzającego. Należy jednak pamiętać, że nie jest to takie rozumienie terminu „warunek prawdziwości”, w jakim warunkami prawdziwości są T-zdania wynikające z teorii prawdy. Na przykład T-zdaniem wynikającym z odpowiedniej teorii prawdy byłoby:

Zdanie wypowiedziane przez X-a do Y-a „To co teraz mówię jest rozkazem o treści podanej przez następne wypowiedzenie”, po którym następuje wypowiedzenie zdania „A” jest prawdziwe, gdy zdanie wypowiedziane przez X-a do Y-a jest rozkazem o treści podanej przez zdanie A.

Warunkiem prawdziwości zdania wprowadzającego w sensie przyjętym przez Davidsona jest po prostu prawa strona powyższej równoważności, natomiast warunkami prawdziwości w sensie potocznym są m.in. samo wypowiedzenie formy „rozkazuję” przez X-a, posiadanie przez X-a uprawnień do wydawania rozkazów Y-owi itp. Dlatego też Davidson pisze, że wypowiedzenie formy „rozkazuję” jest jednym z warunków sukcesu odpowiedniego aktu mowy, chociaż sukces tego aktu polega po prostu na prawdziwości zdania wprowadzającego.

Jak pamiętamy z pierwszego rozdziału sama teoria prawdy nie podaje kryterium prawdziwości⁶³ i nie rozstrzyga, które zdania są prawdziwe, a które fałszywe. W zależności od rodzaju wypowiedzianego zdania zazwyczaj dysponujemy pewnymi – nie zawsze skutecznymi i niezawodnymi – procedurami lub sposobami rozstrzygania jego wartości logicznej. Warunkami (kryteriami) prawdziwości w sensie potocznym są zatem pewne reguły składające się na te procedury, np. kryterium prawdziwości zdania okazjonalnego jest zachodzenie w rzeczywistości fizycznej pewnych faktów lub stanów rzeczy. W przypadku zdań wprowadzających w analizie parataktycznej zdań performatywnych, jedną z reguł pozwalających stwierdzić ich prawdziwość jest użycie w wypowiedzeniu tzw. performatywu wyraźnego, tj. czasownika performatywnego w I osobie czasu teraźniejszego, formy czynnej.

⁶³ Często warunki prawdziwości są po prostu utożsamiane z kryteriami prawdziwości.

Kolejnym zastosowaniem analizy parataktycznej są zdania wyrażające postawy propozycyjalne. W *On Saying That* (1979) Davidson ograniczył się jednak do analizy formy „powiedział, że”. Zanim przedstawię rozwiązanie zaproponowane przez Davidsona warto wspomnieć, że bardzo podobną analizę zdań psychologicznych głosił już w latach 20. Tadeusz Kotarbiński. Co więcej, rozwiązanie to było przez Kotarbińskiego kilkakrotnie przedstawione w różnych publikacjach (por. 1922, 1935, 1958a, 1958b). A zatem analiza proponowana przez Davidsona nie jest całkowicie oryginalna.

Kiedy mówimy, że ktoś coś powiedział, wówczas – w odróżnieniu od przytaczania czyjejś wypowiedzi, które polega na umieszczeniu jej w cudzysłowie – podajemy inną wypowiedź (niekiedy w zupełnie innym języku), która ma to samo znaczenie, co wypowiedź oryginalna. Dlatego też, gdy chcemy stwierdzić, że Galileusz powiedział, że Ziemia się porusza nie musimy cytować go w języku włoskim. Oczywiście zdanie „Galileusz powiedział, że Ziemia się porusza” tworzy kontekst intensjonalny, gdyż zdanie podrzędne nie zawsze da się zastąpić innym zdaniem o tej samej wartości logicznej (w tym przypadku zdaniem prawdziwym), bez zmiany wartości logicznej całości. Wystarczy zamiast „Ziemia się porusza” podstawić inną prawdę, której Galileusz nigdy nie wypowiedział. Zdaniem Davidsona poprawna (tj. ekstensjonalna) analiza tego zdania ma postać następującą:

Ziemia się porusza. Galileusz to powiedział

albo

Galileusz wypowiedział zdanie synonimiczne z następnym zdaniem wypowiedzianym przeze mnie. Ziemia się porusza.

Podobnie jak w poprzednich przypadkach mamy faktycznie do czynienia z wypowiedzeniem dwóch niezależnych logicznie zdań.

Na zakończenie uwag na temat analizy parataktycznej warto zwrócić uwagę na sprawę natury ogólniejszej. W przypadku wyrażen cudzysłowowych wypowiedzenie zdania, na które wskazuje zdanie wprowadzające, ma być identyczne co do kształtu fizycznego z wypowiedzeniem wewnątrz cudzysłowu. W przypadku analizy zdań performatywnych zdanie, na które wskazuje zdanie wprowadzające, ma treść identyczną z treścią rozkazu, prośby lub innego aktu mowy⁶⁴. W ostatnim przypadku zdanie wprowadzające wskazuje na wypowiedzenie, które ma to samo znaczenie, co wypowiedziane przez kogoś zdanie. Wydawać by się zatem mogło, że w ten sam sposób można

⁶⁴ Na pewne trudności związane z analizą parataktyczną zdań performatywnych zwracam uwagę w (Maciaszek 2015).

analizować zdania psychologiczne, np. zdania przypisujące przekonania. W analizie tej natrafiamy jednak na zasadniczą trudność. Aby to zilustrować rozpatrzmy zdanie „Kopernik był przekonany, że Ziemia krąży dookoła Słońca”. Analiza parataktyczna powyższego zdania mogłaby przybrać następującą postać:

Kopernik żywił przekonanie o zawartości podanej przez następujące zdanie.
Ziemia krąży dookoła Słońca.

Problem, jaki się tu pojawia, wiąże się z niejasnym terminem „zawartość przekonania”. Aby nadać temu terminowi w miarę precyzyjne znaczenie należy poczynić pewne – niekiedy daleko idące – założenia filozoficzne, dotyczące naszych stanów mentalnych. Dlatego też Davidson ograniczył się jedynie do wskazania możliwości zastosowania analizy parataktycznej do zdań psychologicznych, jako elementu postulowanego przezeń programu semantycznego⁶⁵, polegającego na podawaniu teorii prawdy dla kolejnych fragmentów języka naturalnego. Pewną wskazówkę stanowi zasugerowana przez niego redukcja postaw propozycjonalnych do aktów mowy, które (przy założeniu szczerości mówiących) aktualizują te postawy. W przypadku przekonania odpowiednimi aktami mowy byłyby stwierdzenia (asercje), gdyż przekonania wyrażamy (aktualizujemy) właśnie za pomocą stwierdzeń.

7. JOHNA FOSTERA ZARZUTY WOBEC TEORII PRAWDY JAKO TEORII ZNACZENIA

Podsumujmy główną tezę sformułowaną w *Prawdzie i znaczeniu*: znaczenia zdań języka podaje cała teoria prawdy. Sformułowanie „podaje znaczenia” pozostawia jednak pewien niedosyt. W jakim sensie teoria Tarskiego podaje znaczenia zdań? W pierwszym rozdziale tymczasowo zaproponowałem, aby owo podawanie znaczeń rozumieć jako przekład z języka przedmiotowego na znany nam metajęzyk. Przekładu dokonuje wprawdzie cała teoria prawdy, ale dla gotowej teorii możemy zastąpić wynikające z niej typowe T-zdania postaci: „Zdanie *Z* jest prawdziwe wtw, gdy *S*”, przez M-zdania postaci:

Zdanie „*Z*” znaczy, że *S*.

gdzie zdanie *S* jest oczywiście tłumaczeniem zdania *Z* języka przedmiotowego na metajęzyk.

⁶⁵ Na trudności związane z analizą zdań psychologicznych proponowaną przez Kotarbińskiego zwracam uwagę w (2016 i 2017), co – jak sądzę –częściowo wyjaśnia również powód, dla którego Davidson ograniczył się jedynie do dokładnej analizy formy „powiedział, że”.

Oczywiście zabieg ten nie wiąże się z budowaniem teorii znaczenia, w której *explicite* wystąpiłby intensjonalny funktor „znaczy, że”. Jest to raczej wyraz naszego przekonania, że teoria prawdy spełnia konwencję T, w której Tarski odwołuje się do przekładu z języka przedmiotowego na metajęzyk (1995/1933, s. 61). Jeżeli pojawią się jakiegokolwiek podstawy, aby zakwestionować teorię prawdy, możemy zastąpić ją inną teorią prawdy, z której będą wynikały (częściowo) inne T-zdania, które z kolei będziemy mogli zastąpić kolejnymi M-zdaniami.

Okazuje się jednak, że opisana wyżej sytuacja nie jest aż tak oczywista i rozwiązanie Davidsona spotkało się z wieloma zastrzeżeniami, które skłoniły go nie tyle do rewizji poglądów, ale do ich doprecyzowania i wyjaśnienia. Spośród tekstów krytycznych wyróżnia się artykuł Johna Fostera *Meaning and Truth Theory* (Foster 1976). Tekstowi temu warto poświęcić więcej uwagi, gdyż został on wysoko oceniony przez samego Davidsona, a ponadto skłonił go do odpowiedzi, która przybrała postać pełnego artykułu, w którym odpowiadając na zarzuty doprecyzowuje swoje stanowisko⁶⁶. W szczególności Davidson wskazuje, że podawanie znaczeń wyrażeń języka naturalnego nie można bez zastrzeżeń utożsamiać z przekładem języka przedmiotowego na metajęzyk.

Artykuł Fostera składa się z czterech części, z których dla naszych celów są interesujące trzy pierwsze⁶⁷. W pierwszym rozdziale Foster stara się podać warunki, jakie powinna spełniać teoria znaczenia:

Teoria ma odkryć semantyczną maszynę, którą wykorzystuje kompetencja językowa, chociaż sama teoria nie określa psychologicznej formy, jaką owa kompetencja przybiera (Forster 1976, s. 2).

Teoria znaczenia ma stanowić zatem reprezentację kompetencji semantycznej użytkownika języka⁶⁸, nie ma zaś być opisem psychologicznych mechanizmów rządzących rozumieniem języka. W tej kwestii między Fosterem i Davidsonem panuje pełna zgoda:

⁶⁶ Zazwyczaj Davidson odpowiadał na zarzuty zdawkowo. W tym przypadku nie tylko odpowiedział pełnym artykułem (1976), ale opublikował go w drugim tomie swoich prac zebranych (1984).

⁶⁷ W czwartej części Foster przedstawia własną propozycję teorii znaczenia, odwołującą się do innego rozumienia terminu „warunki prawdziwości” niż u Davidsona.

⁶⁸ A dokładniej, reprezentację aktualizacji owej kompetencji, która przejawia się jako rozumienie słów drugiego człowieka. Dlatego też stwierdzenia takie jak „Aby zrozumieć słowa drugiego człowieka, należy zbudować dla nich teorię prawdy”, należy rozumieć jedynie jako skróty myślowe.

(...) jeśli teoria poprawnie opisuje kompetencję interpretatora, to musi jej odpowiadać w interpretatorze pewien [psychologiczny] mechanizm (Davidson 2005/1986, s. 96).

Różnica między Fosterem a Davidsonem dotyczy raczej podstawy, na której opiera się teoria znaczenia. O ile bowiem Davidson odwoływał się do wyuczalności języka, czego konsekwencją był wymóg rekurencyjności teorii, to podstawowym warunkiem, jaki Forster narzuca na teorię znaczenia, jest jej interpretacyjność:

Nazwijmy teorię, która wystarcza do opanowania [języka] teorią *interpretacyjną*. Założmy, że dysponujemy językiem L , dla którego poszukujemy teorii znaczenia. Interpretacyjność [teorii znaczenia] zapewnia, że odnoszenie się⁶⁹ [teorii] do wyrażen języka L oraz przyporządkowanie znaczeń wyrażeniom języka L dają się uchwycić we właściwy sposób, tj. że jesteśmy zdolni powiedzieć do jakich wyrażen teoria się odnosi na podstawie kształtu fizycznego owych wyrażen i jakie znaczenia teoria przypisuje owym wyrażeniom poprzez wyrażenie tych znaczeń w naszym języku (*ibidem*).

Dla obydwu warunków nałożonych na interpretacyjną teorię znaczenia Foster przytacza przykłady teorii, które nie spełniają przynajmniej jednego z tych warunków, a zatem interpretacyjne nie są. W pierwszym przypadku rozważa teorię opartą na pewnym ciągu O_i zdań języka L , takim jednak, że nie istnieje procedura, która pozwala określić dla dowolnej liczby naturalnej n , jaki jest kształt fizyczny n -tego wyrazu tego ciągu. Teoria ta pozwala dla każdego zdania S języka L na przyporządkowanie mu znaczenia zgodnie ze schematem „ n -ty element O_i znaczy, że p ”, gdzie n jest pozycją zdania S w O_i , zaś p jest jego przekładem na język angielski. Teoria ta pozwala w pewnym sensie na przyporządkowanie znaczeń zdaniom języka L , ale ponieważ nie możemy powiązać kształtu fizycznego zdań języka przedmiotowego z regułą podającą ich znaczenia, nie możemy użyć tej teorii do interpretowania konkretnych zdań.

Ponieważ podany przez Fostera przykład ma charakter bardzo ogólny, postaramy się go ukonkretnić. Wyobraźmy sobie, że jakiś super komputer

⁶⁹ Teoria znaczenia budowana jest w metajęzyku. Pisząc o odnoszeniu się teorii do wyrażen języka przedmiotowego chodzi mu o to, w jaki sposób wyrażenia języka przedmiotowego reprezentowane są w metajęzyku, tj. jakie wyrażenia metajęzyka odnoszą się do wyrażen języka przedmiotowego. Najprostszym sposobem reprezentowania wyrażen języka przedmiotowego w metajęzyku jest wzięcie ich w cudzysłów. Wówczas – zgodnie ze skrytykowanymi przez Davidsona teoriami Tarskiego oraz Quine, a – wyrażenia cudzysłowowe odnoszą się do wyrażen, które znajdują się wewnątrz cudzysłowu. Oczywiście wyrażenia metajęzyka, które w ten sposób odnoszą się do wyrażen języka przedmiotowego, nie są ich przekładami.

generuje ciąg zdań języka L , podając jednocześnie ich tłumaczenia na język angielski w postaci „ n -ty wyraz generowanego przeze mnie ciągu zdań znaczy, że p ”, ale nie podając jednak kształtu fizycznego zdania, który jest n -tym wyrazem ciągu. Oczywiście taka teoria jest – zdaniem Fostera – bezużyteczna, gdyż nie pozwala nam zrozumieć konkretnych zdań języka L . Wydaje się jednak, że teoria ta jest również bezużyteczna z powodu, o którym Foster nie wspomina. Jeżeli mamy do czynienia z językiem o potencjalnie nieskończonej ilości zdań – a takie są języki naturalne – to nawet gdyby komputer ujawnił kształt fizyczny wyrazów ciągu, to i tak teoria taka nie zagwarantowałaby nam możliwości zinterpretowania każdego zdania języka L , gdyż interesujące nas zdanie po prostu nie zostało jeszcze wygenerowane. Gdyby jednak zostało ono już wygenerowane, to teoria ta byłaby skrajnie niepraktyczna, czy wręcz bezużyteczna, gdyż musielibyśmy jeszcze wyszukać owo zdanie w ciągu, nie wiedząc nawet wcześniej, czy ono się tam znajduje. Przyporządkowanie znaczenia zdaniu o konkretnym kształcie fizycznym byłoby wówczas rezultatem przypadku lub szczęśliwego trafu.

Z przykładu podanego przez Fostera wypływa pewien pozytywny wniosek, który można odnieść do stanowiska Davidsona. Przede wszystkim teoria znaczenia opisana w przykładzie podanym przez Fostera nie jest teorią języka wyuczanego. Ponieważ teoria języka wyuczanego musi być skończenie aksjomatyzowalna oraz rekurencyjna – co zapewnione jest przez wymóg, aby metajęzykowe nazwy wyrażen były strukturalnopolisowe – teoria taka zapewnia właściwe „odnoszenie się do zdań na podstawie ich fizycznego kształtu”⁷⁰. Warunki te spełnia teoria Tarskiego adaptowana do języka naturalnego⁷¹. Mając bowiem dowolne zdanie języka L o konkretnym kształcie fizycznym możemy, odwołując się do skończonej ilości aksjomatów, podać odpowiadające mu T-zdanie, które poda jego przekład. Oczywiście w konkretnym T-zdaniu można zastąpić strukturalnopolisową nazwę zdania wyrażeniem „zdanie będące n -tym wyrazem ciągu C ” – negatywny przykład podany przez Fostera odpowiadałby zatem przypadkowi, gdyby zamiast teorii Tarskiego za teorię znaczenia uznać ciąg „T-zdań” w takiej postaci. Użyłem tutaj cudzysłowu, gdyż faktycznie nie byłyby to T-zdania wynikające z jakiegokolwiek teorii prawdy.

Zauważmy, że uwagi Fostera są zbieżne ze stanowiskiem samego Tarskiego, wyrażonym w *Pojęciu prawdy*, a zaakceptowanym przez Davidsona:

⁷⁰ W tym momencie pomijam fakt, że Foster używa w przykładzie M-zdań. Nieco dalej przyznaje – podobnie jak Davidson – że teorii znaczenia generującej M-zdania nie da się zbudować ze względu na intensjonalny charakter wyrażenia „znaczy, że”.

⁷¹ Skończona aksjomatyzowalność teorii Tarskiego dla języka naturalnego jest konsekwencją skończonego słownika.

W teorii faktycznie interpretacyjnej, wyrażenia języka przedmiotowego są oznaczone przez coś, co jest nazywane *opisami strukturalnymi*, tj. opisami, które są zbudowane z nazw najprostszych symboli języka wraz z predykatem lub terminem funkcyjnym oznaczającym konkatenaację (Foster 1976, s. 3).

Drugi warunek narzucony przez Fostera na interpretacyjną teorię znaczenia dotyczy przyporządkowania zdaniom języka przedmiotowego znaczeń. Jako przykład negatywny Foster podaje teorię w pewnym sensie odwrotną do teorii przytoczonej w poprzednim przykładzie. Niech O_2 będzie ciągiem sądów⁷² wyrażonych przez zdania języka L , przy czym nie istnieje procedura przyporządkowywania owych sądów zdaniom języka angielskiego. Teoria ta pozwala dla każdego zdania S języka L na przyporządkowanie mu znaczenia zgodnie ze schematem „ x wyraża n -ty element ciągu O_2 ”, gdzie x jest strukturalnym opisem (nazwą strukturalnoopisową) zdania S , zaś n jest miejscem sądu wyrażonego przez S w ciągu O_2 . Oczywiście taka teoria nie pozwala użytkownikowi języka angielskiego zrozumieć zdań języka L , chociaż teoria ta w pewnym sensie przyporządkowuje znaczenia (sądy) zdaniom języka L ze względu na ich kształt fizyczny.

Podobnie jak w poprzednim przypadku warto ten przykład ukonkretnić. Przede wszystkim wątpliwość budzi ciąg O_2 , gdyż – jak twierdzi Foster – nie wiadomo czym dokładnie miałyby być ów ciąg sądów. Prościej jest zamiast o ciąg sądów mówić o ciągu reprezentacji sądów. W szczególności reprezentacjami tymi mogą być zdania wyrażające owe sądy. Aby taki ciąg spełniał warunek nałożony w podanym przez Fostera przykładzie, wystarczy założyć, że będzie to ciąg zdań języka nieznanego interpretatorowi. Wówczas nie wiadomo jakie zdanie języka polskiego wyraża sąd wyrażony przez n -te zdanie, czyli n -ty wyraz tego ciągu. W takim przypadku mielibyśmy do czynienia z M-zdaniami postaci „Zdanie x znaczy tyle co $n(O_2)$ ”, gdzie $n(O_2)$ jest n -tym zdaniem ciągu O_2 zdań nieznanego nam języka. Jak łatwo zauważyć teoria spełniająca konwencję T zabezpiecza przed możliwością wyprowadzenia T-zdań, a w konsekwencji M-zdań takiej postaci. Daje temu wyraz Foster stwierdzając, że jedyną znaną mu teorią znaczenia, która – na pierwszy rzut oka – spełnia sformułowane przez niego wymogi⁷³, jest teoria prawdy w ujęciu zaproponowanym przez Davidsona w jego wcześniejszych pracach – szczególnie w *Prawdzie i znaczeniu*.

Zarzuty, które Foster podnosi w drugim rozdziale rozpoczynają się od uwagi, że wynikające z teorii T-zdania podają warunki prawdziwości dla

⁷² Przez sąd rozumie Foster znaczenie zdania, nie charakteryzując tego pojęcia.

⁷³ Poza interpretacyjnością Foster dołącza kilka innych warunków, z których najważniejszym jest ekstensjonalność.

zdań języka przedmiotowego jedynie w słabym sensie. Zdaniem Fostera warunki prawdziwości w sensie standardowym:

(...) są warunkami koniecznymi i wystarczającymi, aby zdanie było prawdziwe: dwa zdania mają te same warunki prawdziwości w przypadku gdy posiadałyby, przy założeniu niezmienności znaczeń, tę samą wartość logiczną we wszystkich możliwych okolicznościach (...) (s. 11).

Innymi słowy – zdaniem Fostera – identyczne warunki prawdziwości powinny gwarantować analityczną równoważność zdań. Należy również zwrócić uwagę, że tak rozumiane warunki prawdziwości w sensie mocnym nie mogą być utożsamiane ze wspomnianymi wcześniej kryteriami prawdziwości zdań⁷⁴. Zdaniem Fostera wynikające z teorii prawdy T-zdania nie są warunkami prawdziwości w sensie mocnym:

(...) T-zdania podają warunki prawdziwości jedynie w tym sensie, że dwa zdania posiadają te same warunki prawdziwości wtedy i tylko wtedy, gdy posiadają te same wartości logiczne (*ibidem*).

Przedstawiona uwaga posłużyła Fosterowi do sformułowania zarzutu, że teoria prawdy faktycznie podaje adekwatną charakterystykę prawdy, ale nie podaje „adekwatnej charakterystyki znaczenia” (s. 13). Aby to zilustrować Foster posłużył się pochodzącym od Tarskiego przykładem semantycznej definicji prawdy, znanym Czytelnikowi z początku rozdziału 1. Jak pamiętamy w przykładzie tym występuje aksjomat spełniania dla dwuargumentowego predykatu \subset :

Ciąg C spełnia formułę „ $x_i \subset x_j$ ” wtw, gdy C^i jest częścią C^j .

Jeżeli weźmiemy predykat \subseteq o tej samej ekstensji co predykat \subset , to powyższy aksjomat możemy zastąpić odpowiednim aksjomatem dla \subseteq bez zmiany wartości logicznych odpowiednich zdań języka przedmiotowego. Takim aksjomatem może być:

Ciąg C spełnia formułę „ $x_i \subseteq x_j$ ” wtw, gdy C^i jest częścią C^j oraz Ziemia się porusza.

⁷⁴ Należy podkreślić, że Davidson również odwołuje się do kryteriów prawdziwości zdań, utożsamiając je ze świadectwami, które pozwalają uznać zdanie za prawdziwe. Tak rozumiane kryteria prawdziwości występują jednak w teorii prawdy, lecz spełniają bardzo ważną rolę w procesie interpretacji. O świadectwach będzie mowa w kolejnym rozdziale.

W konsekwencji z pierwszej teorii wynika T-zdanie:

Zdanie „ $\forall x_i \forall x_j (x_i \subset x_j)$ ” jest prawdziwe wtw, gdy dla każdego a i b , a jest częścią b ,

zaś z drugiej teorii:

Zdanie „ $\forall x_i \forall x_j (x_i \subset x_j)$ ” jest prawdziwe wtw, gdy dla każdego a i b , a jest częścią b oraz Ziemia się porusza.

Zdaniem Fostera obie teorie prawdy podają różne przekłady zdania $\forall x_i \forall x_j (x_i \subset x_j)$, a w konsekwencji, różne znaczenia tego zdania. Ponieważ takich teorii prawdy można zbudować nieograniczenie wiele, żadna z nich nie podaje adekwatnego przekładu, a zatem nie dostarcza adekwatnej charakterystyki znaczenia wyrażen języka przedmiotowego. Zarzut ten – czemu Foster dał wyraz – był nieaktualny w momencie publikacji, gdyż Davidson skorygował swe poglądy z *Prawdy i znaczenia* akcentując proces interpretacji, podlegający ograniczeniom, które blokują możliwość uznania takich teorii prawdy, jak w przykładzie podanym przez Fostera. W *Odpowiedzi Fosterowi* Davidson pisze:

A zatem zaproponowałem kilka empirycznych ograniczeń⁷⁵ na akceptowanie teorii prawdy, które mogą zostać wyrażone bez odwołania się do takich pojęć jak znaczenia, przekład lub synonimiczność, lecz nie bez pewnego rozumienia pojęcia prawdy (Davidson 1984/1976, s. 172).

Ponieważ w przykładzie podanym przez Fostera zdanie po prawej stronie T-zdania nie jest przekładem odpowiedniego zdania języka przedmiotowego, faktycznie nie obowiązuje w nim konwencja T, która ów przekład zakłada⁷⁶. Zdaniem Davidsona zaproponowana przez niego zmiana roli teorii prawdy nie prowadzi do odrzucenia konwencji T, lecz odczytania jej w nowy sposób:

Podobnie jak Tarski pragnę teorii, która spełnia konwencję T, ale podczas gdy on zakłada pojęcie przekładu, aby rzucić światło na pojęcie prawdy, pragnę wyjaśnić pojęcie przekładu zakładając częściowe rozumienie pojęcia prawdy.

Gdy pisałem *Prawdę i znaczenie* zdawałem sobie sprawę z faktu, że do formalnych ograniczeń narzuconych na teorię prawdy, które mogą służyć do interpretacji, należy dodać również ograniczenia empiryczne. Mój błąd nie polegał

⁷⁵ Ogólnym empirycznym ograniczeniem na proces interpretacji jest tzw. Zasada życzliwości, która przejawia się na wiele sposobów. Zasadę tę omówię dokładniej w następnym rozdziale.

⁷⁶ Jak pamiętamy, w semantycznej definicji prawdy Tarski zakładał przekład języka przedmiotowego na metajęzyk, dzięki czemu obowiązywała konwencja T. W przytoczonym przykładzie Foster starał się pokazać, że jeśli potraktujemy teorię prawdy jako teorię znaczenia, to dla teorii prawdy nowej roli konwencja T nie może obowiązywać.

na założeniu, jak zdaje się sugerować Foster, że *każda* teoria, która podaje poprawne warunki prawdziwości może służyć do interpretacji; mój błąd polegał na przeoczeniu faktu, że ktoś może znać właściwą (*sufficiently unique*) teorię prawdy nie wiedząc, że jest to właśnie ta właściwa (*op. cit.*, s. 173).

Owa właściwa (*sufficiently unique*)⁷⁷ teoria jest to taka teoria prawdy, która spełnia konwencję T. Jak pokazał Foster same kryteria formalne nie wystarczają, aby wybrać z pośród konkurencyjnych teorii spełniających owe kryteria tę właściwą:

Foster wskazuje, że chociaż interpretator dysponuje teorią spełniającą konwencję T, to w samej teorii nie znajdzie jakichkolwiek kryteriów, które wskażą na ten fakt (*ibidem*).

Zdaniem Davidsona T-zdania wynikające z teorii prawdy, która spełnia konwencję T, stanowią warunki prawdziwości w mocnym, postulowanym przez Fostera sensie. Aby to uzasadnić, Davidson przytacza analogię do teorii naukowej. Na teorię składają się prawa charakteryzujące się m.in. tym, że uzasadniają kontrfaktyczne okresy warunkowe, np. prawo powszechnego ciężenia uzasadnia „Gdyby to piórko znajdowało się w próżni, to spadałoby ruchem jednostajnie przyspieszonym”. Podniesiony przez Davidsona problem polega na tym, że sam kształt prawa powszechnego ciężenia nie pozwala stwierdzić, czy jest to faktycznie prawo fizyki, czy też jedynie uogólnienie dotychczasowych obserwacji⁷⁸. Różnica jest bardzo istotna, gdyż uogólnienia indukcyjne nie uzasadniają kontrfaktycznych okresów warunkowych. Aby to zilustrować Davidson podaje przykład początkującego fizyka, któremu powiedziano, że dla ciał znajdujących się w próżni masa nie ma wpływu na czas swobodnego spadku. Gdyby poprosić owego adepta, aby wyobraził sobie, że Galileusz spuścił piórko i kulę armatnią z *Empire State Building* oraz, że nie ma atmosfery, a następnie zapytać, który z tych przedmiotów szybciej spadnie na ziemię, to mógłby on odpowiedzieć:

⁷⁷ Angielski termin *sufficiently unique* pozwoliłem sobie przełożyć jako „właściwa”, zdając sobie sprawę z faktu, że przekład ten zubaża treść oryginalnego terminu. Używając terminu *sufficiently unique*, Davidson wskazuje na fakt, że dana teoria może być przyjęta (jako jedyna) na danym etapie interpretacji, gdyż w stopniu wystarczającym spełnia ograniczenia empiryczne narzucone na proces interpretacji, lecz nie wyklucza możliwości, że później zostanie zastąpiona przez inną teorię.

⁷⁸ Podobnego zdania był Kazimierz Ajdukiewicz, formułując dyrektywalną teorię znaczenia. To samo uznawane przez nas zdanie mogło pełnić w teorii (w zależności od przyjętej konwencji) funkcję dyrektywy aksjomatycznej lub zwykłego uogólnienia.

Nie mam pojęcia. Powiedziałeś mi jedynie co się dzieje, gdy rzeczy mają się tak, jak się mają; nie powiedziałeś mi co by się stało, gdyby było inaczej niż jest (s. 174).

Aby wiedzieć czy zdanie ogólne jest jedynie uogólnieniem, czy też prawem, należy odwołać się do zewnętrznych kryteriów. W przypadku teorii naukowych są to zazwyczaj kryteria pragmatyczne, których przyjęcie nadaje uogólnieniom status praw. W przypadku teorii prawdy są to pewne kryteria empiryczne, zgodnie z którymi przebiega proces interpretacji prowadzący do sformułowania odpowiedniej teorii. Wynikające z takiej teorii prawdy warunki prawdziwości (T-zdania) mają wówczas charakter „praw”, czyli dokładnie taki, jaki postulował Foster⁷⁹. Oczywiście T-zdania są warunkami prawdziwości w podanym sensie jedynie w ramach teorii prawdy, z której wynikają.

Przedstawioną przez Davidsona analogię można pociągnąć dalej. Jak wiadomo w procesie rozwoju nauki jedne teorie naukowe bywają zastępowane przez inne – mamy wówczas do czynienia z tzw. rewolucją naukową. Podobnie rzecz się ma z teoriami prawdy dla interpretowanego języka. Na danym etapie procesu interpretacji możemy uznać za właściwą jedną teorię prawdy, ale na kolejnych etapach możemy ją zastąpić zupełnie inną teorią. W toku rozmowy możemy zatem przeżyć wiele „rewolucji” interpretacyjnych.

Chociaż w trzecim rozdziale Foster przyznaje, że przedstawiony zarzut nie dotyczy najnowszych poglądów Davidsona⁸⁰, to ten właśnie – pozornie nieaktualny – zarzut jest szczególnie odpierany w *Reply to Foster* oraz omawiany w literaturze. W tym samym rozdziale Foster wysuwa jednak nowy zarzut pod adresem zmodyfikowanych, tj. wzbogaconych *explicite* o odpowiedni proces interpretacji, poglądów Davidsona. Zarzut ten, skierowany przeciwko Davidsona teorii interpretacji, polega na tym, że pomimo dokonanej korekty, teoria prawdy nadal nie spełnia warunków, jakie powinna spełniać teoria znaczenia. Pomimo tego, że argumenty Fostera są bardzo złożone i opierają się w dużej części na jego własnej rekonstrukcji poglądów Davidsona, warto je w znacznej części przytoczyć, gdyż pozwalają lepiej zrozumieć odpowiedź, jakiej udzielił Davidson w *Reply to Foster*.

⁷⁹ Przedstawiony przez Davidsona kontrargument stawia pod znakiem zapytania podjętą przez Fostera w czwartym rozdziale artykułu próbę zbudowania teorii znaczenia opartej na semantyce światów możliwych. Semantyka ta miała być semantyką warunków prawdziwości w podanym przez niego sensie (zob. Foster 1976, s. 23–32).

⁸⁰ W czasie, gdy Foster pisał swój krytyczny artykuł, poglądy, o których mowa, pojawiły się w niepublikowanych tekstach wygłoszonych przez Davidsona na Uniwersytecie w Oxfordzie w roku 1974.

Punktem wyjścia, podobnie jak poprzednio, jest coś, co Foster nazywa słabością T-zdań, a dokładniej „(...) rozbieżność między słabością T-zdań, a warunków narzuconych na ich konstrukcję” (Foster 1976, s. 17)⁸¹. Owa „słabość” przypisywana przez Fostera T-zdaniom oznacza po prostu, że prawdziwość T-zdania ma „przygodny charakter”, tj. że jest ono prawdziwe, gdy odpowiednie zdanie języka *L* oraz zdanie po prawej stronie mają tę samą wartość logiczną. Jednak, jak twierdzi Foster – przypisując jednocześnie ten pogląd Davidsonowi:

(...) aby równoważność uznać za T-zdanie, czyli za twierdzenie teorii spełniającą konwencję T, konieczne jest dodatkowo, aby zdanie zastępujące *p* [tj. zdanie po prawej stronie równoważności „Zdanie *x* jest prawdziwe wtw, gdy *p*”] było przekładem zdania oznaczonego przez *x* (*ibidem*).

Wprawdzie sama teoria prawdy nie jest jeszcze teorią znaczenia, gdyż jej znajomość – zdaniem Fostera (oraz, jak twierdzi Foster, również zdaniem Davidsona) – nie pozwala jeszcze opanować języka *L*, to jeśli dodatkowo ktoś wie, że ta właśnie teoria jest teorią prawdy (co oznacza wg Fostera, że jest to teoria translacyjna), to owa łączna wiedza wystarczy mu do interpretowania języka *L*. W takim przypadku każde T-zdanie teorii postaci „Zdanie *x* jest prawdziwe wtw, gdy *p*” można odczytać jako „Zdanie *x* stwierdza, że *p*”, z czego Foster wyprowadza następującą konsekwencję:

Mając na myśli owe odczytania [T-zdań] musimy się jedynie cofnąć od T-zdań do aksjomatów [teorii], aby ogarnąć wszystko to, co elementy wnoszą do tego, co stwierdzają *L*-zдания⁸². A ogarnąć to oznacza wiedzieć jak interpretować dowolne *L*-zdanie na podstawie jego struktury (*ibidem*).

W przytoczonym cytacie Foster stara się pokazać, że do zinterpretowania zdań języka *L* nie wystarcza wiedza o tym, co stwierdza teoria prawdy, ale trzeba również wiedzieć, że owa wiedza jest konsekwencją teorii prawdy:

Teza głosząca, że przy założeniu wiedzy, że mamy do czynienia z teorią prawdy, teoria ta służy jako teoria interpretacji jest pierwszą warstwą zmodyfikowanego poglądu Davidsona. Poza nią kryje się nadal druga i starsza warstwa, która

⁸¹ Z tego punktu widzenia przytoczona wcześniej odpowiedź Davidsona może być interpretowana jako unieważnienie drugiego zarzutu.

⁸² Przez *L*-zdanie Foster rozumie po prostu zdanie języka przedmiotowego *L*.

obecnie nabiera nowego znaczenia. Polega ona na empirycznym wyjaśnieniu pojęcia przekładu (s. 17–18)⁸³.

W konsekwencji Foster przypisuje Davidsonowi następujący pogląd:

Wedle nowego poglądu głoszonego przez Davidsona teoria prawdy spełniająca pewne formalne i empiryczne kryteria służy jako teoria znaczenia dla osoby, która wie, że teoria ta owe kryteria spełnia (s. 18).

Aby pokazać błędność poglądu Davidsona Foster proponuje bardzo złożony i raczej sztuczny przykład, w którym dokonywany jest przekład teorii prawdy θ , a dokładniej jej T-zdań „Zdanie x stwierdza, że p ” na inny język:

Załóżmy, że θ' jest dokładnym przeformulowaniem na język francuski wszystkiego, co stwierdza teoria θ . Niech Pierre – Francuz nie znający języka angielskiego – akceptuje θ' na mocy autorytetu swego rodaka François, który jest ekspertem w zakresie językoznawstwa. Jednak François nie mówi Pierrowi, że θ' jest teorią prawdy. A zatem Pierre nie może użyć swojej wiedzy, aby zinterpretować L , chociaż zna fakty, które stwierdza θ' , oraz są to te same fakty, które stwierdza θ . Załóżmy teraz, że Francis, przeceniając możliwości językowe przyjaciela, przepisuje [sformułowaną w języku angielskim] teorię θ , podaje jej reguły inferencji [aksjomaty], informuje Pierra, że one właśnie stanowią teorię prawdy dla L i dodaje formalne kryterium pozwalające rozpoznawać T-zdania. Wyposażony w tę wiedzę Pierre wie, że θ jest teorią prawdy dla L oraz może, wyprowadzając T-zdania, podać angielski przekład dowolnego zdania języka L . A zatem Pierre posiada oba składniki wiedzy, których koniunkcja, jak stwierdzono, wystarcza do opanowania języka – zna fakty stwierdzane przez θ oraz wie, że θ jest teorią prawdy. Jednak nie znając angielskiego oraz nie wiedząc, że θ i θ' stwierdzają te same fakty, nie może wykorzystać własnej wiedzy w rozumieniu L . W tej sytuacji informacja, że θ jest teorią prawdy jest całkowicie zbędna (s. 19).

Przytoczony przykład ma za zadanie pokazać, że można znać fakty, które głosi teoria prawdy, tj. można znać zdania postaci „Zdanie x stwierdza, że p ” (w przykładzie są to zdania w języku francuskim „*x déclare que p*”, gdzie p jest odpowiednim zdaniem w języku francuskim) oraz znać teorię prawdy i dodatkowo wiedzieć, że to jest teoria prawdy (w przykładzie teoria ta sformułowana jest w języku angielskim dla osoby nie znającej tego języka), ale nie móc zinterpretować zdań języka L .

⁸³ W przypisie Foster dodaje, że ów empiryczny opis pojęcia przekładu daje się już dostrzec w *Prawdzie i znaczeniu* (s. 18). Jak pamiętamy owe empiryczne wyjaśnienia odwołują się do Zasady życzliwości.

To, co w tym przykładzie budzi wątpliwości, to założenie, że można przetłumaczyć wszystkie zdania postaci „Zdanie x stwierdza, że p ” na język francuski. Zdania te bowiem są przeformułowanymi T-zdaniami, których może być nieskończenie wiele. Pomińmy jednak to zastrzeżenie i zapytajmy dlaczego same fakty stwierdzane przez teorię nie wystarczają do zrozumienia języka L . Zdaniem Fostera jest tak dlatego, że osoba, która nie wie, że fakty te są konsekwencjami teorii prawdy, nie jest w stanie odtworzyć jej aksjomatów, a w konsekwencji uchwycić znaczenia słów i zwrotów języka L .

Jak się zdaje powyższe stwierdzenie jest w najlepszym wypadku wątpliwe. Wydaje się, że nawet ograniczona lista zdań postaci „zdanie x *déclare que* p ” wystarcza, aby Pierre przynajmniej próbował zrekonstruować stosowną teorię prawdy. Oczywiście Foster mógłby się bronić, iż aby podjąć próbę rekonstrukcji teorii prawdy należy wiedzieć, że ma się do czynienia z teorią prawdy. Oczywiście nie chodzi tu o wiedzę *explicite*, że jest to teoria prawdy, gdyż Pierre – podobnie jak większość śmiertelników – prawdopodobnie nigdy o teorii Davidsona ani o semantycznej definicji prawdy nie słyszał. Chodzi tu raczej o to, że – zdaniem Fostera – samo podanie zdań postaci „Zdanie x *déclare que* p ” nie wystarcza do uruchomienia psychologicznego mechanizmu interpretowania języka, który to mechanizm jest reprezentowany za pomocą teorii prawdy.

Konkluzja podanego przykładu głosi, iż „informacja, że θ jest teorią prawdy jest całkowicie zbędna”. Zgadzając się z tą konkluzją, nie akceptuję natomiast wyprowadzonego przez Fostera stwierdzenia, że teoria prawdy nie jest autentycznie interpretacyjna:

(...) gdyż rozumienie L nie jest wyprowadzone jedynie z tego, co teoria stwierdza. Nowa teza Davidsona jest wadliwa w ten sam sposób. Faktycznie Davidson proponuje nam jako teorię znaczenia kombinację teorii prawdy oraz – przez dołączenie klauzuli stwierdzającej, że ma się z teorią prawdy do czynienia – teorii przekładu ubranej w empirystycznie respektowalne terminy. Przy założeniu, że rozumiemy θ , stwierdzenie, że jest to teoria prawdy, zapewnia podręcznik przekładu z języka L na język, który rozumiemy (...). Jednak koniunkcja zdania θ oraz stwierdzenia, że jest to teoria prawdy, nie jest, jak pokazuje przykład Pierra, adekwatną teorią znaczenia (s. 19–20).

Ostatecznie Foster stwierdza, że zrekonstruowana przez niego nowa teza Davidsona nie jest wcale lepsza od jego poprzedniej tezy, czego przyczyną jest m.in. to, że:

(...) stwierdzenie o byciu teorią prawdy oraz zawartość teorii prawdy nie zależą się (*do not interlock*): nasza wiedza, że θ jest to teorią prawdy nie ma bezpośredniego związku z wiedzą o faktach, które owa teoria stwierdza (s. 20).

Problemowi temu dałoby się – zdaniem Fostera – częściowo zaradzić modyfikując tezę Davidsona do postaci „Pewna teoria prawdy dla języka *L* stwierdza, że...”, gdzie wielokropek zastępuje nam koniunkcję wszystkich zdań wyrażających fakty stwierdzone przez teorię⁸⁴, ale nawet w proponowanej postaci teoria ta nie spełnia wymogów stawianych przed teorią prawdy ze względu na intensjonalność zwrotu „stwierdza, że”. Davidson – jak zauważa Foster – może próbować ratować to sformułowanie, odwołując się do wspomnianej w poprzednim paragrafie analizy parataktycznej i stosując ją do zaproponowanej przez Fostera modyfikacji tezy Davidsona. Analiza taka mogłaby mieć postać:

Pewna teoria prawdy dla języka *L* stwierdza to: (...),

gdzie (...) jest koniunkcją wszystkich zdań wyrażających fakty stwierdzone przez teorię.

Ostatni element argumentu Fostera stanowi próba zakwestionowania analizy parataktycznej. Odwołując się do oryginalnego przykładu z *On Saying That* (1968), gdzie – jak pamiętamy – Davidson proponuje analizę zdania „Galileusz powiedział, że Ziemia się porusza”⁸⁵:

Galileusz wypowiedział zdanie synonimiczne do tego napisu: Ziemia się porusza.

Jeżeli, jak twierdzi Foster, przetłumaczymy zdanie z zaimkiem wskazującym na inny język, to przetłumaczony zaimek powinien wskazywać na ten sam obiekt – w naszym przypadku na napis w języku polskim „Ziemia się porusza”, a nie na napis będący jego tłumaczeniem.

Pozwoliłem sobie przytoczyć (z pewnymi skrótami) długą i zawiłą argumentację Fostera głównie z tego powodu, że składa się na nią wiele uwag, do których musiał ustosunkować się Davidson uściślając tym samym i rozwijając swoje poglądy. Jak pamiętamy drugi argument Fostera opiera się na konstatacji, że warunkiem rozumienia języka *L* jest znajomość faktów stwierdzanych przez teorię prawdy oraz wiedza, że owe fakty z teorii prawdy wynikają. Z poglądem tym Davidson w pełni się zgadza i stwierdza, że zaproponowana przez niego teoria interpretacji spełnia ten wymóg, a zatem zarzuty Fostera są bezzasadne. Nie zgadza się jednak, że z powyższego poglądu wynika, że

⁸⁴ Propozycji Fostera nie przytaczam w szczegółach, gdyż nie ma ona większego znaczenia dla proponowanej przez niego argumentacji (s. 20).

⁸⁵ W tym miejscu bardzo upraszczam analizę tego zdania proponowaną przez Davidsona, gdyż analizie parataktycznej zamierzam poświęcić oddzielne opracowanie. Uważam również, że ten ostatni element argumentu Fostera został przez Davidsona odparty bardzo zdawkowo i wymaga on pogłębionej analizy.

„znajomość języka daje się zredukować do umiejętności przekładu tego języka na inny język” (Davidson 1984/1976, s. 174). Każdy – jak argumentuje Davidson – kto potrafi interpretować język angielski wie, że wypowiedzenie zdania *Snow is white* jest prawdziwe wtedy i tylko wtedy, gdy śnieg jest biały, wie on ponadto, że ten fakt (tj. to właśnie T-zdanie) wynika z teorii prawdy spełniającej konwencję T, nie jest zaś jakimś przypadkowym faktem na temat angielskiego zdania – w konsekwencji osoba ta wie, że ten fakt interpretuje owo zdanie (s. 175). W takim przypadku – na co już wskazywałem – osoba ta może przeformułować owo T-zdanie do odpowiedniego M-zdania:

(...) nie widzę w tym nic groźnego aby przeformułować to, co wie interpretator do formy bardziej intuicyjnej: interpretator wie, że angielskie *Snow is white* znaczy, że śnieg jest biały (*ibidem*).

Oczywiście owa wiedza nie polega jedynie na znajomości faktu wyrażonego przez odpowiednie T-zdanie, ale również na wiedzy, że wynika ono z teorii prawdy spełniającej konwencję T.

Wprawdzie Davidson nazywa wielokrotnie teorię spełniającą konwencję T teorią translacyjną (*translational*), ale nie oznacza to, że wiedza, o której mowa, musi być wyrażona za pomocą konkretnych form językowych. Innymi słowy wiedza, że wypowiedzenie zdania *S* znaczy, że *p*, nie oznacza, że *p* jest przekładem *S*. Aby wyjaśnić to stwierdzenie Davidson odwołuje się do występujących w języku wyrażen okazjonalnych:

W językach naturalnych elementy okazjonalne (*indexical elements*), jak zaimki wskazujące i czas gramatyczny muszą być interpretowane w zależności od okoliczności ich wypowiedzenia. W takim przypadku prawa strona równoważności [T-zdania] nigdy nie jest przekładem zdania, dla którego stanowi warunki prawdziwości (*ibidem*).

Wyobraźmy sobie, że pewna osoba *X* w momencie czasu *T* wypowiada zdanie *I have a headache*. Jest to oczywiście zdanie okazjonalne, a jego wartość logiczna zależy od osoby wypowiadającej to zdanie oraz momentu wypowiedzenia. Przekładem tego zdania na język polski jest oczywiście „Boli mnie głowa”. Jednak odpowiednie T-zdanie wynikające z teorii prawdy obejmującej zdania okazjonalne miałyby postać:

(7.1) Zdanie „*I have a headache*” wypowiedziane przez osobę *X* w momencie czasu *T* jest prawdziwe wtw, gdy osobę *X* w momencie czasu *T* boli głowa.

Aksjomaty teorii prawdy dla zdań okazjonalnych, z których wynika powyższe T-zdanie, muszą uwzględniać kontekst wypowiedzenia, np. dla naszego przykładu aksjomatami takimi byłyby:

Każdy ciąg C przyporządkowuje zaimkowi „I” osobę, która wypowiedziała ten zaimek.

Każdy ciąg C przyporządkowuje wypowiedzeniu zdania w czasie teraźniejszym moment czasu identyczny z momentem czasu wypowiedzenia.

Ciąg C spełnia formułę „*I have a headache*” wypowiedzianą przez osobę X w momencie czasu T wtw, gdy obiekt przyporządkowany przez ciąg C zaimkowi „I” boli głowa w momencie czasu identycznym z momentem czasu przyporządkowanym przez ciąg C wypowiedzeniu tego zdania⁸⁶.

W T-zdaniu (7.1) prawa strona nie jest z pewnością przekładem zdania gramatycznego *I have a headache*, chociaż stanowi jego warunek prawdziwości, a samo T-zdanie może zostać zastąpione przez M-zdanie:

Zdanie *I have a headache* wypowiedziane przez osobę X w momencie czasie T znaczy, że osobę X w momencie czasu T boli głowa.

Oczywiście owo „znaczy, że” nie oznacza bynajmniej – jak to miało miejsce wcześniej – synonimiczności zdania *I have a headache* z prawą stroną równoważności, gdyż nie mamy tutaj do czynienia z przekładem. Użycie tej formy jest jednak – jak stwierdza Davidson – nieszkodliwym przeformułowaniem T-zdania, gdy za pomocą teorii prawdy interpretujemy wypowiedzenia zdań⁸⁷.

⁸⁶ Jedną z funkcji czasu teraźniejszego jest wskazywanie na moment wypowiedzenia, czasu przeszłego na moment poprzedzający moment wypowiedzenia, zaś przyszłego na moment przyszły w stosunku do momentu wypowiedzenia. Zatem dla wypowiedzenia zdania *I had a headache* otrzymalibyśmy aksjomaty:

Ciąg C przyporządkowuje wypowiedzeniu zdania w czasie przeszłym dowolny moment czasu wcześniejszy od momentu czasu wypowiedzenia.

Ciąg C spełnia formułę „*I had a headache*” wypowiedzianą przez osobę X w momencie T wtw, gdy obiekt przyporządkowany przez ciąg C zaimkowi „I” bolała głowa w momencie czasu przyporządkowanemu przez ciąg C wypowiedzeniu tego zdania.

Przedstawiony przeze mnie przykład stanowi jedynie nieprecyzyjnie sformułowany zarys teorii prawdy rozszerzonej na zaimki osobowe oraz czasy gramatyczne. Inne propozycje takiego rozszerzenia można znaleźć w (Lepore i Ludwig 2005 i 2007).

⁸⁷ Zarówno termin „znaczenie”, jak i fraza „znaczy, że” są wieloznaczne. Przegląd znaczeń terminu „znaczenie” można znaleźć w (Ajdukiewicz 1931). Fraza „znaczy, że” w powyższym M-zdaniu nie odpowiada bezpośrednio żadnemu ze znaczeń terminu „znaczenie” zaproponowa-

W tym miejscu warto wspomnieć, że przedstawione wyżej rozwiązanie nawiązuje bardzo wyraźnie do pewnych rozstrzygnięć Kazimierza Twardowskiego oraz Gottloba Fregego, poczynionych jednak w zupełnie innym kontekście. W artykule *O tak zwanych prawdach względnych* (Twardowski 1900) rozróżnia się wypowiedzenia zdań (Twardowski używał terminu „powiedzenia”) od sądów wyrażanych przez zdania. Z powodu okazjonalności, wieloznaczności lub eliptyczności wypowiedzeń, to samo wypowiedzenie może (w różnych) okolicznościach wyrażać różne sądy. Oczywiście Davidson nie akceptowałby Twardowskiego teorii sądu, w której przez sąd rozumie się stwierdzenie istnienia przedmiotu. To co jednak łączy obie koncepcje, to rola rozwiniętej postaci wypowiedzenia. U Davidsona owa rozwinięta postać występuje po prawej stronie odpowiedniego T-zdania i służy do interpretacji danego wypowiedzenia, zaś u Twardowskiego stanowi uzupełnioną formę wypowiedzenia, która odpowiada w sposób jednoznaczny sądowi. Dzięki temu – jak twierdzi Twardowski – nie ma prawd względnych, gdyż zdanie w postaci rozwiniętej jest albo prawdziwe, albo fałszywe. Podobne podejście do zdań okazjonalnych prezentował Frege. W artykule *Myśl. Studium logiczne* (Frege 1918), pojawia się problem czy myśl wyrażona przez zdanie (czyli sens tego zdania) może ulec zmianie w różnych wypowiedzeniach tego zdania. Konkluzja – pomimo pewnych różnic terminologicznych – jest taka sama jak u Twardowskiego: zdanie w formie pełnej, tj. po wyeliminowaniu jawnych lub ukrytych elementów okazjonalnych, wyraża określoną myśl, która jest albo prawdziwa, albo fałszywa.

Podsumowując te uwagi można zauważyć, że postulowana przez Twardowskiego i Fregego pełna forma zdania okazjonalnego lub eliptycznego podaje warunki prawdziwości jego wypowiedzenia. W zależności od okoliczności wypowiedzenia to samo zdanie okazjonalne może być rozwinięte do różnych form pełnych. Pod tym względem stanowisko Davidsona jest całkowicie zgodne ze stanowiskiem Twardowskiego i Fregego. Co więcej, jeśli zdanie w formie pełnej stanowi warunek prawdziwości zdania okazjonalnego⁸⁸, to jest to warunek prawdziwości w sensie Davidsona, nie zaś w sensie podanym przez Fostera. Forma rozwinięta zdania nie pozwala bowiem rozstrzygnąć, czy jego wypowiedzenie jest prawdziwe czy fałszywe, lecz jedynie gwarantuje, że wypowiedzenie zdania eliptycznego lub okazjonalnego w pewnych okolicznościach jest albo bezwzględnie prawdziwe, albo bez-

nemu przez Ajdukiewicza. Być może sens tej frazy można oddać przez „są analitycznie równoważne”. Ową analityczną równoważność zapewnia teoria prawdy, z której wynika (7.1).

⁸⁸ Zdania okazjonalne traktowane są jako eliptyczne formy nieskończonego wielu zdań, gdyż elementy okazjonalne zastępują różne wyrażenia (np. nazwy osób, miejsc i momentów czasu), które pojawiają się w zdaniach dopiero w ich formie rozwiniętej.

względnie fałszywe. Z takim poglądem Davidson bez wątpienia by się zgodził. Różnica między Davidsonem a Fregem i Twardowskim polega na tym, że zdanie w formie rozwiniętej, o ile występuje po prawej stronie T-zdania wynikającego z teorii prawdy, podaje – zdaniem Davidsona – znaczenie zdania okazjonalnego lub eliptycznego po lewej stronie w ramach całej teorii prawdy, natomiast wedle Fregego i Twardowskiego jednoznacznie reprezentuje myśl lub sąd. Z tym stanowiskiem Davidson z pewnością by się nie zgodził, gdyż myśli i sądy były dla niego jedynie hipostazami – zreifikowanymi znaczeniami zdań.

Kolejny element zarzutu Fostera polegał na tym, że w podanej przez niego rekonstrukcji poglądów Davidsona istotną rolę pełni nieekstensjonalny zwrot „teoria prawdy stwierdza, że...” (*T-theory states that...*). Zdaniem Davidsona zarzut ten oparty jest na nieporozumieniu. W procesie interpretacji użytkownik języka odwołuje się wprawdzie do nieredukowalnych pojęć przekonania (*belief*) oraz zamiaru (*intention*), ale jego właściwym celem było „(...) wyjaśnienie znaczenia (interpretacji) bez odwoływania się do niewyjaśnionych pojęć z zakresu językoznawstwa (*linguistic concepts*) (s. 176)”.

W związku z powyższym cytatem konieczne są dwa wyjaśnienia. Po pierwsze, Tarski pragnął zdefiniować pojęcie prawdy bez odwołania się do jakichkolwiek niewyjaśnionych pojęć semantycznych. Podobnie Davidson pragnął podać teorię znaczenia, w której – co brzmi nieco paradoksalnie – nie czyni się żadnego istotnego użytku z takich pojęć jak znaczenie czy stwierdzenie. Teoria prawdy ma nam wyjaśnić czym jest znaczenie bez odwoływania się w tym wyjaśnieniu do przedteoretycznego pojęcia znaczenia lub innych nieekstensjonalnych pojęć wykorzystywanych w popularnej analizie języka (takich jak stwierdzenie). Zupełnie czymś innym jest odwołanie się do pojęcia przekonania i zamiaru, gdyż przypisywanie przekonań i zamiarów pełni niezwykle istotną rolę w procesie interpretowania słów drugiego człowieka. Rezultatem tego procesu jest interpretacja, która przybiera postać teorii prawdy podającej w systematyczny sposób znaczenia interpretowanych słów oraz zdań. Warto mieć na uwadze to rozróżnienie, choćby z tego względu, że termin „interpretacja” może być również użyty w sensie „proces interpretowania”. Niestety ani język polski, ani język angielski nie rozróżniają zbyt precyzyjnie nazw procesów lub czynności od ich wytworów lub rezultatów⁸⁹. W przypadku lektury prac Davidsona może to prowadzić do pewnych nieporozumień.

W teorii prawdy (czyli teorii znaczenia) zawarta jest zatem *implicite* pewna wiedza o przekonaniach i zamiarach rozmówcy, która była wykorzystana

⁸⁹ W dalszym ciągu termin „interpretacja” będę raczej odnosił do procesu interpretowania.

w procesie interpretowania (czego reprezentacją jest dochodzenie do sformułowania teorii prawdy). Oczywiście znajomość samej teorii prawdy jako wytworu tego procesu nie polega na aktualnym posiadaniu owej wiedzy. Jak się zdaje, większość zarzutów Fostera da się sprowadzić do tego, że jego zdaniem znajomość teorii znaczenia powinna obejmować aktualne posiadanie owej wiedzy. Aby to wyjaśnić odwołajmy się do wprowadzonego przez Davidsona pojęcia „bezwładności znaczenia”. Jeżeli dysponujemy teorią znaczenia dla idiolektu drugiego człowieka, to rozumienie jego wypowiedzi wcale nie wymaga znajomości jego rzeczywistych zamiarów ani przekonań w chwili, gdy zostało wypowiedziane konkretne zdanie. Jest to szczególnie wyraźne w przypadku wypowiedzi aktora na scenie. Jego przekonania i zamiary są zupełnie nieistotne dla rozumienia jego słów. W szczególności teoria prawdy niczego nam w przypadku jego wypowiedzi nie stwierdza. Aby sformułować teorię prawdy należało w procesie interpretacji, rozumianym bardzo szeroko jako proces uczenia się języka, wielokrotnie przypisać różnym osobom przekonania i zamiary, które wyrażane bywały w postaci stwierdzeń. Bezwładność znaczenia jest wymogiem natury zarówno pragmatycznej, jak i normatywnej, który zaleca nam, aby nie rewidować przyjętej teorii znaczenia bez bardzo istotnych powodów.

Na zakończenie *Odpowiedzi Fosterowi* Davidson porusza dwie kwestie. Pierwsza dotyczy zarzutu, że analiza parataktyczna nie przechodzi pomyślnie testu przekładu. Druga kwestia dotyczy podejrzenia (wyartykułowanego przez samego Davidsona), że nieekstensjonalny funktor „stwierdza, że” odwołuje się do ukrytych pojęć lingwistycznych.

W pierwszym przypadku Davidson stanowczo deklaruje, że analiza parataktyczna stanowi trzon jego programu semantycznego. Bez umożliwianej przez nią redukcji wypowiedzenia zdania nieekstensjonalnego do wypowiedzenia pary zdań ekstensjonalnych, w którym pierwsze zdanie zawiera element okazjonalny odnoszący się do drugiego, nie jest możliwe zbudowanie teorii języka naturalnego. Zarzut Fostera polegał na spostrzeżeniu, że przekład zaimka wskazującego powinien zachować odniesienie. Na przykład mówiąc o kimś, kto jest obecny w naszym otoczeniu, „On jest filozofem” zaimek „on” odnosi się do tej właśnie osoby. Przekładając to zdanie na język angielski (np. ze względu na obecnego tam obcokrajowca) wypowiemy drugie zdanie *He is a philosopher*. W przypadku drugiego wypowiedzenia, zaimek *he* również będzie odnosił się do tej osoby. Tak rozumiane zachowanie odniesienia zdaje się nie zachodzić w przypadku analizy parataktycznej, gdyż zaimek wskazujący w wypowiedzeniu pary zdań w języku oryginalnym wskazuje na drugi element tej pary, zaś w przypadku przekładu na obcy język, na drugi elementu pary będącej tłumaczeniem. Odniesienie ulega zatem zmianie.

Zarzut ten nie jest – zdaniem Davidsona – poważny, gdyż odwołując się do przykładu analizy parataktycznej zdania „Galileusz powiedział, że Ziemia się porusza”, Foster zakłada, że zachowanie odniesienia wyrażenia okazjonalnego jest warunkiem tego, że dwa wypowiedzenia stwierdzają ten sam fakt lub sąd. W takim jednak przypadku można bez problemu zastosować doskonale znany argument:

(...) jeżeli dwa wypowiedzenia stwierdzają ten sam fakt, gdy zachowane jest odniesienie, trudno jest obronić się przed znanym dowodem, że wszystkie zdania prawdziwe stwierdzają ten sam fakt (s. 177).

Dowodem tym jest oczywiście przytaczany w tym rozdziale argument procy, zaś jego konsekwencja uderza w samo założenie czynione przez Fostera. Wedle tego założenia przekład wypowiedzenia na inny język stwierdzałby inny fakt niż oryginalne wypowiedzenie. Co więcej – jak zauważa Davidson – zupełnie inny fakt stwierdzałoby powtórzenie oryginalnego wypowiedzenia w tym samym języku, w konsekwencji cała sytuacja nie różniłaby się zasadniczo od przekładu. A zatem niezależnie od powołania się *implicite* na Argument procy, nieakceptowaną konsekwencją przyjętego przez Fostera założenia jest również to, że każde wypowiedzenie tego samego zdania stwierdzałoby właściwy sobie i unikalny fakt.

Jak wspomniałem nieco wcześniej, Davidson wykorzystał dokonaną przez Fostera rekonstrukcję własnych poglądów, aby przetestować czy teoria prawdy nie zawierała ukrytych pojęć lingwistycznych. Jednocześnie Davidson wykorzystuje zarzut Fostera do pokazania, że owym ukrytym pojęciem lingwistycznym jest pojęcie znaczenia. Przypomnijmy, że wedle Fostera teoria prawdy stwierdza, że (*states that*) zachodzą fakty opisane za pomocą T-zdań. Davidson proponuje, aby zamiast mówić, że teoria stwierdza, należałoby raczej mówić, że teoria pociąga za sobą (*entails*) T-zdania. Zdaniem Davidsona wypowiadając słowa „Teoria prawdy pociąga za sobą, że zdanie *Śnieg jest biały* jest prawdziwe wtw, gdy śnieg jest biały”, być może odwołujemy się *implicite* do jakiś pojęć lingwistycznych (*linguistic notions*). Pojęcie pociągania za sobą (*entailment*) jest bowiem relacją między teorią, a zdaniem wypowiedzianym przez tego, kto stwierdza, że owo „pociąganie za sobą” zachodzi. Pojęciem tego typu może być – zdaniem Davidsona – pojęcie synonimiczności wypowiedzeń, gdyż:

Jeżeli teoria prawdy pociąga za sobą, że [zdanie] „Śnieg jest biały” w języku angielskim jest prawdziwe wtw, gdy śnieg jest biały, wówczas logiczną konsekwencją teorii jest zdanie synonimiczne z moim wypowiedzeniem „*Śnieg jest biały* w języku angielskim jest prawdziwe wtw, gdy śnieg jest biały” (s. 178).

Zdaniem Davidsona, ponieważ synonimiczność rozumiana jest zazwyczaj jako identyczność znaczeń, istnieje podejrzenie, że teoria prawdy podaje znaczenia wyrażen języka. Nie mamy tu jednak do czynienia z błędnym kołem, gdyż tak rozumiane pojęcie synonimiczności jest faktycznie tożsame z pojęciem przekładu, który staraliśmy się uzyskać nakładając na teorię prawdy pewne warunki. Warunki te były nakładane na proces interpretowania prowadzącego do sformułowania teorii prawdy:

A zatem pojęcie synonimiczności lub przekładu, które jest ukryte w pojęciu pociągania za sobą, może być użyte bez groźby błędnego koła, gdy przedstawimy to, co wie interpretator. Faktycznie, przypisując interpretatorowi teorię translacyjną, czynimy już to założenie (*ibidem*).

Nie chodzi tu oczywiście o przekład wypowiedzeń zdań okazjonalnych, gdyż jak wiemy teoria prawdy nie podaje takiego przekładu. Davidson odwołuje się tutaj do intuicyjnie rozumianej synonimiczności słów dwóch języków oraz ich przekładów, które teoria prawdy niewątpliwie podaje. Można zatem powiedzieć, że teoria prawdy jest oparta na słabszym pojęciu synonimiczności i przekładu, które nie obejmuje wypowiedzeń zdań okazjonalnych. Jest zatem teorią translacyjną. Aby taka teoria mogła zostać sformułowana, należy narzucić na proces interpretowania słów drugiego człowieka pewne warunki, które opiszę dokładniej w następnym rozdziale.

W ostatnim rozdziale pracy przedstawię różne aspekty teorii interpretacji, która jest naturalnym rozwinięciem idei zawartych we wczesnych publikacjach Davidsona, dotyczących teorii znaczenia. Przejście od teorii znaczenia do teorii interpretacji nie było jednak rezultatem gruntownej zmiany poglądów, lecz wiązało się raczej z przesunięciem akcentów, będącym rezultatem dyskusji, które wywołała propozycja Davidsona.

ROZDZIAŁ 4 TEORIA INTERPRETACJI

Teoria prawdy, bez względu na to jak dobrze
została wybrana, nie jest teorią znaczenia (...)
(Davidson 1984/1976, s. 179).

W świetle dotychczasowych ustaleń przytoczony wyżej cytat brzmi paradoksalnie. Ilustruje on jednak wspomniane na końcu poprzedniego rozdziału przesunięcie akcentów, które dokonało się w publikacjach Donalda Davidsona mniej więcej od połowy lat 70. Teoria prawdy będzie nadal nazywana teorią znaczenia, ale z pewnymi zastrzeżeniami. Jak pamiętamy nie jest to teoria, w której termin „znaczenie” pojawia się *explicite* i wskazuje jakim bytem jest znaczenie określonego wyrażenia języka. Teoriami takimi były z kolei tradycyjne teorie znaczenia, np. teoria ideacyjna (Arystoteles oraz empiryści angielscy) lub teoria konotacyjna (Mill), w których zakładano *implicite* istnienie języka jako abstrakcyjnego tworu, pełniącego zarówno funkcję normatywną, jak i umożliwiającego komunikację językową. Wyrażeniom tak pojętego języka teorie znaczenia przypisywały znaczenie (sensy) niezależnie od kontekstu użycia. Z tego punktu widzenia komunikacja językowa była możliwa, gdyż nasze indywidualne idiolekty zbliżały się do owego wyidealizowanego wzorca, zaś wszelkie zakłócenia komunikacji brały się z odstępstw od owego wzorca.

Wprawdzie w *Prawdzie i znaczeniu* Davidson nazywał semantyczną definicję prawdy teorią znaczenia bez znaczeń, ale w swych wczesnych pracach twierdził, że semantyczna definicja prawdy podaje znaczenia wyrażen języka, co mogło niekiedy sugerować, że akceptował istnienie języka w sensie tradycyjnym. Jednak w późniejszych pracach Davidson zrewidował swój sposób mówienia o semantycznej definicji prawdy, kładąc akcent na fakt, że jest to teoria, która przede wszystkim służy do interpretowania wypowiedzi niepowtarzalnego i dynamicznie zmieniającego się idiolektu drugiego

człowieka, nie zaś wypowiedzeń należących do abstrakcyjnego tworu zwanego językiem. A zatem w ścisłym tego słowa znaczeniu teoria prawdy stanowi bardzo istotny element teorii interpretacji indywidualnych wypowiedzeń rozmówcy.

Podawanie znaczenia danego wyrażenia wiąże się zawsze z podaniem jego synonimu. Jest tak nawet w przypadku teorii przypisujących wyrażeniom określone byty jako znaczenia. Na przykład w teorii konotacyjnej Milla znaczeniem nazwy „człowiek” jest jego konotacja, na którą składają się dwie cechy: rozumność i ludzki kształt. W konsekwencji synonimem nazwy „człowiek” jest „istota, która jest rozumna i posiada ludzki kształt”. Można to potraktować jako przekład z języka naturalnego na specyficzny język, w którym formułowana jest konotacyjna teoria znaczenia. W przypadku teorii prawdy w ogólnym przypadku nie mamy nawet do czynienia z przekładem, gdyż dla zdań okazjonalnych teoria ta nie podaje nam ich synonimów. Trudno zatem mówić, aby podawała znaczenia tych zdań niezależnie od kontekstu ich użycia. Niemniej jednak teoria prawdy podaje znaczenia wypowiedzeń wyrazów, rozumiane jako wkład wyrazów do warunków prawdziwości wypowiedzeń zdań, w których występują. Co więcej, w ostatecznie przyjętej teorii prawdy zawarte są *implicite* postawy propozycyjalne przypisywane mówiącemu, gdyż przypisanie postaw stanowi warunek zbudowania teorii prawdy. Mimo przedstawionych zastrzeżeń, tam gdzie nie będzie prowadziło to do nieporozumień, nadal używał będę terminu „język” oraz pisał, że teoria prawdy podaje znaczenia, zdając sobie jednak sprawę, że zinterpretowanie wypowiedzenia wyrażenia to faktycznie coś innego niż podanie znaczenia tego wyrażenia.

W rozdziale tym przyjrzymy się, jakie warunki muszą być spełnione, aby teoria prawdy budowana dla danego idiolektu faktycznie interpretowała wypowiedzenia tego idiolektu, co stanowi rozwinięcie odpowiedzi, jakiej udzielił Davidson na zarzuty Fostera. W tym celu przyjrzymy się różnym aspektom procesu interpretacji, z których część została już zasygnalizowana w poprzednim rozdziale. Warto również jeszcze raz podkreślić, iż w tradycyjnym rozumieniu semantyka podaje znaczenia wyrażen języka rozumianego jako pewien abstrakt i niezależnie od kontekstu użycia jego wypowiedzeń. W przypadku interpretacji mamy do czynienia z wypowiedzeniami, które miały miejsce w określonym kontekście czasowym i sytuacyjnym. To jedynie bezwładność znaczenia, o której była mowa w poprzednim rozdziale, wywołuje wrażenie, że słowa posiadają znaczenia niezależnie od kontekstu.

Interpretując wypowiedzenia drugiego człowieka zawsze odwołujemy się do kontekstu tych wypowiedzeń. Z tego punktu widzenia teoria interpretacji powinna być traktowana jako teoria pragmatyczna, gdyż przez pragma-

tykę rozumie się zazwyczaj naukę o użyciu wyrażen językowych. Widać to wyraźnie u Grice'a, który dokonywał rozróżnienia między znaczeniem konwencjonalnym wyrażenia (*conventional meaning*), a tym, co mówiący w danym kontekście miał na myśli (*speakers meaning*). Pierwsze należałoby do semantyki, zaś drugie do pragmatyki. W przypadku Davidsonowskiej teorii interpretacji takie rozumienie pragmatyki zawodzi i teoria ta zaliczana jest przez niego do semantyki¹. To niestandardowe potraktowanie semantyki jest konsekwencją jego poglądów na interpretację, w której rezygnuje się *de facto* z operowania znaczeniem przypisanym do wyrażen językowych niezależnie od okoliczności ich użycia. Tradycyjnie rozumiana semantyka, podobnie jak tradycyjnie rozumiany język, straciła u Davidsona rację bytu.

Od tej chwili, wzorem późniejszych prac Davidsona, zamiast terminu „język” będę – o ile nie będzie to prowadziło do niezręczności – stosował termin „idiolekt”. Za terminem tym kryje się deklarowany przez Davidsona antykonwencjonalizm, zgodnie z którym odrzuca on tradycyjne pojęcie języka jako abstrakcyjnego tworu konstytuowanego przez konwencje. Nie oznacza to jednak, że Davidson całkowicie zarzucił używania terminu „język” – rozumiał go jednak w inny od tradycyjnego sposób jako klasę podobnych idiolektów². Mówiąc o idiolektcie danej osoby nie będę miał na myśli jej indywidualnego języka stanowiącego osobniczy wariant języka etnicznego, lecz ciągle ewoluujący zespół aktualnych lub potencjalnych wypowiedzi w danym miejscu i danym czasie, gdyż wypowiedzenia danej osoby nagrane lub spisane przed laty wymagają interpretacji nawet przez nią samą.

Podobieństwo idiolektów wchodzących w skład języka rozumianego w sensie Davidsona przejawia się na trzy sposoby. Po pierwsze, podobieństwem lub identycznością kształtów fizycznych poszczególnych słów. Po drugie, podobieństwem lub identycznością reguł gramatycznych. Po trzecie wreszcie, identycznością znaczeń przeważającej ilości podobnych lub identycznych co do fizycznego kształtu słów wchodzących do zasobów różnych idiolektów oraz identycznością znaczeń odpowiednich wyrażen złożonych. Gdybyśmy jednak chcieli wyeliminować z tego stwierdzenia termin „znaczenie” należałoby napisać, że trzeci przejaw podobieństwa idiolektów polega na

¹ Przez pragmatykę Davidson rozumie badanie przyczynowego wpływu zinterpretowanych uprzednio wypowiedzi na słuchacza. Jest to szczególnie wyraźne w przypadku jego przyczynowej teorii metafory. Oczywiście wpływ ten zależy również od kontekstu użycia.

² W tym miejscu można się posłużyć analogią do znanego sporu między realistami a nominalistami. Tradycyjne, „realistyczne” rozumienie języka przypisuje mu status idealnego wzorca dla indywidualnych idiolektów, które w większym lub mniejszym stopniu się do niego zbliżają. W podejściu nominalistycznym, reprezentowanym przez Davidsona, językiem nazywa się po prostu zbiór podobnych do siebie idiolektów.

identyczności (przeważającej części) aksjomatów różnych teorii prawdy budowanych dla owych idiolektów w tym samym metajęzyku. Z kolei teorie prawdy budowane w jednym z idiolektów języka dla drugiego idiolektu tego samego języka miałyby trywialny lub prawie trywialny kształt, gdyż wszystkie – lub niemal wszystkie – wynikające z nich T-zdania mają postać:

Zdanie „*p*” jest prawdziwe wtedy i tylko wtedy, gdy *p*.

Przedstawione powyżej uwagi mają jedynie charakter wstępny i zostaną rozwinięte w dalszej części rozdziału, gdzie omówię kilka najistotniejszych aspektów teorii interpretacji. Ponieważ przedstawiana problematyka uwikłana jest w znacznym stopniu w inne obszary systemu filozoficznego Davidsona, będę w większym niż dotychczas stopniu musiał nawiązywać do jego poglądów wykraczających poza semantykę, a w szczególności do epistemologii. Nie ma w tym jednak nic dziwnego, gdyż w przypadku teorii interpretacji liczy się proces budowania teorii prawdy, który w dużej mierze polega na interakcji interpretatora z innymi ludźmi oraz otoczeniem fizycznym, co nieodłącznie wiąże się z poznaniem.

1. PERSPEKTYWA TRZECIOOSOBOWA

Teoria prawdy budowana w moim własnym idiolektcie dla mojego własnego idiolektu ma zupełnie inny status od teorii prawdy, którą formułuję we własnym idiolektcie dla idiolektu innej osoby. Pierwsza z nich jest trywialna, a owa trywialność przejawia się tym, że aksjomaty teorii prawdy oraz wynikające z niej T-zdania mają status zdań przyjętych *a priori*. W przeciwieństwie do tego, teoria prawdy budowana dla idiolektu drugiego człowieka nie ma już trywialnego charakteru w powyższym sensie, nawet jeśli wszystkie aksjomaty dla teorii prawdy jego idiolektu będą identyczne z trywialnymi aksjomatami teorii prawdy dla naszego własnego idiolektu. Taka sytuacja oznacza jedynie, że budując ową teorię nie stwierdziłem do tej pory żadnej różnicy znaczeniowej między idiolektem osoby interpretowanej a moim własnym idiolektem. Teorię tę można jednak uznać za trywialną w innym sensie – ze względu na trywialny kształt aksjomatów.

Mimo że obie teorie prawdy nie będą się różniły pod względem kształtu aksjomatów oraz wynikających z nich T-zdań, to różnica między nimi jest spowodowana odmiennym charakterem aksjomatów, a w konsekwencji odpowiednich T-zdań. W przypadku drugiej teorii trywialny kształt aksjomatów oraz T-zdań nie jest przyjmowany *a priori*, lecz jest rezultatem nietrywialnego procesu interpretacji. W przypadku T-zdań różnicę tę można

wskazać podając w nawiasie, kto wypowiedział dane zdanie. I tak z trywialnej teorii prawdy mojego własnego idiolektu wynika m.in. T-zdanie:

(Moje wypowiedzenie) „Śnieg jest biały” jest prawdziwe wtw, gdy śnieg jest biały,

zaś z równokształtnej, lecz nietrywialnej teorii prawdy dla idiolektu X-a:

(Wypowiedzenie X-a) „Śnieg jest biały” jest prawdziwe wtw, gdy śnieg jest biały.

Celem tego paragrafu jest między innymi wyjaśnienie różnicy charakteru obu równokształtnych teorii odwołując się do różnicy między perspektywą pierwszoosobową a perspektywą trzeciosobową. Perspektywa pierwszoosobowa może być rozumiana jako uprzywilejowana i bezpośrednia pozycja, jaką mamy w dostępie do znaczeń wyrażeń naszego własnego idiolektu oraz do naszych własnych przekonań oraz innych postaw propozycyjalnych. Owa wyjątkowa pozycja polega, wyrażając się niezbyt ściśle, na całkowitej zgodności znaczeń, które przypisuję własnym wypowiedzeniom zdań, z postawami propozycyjnymi (szczególnie przekonaniem), które są aktualizowane przez odpowiednie wypowiedzenia. Jeśli jestem przekonany, że Ziemia jest okrągła, to wypowiedziane przeze mnie asercja „Ziemia jest okrągła” aktualizuje – przy znaczeniu, które przypisuję występującym w nim słowom – owo przekonanie. Fenomen ten Davidson próbował wyjaśnić w artykule *First Person Authority* (1984a), gdzie zwraca uwagę, że pozycja pierwszoosobowa posiada dwa, nierozzerwalnie związane ze sobą aspekty – językowy oraz epistemologiczny. W przypadku słów wypowiedzianych przeze mnie w moim własnym idiolekcie nie muszą przypisywać im znaczeń na drodze interpretacji, gdyż słowa te po prostu rozumiem i potrafię ich użyć. Słowa te – jeśli wypowiedziane szczerze – aktualizują moje przekonania oraz inne postawy propozycyjne. Na ten temat Davidson miał stosunkowo niewiele do powiedzenia, przede wszystkim zaś nie wyjaśnił szczegółowo na czym polega przypisywanie znaczeń wyrażeniom własnego idiolektu, co przejawia się specyficznym rozumieniem i użyciem owych wyrażeń. Zamiast tego Davidson skupił się na „faktach epistemologicznych”, tj. na szczególnej pozycji pierwszej osoby w przypisywaniu samej sobie przekonań.

Oczywiście użycie wyrażeń własnego idiolektu przy aktualizowaniu własnych postaw nie musi być całkowicie zgodne z konwencją językową obowiązującą dla danej wspólnoty językowej. Każdy z nas mówi to, co chce powiedzieć oraz wyraża to, co chce wyrazić, i czyni to we własnym idiolekcie, co stanowi wyzwanie dla interpretatora.

Aby zinterpretować wypowiedzenia drugiego człowieka interpretator musi zbudować teorię prawdy tak, jak buduje się teorię empiryczną. Jeżeli jest to idiolekt użytkownika naszej wspólnoty językowej zakładamy początkowo najprostszą teorię z aksjomatami o trywialnym kształcie. Zakładamy zatem, że nasz rozmówca aktualizuje swoje przekonania za pomocą tych samych wypowiedzeń, za pomocą których sami byśmy aktualizowali te same postawy. Jednak w takim przypadku jesteśmy nieustannie gotowi weryfikować teorię prawdy. Jest tak, gdyż rozmawiając z drugim człowiekiem nieustannie zestawiamy przypisywane mu przekonania oraz inne postawy propozycyjalne z naszymi własnymi. Owa gotowość do weryfikowania teorii prawdy aktualizowana jest wtedy, gdy podczas rozmowy musimy przypisać rozmówcy przekonania, których sami nie podzielamy. Nie chodzi tu oczywiście o wszystkie możliwe przekonania – wielokrotnie różnimy się w sprawie tego, czy jakieś fakty mają lub miały miejsce³. Jednak istnieje cała grupa przekonań, o których rozsądnie zakładamy, że musimy je dzielić z naszym interlokutorem. Chodzi tu przede wszystkim o tzw. przekonania empiryczne oraz przekonania, które można określić niezbyt ściśle jako analityczne lub jako postulaty znaczeniowe. Pierwsze z nich aktualizowane są uznaniem zdań typu „Oto królik” w obecności odpowiednich bodźców, drugie – uznaniem odpowiednich zdań bez względu na okoliczności (bodźce), np. „Wszystkie króliki są zwierzętami”. Stwierdzenie różnicy w takich przekonaniach skłania nas zazwyczaj, o ile nie podejrzewamy kłamstwa, pomyłki lub żartu, do zweryfikowania teorii prawdy.

Pojawia się zatem problem, jak przypisujemy postawy sobie i innym. Jest rzeczą oczywistą, że istnieje wyraźna asymetria w przypisywaniu postaw propozycyjalnych samemu sobie i innym osobom. W przypadku naszych własnych postaw i przekonań czynimy to bez jakichkolwiek świadectw⁴ – po prostu o czymś jesteśmy przekonani, w coś wierzymy, w coś wątpimy, coś lub kogoś oceniamy dobrze albo źle, coś wolimy od czegoś innego. Nie oznacza to oczywiście, że za każdym razem potrafimy od razu i bez zastanowienia wyartykułować własne postawy – bywa, że mamy wątpliwości. Jednak odwoływanie się do jakichkolwiek świadectw w żaden sposób w tym nam nie pomoże. Oczywiście można tu podnieść zarzut, że rozmowa z inną osobą, np. z psychologiem, prowadzi do tego, że na podstawie świadectw dostępnych

³ W tym miejscu, podobnie jak Davidson, odwołuję się jedynie do przekonań. W przypadku innych postaw propozycyjalnych, np. ocen i preferencji, tolerowane różnice mogą być o wiele wyraźniejsze.

⁴ Oczywiście świadectw na rzecz przypisania komuś przekonań nie należy mylić z przyczynami tych przekonań, chociaż w przypadku przekonań empirycznych przyczyny przekonań mogą być jednocześnie świadectwami na rzecz ich przypisania drugiemu człowiekowi.

rozmówcy sami dowiadujemy się o własnych, często ukrytych przekonaniach, pragnieniach, wartościach i preferencjach. Jednak w takim przypadku nabyta wiedza o własnych postawach zapośredniczona jest w perspektywie trzeciosobowej, czy też sami siebie traktujemy trzeciosobowo.

Perspektywa trzeciosobowa pozwala nam interpretować rozmówcę, zaś my sami jesteśmy również interpretowani z tej perspektywy przez naszych rozmówców. Sformułowanie to zostało tu użyte celowo, gdyż interpretowanie wykracza poza interpretowanie wypowiedzi i jest interpretowaniem zachowania werbalnego i pozawerbalnego. Polega ono na dwóch nierozłącznych czynnościach: na przypisywaniu interpretowanej osobie postaw propozycyjalnych na podstawie pewnych świadectw, którymi mogą być działania, zachowania, gesty, grymasy twarzy, ale przede wszystkim wypowiedzenia, czyli świadectwa werbalne, oraz na budowaniu teorii prawdy dla jego wypowiedzi, co można – przy podanych wcześniej zastrzeżeniach – nazwać przypisywaniem znaczeń jego wypowiedziom. Obie czynności są nierozłączne – przypisanie znaczeń w dużej mierze zależy od okoliczności wypowiedzenia, czyli od świadectw empirycznych towarzyszących wypowiedzeniu wyrażającemu przekonanie lub inną postawę żywioną przez mówiącego⁵. Jednak warunkiem przypisania mówiącemu stosownego przekonania jest uznanie, że mówiący wypowiedział swe słowa szczerze, nie pomylił się, mówił dosłownie oraz jest istotą myślącą nie zaś automatem itp. Davidson wielokrotnie podkreślał, że interpretować jest bardzo łatwo, lecz opisać szczegółowo proces interpretacji jest niezwykle trudno.

Interpretowanie nie jest aktem jednorazowym. Jest raczej pewnym procesem opartym na świadectwach empirycznych dostępnych w obliczu kolejnych wypowiedzi naszego interlokutora oraz na jego odpowiedziach na pytania, które do niego kierujemy. W procesie tym stawiamy hipotezy zarówno co do znaczeń słów używanych przez naszego rozmówcę, jak i postaw przez niego żywionych. Podany tu szkic opisu procesu interpretacji jest bardzo ogólnikowy i w kolejnych paragrafach postaram się rzucić na niego więcej światła. Oczywiście sama interpretacja przebiega zazwyczaj niezwykle sprawnie. Umiejętność interpretowania słów drugiego człowieka jest bowiem dyspozycją, którą nabywamy ucząc się języka⁶. Davidson uważał, że szczegó-

⁵ W tym miejscu warto zauważyć, że w opisie procesu interpretacji szczególną rolę pełni wspomniana w poprzednim rozdziale analiza parataktyczna zdań psychologicznych. Przypisanie komuś postawy propozycyjalnej polega bowiem na rozpoznaniu typu tej postawy (przekonanie, wątpliwość, życzenie, preferencja itp.) oraz jej zawartości.

⁶ Nie jest to sytuacja wyjątkowa. Umiejętność chodzenia i utrzymywania równowagi jest również nabywana w początkowym okresie naszego życia i podobnie jak interpretowanie, dla większości osób chodzenie nie stanowi większego problemu. Jednak dokładne i niebanalne

łowy opis procesu interpretacji musiałby być dokonany w kategoriach psychologicznych lub neurofizjologicznych i jako taki nie należy do filozofii języka, która musi zadowolić się podaniem ogólnych warunków umożliwiających zachodzenie tego procesu.

Pomijając problem rozpoznawania prawdomówności rozmówcy oraz pewne problemy pochodne (rozpoznawanie samooszukiwania się mówiącego, hipokryzji, mówienia metaforycznego i ironicznego), interpretując wypowiedzi drugiego człowieka mamy do czynienia z dwoma zależnymi od siebie i nierozłącznymi czynnościami – procesem przypisywania znaczeń oraz procesem przypisywania przekonań. Wszelka rozmowa, wszelki kontakt werbalny z innym człowiekiem jest – wedle Davidsona – procesem wzajemnej interpretacji. Oczywiście nie należy zapominać o zasadzie bezwładności znaczenia, która nakazuje nam akceptować przyjętą teorię prawdy, dopóki nie zajdą okoliczności, które zmuszą nas do jej rewizji.

Aby uściślić poruszone kwestie powróćmy do eseju *First Person Authority*, w którym Davidson rozróżnia dwie asymetrie zachodzące w stosunku perspektywy pierwszoosobowej do perspektywy trzecioosobowej. Pierwsza asymetria zachodzi między przypisaniem postawy propozycjonalnej samemu sobie przeze mnie, a przypisaniem mi tej postawy przez kogoś innego. Druga asymetria dotyczy mojej gwarancji prawdziwości wypowiedzianego przeze mnie zdania przypisującego mnie samemu ową postawę i czyjejś gwarancji prawdziwości tego samego wypowiedzenia (tj. wypowiedzianego przeze mnie zdania przypisującego mnie samemu ową postawę).

Z punktu widzenia budowania teorii prawdy bardziej interesująca jest asymetria druga. W tym kontekście Davidson wyjaśnia uprzywilejowaną pozycję pierwszej osoby odwołując się do koniecznych cech interpretacji mowy. Wspomniana asymetria gwarancji prawdziwości wynika z tego, że wiedza o tym, co chciałem powiedzieć (*what I meant*), daje mi z koniecznością wiedzę o przekonaniu, które wyraziłem (aktualizowałem) swym wypowiedzeniem. Innymi słowy mówiący nie myli się w sprawie znaczenia wypowiedzianych słów, zaś interpretator nie posiada gwarancji nieomyślności przypisując znaczenia wypowiedzeniom mówiącego. Zdaniem Davidsona stwierdzenie to jest zasadnicze dla natury interpretacji, gdyż dotyczy procesu dochodzenia do rozumienia wypowiedzeń. Ten proces nie może być taki sam dla mówiącego i interpretatora (zob. 2001a/1984a, s. 12). W konsekwencji nie ma gwarancji, że słuchacz właściwie zinterpretuje słowa mówiącego, ale sam

opisanie czynności chodzenia i utrzymywania przy tym równowagi – podobnie jak czynności interpretowania – jest niezwykle trudne. Oczywiście, aby dyspozycje te mogły się rozwinąć, istota ludzka musi posiadać pewną wrodzoną zdolność.

mówiący nie może w sposób niewłaściwy interpretować swych własnych słów. Dla mówiącego jego własne wypowiedzenia mają po prostu takie znaczenia, jakie im on sam przypisuje, zaś zbudowana dla jego idiolektu teoria prawdy byłaby – o czym już była mowa – całkowicie trywialna. Źródłem konieczności czy aprioryczności warunkującej trywialność teorii prawdy dla własnego idiolektu jest to, że ustanawiamy ją własną praktyką językową.

W przeciwieństwie do teorii prawdy dla własnego idiolektu, nietrywialność teorii prawdy dla interpretowanego idiolektu jest rezultatem tego, że konstruujemy ją w procesie interpretacji, chociaż często zakładamy ją na początku tego procesu. Z przypadkiem tym mamy do czynienia, gdy interpretujemy słowa osoby, która posługuje się tym samym językiem co my. Jak już pisałem, przyjęta na początku teoria prawdy może jedynie wydawać się trywialna, gdyż jest równokształtna z trywialną teorią naszego własnego idiolektu. Jeśli zajdzie taka konieczność, to w procesie interpretacji możemy formułować kolejne hipotezy na temat T-zdań wynikających z teorii prawdy, a następnie weryfikować je w oparciu o kolejne świadectwa. Świadectwami tymi są nie tylko okoliczności wypowiedzenia, ale również świadectwa werbalne, pozwalające przypisać mówiącemu w danych okolicznościach postawy propozycyjalne, a szczególnie przekonania. Z tego punktu widzenia teoria interpretacji idiolektu drugiego człowieka powinna opisywać dynamiczny proces przechodzenia od jednej teorii prawdy do kolejnej teorii prawdy.

W tym miejscu można zadać pytanie, czy (trywialna) teoria prawdy własnego idiolektu również może ulec zmianie. Wszak znaczenia wyrażen w naszym własnym idiolecie ulegają zmianie wraz z upływem czasu. Można się o tym przekonać czytając własne teksty i zapiski lub słuchając własne nagrania sprzed lat. W takim przypadku musimy na nowo interpretować własne słowa, co oznacza, że traktujemy je z perspektywy trzecioosobowej. Jednak z perspektywy pierwszoosobowej, jeśli pominiemy dla uproszczenia ewentualne zmiany zasobu używanych słów (słownika), trywialna teoria prawdy dla naszego własnego idiolektu nie ulega zmianie – zarówno przed laty, jak i obecnie, budowane przez nas teorie prawdy naszego własnego idiolektu posiadają trywialne aksjomaty i wynikają z nich trywialne T-zdania. Jednak znaczenia moich wypowiedzeń mogły ulec zmianie. Przed laty wyznaczała je teoria prawdy identyczna z obecną; ta z kolei wyznacza znaczenia obecnego idiolektu. Jednak znaczenia wypowiedzeń sprzed lat wyznacza dla mnie (w chwili obecnej) nowa i nietrywialna teoria prawdy. To co decyduje o zmianie znaczeń to przekonania, które w tym czasie uległy zmianie.

Stanowi to – jak sądzę – argument na rzecz tezy, że przynajmniej w przypadku naszego własnego idiolektu trywialna teoria prawdy nie podaje znaczeń

występujących w nim wyrażen. Znaczenia te współwyznacza nasz własny układ przekonań i postaw propozycyjalnych. Układ ten nieustannie ewoluuje wraz ze zdobywaniem nowej wiedzy, zmianą poglądów, opinii oraz preferencji itp. Ponieważ owa ewolucja znaczeń naszego własnego idiolektu nie wiąże się ze zmianą teorii prawdy, to pozostaje często niezauważona. Dopiero interpretacja własnych słów zapisanych lub wypowiedzianych przed laty, tj. potraktowanie siebie samego z perspektywy trzecioosobowej, pozwala na dostrzeżenie owych zmian⁷.

Z perspektywy pierwszoosobowej mój własny idiolekt jest dla mnie całkowicie przezroczysty. Dlatego też – jak zauważa Davidson – słowa bezpośrednio odnosimy do rzeczy, a zdania do zdarzeń, faktów i sytuacji, co odpowiada popularnemu pogładowi, który leży u podstaw semantyk odniesieniowych. Szczególna pozycja pierwszej osoby przejawiająca się ową przezroczystością znaczeń wypowiedzi własnego idiolektu mylnie sugeruje, że istnieje bezpośredni i niezapośredniczony w języku (a zarazem w znaczeniach jego wyrażen) dostęp do rzeczywistości takiej, jaka nam się zdaje, że jest sama w sobie. Zdaniem Davidsona pogląd ten jest mylny, gdyż rzeczywistość jest przez język kategoryzowana, z czego nie zawsze zdajemy sobie sprawę. Teoria prawdy dla mojego własnego idiolektu jest dla mnie samego trywialna, zaś moje własne przekonania i postawy są również w pewien sposób trywialne, gdyż nie są za pośrednictwem żadnych świadectw⁸. Fakt ten sprawia złudne wrażenie, że żywione przeze mnie przekonania bezpośrednio odnoszą się do faktów⁹.

Do sprawy tej powraca Davidson w eseju *What is present to the Mind* (1989). U podstaw prób scharakteryzowania rzekomo istniejącej relacji między umysłem a obiektem, zdarzeniem, faktem lub stanem rzeczy leży – jak twierdzi Davidson – kartezjańska tendencja do szukania wiedzy pewnej, która jest bezpieczna od błędu. Jeżeli jest to powiązane z założeniem, że wszelka wiedza polega na psychologicznym kontakcie umysłu z obiektem, to należy znaleźć takie obiekty, dla których niemożliwy jest błąd – „obiekty, które są, jakimi się zdają i wydają się, jakimi są” (2001a/1989, s. 56). Filozofowie różnych orientacji zdawali sobie sprawę, że obiekty, które obserwujemy na co dzień takimi nie są, bo być może są inne niż nam się zdają. Jednak, jak

⁷ Przedstawiony tutaj argument na rzecz znaczeniowej roli przekonań pochodzi ode mnie. Teoria znaczenia oparta o przekonania pełniących funkcję postulatów (dyrektyw) znaczeniowych pochodzi od Ajdukiewicza (por. 1931, 1934 i 1934a).

⁸ Aktualizowanie naszych postaw za pomocą zdań (wypowiedzianych lub tylko pomyślanych) jest również dla nas trywialne.

⁹ Jak pamiętamy, przeciw obiektywnemu istnieniu faktów jako korelatów semantycznych zdań wymierzony jest Argument procy.

twierdzi Davidson, próbowano je wymyślić – taką próbę stanowią choćby impresje, które stanowią u Hume’a źródło idei, pełniących również funkcję znaczeń nazw.

We wspomnianym artykule Davidson argumentuje, że przekonanie nie jest aktem psychicznym skierowanym na jakikolwiek obiekt lub fakt pozajęzykowy i stara się to pogodzić z uprzywilejowaną pozycją pierwszej osoby. Uprzywilejowana pozycja pierwszej osoby polega na tym, że myślący w sposób trywialny wie, o czym w danym momencie myśli. Jednocześnie jego myśli są przyczynowo powiązane z pewnym obiektem. Oczywiście stwierdzenie, że obiekt jest przyczyną przekonania, stanowi jedynie skrót myślowy. Związek przyczynowo skutkowy zachodzi bowiem jedynie między zdarzeniami¹⁰. Nie wynika jednak z tego, że myślący musi coś na temat tego obiektu wiedzieć. Stanowisko tradycyjne, które upatruje pewności poznania w wiedzy pierwszoosobowej (dotyczy to m.in. Arystotelesa, Kartezusza oraz empirystów, np. Hume’a) prowadzi do dylematu, który polega na tym, że z jednej strony żywienie przekonania to pozostawanie w bezpośredniej relacji podmiotu z pewnym „obiektem”, zaś z drugiej strony, relacji tej nie daje się w żaden satysfakcjonujący sposób wyjaśnić. Wyjściem z tego dylematu jest porzucenie stanowiska, że perspektywa pierwszoosobowa dostarcza nam wiedzy pewnej, która stanowi uzasadnienie wszelkiej pozostałej wiedzy. Między przekonaniem a obiektem zachodzi związek przyczynowy, ale nie stanowi on żadnego uzasadnienia wiedzy o tym obiekcie. Jest dokładnie odwrotnie – związek przyczynowy zachodzący między zdarzeniami a przekonaniem może stanowić zarówno źródło wiedzy, jak i źródło błędnych przekonań¹¹.

Każdy obiekt będący przedmiotem naszej wiedzy posiada nieskończenie wiele własności, których nie jest w stanie wyczerpać żaden opis. Wiedza na temat obiektu, nawet subiektywna wiedza z perspektywy pierwszej osoby, dotyczy jedynie pewnych jego aspektów. Stwierdzenie, że świat oddziałuje na nas przyczynowo nie wystarczy, aby wyjaśnić, dlaczego wiedza jest selektyw-

¹⁰ Wspomniany związek przyczynowy zachodzi – wedle Davidsona – między pewnych zdarzeniem, w którym ów obiekt uczestniczy (np. percepcja tego obiektu), a zdarzeniem psychicznym (np. aktualizacją przekonania o istnieniu tego obiektu).

¹¹ Termin „wiedza” jest często używany przez Davidsona w sposób dosyć luźny jako synonim terminu „przekonanie”. Dlatego można mówić o błędnej wiedzy. Wiedza o istnieniu yeti ma prawdopodobnie swoją przyczynę w zdarzeniach zachodzących w rzeczywistości. Nie stanowią jednak jej uzasadnienia, gdyż te same zdarzenia mogą wywołać jako skutki zupełnie inne przekonania (np. przekonanie, że w oddali widać niedźwiedzia). W rozdziale drugim argumentowałem, że porządek odkrycia oparty jest na związkach przyczynowych, zaś porządek uzasadniania na związkach logicznych. Związek przyczynowy między zdarzeniami w świecie zewnętrznym a przekonaniem na temat tego świata nie gwarantuje nam prawdziwości owych przekonań.

na, ani nie jesteśmy w stanie uzasadnić nawet tej selektywnej wiedzy, którą posiadamy. Davidson posuwa się dalej – gdybyśmy dysponowali jedynie perspektywą pierwszoosobową, to jakkolwiek wiedza faktycznie byłaby niemożliwa. Wiedzę o świecie zewnętrznym można sformułować, uzgadniać, weryfikować i uzasadniać jedynie z perspektywy trzeciej osoby, interpretując słowa innych ludzi i przypisując im przekonania w obecności dostępnych świadectw, których źródłem są oddziaływania przyczynowe rzeczywistości. Innymi słowy, gdyby nie było perspektywy trzeciosobowej, nie byłoby żadnej wiedzy – nawet wiedzy, która określana jest mianem pierwszoosobowej. Nie byłoby również języka. Obie perspektywy są zatem nierozłączne.

Nieco wcześniej była mowa, że z perspektywy pierwszoosobowej nasze przekonania (lub zdania, które je wyrażają) zdają się odnosić bezpośrednio do pewnych obiektów. Perspektywa trzeciosobowa niczego tutaj nie zmienia. Istnieje bowiem wyraźna tendencja, aby przypisywać przekonaniom drugiego człowieka (lub zdaniom, które je wyrażają) pewne pozajęzykowe byty, jak fakty i sytuacje. Wedle Davidsona jest to jednak jedynie reifikujący sposób mówienia. Aby to uzasadnić Davidson odwołuje się do następującej analogii. Liczby określają ciężary, ale nie stanowią „części” ciężarów. Ponadto liczby nie określają ciężarów bezwzględnie, lecz możliwe jest każde przyporządkowanie liczb ciężarom, które zachowuje stosunek między ważonymi ciałami. Podobnie jak w przypadku ważenia potrzebujemy zbioru postulowanych obiektów (czyli „ciężarów”), zbioru posiadającego strukturę, na którą możemy odwzorować ważne obiekty, tak i w przypadku przypisywania postaw propozycyjalnych (przekonań) potrzebujemy zbioru postulowanych obiektów, w którego strukturze możemy odwzorować relacje i własności odpowiednich stanów psychicznych. Podobnie jak nie ma obiektywnie istniejących wyrażonych w liczbach ciężarów, tak i nie trzeba zakładać obiektywnego istnienia obiektów zwanych „obiettami przekonań”, np. stanów rzeczy lub faktów. Skoro liczby nie pełnią żadnej roli fizycznej, tak i „obiekty przekonań” nie pełnią dla myślącego żadnej roli psychologicznej bądź epistemologicznej. Czy zatem możemy się bez nich obyć? Zdaniem Davidsona na poziomie języka jest to niemożliwe, a co za tym idzie jesteśmy w pewien sposób skazani na to, aby nasza wiedza o rzeczywistości formułowana była za pomocą różnego rodzaju hipostaz, których źródłem jest język. Rzeczywistość sama w sobie pozostaje w pewnym sensie nieprzenikniona.

Aby poprzeć powyższe stwierdzenia przywołajmy – podobnie jak uczynił to Davidson – znany eksperyment myślowy Ziemi bliźniaczej, który został podany przez Hilarego Putnama w znanym artykule *Znaczenie terminu „znaczenie”* (1975). Przykład ten pokazuje, że interpretowany może w pewnym sensie nie znać obiektu, którego użyliśmy, aby wyznaczyć

jego myśli. Wyglądająca i smakująca zupełnie jak woda substancja *XYZ*, która wypełnia rzeki, jeziora i morza Ziemi bliźniaczej nazywana jest na tamtej planecie również wodą. Ziemiński przybysz może zatem przypisać swemu sobowtórowi z Ziemi bliźniaczej przekonanie, że w tamtejszych rzekach, jeziorach i morzach znajduje się woda w „ziemskim” tego słowa znaczeniu. To jednak substancja *XYX*, nie zaś *H₂O* oddziałuje przyczynowo na myśli i przekonania naszego sobowtóra. Nie wynika jednak z tego, że interpretowany nie wie co myśli. Przekonanie nie może jednak służyć do identyfikacji obiektu, gdyż interesująca nas różnica fizyczna (np. różnica między wodą a substancją *XYZ*) może nie znajdować się „w głowie”, lecz w zewnętrznym świecie. Różnicy tej przekonanie nie jest w stanie uchwycić, gdyż przekonanie nie ujmuje bezpośrednio obiektu, o którym się zwykło mówić, że jest przedmiotem przekonania.

2. ONTOGENEZA PROCESU INTERPRETACJI – TRZY RODZAJE WIEDZY I TRIANGULACJA

Kontynuując rozważania na temat roli perspektywy pierwszoosobowej oraz perspektywy trzecioosobowej w procesie interpretacji spróbujemy rzucić nieco światła na genezę procesu interpretacji. Nie chodzi tutaj o filogenezę interpretacji (czyli historię nabywania kompetencji językowych przez rodzaj ludzki), lecz o jej ontogenezę, czyli o to, jak każdy z nas, po stosunkowo krótkim okresie nauki, stał się istotą zdolną do komunikacji językowej¹². W paragrafie tym omówię przede wszystkim proces triangulacji umożliwiający uchwycenie pojęcia prawdy, które jest niezbędne w procesie interpretacji¹³. W artykule *Three Varieties of Knowledge* (1991) Davidson wprowadził trzy rodzaje wiedzy, których istnienie warunkowane jest istnieniem omówionych w poprzednim paragrafie perspektyw poznawczych. Symultaniczne nabywanie owych trzech rodzajów, czego warunkiem jest wzajemna przyczynowa interakcja dwóch osób oraz obiektywnej rzeczywistości, Davidson nazywa triangulacją.

W rozdziale drugim przedstawiłem krytykę, którą Davidson wysunął przeciw empirystycznym teoriom znaczenia. Zdaniem Davidsona błędem

¹² W *Theories of Meaning* Davidson wiązał możliwość nauczania się języka z jego wyuczalnością oraz jej konsekwencjami dla teorii języka. Obecnie zajmujemy się użytkownikiem języka i przyjrzymy się, jakie są warunki nauczania się języka wyuczanego.

¹³ Jak pamiętamy z poprzednich rozdziałów, prawda była dla Davidsona pojęciem pierwotnym i niedefiniowalnym, co pozwoliło mu odwrócić rolę semantycznej definicji prawdy. Obecnie wyjaśnimy, w jaki sposób każdy z nas dochodzi do tego pojęcia.

empirystów było utożsamienie psychologicznego porządku uczenia się języka z epistemologicznym porządkiem uzasadniania. Utożsamienie tych porządków prowadziło do błędnych poglądów na temat znaczenia. O ile jednak proces uczenia się języka był obojętny dla głoszonej przez Davidsona tezy, że funkcję teorii znaczenia pełni teoria prawdy w stylu Tarskiego, to analizując sam proces budowania konkretnej teorii prawdy, czyli interpretowania słów drugiego człowieka, nie sposób jest nie dostrzec pewnej analogii do procesu uczenia się języka. Jest tak, gdyż interpretowanie słów drugiego człowieka, jak i uczenie się własnego języka, polega – w sensie, który postaram się wyjaśnić – na nabywaniu różnych rodzajów wiedzy. Widać tu wyraźnie, że w przypadku triangulacji pojęcie języka rozumianego jako zbiór idiolektów, dla których interpretator buduje identyczne lub niemal identyczne teorie prawdy jest niezwykle użyteczne. Nie zmienia to jednak faktu, że w procesie triangulacji teoretycznie wystarczy nam jeden nauczyciel. W takim przypadku tworzymy symultanicznie własny idiolekt i interpretujemy idiolekt nauczyciela¹⁴.

Oba procesy, tj. uczenie się pierwszego języka oraz interpretowanie wypowiedzi rozmówcy, wiążą się ze wzajemną interakcją interpretatora lub uczącego się, interpretowanego lub nauczyciela oraz obiektywnie istniejącego i zewnętrznego wobec nich świata zdarzeń fizycznych. W przypadku, gdy język już został opanowany, interpretator dysponuje teorią prawdy własnego idiolektu oraz układem własnych postaw propozycyjalnych, w tym przekonaniach, zaś jego celem jest odtworzenie teorii prawdy dla idiolektu rozmówcy oraz odtworzenie układu jego postaw propozycyjalnych. W przypadku akwizycji pierwszego języka układ postaw propozycyjalnych ma dopiero powstać, gdyż w początkowym okresie niemowlę nie żywi żadnych przekonań, lecz jedynie reaguje na bodźce¹⁵. Jednak uwzględniając społeczny charakter uczenia się języka, nie sposób oddzielić go – przynajmniej od pewnego momentu – od procesu interpretacji, gdyż ucząc się ojczystego języka każdy z nas faktycznie interpretuje słowa swego nauczyciela oraz buduje kolejne teorie prawdy dla jego słów, weryfikując je wobec świadectw będących oddziaływaniami obiektywnie istniejącego fizycznego otoczenia. Czynności te polegają na jednoczesnym zdobywaniu trzech rodzajów wiedzy: subiektywnej, intersubiektywnej i obiektywnej¹⁶. Oczywiście termin „wiedza”

¹⁴ Proces triangulacji stanowi warunek opanowania języka, zaś opis tego procesu nie jest opisem psychologicznego mechanizmu uczenia się języka.

¹⁵ W swych pismach psychologicznych zebranych w (2001) oraz (1992), Davidson podkreśla, że warunkiem posiadania postaw propozycyjalnych (myślenia) jest znajomość języka. Nie oznacza to jednak, że odmawia on istotom, które nie znają języka, przeżyć świadomych.

¹⁶ Różnica między procesem uczenia się języka a interpretacją polega głównie na proporcjach między nabywanymi rodzajami wiedzy.

jest tu rozumiany w sensie potocznym i być może bardziej na miejscu byłoby określać wymienione trzy rodzaje wiedzy jako trzy rodzaje przekonania¹⁷.

Wiedza subiektywna to wiedza o własnych postawach propozycjonalnych – przekonaniach, pragnieniach, preferencjach, obawach itp., do których każdy z nas ma uprzywilejowany dostęp z perspektywy pierwszej osoby. Wiedza ta aktualizowana jest za pomocą wypowiedzeń zdań, które jeśli są wypowiedziane na serio i szczerze, są dla nas samych trywialnie prawdziwe. Nie oznacza to, że zawsze i bez wahania jesteśmy w stanie wyrazić własne postawy propozycjonalne. W wielu przypadkach brak nam bowiem w języku odpowiedniego czasownika psychologicznego lub zdajemy sobie sprawę z własnej niepewności co do rodzaju postawy lub jej natężenia. W takich przypadkach zazwyczaj modyfikujemy czasownik psychologiczny odpowiednimi określeniami, np. „jestem prawie przekonany”, „chyba wolę to niż tamto” itp. Sytuacja ta nie odbiega jednak zasadniczo od opisu zewnętrznej wobec nas rzeczywistości. Bardzo często nie jesteśmy w stanie dokładnie opisać koloru, zdarzenia czy obiektu. Jest to spowodowane ubóstwem zarówno języka obiektywistycznego (nazwijmy tak ten fragment języka, który służy do opisu świata realnego), jak i języka mentalistycznego (opisującego postawy propozycjonalne). Wiedzę subiektywną aktualizuje się za pomocą zdań dających się analizować za pomocą analizy parataktycznej. W zdaniu wprowadzającym pojawia się czasownik psychologiczny (mentalistyczny) odnoszący się do rodzaju postawy, zaś w zdaniu podającym treść postawy pojawiają się zazwyczaj terminy opisujące świat zewnętrzny (obiektywistyczne).

Drugi rodzaj wiedzy to wiedza obiektywna, tj. wiedza na temat obiektywnie istniejącego świata zdarzeń fizycznych. Jak twierdzi Davidson zdarzenia w świecie są przyczynami naszych przekonań, szczególnie przekonań percepcyjnych, które są aktualizowane za pomocą zdań okazjonalnych:

Wiele naszych prostych percepcji na temat tego, co się dzieje w świecie, nie jest opartych na dalszych świadectwach; nasze przekonania percepcyjne są po prostu skutkami zdarzeń i obiektów znajdujących się wokół nas (Davidson 2001a/1991, s. 205).

¹⁷ W szczególności nie gwarantuje się tutaj prawdziwości, czego zazwyczaj wymaga się od wiedzy w sensie ścisłym. Wymóg ten wyrażany jest przez aksjomat „Jeżeli x wie, że p , to p ”, który obowiązuje jedynie w przypadku wiedzy subiektywnej czyli pierwszoosobowej. Jeżeli bowiem wiem, że żywię pewne przekonanie, to je żywię. Oczywiście samo przekonanie nie musi być słuszne, zaś zdanie, które je aktualizuje nie musi być prawdziwe.

Wiedza obiektywna podatna jest na nieustanną rewizję. Jest to bowiem wiedza zależna od:

(...) funkcjonowania organów zmysłów, a ta przyczynowa zależność od zmysłów sprawia, że moje przekonania na temat świata natury są podatne na ten rodzaj niepewności, który pojawia jedynie sporadycznie w przypadku przekonań na temat mych własnych stanów umysłu (*ibidem*).

Owa niepewność bierze się stąd, że związek przyczynowy nie uzasadnia wiedzy i przekonań, gdyż wiedza i przekonanie zawsze są uwikłane w język.

Wyjaśnijmy, na czym polega owo uwikłanie w język. Zwierzęta reagują na bodźce – często są to reakcje wyuczone i podobne do zachowań ludzkich. Jednak – zdaniem Davidsona – nie posiadają one postaw propozycyjalnych. Jeżeli przypisujemy im postawy propozycyjalne, to stanowi to przykład homomorfizacji. Prawdopodobnie w ten sam sposób da się wyjaśnić pierwsze zachowania werbalne niemowląt, przejawiające się wypowiedzianiem pojedynczych słów lub reagowaniem na pewne słowa. Jednak człowiek, który zna język, poza reakcjami na bodźce, musi posiadać postawy propozycyjalne, a te – zdaniem Davidsona – zawsze aktualizują się w języku. Ponieważ wypowiedzenia aktualizujące postawy zawierają słowa, postawy tworzą sieć. W konsekwencji nie można żywić jednej postawy, pozornie aktualizując ją za pomocą wypowiedzenia. W szczególności, aby żywić jedno przekonanie, należy żywić wiele innych przekonań¹⁸.

Trzeci rodzaj wiedzy to wiedza intersubiektywna. Wiedza ta polega na przypisywaniu postaw propozycyjalnych innym ludziom¹⁹ i aktualizuje się ją stosując tę samą terminologię co w przypadku wiedzy subiektywnej (psychologiczną lub mentalistyczną w zdaniu wprowadzającym oraz obiektywistyczną w zdaniu podającym zawartość postawy propozycyjalnej). Różnica polega na tym, że w przypadku wiedzy subiektywnej obywamy się bez jakichkolwiek świadectw, to w przypadku wiedzy intersubiektywnej nie możemy się bez nich obyć. Świadectwa, które pozwalają na przypisywanie postaw propozycyjalnych innym ludziom nie są jednak tak jednorodne, jak w przypadku wiedzy obiektywnej. Składają się na nie zarówno wypowiedzi innych ludzi, które musimy interpretować, ich zachowania pozajęzykowe, oraz zwykłe dane

¹⁸ Zarówno metafora sieci, jak i holizm przekonaniowy (postaw propozycyjalnych) zostały dokładnie omówione w pismach z zakresu filozofii psychologii.

¹⁹ Należy odnotować, że Davidsonowskie rozumienie wiedzy intersubiektywnej odbiega zdecydowanie od powszechnego sposobu rozumienia terminu „intersubiektywność”. Przez to co intersubiektywne, rozumie się zazwyczaj to, co jest wspólne, dostępne poznaniu wszystkich. W tym sensie mianem wiedzy intersubiektywnej należałoby określić wiedzę obiektywną w sensie podanym przez Davidsona.

wrażeniowe – musimy bowiem konfrontować ich zachowania i wypowiedzenia z tym, co dzieje się w otoczeniu.

Wszystkie wymienione odmiany wiedzy Davidson określa mianem wiedzy empirycznej. Wiedza subiektywna byłaby to zatem empiryczna wiedza o mnie samym, intersubiektywna – empiryczna wiedza o umysłach innych ludzi, zaś obiektywna – empiryczna wiedza o świecie. Ponadto – zdaniem Davidsona – wymienione rodzaje wiedzy „zajmują się aspektami tej samej rzeczywistości; różnią się jedynie sposobem dostępu do owej rzeczywistości” (*ibidem*). W cytacie tym można dostrzec charakterystyczny dla Davidsona monizm, wedle którego wszelkie zdarzenia, nawet te opisywane jako psychiczne, są faktycznie zdarzeniami fizycznymi²⁰.

Zdaniem Davidsona filozofowie akceptowali zazwyczaj istnienie przynajmniej niektórych z wymienionych przez niego trzech rodzajów wiedzy. Nie była natomiast powszechnie akceptowana teza o ich wzajemnej niezależności. W różnych doktrynach filozoficznych daje się bowiem dostrzec wyraźne tendencje redukcjonistyczne, które przejawiają się próbą wyprowadzenia jednych typów z innych, czyli *de facto* zredukowaniem pewnych typów wiedzy. W tradycji kartezjańskiej i fenomenologicznej jako podstawową traktuje się wiedzę subiektywną, zaś pozostałe dwa rodzaje wiedzy mają jedynie charakter wtórny. Tendencja ta jest widoczna również w niektórych doktrynach empirystycznych, np. w filozofii Hume’a. W realistycznych formach empiryzmu, podstawowym typem wiedzy jest wiedza obiektywna. Wiedza intersubiektywna jest natomiast wyróżniana przez takie nurty filozofii współczesnej, jak hermeneutyka lub filozofia dialogu²¹. W obu daje się bowiem zauważyć pierwszeństwo interpretacji tekstu lub słów drugiego człowieka. Poznanie przekonań, wartości, preferencji i obaw drugiego człowieka umożliwia, wedle przedstawicieli tych kierunków, poznanie samego siebie oraz poznanie świata²².

Podana tutaj charakterystyka trzech rodzajów wiedzy ze względu na „sposób dostępu do rzeczywistości”, czyli ze względu na rodzaj świadectw (lub ich brak, jak w przypadku wiedzy subiektywnej) używanych do uzasadniania twierdzeń, oraz pierwszeństwo, jakie przyznaje się w ramach pewnych kierunków wyróżnionemu rodzajowi wiedzy, rzutuje na ich charakter. Kartezjanizm oraz fenomenologia, jak również internalistyczne formy empiryzmu, odwołują się przede wszystkim do oczywistości i aprioryczności oraz do

²⁰ Zainteresowanego czytelnika odsyłam do pism psychologicznych Davidsona.

²¹ Przykłady te pochodzą ode mnie, sam Davidson o nich nie wspomina.

²² W przypadku tych kierunków (przedstawionych przeze mnie w sposób skrajnie uproszczony) nie mamy jednak do czynienia z redukcją wiedzy subiektywnej i obiektywnej do wiedzy intersubiektywnej, lecz jedynie z uprzywilejowaniem roli wiedzy intersubiektywnej.

zawartości danych wrażeniowych (impresji), których nie zalicza się jednak do wiedzy obiektywnej. Realisci odwołują się do wiedzy obiektywnej o istniejącej niezależnie od podmiotu poznającego rzeczywistości. Hermeneutyka i filozofia dialogu postrzegane są w tradycji analitycznej jako doktryny niejasne. Nie ma w tym nic dziwnego – twierdzenia, które są w nich formułowane nie są ani oczywiste same przez się, ani oparte na danych doświadczenia. Są one oparte na niejednorodnych i wzajemnie uwikłanych świadectwach – wypowiedziach innych ludzi, ich czynach, warunkach środowiskowych, społecznych oraz historycznych, w których żyją lub żyli itp. Paradygmatycznymi naukami dla hermeneutyki są nauki humanistyczne²³.

W kwestii wzajemnych związków między trzema typami wiedzy Davidson zajmuje stanowisko odmienne od przedstawicieli wielu tradycyjnych kierunków filozoficznych. Przede wszystkim uważa on, że wymienione typy wiedzy faktycznie ze sobą współistnieją i każdy z nas posiada, oraz musi posiadać, wiedzę z każdego z wymienionych rodzajów, zaś wiele odmian sceptycyzmu w filozofii ma swoje źródło w trudności z wyprowadzeniem jednego rodzaju wiedzy z pozostałych. Na przykład wszelkie scjentystyczne formy empiryzmu, np. behawioryzm, charakteryzował sceptycyzm lub wręcz nihilizm w kwestii wiedzy o postawach propozycyjalnych drugiego człowieka. Było to spowodowane trudnością z wyprowadzeniem wiedzy o postawach propozycyjalnych innych ludzi z wiedzy obiektywnej. Inni ludzie byli bowiem opisywani w kategoriach bodźców i reakcji, czyli jedynie jako części obiektywnie istniejącego świata. Z kolei sceptycyzm w kwestii istnienia świata (który niekiedy przybierał skrajną formę solipsyzmu) może być traktowany jako rezultat trudności wyprowadzenia wiedzy obiektywnej na temat świata z wiedzy subiektywnej (w przypadku solipsyzmu również trudności wyprowadzenia wiedzy intersubiektywnej z wiedzy subiektywnej).

Wedle Davidsona wyróżnione przez niego trzy rodzaje wiedzy są wzajemnie nieredukowalne. Zatrzymajmy się na chwilę na nieredukowalności wiedzy subiektywnej do wiedzy obiektywnej. Nieredukowalność ta – zdaniem Davidsona – ma charakter logiczny. Prawdziwość zdania głoszącego, że jestem przekonany, że *p*, jest logicznie niezależna od prawdziwości zdania *p*. Wyjaśnia to doskonale analiza parataktyczna zdań psychologicznych, gdyż wartość logiczna zdania wprowadzającego nie zależy od wartości logicznej zdania podającego treść postawy propozycyjalnej. Nasze przekonania tworzą

²³ Do podobnych poglądów doszedł Kazimierz Ajdukiewicz w artykule *Metodologiczne typy nauk* (1938). Zwrócił on uwagę na niejednorodność świadectw w przypadku nauk humanistycznych. Są to bezpośrednia oczywistość w przypadku twierdzeń *a priori*, dane doświadczenia oraz rozumienie innych ludzi.

holistyczny i koherentny system. Znajomość tego systemu przez każdego z nas, czyli wiedza subiektywna, nie zapewnia prawdziwości żadnego z tych przekonań, czyli nie zapewnia wiedzy obiektywnej. Stwierdzenie to zdaje się kłócić z Zasadą życzliwości, która w jednym ze sformułowań głosi, że przekonania o świecie, żywione przez interpretatora i interpretowanego, są w przeważającej ilości przypadków prawdziwe. Jest to jednak niezgodność jedynie pozorna, gdyż – wedle Davidsona – dla żadnego konkretnego przypadku przekonania empirycznego nie można uzasadnić jego prawdziwości. Mimo to masowy błąd jest niemożliwy. Logiczna niezależność wiedzy subiektywnej od obiektywnej przejawia się również tym, że żadna wiedza o świecie nie pociąga za sobą wiedzy o naszych własnych postawach propozycyjalnych: „między umysłem a naturą rozciąga się logiczna lub epistemiczna bariera” (s. 207). Każdy z wyróżnionych rodzajów wiedzy zależy jednak od pozostałych w sensie, który można określić jako genetyczny. Oznacza to, że nie może istnieć oraz nie mógł powstać niezależnie od pozostałych.

Celem systemu filozoficznego Davidsona – systemu, w którym semantyka stanowi jedynie fragment – jest zarysowanie ogólnego obrazu, obejmującego wszystkie rodzaje wiedzy i wyjaśniającego zachodzące między nimi związki. System, który Davidson konsekwentnie rozwijał od połowy lat 60., miał pozwolić na udzielenie odpowiedzi na trzy wzajemnie zależne od siebie pytania: „W jaki sposób umysł może poznać świat?”, „W jaki sposób umysł może poznać umysły innych? i „W jaki sposób możemy poznać zawartość własnych umysłów nie odwołując się do żadnych świadectw?”. Wzajemna nieredukowalność trzech rodzajów wiedzy sprawia, że powyższe pytania są od siebie niezależne, zaś odpowiedź na nie pozwoli nam na zrozumieć, dlaczego wyróżnione rodzaje wiedzy są od siebie nieredukowalne:

Starając się przedstawić związki między trzema rodzajami wiedzy należy uczynić znacznie więcej niż pokazać, że są one wzajemnie nieredukowalne; musimy się dowiedzieć *dłaczego* są wzajemnie nieredukowalne (s. 208).

Odpowiedź na to pytanie dotyczy z kolei niezbywalnej i niczym niezastąpionej roli, jaką pełni z osobna każdy z wyróżnionych rodzajów wiedzy w kształtowaniu się naszych pojęć.

Pierwszy problem dotyczy pozornej niezależności (we wskazanym powyżej sensie genetycznym) obiektywnej wiedzy o świecie od wiedzy subiektywnej oraz wiedzy intersubiektywnej. Podstawowy argument Davidsona przeciwko stanowisku głoszącemu taką niezależność opiera się na konstatacji, że wszelka wiedza o świecie musi być wyrażona (zaktualizowana) i zakomunikowana w języku. Jednak wszelka komunikacja językowa zakłada wzajemne

przypisywanie przekonań i myśli. Bez tego nie bylibyśmy w stanie skonstruować teorii prawdy, tj. przypisać znaczeń wypowiedzeniom własnym oraz wypowiedzeniom naszych rozmówców:

Aby żyć przekonanie nie wystarczy rozróżniać między aspektami rzeczywistości i zachowywać się w różny sposób w różnych okolicznościach; (...) Posiadanie przekonania wymaga ponadto dostrzeżenie kontrastu między przekonaniem prawdziwym i fałszywym, między wyobrażeniem a realnością, między zdawaniem się a byciem. (...) Ktoś, kto posiada przekonanie o świecie – lub o czymkolwiek innym – musi uchwycić pojęcie prawdy obiektywnej, tego co zachodzi niezależnie od tego co myśli (s. 209).

Zdaniem Davidsona źródłem pojęcia prawdy obiektywnej jest komunikacja językowa zapoczątkowana w procesie triangulacji, gdyż tylko komunikacja językowa stanowi probierz adekwatnego użycia słów i zdań:

Nie mamy podstaw, aby przypisać komukolwiek zdolność do rozróżnienia między tym, co myśli, że ma miejsce, a tym co faktycznie ma miejsce, o ile ta osoba nie dzieli z nami języka; bez tego rozróżnienia nie mamy niczego, co można by nazwać myślą (s. 210).

Odpowiedź na pytanie, w jaki sposób jest możliwe rozumienie słów drugiego człowieka, można oprzeć na opisie procesu uczenia się pierwszego języka lub na opisie procesu interpretacji słów drugiego człowieka, który posługuje się całkowicie nieznanym idiolektem. W drugim przypadku mamy do czynienia z interpretacją radykalną, która została zainspirowana pojęciem przekładu radykalnego, wprowadzonego przez Quine'a w *Słowie i rzeczy*²⁴. W procesie tym nie da się, przynajmniej na początku, bezpośrednio uchwycić postaw propozycyjalnych drugiego człowieka, nie można zindywidualizować przekonań, pragnień, obaw, tj. odtworzyć aktualizujących je zdań. Można natomiast – jak twierdzi Davidson – wykryć postawy, które nie są aktualizowane za pomocą zdań, np. uznanie jakiegoś zdania za prawdziwe, pragnienie, aby zdanie było prawdziwe lub preferowanie, aby raczej to, a nie drugie zdanie było prawdziwe. Postawy te możemy wykryć nie wiedząc, co dane zdania znaczą.

Proces interpretacji (jak również proces uczenia się pierwszego języka oraz proces interpretacji radykalnej) opiera się na Zasadzie życzliwości, której nieobowiązywanie uniemożliwiłoby jakiekolwiek porozumienie i wykluczyłoby możliwość nauczania się pierwszego języka. Interpretując słowa drugiego

²⁴ Do sprawy interpretacji radykalnej powróć nieco dalej.

człowieka musimy przypisać mu niesprzeczność przekonań²⁵ i szczerość, oraz utożsamiać znaczną część jego przekonań, szczególnie na temat obiektywnie istniejącej rzeczywistości, z naszymi własnymi przekonaniem. Mówiąc o obiektywnej rzeczywistości zakłada się *implicite* pojęcie obiektywnej prawdy, które tworzy się w procesie interakcji dwóch osób i obiektywnego świata, czyli w procesie triangulacji. Jak pamiętamy, Davidson odrzucił prawdę rozumianą jako ścisłą korespondencję, gdyż ta zakłada izomorficzność rzeczywistości, myśli i języka, co pociąga za sobą konieczność wprowadzenia zreifikowanych znaczeń, które pośredniczą między naszymi przekonaniem lub aktualizującymi je zdaniem a rzeczywistością.

Przedstawmy pokrótce, w jaki sposób triangulacja pozwala na pojawienie się pojęcia prawdy obiektywnej. Argument Davidsona ma następującą postać. Każdy z nas „klasyfikuje obiekty i aspekty rzeczywistości w ten sposób, że traktuje pewne bodźce jako bardziej podobne do innych” (s. 212). Zasadę tej klasyfikacji należy jednak sprowadzić do podobieństwa reakcji, przy czym dotyczy to również reakcji werbalnych. Te bodźce, na które reagujemy podobnie, są uznawane za podobne do siebie, chociaż zdarzenia i rzeczy bywają do siebie podobne na nieograniczenie wiele sposobów. A zatem pewne grupy bodźców są wyróżnione ze względu na układy (*patterns*) reakcji, które wywołują. Układy te zostały ukształtowane ewolucyjnie lub też, w przypadku reakcji werbalnych, w procesie ewolucji mowy, która zachodziła w określonym środowisku przyrodniczym. W rezultacie powstają określone układy podobnych zachowań, w tym również zachowań językowych, zaś pewne rodzaje bodźców, za którymi kryją się zdarzenia w świecie zewnętrznym, wyróżnione są ze względu na odpowiadający im układ czy szablon podobnych do siebie zachowań.

Następnie Davidson stawia pytanie o społeczne kryterium podobieństwa reakcji na bodźce. Podobieństwo to może być wykryte, gdy obserwator koreluje reakcje drugiego człowieka z obiektami lub zdarzeniami zachodzącymi w świecie, czyli w procesie triangulacji. Obserwując zdarzenia oraz reakcje na owe zdarzenia możemy wyróżnić typy bodźców, odpowiadające typom reakcji językowych. Na przykład możemy skorelować pewien typ (podobnych do siebie) bodźców z typem reakcji polegających na wypowiedzaniu przez innych ludzi podobnych do siebie dźwięków, np. „śnieg”. W ten sposób kategoryzujemy zarówno zdarzenia w świecie rzeczywistym, jak i dźwięki (wypowiedzenia), będące reakcjami na owe zdarzenia. W innej społeczności językowej, np. w społeczności Eskimosów, kategoryzacja może przybrać inną postać.

²⁵ W sensie braku jawnej sprzeczności.

Triangulacyjny model odpowiadający, w najprostszym przypadku, trójkątowi złożonemu z dwóch osób i zewnętrznej wobec nich rzeczywistości jest wszechobecny i nie dotyczy jedynie procesowi uczenia się pierwszego języka oraz interpretacji radykalnej. W ramach tego modelu można opisać standardową rolę interpretatora i interpretowanego w komunikacji językowej w ramach tego samego języka. W procesie triangulacji ustalane są bowiem wspólne reakcje na bodźce, oraz wyznaczane są same bodźce.

Trzeba mieć dwa punkty widzenia, aby zlokalizować przyczynę myśli (...). Możemy to traktować jako formę triangulacji: każdy z dwóch ludzi reaguje inaczej na bodźce zmysłowe płynące z pewnego kierunku. Przecięcie tych linii wyznacza wspólną przyczynę (s. 213).

Tak rozumiana triangulacja nie jest możliwa bez znajomości umysłu drugiego człowieka, jego przekonań i postaw propozycyjalnych. Bez wiedzy intersubiektywnej nie jesteśmy w stanie osiągnąć wiedzy obiektywnej, zaś bez wiedzy obiektywnej nie jest możliwa wiedza intersubiektywna. Podobnie jest w przypadku wiedzy subiektywnej. Wiedza o własnych postawach propozycyjnych jest możliwa jedynie wtedy, gdy posiadamy pozostałe rodzaje wiedzy. Wprawdzie proces triangulacji polega na jednoczesnym zdobywaniu trzech rodzajów wiedzy, między triangulacją a interpretacją nie istnieje ostra granica. Różnica polega na tym, że triangulacja opisuje przede wszystkim genezę wszelkiej wiedzy, natomiast interpretacja – rozwój i modyfikację wiedzy we wszystkich wymienionych aspektach. W oba procesy uwikłany jest bowiem zmysłowy, a w istocie przyczynowy kontakt ze światem zewnętrznym, przypisywanie postaw (szczególnie przekonań) drugiemu człowiekowi oraz zestawianie owych postaw z własnymi postawami. Zwornikiem tych trzech procesów jest konstytuująca się w nich komunikacja językowa możliwa dzięki przypisywaniu znaczeń wypowiedziom własnym oraz drugiego człowieka.

Z opisu procesu triangulacji widać wyraźnie, że żaden z wyróżnionych trzech rodzajów wiedzy nie jest sprowadzalny do pozostałych, zaś warunkiem posiadania wiedzy określonego typu jest posiadanie wiedzy pozostałych typów. Co więcej, warunkiem posiadania wiedzy jest posiadanie języka, zaś warunkiem posiadania języka jest posiadanie wiedzy. Nieco wcześniej była mowa o tym, że znajomość języka, tj. znaczeń występujących w nim zdań i wyrażeń, sprowadza się do znajomości teorii prawdy tego języka i posiadania pewnego układu przekonań. Obecnie widać, że znajomość języka i wiedza wymienionych trzech typów jest nieseparowalna. Davidsonowski model triangulacyjny jest zatem najbardziej podstawowym i najogólniejszym modelem komunikacji językowej.

Teza głosząca, że warunkiem znajomości języka jest wiedza, nie jest nowa. Frege uważał, że sensy wyrażeń są nośnikami wiedzy, zaś Ajdukiewicz oparł dyrektywalną teorię znaczenia na dyrektywach znaczeniowych odpowiadających pewnym szczególnym rodzajom przekonań. Davidson zwrócił jednak uwagę na wzajemną nieredukowalność trzech rodzajów wiedzy oraz swoiste sprzężenie zwrotne między posiadaniem wiedzy a znajomością języka.

Na zakończenie tych rozważań zwróćmy uwagę, że z punktu widzenia modelu triangulacyjnego staje się jasne, że interpretowany przez nas człowiek nie jest jedynie przedmiotem, lecz jest podmiotem, któremu musimy przypisać postawy propozycjonalne. Jest tak, gdyż nieprzypisanie drugiemu człowiekowi postaw propozycjonalnych uniemożliwiłoby wyjaśnienie istnienia wiedzy obiektywnej, a w konsekwencji wyjaśnienie wytworzenia się pojęcia przedmiotu oraz prawdy obiektywnej. Uniemożliwiłoby zatem jakiegokolwiek opis fizyczny świata zewnętrznego, gdyż nie dysponowalibyśmy językiem stosownym, aby wyrazić nasze przekonania o świecie zewnętrznym. Różnica między przedmiotem a podmiotem jest to przede wszystkim różnica świadectw pozwalających na budowanie wiedzy obiektywnej i świadectw pozwalających na budowanie wiedzy intersubiektywnej. Nie oznacza to, że nie możemy drugiego człowieka, a nawet siebie samego, opisywać jak przedmiot. Opisywanie drugiego człowieka w kategoriach przedmiotowych oznacza jedynie zmianę języka opisu z psychologicznego na fizyczny oraz zmianę stosownych do języka opisu świadectw²⁶. Opis ten jest możliwy również dlatego, że opisujący pozostaje lub pozostawał w relacji komunikacyjnej z innymi ludźmi jako podmiotami, co Davidson wyraził w cytacie nawiązującym wyraźnie do Protagorasa:

Podstawą wszelkiej wiedzy jest wspólnota umysłów; to ona jest miarą wszystkich rzeczy. Nie ma sensu kwestionować adekwatności tej miary lub szukać standardów bardziej podstawowych (s. 218).

3. WARUNKI MOŻLIWOŚCI INTERPRETACJI – KRYTYKA SCHEMATU POJĘCIOWEGO

Aby komunikacja językowa między dwoma osobami przebiegała z sukcesem, rozmówcy powinni – pomimo wszelkich różnic – móc się w pewien sposób do siebie „dostroić”. Oznacza to, że nie powinno być między nimi nieprzekraczalnych różnic. O jakie różnice może tu jednak chodzić? Wydaje się, że

²⁶ Drugi człowiek jest bowiem dla nas również obiektem świata zewnętrznego.

dobrym kandydatem na ową nieprzekraczalną różnicę jest tzw. schemat pojęciowy.

Pojęcie schematu pojęciowego zrobiło karierę w latach 30. XX stulecia za sprawą amerykańskich antropologów F. Boasa, E. Sapira oraz B. Whorfa, którzy badając egzotyczne języki, głównie Indian Ameryki Północnej, głosili pogląd, wedle którego obrazy świata użytkowników tych języków są niedostępne dla przedstawicieli innych kultur i języków²⁷. Pogląd ten, a właściwie cała gama zbliżonych poglądów, nosi zazwyczaj etykietę hipotezy Sapira-Whorfa. Konsekwencją owej hipotezy jest zasadnicza niemożliwość zinterpretowania wypowiedzi przedstawicieli egzotycznych kultur. W świetle poprzedniego paragrafu jest to niebezpieczeństwo bardzo realne, gdyż jeżeli przez metaforycznie brzmiący termin „obraz świata” uznamy charakteryzowane tam trzy rodzaje wiedzy, to można uznać, że wiedza o świecie użytkowników owych języków jest niewspółmierna z naszą wiedzą i nigdy nie zdołamy jej wyrazić za pomocą pojęć dostępnych w językach europejskich. Zagadnienie to podjął Davidson w artykule *O pojęciu schematu pojęciowego* (1974).

Zdaniem Davidsona uznanie istnienia schematu pojęciowego wiąże się z mocną wersją relatywizmu pojęciowego. W wersji ten relatywizm nie głosi jedynie dość oczywistej tezy, że istoty ludzkie – w zależności od przynależności kulturowej – kategoryzują rzeczywistość za pomocą różnych pojęć (tę tezę Davidson całkowicie akceptował), lecz, że systemy owych pojęć są wobec siebie w pewien sposób niewspółmierne lub wzajemnie niewyraźne. Stanowisko to Davidson przypisuje również pewnym filozofom, m.in. Kuhnowi oraz (nie wprost) Quine'owi²⁸:

Filozofowie różnych orientacji lubią mówić o schematach pojęciowych. Schematy pojęciowe są, jak się twierdzi, sposobami organizowania doświadczenia; są systemami kategorii nadającymi formę danym zmysłowym; są punktami widzenia, z których ludzie, kultury czy epoki obserwują bieg rzeczy. Dwa schema-

²⁷ Podobne poglądy głosił już na początku XIX wieku niemiecki filozof i antropolog Wilhelm von Humboldt.

²⁸ Istnienie schematów pojęciowych uznawał Rudolf Carnap. Na gruncie polskim stanowisko to głosił Kazimierz Ajdukiewicz w dyrektywnej teorii znaczenia. Teoria ta – pozostająca pod wpływem francuskiego konwencjonalizmu – dopuszczała wielość wzajemnie nieprzetłumaczalnych języków teorii naukowych. Oczywiście układ pojęć leżący u podstaw poszczególnych teorii był dostępny naszemu poznaniu, ale jedynie w ramach danej teorii i jej języka. Jednak pojęć tych nie dawało się wyrazić w innym znanym nam języku, ani zakomunikować komuś, kto nie znał danej teorii i jej języka. Wedle antropologów amerykańskich również możliwe było uchwycenie pojęć egzotycznego języka, należało jednak tego języka się nauczyć. Nie można było jednak owych pojęć wyrazić w języku angielskim, a w konsekwencji ów egzotyczny język nie był interpretowany.

ty mogą być nieprzekładalne; wówczas przekonania, pragnienia, nadzieje i wiedza charakteryzujące jednego człowieka nie mają właściwych odpowiedników u tego, kto reprezentuje inny schemat. Sama rzeczywistość jest względna wobec schematu: to, co uważa się za realne w jednym systemie, może nie być takie w innym (1991/1974, s. 280).

Schemat pojęciowy nie został nigdy ściśle zdefiniowany, co wpłynęło na sposób argumentacji Davidsona, który musiał odwoływać się do pojęcia intuicyjnie uchwytnego, lecz przy bliższej analizie – niejasnego. Dlatego też podana w powyższym cytacie charakterystyka schematu pojęciowego ma charakter mocno metaforyczny, chociaż kryją się w niej stosunkowo łatwe do uchwycenia intuicje. W zależności bowiem od wersji relatywizmu, schemat pojęciowy jest jednoznacznie przypisany do konkretnej osoby, do pewnej wspólnoty językowej lub do społeczności żyjącej w danym okresie historycznym. Szczególnie istotne dla tego poglądu wydaje się jednak to, że różne obrazy świata rozumiane jako różne schematy pojęciowe różnych ludzi (lub społeczności), chociaż są wypełnione tą samą i dostępną dla wszystkich zawartością empiryczną (*empirical content*), są wzajemnie niedostępne. Na poziomie języka objawia się to całkowitą nieprzekładalnością pojęć z jednego języka na drugi. Gdyby stanowisko to było słuszne, oznaczałoby to całkowitą niemożliwość interpretacji słów przedstawicieli innych społeczności, a zakładając, że schemat pojęciowy jest domeną indywidualną – również członków tej samej wspólnoty językowej. W kategoriach semantyki Davidsona oznaczałoby to zasadniczą niemożność zbudowania teorii prawdy dla ich języka oraz niemożność przypisania jego użytkownikom przekonań i pragnień²⁹.

Krytykując niejasne pojęcie schematu pojęciowego i nawiązując do *Dwóch dogmatów*, Davidson starał się obalić trzeci i – jak stwierdził na końcu artykułu – ostatni dogmat empiryzmu, który określił mianem dogmatu o dualizmie schematu pojęciowego i zawartości empirycznej³⁰. Jednocześnie pośre-

²⁹ Podobne do Davidsona poglądy na temat pojęcia schematu pojęciowego głosił Karl Popper (zob. Popper 1994). Porównanie obu rodzajów krytyki przeprowadziłem w (Maciaszek 2014).

³⁰ Wbrew tłumaczeniu polskiemu, angielskie słowo *content* oddaję jako „zawartość”, nie zaś jako „treść”. Treść utożsamia się bowiem niekiedy ze znaczeniem wyrażenia, czego chciałbym tutaj uniknąć. Termin „zawartość empiryczna” może być natomiast rozumiany szeroko, również w kategoriach skutków przyczynowych oddziaływań świata zewnętrznego na użytkowników języka. Za takim rozwiązaniem przemawia również to, że w *Dwóch dogmatach* Quine pośrednio obalił utrzymywane przez filozofów Koła wiedeńskiego przekonanie o istnieniu znaczenia empirycznego. Drugi z krytykowanych przez Quine’a dogmatów dotyczył bowiem istnienia tzw. zdań protokolarnych rozumianych jako zdania odpowiadające bezpośrednio danym doświadczenia. Skoro takich zdań nie ma, nie ma zatem sensu mówić o znaczeniu empirycznym.

dnio krytykował Quine'a, który w *Słowie i rzeczy* faktycznie ów dogmat utrzymywał. Konsekwencją uznania wspomnianego dualizmu jest – zdaniem Davidsona – przekonanie o wielości różnych i niewspółmiernych schematów pojęciowych, czyli teza relatywizmu pojęciowego. Istota Davidsonowskiej krytyki tej tezy sprowadza się do pokazania, że nie da się jej sensownie wyeksplikować, ani nawet sformułować. A zatem nie ma możliwości wykazania tezy o istnieniu niewspółmiernych schematów pojęciowych, podobnie jak nie ma możliwości wykazania, że schematów takich nie ma.

W świetle omówionej metodologii interpretacji nie jesteśmy w stanie ocenić, czy pojęcia lub przekonania innych radykalnie różnią się od naszych. (...) Nie znaleźliśmy bowiem żadnych podstaw, na których można by twierdzić, że dwa schematy są różne. Równie nieuzasadnione byłoby ogłoszenie wspaniałej wieści, że cała ludzkość – a przynajmniej wszyscy użytkownicy języka – posiadają wspólny schemat i ontologię (s. 297).

Argumentacja Davidsona przeprowadzana jest zasadniczo na płaszczyźnie języka. Ten sposób argumentacji opiera się na wielokrotnie już przytaczanym poglądzie Davidsona, że wszelkie przekonania muszą dać się zwербalizować. Relacja przyczynowa między zdarzeniami w świecie zewnętrznym a przekonaniem jest wprawdzie zapośredniczona w języku, ale zdarzenia w świecie rzeczywistym nie uzasadniają przekonań. Dlatego też te same zdarzenia mogą u różnych ludzi, najczęściej przedstawicieli różnych kultur, wywoływać różne przekonania. W akcie przekonania ujmujemy bowiem przedmiot będący przedmiotem doświadczenia jedynie pod pewnym względem – zawsze jakby jednostronnie, co jest uwarunkowane kategoryzacją narzuconą nam przez język³¹.

Samo uznanie, że istnieją różne kategoryzacje rzeczywistości, nie jest groźne. Przedmiotem krytyki Davidsona jest bowiem teza, że schematy pojęciowe, które przejawiają się w różnych kategoryzacjach rzeczywistości, są nieprzekładalne. Co więcej, jeśli schemat pojęciowy utożsamimy – co czynią

³¹ Doprowadziło to Quine'a do tezy o niezdeterminowaniu przekładu i nieokreśloności odniesienia, które można traktować jako konsekwencje trzeciego dogmatu. Oczywiście stanowisko Quine'a jest bardzo subtelne, złożone i podatne na wiele interpretacji. Dlatego też argumentacja Davidsona może uderzać w Quine'a jedynie pośrednio. Bezpośrednio Davidson atakuje te koncepcje, które za konsekwencję przyjęcia dualizmu schemat pojęciowy – zawartość empiryczna uznają całkowitą nieprzekładalność języków. Stosunek Davidsona do Quine'a nie jest tutaj jednoznaczny. Traktując termin „nieprzekładalność” dosłownie, krytyka ta Quine'a bezpośrednio nie dotyczy. Quine mówił przecież o podręcznikach przekładu, a zatem uznawał, że przekład jest możliwy, a co więcej każdy przekład miał być zgodny z ogółem dostępnych bodźców.

implicite relatywisci – ze zbiorem języków wzajemnie przekładalnych, to nieprzekładalność może być traktowana jako warunek konieczny odróżnienia różnych schematów pojęciowych. Davidson pokazuje zatem, że terminowi „nieprzekładalność” nie daje się nadać żadnego jasnego sensu.

Jak stwierdziłem na początku, termin „schemat pojęciowy” nie doczekał się ścisłej definicji i zazwyczaj charakteryzowany jest w sposób metaforyczny³². Davidson rozpoczyna argumentację od wyliczenia dwóch grup metafor, które pojawiają się w literaturze poświęconej schematowi pojęciowemu. Do pierwszej grupy należą wyrażenia: „organizuje”, „systematyzuje” i „dzieli”, które odnoszą się do świata lub do danych doświadczenia. Należy w tym miejscu podkreślić, że Davidson nie traktuje metafory jako nośnika sensu, lecz jako środek oddziaływania przyczynowego na słuchacza, i nie przypisuje metaforze żadnych znaczeń poza znaczeniami dosłownymi (por. Davidson 1978). W konsekwencji odniesienie wyżej wymienionych terminów w ich dosłownych sensach do schematów pojęciowych można uznać z jego punktu widzenia za przedsięwzięcie absurda³³, co widać na przykładzie przytoczonego poniżej argumentu:

Pojęcie organizowania pojedynczego przedmiotu (świata, przyrody itp.) jest niejasne, o ile przez ten przedmiot nie rozumie się czegoś, co składa się z jeszcze innych przedmiotów. Ktoś, kto postanawia zorganizować szafę, rozmieszcza w niej rzeczy. Gdyby powiedziano mu, aby zorganizował nie buty i koszule, ale samą szafę, byłby zdezorientowany. Na czym polegałoby organizowanie Pacyfiku? – zapewne na wyprostowaniu brzegów, przemieszczeniu wysp albo zgłądzeniu ryb (1991/1974, s. 290).

Do drugiej grupy metafor Davidson zalicza wyrażenia „odpowiadają (czemuś)”, „harmonizują (z czymś)”, „przewidują (coś)”, „wyjaśniają (coś)”

³² Nie dotyczy to wspomnianej już dyrektywnej teorii znaczenia autorstwa Kazimierza Ajdukiewicza. W tym przypadku schemat pojęciowy można utożsamić z zespołem dyrektyw znaczeniowych danego języka (sam Ajdukiewicz, rzecz jasna, terminu „schemat pojęciowy” nie używa). Nieprzekładalność schematów pojęciowych w tym sensie jest konsekwencją bardzo restrykcyjnego rozumienia przez Ajdukiewicza terminu „przekład”. Interpretacja w sensie Davidsona jest pojęciem o wiele „luźniejszym”, co wiąże się również z nieobecnością u Ajdukiewicza rekurencyjnością teorii prawdy. Sprawę tę poruszam w polemice z prof. Adamem Nowaczykiem (2015a).

³³ Davidson był zwolennikiem tzw. przyczynowej teorii metafory (por. Maciaszek 2016a), którą zamierzam dokładnie omówić w innej monografii. W *O pojęciu schematu pojęciowego* Davidson odwołuje się *implicite* do tej teorii metafory, chociaż jej *explicite* nie wymienia. Metafora nie jest bowiem – jego zdaniem – nośnikiem treści pojęciowej (*conceptual content*). Zauważmy, że w tym przypadku, w odróżnieniu od wyrażenia *empirical content*, (zawartość empiryczna), termin *content* oddają jako „treść”.

i „są zgodne z (czymś)”. Wedle Davidsona są to jedynie różne określenia na stwierdzenie prawdziwości zdań, co prowadzi do konkluzji:

(...) coś jest akceptowalnym schematem pojęciowym lub teorią, jeśli jest prawdziwe. (...) Kryterium schematu pojęciowego, który różni się od naszego, byłoby zatem następujące: w dużej mierze prawdziwy, ale nieprzetłumaczalny (s. 293).

W powyższym cytacie Davidson nawiązuje *implicite* do teorii interpretacji. Dokonanie przekładu (z zastrzeżeniami przedstawionymi na początku tego rozdziału) stanowi w istocie część interpretacji obcego języka i wiąże się ze zbudowaniem teorii prawdy dla zdań obcego idiolektu w metajęzyku stanowiącym nasz własny idiolekt³⁴. Jeżeli zatem zinterpretujemy czyjeś słowa, to nie ma podstaw do przypisania mu odmiennego schematu pojęciowego. Jeśli jednak nie zdołamy zinterpretować czyichś słów, to „tym samym stanowi to dowód na to, że nie jest to zachowanie językowe” (s. 283).

W przytoczonej argumentacji Davidson atakuje przede wszystkim poglądy strukturalistów amerykańskich, którzy twierdzili, że każda społeczność językowa, szczególnie społeczność odizolowana od pozostałych, posiada unikalny i nieprzetłumaczalny schemat pojęciowy. W konsekwencji utrzymywali, że schematy pojęciowe są niedostępne dla osób odmiennych kultur i nieprzekładalne. Schemat pojęciowy miał tworzyć unikalny obraz świata, który jest niewyrażalny w innych językach i jest z zasady niepoznawalny przez ludzi należących do innych społeczności lub kultur, o ile nie wnikną do danej społeczności i nie nauczą się od podstaw jej języka tak, jak uczyli się swego pierwszego języka. Języka takiej społeczności nie dałoby się zatem zinterpretować. Krytykując to stanowisko Davidson twierdzi, że interpretacja słów drugiego człowieka jest możliwa zawsze. Przekład, który w niej występuje, może być niezgrabny lub zawiły, ale jest zawsze możliwy. Bardzo często przekład taki nie nadaje się do celów literackich³⁵. Z tego punktu widzenia teza, że każdy człowiek lub społeczność językowa posiada własny nieprzetłumaczalny schemat pojęciowy, który wypełniany jest zawartością empiryczną, nie jest niczym uzasadniona.

Interpretacja radykalna w sensie Davidsona miała zastąpić przekład radykalny. W obu teoriach istotną rolę pełni pojęcie przekładu. Jednak między

³⁴ Jak pamiętamy, interpretacja jest pojęciem szerszym od przekładu. W tym momencie można jednak pozwolić sobie na uproszczenie.

³⁵ Za ilustrację może posłużyć nam trudność, na którą natrafiają tłumacze poezji z języków romańskich na język polski, gdyż w językach romańskich „księżyc” jest rodzaju żeńskiego, zaś w polskim rodzaju męskiego. Trudno jest zatem oddać na język polski w „zgrabny” sposób skojarzenia księżyc z kobietą. Nie oznacza to jednak, że interpretacja wiersza nie jest możliwa.

przekładem u Quine'a i u Davidsona zachodzą istotne różnice. Dotyczą one w szczególności sposobu doboru świadectw pozwalających na dokonanie przekładu. Jak pamiętamy, teoria prawdy dla języka naturalnego zawierającego zdania okazjonalne nie dokonuje przekładu w ścisłym tego słowa znaczeniu, lecz na przekładzie jedynie jest oparta. W aksjomatach teorii prawdy pojawiają się parametry osoby, miejsca i czasu. Z kolei Quine ograniczał się do świadectw behawioralnych – uznawania lub odrzucania zdań wobec dostępnych bodźców, co pozwoliło mu na zdefiniowanie tzw. znaczeń bodźcowych zdań, umożliwiających dokonanie przekładu zdań okazjonalnych. Ponieważ znaczenia bodźcowe nie wystarczają do dokonania przekładu bardzo wielu zdań języka, Quine musiał założyć istnienie schematów pojęciowych. Wyrazem schematów pojęciowych są tzw. hipotezy analityczne, które pozwalają na dokonanie przekładu tych zdań, dla których przekład na podstawie świadectw empirycznych jest niemożliwy. Ponieważ dwaj tłumacze posiadają różne schematy pojęciowe, stworzone przez nich podręczniki przekładu są niezgodne, co stanowi jedno z możliwych sformułowań tezy o niezdeteminowaniu przekładu.

Davidson, poza świadectwami empirycznymi (bodźcowymi), odwoływał się do różnorodnych świadectw empirycznych związanych z trzema rodzajami wiedzy. W ramach procedury interpretacji opartej na tak szerokich podstawach nie da się sensownie mówić o schemacie pojęciowym, zatem nie da się w żaden sposób nadać sensu takiej wersji tezy relatywizmu pojęciowego, która implikuje dualizm schematu pojęciowego i zawartości empirycznej. Relatywizm pojęciowy jest, jak pisze na początku, „doktryną egzotyczną i atrakcyjną” (s. 280), lecz przy bliższej analizie okazuje się być pozbawiona zarówno podstaw, jak i treści. Atrakcyjność doktryny relatywizmu pojęciowego ujęli zwięźle Lepore i Ludwig:

Zafascynowanie doktryną relatywizmu pojęciowego polega na myśli, że przechodząc od punktu widzenia reprezentującego jeden schemat pojęciowy, do punktu widzenia reprezentującego inny schemat pojęciowy, zdanie prawdziwe może stać się zdaniem fałszywym, przy zachowaniu dotychczasowego znaczenia (Lepore i Ludwig 2015, s. 318).

Z punktu widzenia Davidsonowskiej teorii interpretacji nie można w takim przypadku mówić o zachowaniu dotychczasowego znaczenia. Zmiana wartości logicznej zdania, przy zachowaniu dotychczasowego znaczenia, jest bowiem nie do pogodzenia z Tarskiego teorią prawdy. Co więcej stanowisko to zakłada *implicite* istnienie zreifikowanych znaczeń, które są w pewnym sensie niezależne od przekonań. W zakończeniu *O pojęciu schematu pojęciowego* Davidson pisze:

Odrzucając pojęcie niezinterpretowanej rzeczywistości, czegoś poza wszystkimi schematami i nauką, nie rezygnujemy bynajmniej z pojęcia prawdy obiektywnej. (...) Prawdziwość zdań pozostaje, oczywiście, względna wobec języka, ale ten jest wystarczająco obiektywny (Davidson 1991/1974, s. 297).

Język jest wystarczająco obiektywny, gdyż musi być w zgodzie z oddziaływaniami przyczynowymi między zdarzeniami zachodzącymi w obiektywnie istniejącym świecie fizycznym a naszymi przekonaniem. Pozostaje oczywiście zrelatywizowanie prawdziwości zdań do języka, co jest oczywiste w świetle Tarskiego semantycznej definicji prawdy. Oczywiście relatywizowanie prawdy do języka nie jest w żadnym przypadku przejawem relatywizmu pojęciowego. Co więcej, dwa różne języki mogą być (oczywiście tylko teoretycznie) identyczne co do fizycznych kształtów słów i zdań, lecz różnić się znaczeniami. W takich językach, dwa równokształtne zdania mogą mieć faktycznie różne wartości logiczne – nie są to jednak te same zdania, gdyż różnią się znaczeniami.

4. WARUNKI MOŻLIWOŚCI INTERPRETACJI – ZASADA ŻYCZLIWOŚCI

O Zasadzie życzliwości była już mowa kilka razy. Posiada ona jednocześnie charakter normatywny oraz – co wydaje się tu szczególnie istotne – stanowi swoiste „prawo”, którego zachodzenie stanowi warunek konieczny interpretacji. Zasada pojawia się w nieco odmiennych sformułowaniach w tekstach pochodzących z różnego okresu twórczości Davidsona.

W *Three Varieties of Knowledge* Davidson dokonuje syntezy wielu sformułowań Zasady życzliwości, wyróżniając jej dwie podstawowe wersje, które nazywa Zasadą koherencji oraz Zasadą korespondencji (2001a/1991, s. 211). Pierwsza z tych zasad głosi, że interpretując słowa drugiego człowieka zakładamy, że jego przekonania tworzą w miarę spójny system. Przyjmujemy bowiem, że nasz interlokutor nie żywi jawnie sprzecznych przekonań. Oczywiście wymóg logicznej niesprzeczności jest w przypadku naszych potocznych przekonań zbyt wygórowany. Chodzi tu raczej o niesprzeczność lokalną – zakładamy bowiem, że nikt nie głosi na dany temat w danym czasie jawnych i oczywistych sprzeczności. Oczywiście tak rozumiana zasada nie jest bezwzględnie regułą. Niekiedy, co każdy z nas wie z własnego doświadczenia, przypisujemy innym ludziom głoszenie sprzecznych poglądów. Często zarzut ten pojawia się w debatach oraz walce politycznej. Jest on jednak możliwy jedynie dzięki temu, że interpretując wypowiedzenia drugiego człowieka zakłada się obowiązywanie Zasady koherencji. Jej nieobowiązywanie, czyli

całkowita niespójność przekonań interpretowanego, sprawiłoby, że nie dałoby się zinterpretować jego słów, a w konsekwencji nie miałyby sensu zarzucanie mu niekoherencji. Aby można było mówić o odstępstwie od reguły, reguła powinna w większości przypadków obowiązywać. To, że możemy niekiedy zarzucić komuś niespójność niektórych przekonań, jest możliwe jedynie dzięki temu, że jego przekonania traktujemy w przeważającej ilości przypadków za spójne. Jak pamiętamy Davidson był holistą przekonaniowym. Z tego punktu widzenia o sprzeczności (niekoherencji lub niespójności) można mówić jedynie w odniesieniu do niesprzecznego układu przekonań, który przypisujemy interpretowanej osobie.

Drugą zasadę, która również stanowi warunek możliwości interpretacji, Davidson określił mianem Zasady korespondencji³⁶. Dostępne nam świadectwa, tj. wypowiedzenia i zachowania interpretowanego, zawsze mają miejsce w kontekście obiektywnie zachodzących zdarzeń, które oddziałują przyczynowo na nasze zachowania werbalne i pozawerbalne. Zasada korespondencji głosi, że reakcje interpretowanego i interpretatora są skutkami tych samych zdarzeń w świecie zewnętrznym, co stanowi jeden z warunków możliwości wzajemnej interpretacji. Wobec dostępnych naszym zmysłom zdarzeń zachodzących w świecie zewnętrznym, np. ulewnego deszczu, zakładamy, że interpretowany uznaje – podobnie jak my – że pada deszcz, czemu może dać wyraz odpowiednim zachowaniem językowym i pozajęzykowym. Oczywiście w poszczególnych przypadkach możemy się mylić, ale stwierdzenie pomyłki jest możliwe jedynie dzięki temu, że w przeważającej większości przypadków Zasada korespondencji obowiązuje, co wiąże się z założeniem, że nasze zmysły funkcjonują w podobny sposób.

Davidson był zwolennikiem tzw. teorii dystalnej³⁷ – uważał, że zdarzenia świata zewnętrznego istnieją niezależnie od nas i są przyczynami naszych zachowań językowych i pozajęzykowych oraz ostatecznymi przyczynami wszystkich postaw propozycyjalnych, również tych, które nie są bezpośrednimi skutkami zdarzeń w świecie fizycznym³⁸. Zasada korespondencji pomaga wykryć przekonania w pewnych, szczególnych sytuacjach. Na przykład obfity opad śniegu pozwala przypisać każdemu przekonanie (empiryczne), że pada

³⁶ Zasady korespondencji nie należy utożsamiać z korespondencyjną teorią prawdy. Korespondencja rozumiana jest przez Davidsona w sposób przyczynowy – świat oddziałuje przyczynowo na nasze przekonania oraz inne postawy, które aktualizujemy za pomocą wypowiedzeń zdań.

³⁷ W odróżnieniu od Quine'a, który pisał jedynie o bodźcach, nie wypowiadając się o ich źródle. Dlatego też teoria Quine'a nosi nazwę proksymalnej.

³⁸ W większości przypadków nie chodzi o bezpośredni związek przyczynowy, lecz o historię przyczynową opisaną jako triangulacja interakcji ze światem oraz innymi ludźmi.

śnieg³⁹. Jak już była mowa, podobne zdarzenia w świecie fizycznym wywołują podobne przekonania. Dzięki obowiązywaniu Zasady korespondencji można wykryć odstępstwa: pomyłki osoby interpretowanej lub przejawy nieszczerości oraz przypadki wątpliwe. Oczywiście nie chodzi jedynie o pewną korelację zdarzeń i zachowań – bodźców i reakcji na bodźce. W wielu przypadkach, aby przypisać komuś przekonanie nie wystarczy bowiem obserwacja jego zachowań niewerbalnych lub nawet zachowań werbalnych obserwowanych jedynie w pewnych, szczególnych okolicznościach. W wielu przypadkach interpretator przypisuje interpretowanemu przekonania różne od tych, które sam żywi. Jednak w przypadku zdań okazjonalnych, np. „Pada śnieg”, aktualizujących proste przekonania empiryczne, interpretator ma podstawy, aby utożsamić własne przekonania z przekonaniami interpretatora. A zatem konsekwencją zachodzenia Zasady korespondencji jest możliwość utożsamienia dużej części przekonań interpretowanego z przekonaniami interpretatora, co pozwala na zinterpretowanie jego słów: „(...) jedna zasada wyposaża mówiącego w namiastkę logiki, druga wyposaża go w to, co interpretator bierze za prawdziwe przekonanie na temat świata” (2001a/1991, s. 211). Chodzi tu przede wszystkim o przekonania empiryczne wyrażane przez zdania okazjonalne, np. „Pada deszcz”, „To jest zielone”, „Oto królik” lub „Pada śnieg”.

Podane przykłady są faktycznie zdaniami obserwacyjnymi, które stanowią szczególny przypadek zdań okazjonalnych. Zdania obserwacyjne są bezpośrednio związane przyczynowo z rzeczywistością, natomiast asertywne wypowiedzenia pozostałych zdań obserwacyjnych (np. tzw. zdań nieprzejrzystych poznawczo, np. „Oto dziekan” czy „Oto nauczyciel”) są wprawdzie skutkami pewnych zdarzeń, ale jako zapośredniczone w istniejącej sieci przekonań, nie mogą stanowić początku procesu interpretacji. Z kolei wypowiedzenia zdań nieokazjonalnych są związane przyczynowo z obiektywnie istniejącą rzeczywistością jedynie pośrednio. Do sprawy tej powrócę w paragrafie poświęconym interpretacji radykalnej.

Zanim Zasada życzliwości została sformułowana w swej ostatecznej postaci w *Three Varieties of Knowledge* była artykułowana przez Davidsona w różny sposób. Niekiedy przybierała postać wymogu, aby interpretator i interpretowany posiadali wiele wspólnych przekonań, czyli wymogu mak-

³⁹ Jest to rzecz jasna pewne uproszczenie. Osobie, która nie zna nazwy „śnieg” (w jakimkolwiek języku), nigdy wcześniej nie widziała śniegu (w rzeczywistości, na ekranie, na zdjęciu lub nawet na obrazku) nie jesteśmy w pierwszym momencie przypisać przekonania, że pada śnieg, gdyż przekonania muszą być zawsze zwerbalizowane. Być może osoba taka pomyśli, że z nieba pada mąka. Przekonanie, że pada śnieg, będziemy mogli jej przypisać, gdy w wyniku procesu triangulacji odrzuci przekonanie o mące padającej z nieba. Nie oznacza to, że musimy się z nią zgadzać w każdej kwestii dotyczącej natury śniegu.

symalizacji zgody między uczestnikami komunikacji językowej. W innych miejscach Zasada życzliwości głosiła, że masowy błąd jest niemożliwy, czyli przeważająca większość przekonań uczestników komunikacji językowej jest prawdziwa. Bez względu na sformułowanie, Zasada życzliwości wiąże się z przekonaniem o podobieństwie możliwości poznawczych oraz przekonaniem o racjonalności użytkowników języka. Pojęcie racjonalności jest jednak rozumiane u Davidsona w sposób bardzo szeroki. Zasada koherencji wiąże się z rozumieniem racjonalności jako dyspozycji do praktycznego stosowania reguł logiki, przynajmniej w minimalnym zakresie. Przejawia się ona ograniczeniami narzuconymi na żywione przekonania. Ograniczenia te to nic innego jak praktyczne stosowanie zasady niesprzeczności oraz umiejętność stosowania elementarnych reguł wynikania logicznego. Termin „racjonalność” w tym sensie pochodzi od angielskiego terminu *reason* rozumianego jako rozum.

Zasada korespondencji odwołuje się natomiast do Davidsonowskiego pojęcia działania, gdyż akty wypowiedzania lub uznawania zdań są szczególnymi działaniami. Przez działanie, w odróżnieniu od zdarzenia, w którym po prostu uczestniczymy, Davidson rozumie zdarzenie, które ma rację lub powód (*reason*). Dlatego też każde działanie jest z definicji racjonalne (*rational*)⁴⁰. Na powód działania składają się – w wielkim uproszczeniu – tzw. przednastawienie do działania pewnego typu (np. chęć, uznanie za słuszne, poczucie obowiązku itp.) oraz przekonanie, że faktycznie podjęte działanie jest działaniem właśnie tego typu. Można zatem powiedzieć, że przednastawienie wyznacza cel działania, zaś przekonanie – środki, które zapewniają jego realizację. Można zatem uznać, że wspólnota i prawdziwość większości przekonań gwarantuje skuteczność ludzkich działań zarówno językowych, jak i pozajęzykowych⁴¹.

Na zakończenie uwag na temat Zasady życzliwości warto zauważyć związek tej zasady ze sformułowaną przez H. P. Grice’a Zasadą kooperacji (por. Grice 1975). Zasada ta głosi, że rozmowa jest rodzajem współpracy i (podobnie jak w przypadku wszelkiej innej współpracy) należy wnosić do konwersacji tyle, ile na danym etapie rozmowy jest wymagane. Ponieważ sformułowanie to jest niezwykle ogólne, Grice formułuje tzw. maksymy

⁴⁰ Oczywiście chodzi tu o racjonalność w jednym z wielu znaczeń tego terminu. Działanie, które jest racjonalne w tym sensie, może być określone jako całkowicie irracjonalne w normatywnym znaczeniu terminu „racjonalność”.

⁴¹ W filozofii działania Davidsona da się dostrzec bardzo wyraźny wpływ pragmatyzmu amerykańskiego. Filozofia działania stanowi bardzo ważny fragment systemu filozoficznego Davidsona, który wymaga osobnego opracowania. Definicję działania można znaleźć w (1963). Prace poświęcone filozofii działania zostały zebrane przede wszystkim w pierwszej części (2001).

konwersacyjne, które mają uszczegółowić Zasadę kooperacji. Są to kolejno Maksyma ilości, która domaga się przekazywania odpowiedniej w danych okolicznościach ilości informacji, Maksyma jakości (nie kłam i nie wygłaszaj bezpodstawnych sądów), Maksyma relacji (mów na temat) oraz Maksyma sposobu (mów w sposób jasny, zwięzły, jednoznaczny itp.).

Chociaż Grice sformułował Zasadę kooperacji oraz maksymy konwersacyjne w postaci normatywnej, ich faktyczne nieobowiązywanie – podobnie jak nieobowiązywanie Zasady życzliwości – uniemożliwiłoby praktycznie wszelką komunikację językową. W szczególności nieobowiązywanie Maksymy ilości (informacji) może skutkować w szczególnym przypadku podawaniem nieskończonej ilości informacji. Przejawia się to głoszeniem sprzeczności⁴² lub ewidentnych fałszów stojących w sprzeczności z przyjętymi przekonaniem, co u Davidsona zakazane jest przez Zasadę koherencji. Zarówno Grice, jak i Davidson uznawali użytkowników języka za istoty racjonalne, które – poza sytuacjami wyjątkowymi – starają się przestrzegać reguły komunikacji językowej oraz faktycznie je przestrzegają. Dlatego też zarówno u Grice’a, jak i Davidsona wygłaszanie w tym samym czasie zdań sprzecznych lub zdań sprzecznych z przyjętym przekonaniem jest traktowane jako pozorne łamanie reguł i stał się środek służący do rozpoznawania wypowiedzi metaforycznych⁴³. Oczywisty jest również związek Maksymy jakości z Davidsonowską Zasadą korespondencji, bowiem prawdziwość przeważającej ilości przekonań jest zagwarantowany przez ich bezpośredni lub chociaż pośredni związek przyczynowy z obiektywnie istniejącą rzeczywistością.

5. HOLISTYCZNY ASPEKT INTERPRETACJI

Interpretacja idiolektu innego człowieka, która realizuje się poprzez budowanie teorii prawdy, stanowi projekcję na cały idiolekt interpretacji tych wypowiedzi, z którymi do tej pory zetknął się interpretator, gdyż – jak pamiętamy z rozdziału drugiego – nie można znać części języka, lecz jedynie można znać częściowo cały język. Stwierdzenie to ma wiele aspektów oraz pewne konsekwencje, które omówimy w tym krótkim paragrafie.

⁴² Wygłoszenie zdań sprzecznych oznacza podanie nieskończonej ilości informacji.

⁴³ Różnica między Gricem a Davidsonem polegała na tym, że pierwszy był zwolennikiem teorii komunikowania metaforycznego i uważał, że w kontekście wypowiedzenia słuchacz może implikować metaforycznie co mówiący miał na myśli (*speaker's meaning*), zaś Davidson, który był zwolennikiem przyczynowej teorii metafory, uważał, że wypowiedzenie metaforyczne nie niesie żadnej treści, lecz oddziałuje na słuchacza przyczynowo.

Interpretacja wypowiedzeń idiolektu naszego rozmówcy jest ściśle związana z przypisaniem mu sieci postaw propozycyjalnych, a w szczególności sieci przekonań (holizm przekonaniowy), na podstawie obserwowalnych działań i zachowań we wspólnej, oddziałującej na interpretatora oraz interpretowanego rzeczywistości. Oczywiście słowo „zbudować” w kontekście teorii prawdy należy rozumieć analogicznie do słowa „znać”, gdy mówi się o kimś, że zna reguły gramatyczne własnego języka. Zbudować teorię prawdy dla idiolektu drugiego człowieka, to po prostu zrozumieć jego wypowiedzenia, czyli odtworzyć znaczenia jego słów wraz ze wszystkimi zależnościami, które między nimi zachodzą. Układ owych zależności znaczeniowych da się bowiem wyrazić w kategoriach teorii prawdy, gdyż teoria ta jest rekurencyjna, ekstensjonalna oraz ma charakter holistyczny. Przejawia się to tym, że te same wyrażenia występują w różnych aksjomatach teorii. Podobnie jak używanie własnego idiolektu nie wymaga posiadania umiejętności artykułowania reguł gramatycznych, lecz może być opisane za pomocą owych reguł, tak interpretacja idiolektu innego człowieka nie wymaga umiejętności wyartykułowania aksjomatów teorii prawdy, ale może być za ich pomocą opisana, gdyż stanowi jedynie reprezentację lub model pewnego mechanizmu psychologicznego lub neurofizjologicznego, który aktualizuje nasze dyspozycje językowe.

W artykule *A Nice Derangements of Epitaphs* (1986), do którego powrócimy w paragrafie 8, Davidson porównuje teorię prawdy do „maszyny służącej do interpretowania wypowiedzeń”:

Interpretator posiada zatem system do interpretowania tego co słyszy lub mówi. Można traktować ten system jako maszynę, która dla dowolnego wypowiedzenia (oraz pewnych parametrów określonych przez okoliczności wypowiedzenia), produkuje interpretację. Jednym z modeli takiej maszyny jest teoria prawdy (...) (2005/1986, s. 95).

Należy jednak zwrócić uwagę, że nie chodzi tutaj o jedną, konkretną teorię prawdy, ale o pewien schemat teorii prawdy. Konkretna teoria prawdy jest bowiem modelem dla rezultatu konkretnego procesu interpretacji, zaś teoria prawdy, rozumiana jako ogólny schemat, pełni u Davidsona funkcję modelu kompetencji językowej interpretatora⁴⁴, która w danym przypadku aktualizuje się za pomocą konkretnej teorii:

W każdym razie stwierdzenia na temat tego, co konstituuje satysfakcjonującą teorię prawdy nie są ani stwierdzeniami na temat pojęciowej wiedzy (*proposi-*

⁴⁴ Nasuwa się tutaj analogia do użycia terminu „maszyna Turinga”. Niekiedy terminem tym określa się konkretny ciąg instrukcji, a niekiedy pewien ogólny opis, jak ów ciąg instrukcji powinien być zbudowany.

tional knowledge) interpretatora, ani stwierdzeniami na temat szczegółów pracy pewnych części jego mózgu. (...) jeśli jednak teoria poprawnie opisuje kompetencję interpretatora, to interpretator musi posiadać mechanizm, który jej odpowiada (s. 96).

Nieco wcześniej była mowa o tym, że teoria prawdy powinna być teorią dla całego idiolektu obejmującego interpretowanie zasłyszanych wypowiedzi, aproksymowaną na cały idiolekt. Interpretator z oczywistych względów nie ma bowiem dostępu do wszystkich możliwych wypowiedzi interpretowanego. Można zatem powiedzieć, że teorię prawdy budujemy zawsze dla hipotetycznego idiolektu. A zatem konstruując teorię prawdy wykraczamy poza interpretowane wypowiedzenie lub wypowiedzenia – interpretacja ma charakter holistyczny również w tym sensie, że dotyczy całego idiolektu.

W tym miejscu warto zwrócić uwagę na szczególny przypadek. Jak bowiem należy podejść do interpretacji oderwanych cytatów należących do nieznanego nam języka, dla których dysponujemy przekładem na język polski? Czy cytaty takie faktycznie interpretujemy w sensie podanym przez Davidsona? Rozpatrzmy słynne zdanie *Eppur si muove* wypowiedziane ongiś przez Galileusza⁴⁵. Wiele osób, które nie zna włoskiego, ale zna historię i okoliczności tego wypowiedzenia, powiedziałoby, że zdanie to znaczy, że Ziemia się porusza. Czy za przykładem tym stoi jakaś teoria prawdy, z której wynikałoby T-zdanie:

Zdanie „*Eppur si muove*” jest prawdziwe wtw, gdy Ziemia się porusza,
które z kolei można przekształcić do M-zdania:

Zdanie „*Eppur si muove*” znaczy, że Ziemia się porusza.

Ponieważ poprawnym tłumaczeniem tego zdania jest okazjonalne „Jednak się porusza”, z owej hipotetycznej teorii wynikałoby raczej T-zdanie:

(4.1) Wypowiedziane przez Galileusza zdanie na temat Ziemi „*Eppur si muove*” jest prawdziwe wtw, gdy Ziemia się porusza.

Na tak postawione pytanie możliwe są dwie odpowiedzi. Według pierwszej (4.1) jest konsekwencją teorii prawdy dla hipotetycznego idiolektu powstałego ze zdania *Eppur si muove* przez uzupełnienie go gramatyką oraz zasobem słów języka polskiego. Idiolekt ten jest jedynie aproksymacją zdania *Eppur si muove*, ale w razie potrzeby możemy go stopniowo modyfikować za pomocą reguł gramatycznych oraz słów języka włoskiego. Każde uzupełnienie

⁴⁵ Zdanie to przytaczam również dlatego, że posłużył się nim Davidson w artykule *On Saying That* (1968).

wymaga oczywiście rewizji odpowiedniej teorii prawdy. Druga możliwość została zasugerowana przez Kazimierza Ajdukiewicza w *O znaczeniu wyrażen* (1931). Osoba stojąca (w Polsce) przed sklepem z koronkami z napisem *Dentelles* faktycznie nie rozumie – zdaniem Ajdukiewicza – słowa *dentelles* jako słowa języka francuskiego. A zatem z punktu widzenia teorii interpretacji zdanie okazjonalne *Dentelles* w ogóle nie jest interpretowane jako zdanie języka francuskiego, lecz jako zdanie języka polskiego uzupełnionego wyrażeniem *dentelles* jako synonimem terminu „koronki”. Z teorii prawdy dla tak uzupełnionego języka polskiego wynikałoby T-zdanie:

Zdanie „*Dentelles*” jest prawdziwe wtw, gdy to są koronki.

Jak argumentuje Ajdukiewicz, samo skojarzenie tego napisu z koronkami nie wystarcza bowiem do uchwycenia jego znaczenia w języku francuskim, gdyż osoba ta nie uchwyciła związków znaczeniowych tego słowa z innymi wyrażeniami języka francuskiego. W przypadku zdania *Eppur si muove* sytuacja zdaje się być analogiczna. Zdanie to można bowiem potraktować jako włączony do języka polskiego, a dokładniej do idiolektu interpretatora, włoski synonim zdania okazjonalnego „Jednak się porusza” wypowiedzianego przez Galileusza.

Pierwsze rozwiązanie ma jednak przewagę nad drugim, gdyż o ile jesteśmy w stanie uznać, że pewne aksjomaty dla słów *eppur* oraz *moverse* na skutek przyzwyczajenia stają się częścią trywialnej teorii prawdy naszego własnego idiolektu, to trudno wyjaśnić w jaki sposób można rozszerzyć owe aksjomaty do teorii prawdy dla języka włoskiego, gdybyśmy postanowili się go nauczyć.

Holistyczny aspekt interpretacji zapewnia nam możliwość korygowania teorii prawdy, gdyż rekonstruując tę teorię, odkrywamy związki między znaczeniami nowych wyrazów i rolą tych wyrazów w warunkach prawdziwości zdań. Oczywiście nie zawsze jest to praktycznie możliwe. W przypadku tłumaczenia kilku ocalałych wypowiedzi (np. inskrypcji) zapomnianego już języka, dla których jednak posiadamy tłumaczenia na znany język z tej samej epoki, możemy nigdy nie móc zweryfikować odpowiedniej teorii prawdy. Nawet jednak w takim przypadku owa teoria prawdy jest zbudowana dla całego hipotetycznego idiolektu. Teoria ta ma faktycznie charakter szczątkowy, ale tłumacz zakłada, że nieliczne aksjomaty i wynikające z nich T-zdania, które jest w stanie z nich wyprowadzić, stanowią część aksjomatów i wynikających z nich T-zdań, które stanowiłyby teorię prawdy dla całego badanego języka, który z braku innej możliwości musiał jakoś uzupełnić odpowiednimi fragmentami własnego idiolektu. W przypadku odkrycia kolejnych napisów, mogą mieć miejsce dwie możliwości. Może się zdarzyć, że nowe aksjomaty

i wynikające z nich T-zdania „pasują” do istniejącej teorii i wystarczy zamienić trywialne aksjomaty dla polskich wyrażeń na nietrywialnie aksjomaty dla wyrażeń badanego języka. Może się zdarzyć, że dotychczasową teorię prawdy trzeba będzie zrewidować, tj. zmienić przynajmniej niektóre aksjomaty dla wyrażeń badanego języka. W obu przypadkach następuje jednak zmiana teorii prawdy.

Opisane wyżej przykłady mogą wydawać się sztuczne. Nie jest to jednak sytuacja, która zachodzi sporadycznie. Interpretując słowa drugiego człowieka, nawet użytkownika tego samego języka co my, budujemy teorię prawdy dla całego hipotetycznego idiolektu, chociaż faktycznie nie mamy dostępu do wszystkich możliwych wypowiedzi osoby interpretowanej. Dostępną nam próbkę wypowiedzi aproksymujemy na cały idiolekt tej osoby, uzupełniając go fragmentami naszego własnego idiolektu. W przeciwieństwie do opisanej poprzednio sytuacji kolejne próbki wypowiedzi osoby interpretowanej nie skłaniają nas za każdym razem do zmiany teorii prawdy, gdyż nasze idiolekty w dużej części po prostu się pokrywają. W takim przypadku odpowiednie aksjomaty teorii prawdy, jak również niektóre z wynikających z niej T-zdań, pozostają trywialne. Sprawę komplikuje fakt, że każdy z nas niekiedy się myli, lub przekręca słowa. Do sprawy interpretowania błędów, przejęczyń oraz pomyłek w idiolektie znanego nam języka powróć w paragrafie 7.

6. PRZEKŁAD RADYKALNY

W podanym wcześniej przykładzie budowanie teorii prawdy dla idiolektu języka włoskiego jest niezwykle ułatwione dzięki słownikom i podręcznikom będącym rezultatem pracy pokoleń filologów. Czy można jednak budować od podstaw teorię prawdy dla idiolektu zupełnie obcego języka? Aby odpowiedzieć na to pytanie Davidson odwołuje się do przekładu radykalnego, wprowadzając pojęcie interpretacji radykalnej. Pojęcie interpretacji radykalnej stanowi u Davidsona nie tylko centralne pojęcie jego późniejszej filozofii języka, ale jest uwikłane w poglądy epistemologiczne (omówione w paragrafach o perspektywie trzecioosobowej oraz triangulacji), filozofię umysłu, filozofię działania, teorię racjonalności, a nawet metafizykę.

Zanim zajmiemy się pojęciem interpretacji radykalnej, warto przyjrzeć się dokładniej niż do tej pory wprowadzonemu przez Quine’a w *Słowie i rzeczy* (1960) pojęciu przekładu radykalnego oraz teorii znaczenia w kategoriach podręcznika przekładu. Sam Davidson uważał, że interpretacja radykalna stanowi pewną wersję przekładu radykalnego, zaś teoria prawdy stanowi rekurencyjną wersję podręcznika przekładu. Przez przekład radykalny będziemy rozumieć przekład od podstaw lub też zgodnie z etymologią – przekład

„od korzeni”, czyli przekład z zupełnie nieznanego i niepodobnego do innych języka, który Quine nazywa językiem z dżungli. Pojęcie przekładu radykalnego zostało wprowadzone po to, aby opisując proces tworzenia podręcznika przekładu języka z dżungli można się było odwoływać jedynie do obserwowalnych bodźców oraz obserwowalnych reakcji użytkowników tego języka na owe bodźce, co pozostawało w zgodzie z empirystycznym, a dokładniej behawioralnym nastawieniem do językoznawstwa. Tworzący podręcznik przekładu dla języka z dżungli tłumacz skazany jest na posługiwanie się jedynie behawioralnymi kategoriami bodźca oraz reakcji na bodziec. Tym samym Quine pokazuje, że można – konstruując teorię znaczenia – zrezygnować z tradycyjnych pojęć językoznawczych, w szczególności zaś z pojęcia znaczenia jako intersubiektywnie dostępnego bytu, którego uchwycenie umożliwi komunikację językową⁴⁶. W tym sensie teoria znaczenia w kategoriach podręcznika przekładu staje się – podobnie jak semantyczna definicja prawdy u Davidsona – „teorią znaczenia bez znaczeń”.

Jednym z najbardziej znanych, a zarazem najbardziej kontrowersyjnych twierdzeń jest słynna teza o niezdeterminowaniu przekładu. Intuicyjnie można tę tezę rozumieć, że znaczenia przypisywane przez tłumacza zdaniom nieznanego języka z dżungli są niedookreślone przez ogół obserwowalnych bodźców. Jak pamiętamy z paragrafu poświęconego krytyce schematu pojęciowego, znaczenia zdań zależą u Quine’a od hipotez analitycznych stosowanych przez tłumacza, będących wyrazem jego indywidualnych schematów pojęciowych. Ponieważ Quine – podobnie jak Davidson – nie korzysta *explicite* z pojęcia znaczenia, formułuje tezę o niezdeterminowaniu, odwołując się do pojęcia podręcznika przekładu.

Oczywiście tworzenie podręcznika przekładu ma charakter eksperymentu myślowego i nie należy doszukiwać się jakichkolwiek realnych przykładów takich podręczników. Podręcznik przekładu, podobnie jak teoria prawdy, jest z jednej strony „nośnikiem” znaczenia, a z drugiej modelem pewnej kompetencji językowej. Aby budować podręczniki przekładu radykalnego tłumacze skazani są na dwa sposoby nawiązania kontaktu językowego z użytkownikami języka z dżungli. Po pierwsze, na kojarzeniu zdań wypowiedzianych przez tubylców z obserwowalnymi bodźcami; po drugie zaś, rozpoznawaniu oznak uznania lub odrzucenia przez tubylców zdań wypowiedzianych przez badacza w obecności obserwowalnych bodźców. Ponieważ drugi sposób nie jest wcale oczywisty, dodatkowo należy przyjąć, że użytkownicy języka faktycznie jednoznacznie uznają lub odrzucają znaczną większość zdań swego języka

⁴⁶ Warto w tym miejscu przypomnieć słowa Arystotelesa z *Hermeneutyki*: wrażenie (przedmiotu) doznawane w duszy jest takie samo dla wszystkich.

w określonych okolicznościach, że w większości przypadków nie kłamią, nie spiskują przeciw tłumaczowi i nie żartują z niego – innymi słowy Quine uznaje również pewną wersję Zasady życzliwości.

W znanym przykładzie ze *Słowa i rzeczy* tłumacz, kojarząc obecność białego królika z wypowiedzeniem przez tubylca zdania *Gavagai*, tłumaczy to wypowiedzenia na zdanie „(Oto) królik”. Nie wyklucza to jednak przekładów konkurencyjnych, na przykład „(Oto) zwierzę”, „(Oto) białe” lub „(Oto) białe zwierzę” itd. Aby dokonać właściwego wyboru tłumacz musi powtórzyć *Gavagai* w obecności tubylca, wskazując jednocześnie na inne zwierzęta, białe przedmioty oraz króliki o innym niż białe zabarwieniu sierści, obserwując bacznie oznaki uznania lub odrzucenia zdania *Gavagai*. Tłumacz formułuje zatem pewne hipotezy robocze, a następnie weryfikuje je w oparciu o dostępny materiał empiryczny. Początkowo procedura tłumaczenia ograniczona jest jedynie do pewnych rodzajów zdań, w ramach tzw. eksperymentalnej bazy przekładu, na którą składa się:

1. Tłumaczenie zdań obserwacyjnych;
2. Tłumaczenie środków językowych pełniących funkcję logiczną spójników prawdziwościowych;
3. Rozpoznanie, lecz nie tłumaczenie, zdań analitycznych bodźcowo;
4. Rozpoznanie, lecz nie tłumaczenie, par zdań synonimicznych bodźcowo.

Pojawiające się tu pojęcia składają się na wprowadzoną przez behawioralną namiastkę semantyki, którą Quine nazywa „semantyką wegetariańską”. Podstawowym pojęciem semantyki wegetariańskiej, które pozwala na stworzenie teorii znaczenia w kategoriach podręcznika przekładu, jest znaczenie bodźcowe zdania, które jest parą złożoną ze znaczenia bodźcowego pozytywnego oraz znaczenia bodźcowego negatywnego. Przez pierwsze z nich rozumie się ogół bodźców, wobec których zdanie jest uznawane; przez drugie ogół bodźców, wobec których zdanie jest odrzucane przez wyidealizowanego, lecz typowego i kompetentnego użytkownika języka, który nie kłamie, nie wprowadza badacza w błąd oraz nie uczestniczy w spisku, którego celem jest oszukanie tłumacza itd. Oczywiście eksperymentalna baza przekładu nie obejmuje wszystkich rodzajów zdań występujących w języku z dżungli.

Zdania oznajmujące dzieli Quine na zdania okazjonalne, np. „Oto królik”, zdania wieczne (*standing sentences*), np. „Wiosna się już zaczęła”, oraz inne rodzaje zdań, o których będzie mowa dalej. Zdania wieczne mogą być potwierdzane lub odrzucane niezależnie od bodźców dostępnych w momencie ich wypowiedzenia, np. przytoczone wyżej zdanie wieczne przez pewien określony czas jest powszechnie uznawane (niezależnie od bodźców) a przez inny określony czas powszechnie odrzucane (również niezależnie od bodź-

ców). Z kolei zdania okazjonalne są uznawane lub odrzucane w zależności od bodźców dostępnych w momencie ich wypowiedzenia. Spośród zdań okazjonalnych wyróżnia Quine zdania obserwacyjne, które określa jako zdania, dla aprobaty lub dezaprobaty których, znający język użytkownik potrzebuje jedynie obecności bodźca, nie potrzebuje natomiast żadnej dodatkowej wiedzy np. „(Oto) królik”, „(Oto) czerwone”, „Boli mnie”, „On ma brudną twarz”. Pozostałe zdania okazjonalne, tj. zdania których uznanie lub odrzucenie nie zależy jedynie od bodźców, lecz od pewnej wiedzy pozajęzykowej, to tzw. zdania nieprzejryste poznawczo, np. „(Oto) dziekan”⁴⁷. Dlatego wśród wyróżnionych do tej pory zdań, jedynie zdaniom obserwacyjnym można przypisać określone znaczenia bodźcowe, co umożliwia ich bezpośredni przekład z jednego języka na drugi. Zdaniom wiecznym i zdaniom nieprzejrystym poznawczo nie da się przypisać znaczeń bodźcowych i dlatego nie mogą być przetłumaczalne jedynie na podstawie obserwacji zachowań językowych w obecności dostępnych bodźców. O przekładzie musi zatem decydować coś innego – jakiś inny czynnik, który nie może zostać wyrażony w kategoriach behawioralnych.

Wśród zdań wiecznych Quine wyróżnia dodatkowo tzw. zdania analityczne bodźcowo oraz kontradyktoryczne bodźcowo. Pierwsze są uznawane, a drugie odrzucane przez wszystkich kompetentnych użytkowników języka (przy wymienionych poprzednio zastrzeżeniach) zawsze, wszędzie i wobec wszystkich dostępnych bodźców. Rzecz jasna zdaniami analitycznymi bodźcowo będą nie tylko zdania analityczne w sensie tradycyjnym, ale również wszystkie zdania, których prawdziwość nie jest w danej społeczności przez nikogo kwestionowana, np. „Króliki są wcieleniami dusz zmarłych przodków”. Jeśli jest to wierzenie głęboko ugruntowane i niekwestionowane, to nie da się powyższego zdania odróżnić od zdań uważanych po prostu za analityczne. W oparciu o intersubiektywnie dostępne bodźce tłumacz jest zazwyczaj w stanie jedynie rozpoznać⁴⁸, że pewne zdania są analityczne

⁴⁷ Aby wyjaśnić osobie, która nie zna organizacji szkolnictwa wyższego, znaczenie zdania „(Oto) dziekan”, nie wystarczy jedynie wskazać osobę, która pełni tę funkcję, gdyż jeśli pokazywani jej do tej pory dziekani posiadali brody, może dojść do wniosku, że „dziekan” oznacza brodatego mężczyznę.

⁴⁸ Nie wnikając w szczegóły oraz pomijając krytykę zdań analitycznych przeprowadzoną przez Quine’a w *Dwóch dogmatach*, zdania analityczne są to zdania, których prawdziwość zależy jedynie od znaczeń występujących w nich terminów, nie zaś od przygodnych faktów. Zdanie wieczne „Krokusy już wzeszły” nie jest analityczne, gdyż ich prawdziwość zależy od pewnego przygodnego faktu, tj. faktu pojawienia się krokusów. Oczywiście zdanie to może być zakwalifikowane jako analityczne bodźcowe, gdyż użytkownicy mogą je uznawać w obecności wszelkich bodźców. Oczywiście tak być nie musi, gdyż w pewnym momencie, np. zimą, zdanie to zacznie z powrotem być odrzucane. Przykład ten świadczy o tym, że nawet kwalifikacja

bodźcowo lub kontrydiktoryczne bodźcowo. Nie jest jednak w stanie dokonać ich tłumaczenia przez odwołanie się jedynie do znaczeń bodźcowych. Początkowo może bowiem wysunąć hipotezę, że zdanie „Króliki są wcieleniami dusz zmarłych przodków” powinno być tłumaczone za pomocą zdania „Dwa plus dwa równa się cztery”. Co zatem decyduje o tym, że ta hipoteza ostatecznie zostanie odrzucona w podręczniku przekładu? Odpowiedź na to pytanie jest oczywista. Podręcznik przekładu podaje tłumaczenia zdań, ale zdania te nie są wyrwane z kontekstu całego języka, gdyż przekład ma charakter holistyczny. Jest tak dlatego, że różne zdania zawierają, jako swe części, te same słowa, co musi być w przekładzie uwzględnione. Tak rozumiane związki znaczeniowe między zdaniami pozwalają na wykluczenie bardzo wielu początkowych hipotez dotyczących przekładu.

Dla zdań okazjonalnych nieobserwacyjnych oraz zdań wiecznych nieanalitycznych i niekontrydiktorycznych bodźcowo, tłumacz może, w ramach eksperymentalnej bazy przekładu, stwierdzić jedynie synonimiczność bodźcową, czyli identyczność pozytywnych i negatywnych znaczeń bodźcowych. Przykładami zdań synonimicznych bodźcowo mogą być „Oto kawaler” i „Oto mężczyzna nieżonaty”, mimo że zdania te są w ramach bazy eksperymentalnej przekładu nieprzetłumaczalne. Ponadto obserwując zachowania językowe tubylców, tłumacz może znaleźć w ich języku wyrazy bądź konstrukcje, które pełnią funkcje spójników logicznych: negacji, koniunkcji i alternatywy. Umożliwia to tłumaczenie „struktury” logicznej zdań, a w szczególności praw logiki.

Wprawdzie eksperymentalna baza przekładu nie wystarcza do tego, aby dokonać przekładu z jednego języka naturalnego na drugi, stanowi ona jednak podstawę przekładu oraz jedyne dostępne intersubiektywne kryterium jego poprawności. Warto w tym miejscu odwołać się do metafory teorii jako pola sił z *Dwóch dogmatów empiryzmu*, która miała ilustrować tezę holizmu konfirmacyjnego. Jest to uzasadnione tym, że język naturalny jest faktycznie namiastką teorii, w którą uwikłane są elementarne, przednaukowe przekonania o świecie. Eksperymentalna baza przekładu umożliwia tłumaczenie zdań peryferyjnych, czyli zdań obserwacyjnych, oraz zdań centralnych, czyli praw logiki⁴⁹. Tłumaczenie „środka” owego pola sił dodatkowo zależy od nieuwarunkowanych empirycznie schematów pojęciowych tłumacza, które znajdują wyraz w hipotezach analitycznych. W tym momencie przecho-

zdania przez tłumacza do określonej grupy może być przedmiotem rewizji w ramach procesu przekładu radykalnego.

⁴⁹ Faktycznie pozwala utożsamić schematy praw logiki języka z dżungli ze schematami praw logiki języka tłumacza oraz formę logiczną zdań języka z dżungli z formą logiczną zdań języka tłumacza.

dzi się z poziomu niższego – behawioralnego – na poziom podręcznika przekładu, który można określić mianem poziomu znaczeniowego.

Kontynuując pracę nad podręcznikiem przekładu, tłumacz może weryfikować rezultaty swojej pracy, formułując kolejne zdania w obecności informatora i obserwując jego reakcje wobec dostępnych bodźców. Jednak w miarę oddalania się weryfikowanych zdań od brzegu „pola sił”, możliwość eksperymentalnego korygowania przekładu traci na znaczeniu. W miarę testowania coraz to nowych hipotez dotyczących przekładu następuje dopasowywanie znaczeń przypisywanych wyrażeniom tłumaczonego języka do zdań już przetłumaczonych, przy czym dopasowanie to zależy w coraz większym stopniu od schematów pojęciowych tłumacza. Znaczenia przypisane wyrażeniom języka z dżungli są zatem wypadkową hipotez analitycznych tłumacza oraz obserwowalnych zachowań językowych użytkowników wobec dostępnych bodźców. Tak rozumiany układ znaczeń ma charakter zdecydowanie holistyczny. Co więcej, tak konstruowany podręcznik przekładu zakłada istnienie związków znaczeniowych między wyrazami tłumaczonego języka, chociaż związki te – w odróżnieniu od teorii prawdy – nie są reprezentowane przez rekurencyjne aksjomaty.

Punktem wyjścia teorii znaczenia w kategoriach podręcznika przekładu jest semantyczny odpowiednik holizmu konfirmacyjnego. Zdania wyrażające nasze przekonania są potwierdzane lub obalane nie pojedynczo, lecz w ramach całego układu przekonań ukształtowanego przez schemat pojęciowy tłumacza. Stąd już blisko do poglądu, że przekonania drugiego człowieka są faktycznie dla nas niedostępne, co uniemożliwiałoby interpretację jego wypowiedzi. Wobec powyższej konsekwencji przeprowadzona przez Davidsona krytyka schematu pojęciowego „wypełnionego” empiryczną zawartością staje się całkowicie zrozumiała.

Przyjmowane przez Quine’a założenia prowadzą wprost do tezy o niezdeterminowaniu przekładu, w myśl której konkurencyjne systemy hipotez analitycznych dwóch tłumaczy mogą być zgodne z bazą eksperymentalną przekładu, ale niezgodne między sobą. Innymi słowy dwa systemy hipotez analitycznych mogą być zgodne ze wszystkimi dyspozycjami do zachowań językowych (tj. werbalnego reagowanie na bodźce), a mimo to mogą prowadzić do dwóch podręczników przekładu niezgodnych, a nawet wzajemnie sprzecznych w przypadku zdań, dla których nie istnieją bodźce pozwalające na ich uznanie lub odrzucenie.

Problematyce niezdeterminowania przekładu poświęcona jest bardzo bogata literatura, co sprawia, że jest to jeden z najpopularniejszych tematów współczesnej filozofii języka. Nie zmienia to faktu, że teza o niezdeterminowaniu jest bardzo trudna do sformułowania i sam Quine wielokrotnie ją

precyzował, często w reakcji na słowa krytyki. Spośród wielu wersji tezy za stosunkowo najmniej kontrowersyjne uchodzi sformułowanie w kategoriach podręcznika przekładu:

System zasad (podręcznik) przekładu języka z dżungli na nasz język jest indukcyjną definicją relacji przekładu wraz z twierdzeniem, że koreluje ona zdania w zgodzie z odpowiednimi zachowaniami. Teza o niezdeterminowaniu przekładu głosi, że takie twierdzenia o dwóch podręcznikach mogą być równocześnie prawdziwe, podczas gdy ustanowionych przez nie relacji przekładu nie można używać na przemian bez otrzymania zdań niespójnych. Mówiąc inaczej: zdania naszego własnego języka wskazane przez dwa konkurencyjne podręczniki przekładu jako przekłady danego zdania języka z dżungli nie mogą być wzajemnie wymienne we wszystkich kontekstach naszego języka (Quine 1993/1990, s. 125–126).

Jeżeli odrzucimy – jak czynili to Quine i Davidson – zreifikowane znaczenia, które mogą być w jakiś sposób uchwycone lub poznane, to znaczenie, które przestaje być bytem, staje się znacznie mniej uchwytnie gdyż musi być podane w ramach „czegoś innego”, np. podręcznika przekładu lub teorii prawdy. W takim przypadku nie możemy jednak bezpośrednio porównywać znaczeń. Niezdeterminowanie, zarówno w wersji Quine’a, jak i Davidsona, jest konsekwencją sposobu konstruowania teorii znaczenia – jak pamiętamy, Davidson uznawał bowiem tezę o niezdeterminowaniu interpretacji. Ponieważ Quine starał się wyeksplikować tezę o niezdeterminowaniu, odwołując się do pojęcia przekładu radykalnego, przyjrzymy się teraz wprowadzonemu przez Davidsona pojęciu interpretacji radykalnej.

7. INTERPRETACJA RADYKALNA

Wbrew pozorom interpretacja radykalna nie jest jedynie eksperymentem myślowym⁵⁰. W przeszłości bardzo często, obecnie z powodu postępującej globalizacji coraz rzadziej, mieliśmy do czynienia z sytuacją pierwszego spotkania przedstawicieli zupełnie różnych społeczności językowych. Jeśli do takiego spotkania jednak dochodzi, to pomimo braku jakichkolwiek podobieństw między językami, braku tłumaczy, słowników czy jakichkolwiek innych środków ułatwiających komunikację, interpretacja okazuje się jednak możliwa. Po pewnym czasie pojawiają się bowiem tłumacze, którzy mogą przekładać wypowiedzenia jednego języka na drugi. Interpretacja radykalna

⁵⁰ Eksperymentem myślowym jest natomiast podręcznik przekładu.

jest zatem faktem. Problemem pozostaje natomiast jej opis w kategoriach Davidsonowskiej semantyki.

W przełomowym artykule *Interpretacja radykalna* (1973) Davidson podejmuje problemy, które można dostrzec w załączkowej postaci już w *Prawdzie i znaczeniu*, a które następnie rozwija w *Reply to Foster* (1976). W *Interpretacji radykalnej* Davidson zawarł również krytykę teorii znaczenia w kategoriach podręcznika przekładu oraz, co interesuje nas w tym paragrafie najbardziej, opisał jak interpretacja radykalna jest w ogóle możliwa. Krytyka poglądów Quine'a⁵¹ rozpoczyna się od uwagi, że teoria interpretacji nie może być – w ścisłym tego słowa znaczeniu – teorią przekładu. Teoria przekładu angażuje bowiem:

(...) trzy języki, czyli język przedmiotowy, język podmiotowy i metajęzyk (język z którego [tłumaczymy], język na który tłumaczymy, i język teorii, która mówi, które wyrażenia języka przedmiotowego są przekładami jakich wyrażeń języka przedmiotowego). W ogólnym przypadku można wiedzieć, jakie zdania języka podmiotowego są przekładami, których zdań języka przedmiotowego, nie wiedząc, co którekolwiek zdanie każdego z tych języków znaczy (w jakimkolwiek sensie, który pozwalałby komuś, kto rozumie tę teorię, interpretować zdania języka przedmiotowego) (1992/1973, s. 101).

Jeśli język podmiotowy jest identyczny z językiem teorii, to ktoś, kto rozumie teorię przekładu może użyć podręcznika przekładu do interpretowania wypowiedzi języka przedmiotowego. Okazuje się jednak, że w takim przypadku sama teoria przekładu (podręcznik przekładu) nie wystarcza do skutecznej interpretacji. Interpretator korzysta bowiem z dwóch rzeczy, których sama teoria przekładu nie stwierdza: „(...) z faktu, że język podmiotowy jest jego własnym językiem i ze swojej wiedzy o tym, jak interpretować wypowiedzi własnego języka” (s. 102).

W konsekwencji teoria przekładu nie nadaje się na teorię interpretacji, gdyż założenie, że przytoczone zdanie należy do czyjegoś języka „nie może mieć miejsca w teorii, która powinna być przydatna dla każdego, kto dokonuje tłumaczenia” (*ibidem*). Dodajmy również, że interpretacja wypowiedzi musi uwzględniać składalność znaczeń, której nie daje nam żadna teoria przekładu (w sensie, w jakim teorię przekładu rozumiał Davidson). W przypadku interpretacji języka przedmiotowego konieczne jest bowiem „odsłonięcie jego struktury semantycznej”. A zatem stwierdzenie, że teoria przekładu pozwala na interpretowanie wypowiedzi w języku przedmiotowym, nawet w poda-

⁵¹ Krytyka poglądów Quine'a jest bardzo subtelna i ma zazwyczaj postać polemiki z przypisywanymi mu poglądami, nie wymieniając go z nazwiska.

nym przypadku, jest mylące. Faktycznie podręcznik przekładu może jedynie wspomóc proces właściwej interpretacji:

Teorię interpretacji języka przedmiotowego można zatem uważać za rezultat połączenia teorii interpretacji znanego języka, która odsłania jego strukturę, oraz systemu przekładu z nieznanego języka na obcy (s. 103).

Owo połączenie – jak stwierdza Davidson – pozwala na odrzucenie odwołania do znanego języka. Wówczas otrzymamy teorię interpretacji, która ujawnia strukturę języka przedmiotowego. Oczywiście teoria ta musi być wyrażona w jakimś języku i przybiera postać Tarskiego definicji prawdy⁵². Reszta omawianego artykułu poświęcona jest uzasadnieniu, że odpowiednio dostosowana do potrzeb języka naturalnego teoria prawdy może służyć jako teoria interpretacji. Argument na rzecz powyższego stwierdzenia przybiera postać odpowiedzi na trzy pytania:

1. Czy rozsądnie jest sądzić, że dla języka naturalnego można podać teorię prawdy tego rodzaju, jaką opisaliśmy?
2. Czy można byłoby stwierdzić, że taka teoria jest poprawna, na podstawie wiarygodnych świadectw osiągalnych dla interpretatora nie znającego uprzednio języka, który ma być zinterpretowany?
3. Gdybyśmy wiedzieli, że teoria jest prawdziwa, czy byłoby możliwe interpretowanie wypowiedzi osób mówiących tym językiem? (s. 105).

Ponieważ na temat problemu poruszonego w pierwszym pytaniu była już wielokrotnie mowa, w dalszym ciągu ograniczę się do odpowiedzi, które udzielił Davidson na pytania drugie i trzecie.

Czy można zatem zweryfikować teorię prawdy, odwołując się do świadectw dostępnych nam zanim rozpoczęliśmy proces interpretacji? Pytanie to jest faktycznie pytaniem o sam początek interpretacji radykalnej, a dokładniej – pytaniem o pozajęzykowe świadectwa, które pozwalają rozpocząć proces interpretowania słów drugiego człowieka posługującego się językiem całkowicie nam nieznanym. To, czym dysponuje interpretator zgodnie ze swymi dyspozycjami językowymi, to możliwość interpretowania słów drugiego człowieka za pomocą teorii prawdy dla jego idiolektu. Teoria prawdy rozumiana jako „teoretyczna maszyna”, będąca modelem naszej zdolności

⁵² Celem tego dość zawiłego i na pierwszy rzut oka niejasnego argumentu jest, jak sądzę, pokazanie, że interpretowanie wypowiedzi języka przedmiotowego za pomocą podręcznika przekładu nie jest możliwe. Jego warunkiem jest zbudowanie teorii prawdy. Podstawowy powód jest następujący: podręcznik przekładu nie podaje struktury semantycznej zdań języka przedmiotowego, gdyż nie jest rekurencyjny.

do interpretacji, nie może być jednak weryfikowana bezpośrednio – to co podlega weryfikacji, to wynikające z niej T-zdania, które jednak nie zawierają informacji na temat owej „maszyny”, tj. aksjomatów (s. 108). Pozostaje zatem otwarte pytanie o to, w jaki sposób konstruujemy pierwszą wersję teorii prawdy dla nieznanego nam języka, którą następnie możemy weryfikować w oparciu o wynikające z niej T-zdania.

Jak pamiętamy, w zależności od sformułowania konwencja T zakładała wymóg, aby metajęzyk obejmował język przedmiotowy albo przekład z języka przedmiotowego na metajęzyk:

W pracy Tarskiego T-zdania uznawane są za prawdziwe, ponieważ zakłada się, że prawa część równoważności jest przekładem zdania, dla którego podaje się warunek prawdziwości (*ibidem*).

W konsekwencji T-zdania uznawane są za twierdzenia teorii prawdy, które pozwalają określić zakres predykatu „prawdziwy”: jeśli na podstawie jakichś kryteriów uznajemy prawą stronę równoważności, to zdanie po lewej stronie uznajemy za prawdziwe. Procedura ta służy, jak chciał tego Tarski, do wyznaczania zakresu predykatu „prawdziwy” w języku przedmiotowym, lecz nie jest przydatna w przypadku interpretacji radykalnej:

Nie możemy jednak z góry zakładać, że przekład da się uznać za poprawny bez przesądzenia kwestii radykalnej interpretacji; w zastosowaniach empirycznych musimy zrezygnować z tego założenia (s. 108–109).

Dlatego też – o czym była już wielokrotnie mowa – Davidson potraktował prawdę jako pojęcie proste i niedefiniowalne, co dało mu możliwość wyjaśnienia przekładu i interpretacji. W takim przypadku T-zdania nie służą już do ustalania zakresu terminu „prawdziwy”, lecz do weryfikowania teorii prawdy. Może się bowiem okazać w świetle dostępnych świadectw, że wynikające z teorii T-zdanie, czyli twierdzenie teorii prawdy, jest fałszywe, co stanowi powód do rewizji całej teorii, nie zaś do wykluczenia zdania po lewej stronie równoważności z zakresu predykatu „prawdziwy”. Rewizja teorii stanowi jednocześnie rewizję przekładu⁵³. Dlatego też już w *Prawdzie i znaczeniu* Davidson utrzymywał, że teoria prawdy w jego ujęciu jest teorią empiryczną.

⁵³ W takim przypadku nastąpi zmiana znaczenia zdania po lewej stronie równoważności. Przy nowym znaczeniu (równokształtne) zdanie może okazać się prawdziwe.

Zaproponowane przez Davidsona przeformułowanie teorii prawdy wymaga również przeformułowania konwencji T, aby ta nie odwoływała się do pojęcia przekładu:

(...) zadowalająca teoria prawdy musi pociągać, dla każdego zdania *s* języka przedmiotowego, zdanie postaci: *s* jest prawdą wtedy i tylko wtedy, gdy *p*, gdzie „*p*” można zastąpić jakimkolwiek zdaniem, które jest prawdziwe wtedy i tylko wtedy, gdy prawdziwe jest *s* (s. 109).

Zauważmy, że w sformułowaniu tym, podobnie jak w oryginalnym sformułowaniu pochodzącym od Tarskiego, zawarte jest odwołanie do holistycznego charakteru teorii prawdy. Możemy bowiem wyobrazić sobie teorię prawdy, z której wynikają następujące T-zdanie:

Zdanie „*Snow is white*” jest prawdziwe wtw, gdy trawa jest zielona.

Zdanie to w izolacji nie jest podatne na weryfikację prowadzącą do rewizji teorii prawdy, gdyż trawa faktycznie jest zielona, zaś anglojęzyczny użytkownik języka przedmiotowego zapewne uzna zdanie *Snow is white* za prawdziwe. Przeformułowana konwencja T głosi jednak, że z zadowalającej teorii prawdy muszą wynikać wszystkie zdania spełniające podany warunek. Jeżeli zatem znajdziemy T-zdanie, które uznamy za fałszywe, to musimy zrewidować całą teorię prawdy. Takie T-zdania mogą odpowiadać albo zdaniom obserwacyjnym, albo – o czym Davidson w *Interpretacji radykalnej* wyraźnie nie wspomina – zdaniom mającym charakter powszechnie akceptowanych prawd, np. zdaniom analitycznym, prawdom matematycznym lub aksjomatycznym dyrektywom znaczeniowym w sensie Ajdukiewicza. Aby uznać T-zdanie za fałszywe, musimy bowiem założyć zgodę między mówiącym a interpretatorem. W przypadku zdań odpowiadającym danym wrażeniowym sytuacja jest oczywista. Rozpatrzmy następujące T-zdanie, które wynika z tej teorii:

Zdanie w idiolekcie *x*-a „*This is snow*” (wypowiedziane przez *x*-a lub interpretatora) w momencie *t* jest prawdziwe wtw, gdy w momencie *t* osoba *x* lub interpretator wskazuje na trawę⁵⁴.

To, czy przytoczone T-zdanie jest prawdziwe czy fałszywe zależy od tego, czy interpretowany zaakceptuje zdanie *This is snow*. Jeśli tak, to powyższe T-zdanie jest prawdziwe i musimy uznać – choćby tymczasowo – że *snow* oznacza w jego idiolekcie trawę. Jeśli nie, będzie to oznaczało, że teorię

⁵⁴ Postać przytoczonego T-zdania ma uwzględnić możliwość testowania teorii prawdy przez interpretatora.

prawdy należy zrewidować, gdyż wynikające z niej T-zdanie okazało się fałszywe.

W przypadku powszechnie uznanych prawd sprawa nie jest aż tak prosta, gdyż bywają prawdy powszechnie uznawane w jednej społeczności, lecz nie uznawane w innych, np. „Króliki są wcieleniami dusz przodków”. Wyobraźmy sobie jednak, że z teorii prawdy wynika następujące T-zdanie:

Zdanie „*Snow melts in heat*” jest prawdziwe wtw, gdy trawa się topi w cieple.

Zdanie to pozwala na rewizję teorii prawdy. Podobnie jak w poprzednim przypadku decydujące jest uznanie lub odrzucenie przez rozmówcę zdania *Snow melts in heat*. W pierwszym przypadku powyższe T-zdanie jest fałszywe (gdyż interpretowany uznaje zdanie *Snow melts in heat*, lecz interpretator nie uznaje prawej strony), w drugim przypadku zaś prawdziwe.

Dzięki temu teoria prawdy w nowej postaci może nie tylko podawać przykład wypowiedzeń zdań języka przedmiotowego na metajęzyk, lecz może faktycznie służyć do czegoś więcej – do interpretowania wypowiedzeń. W szczególnym przypadku, np. zdań bez elementów okazjonalnych, interpretacja sprowadza się do przekładu. Jednak w przypadku wypowiedzeń zdań okazjonalnych, a te w języku naturalnym zdecydowanie przeważają, interpretacja jest czymś więcej niż przekład – jest wyrazem tego, co mówiący powiedział. W artykule *Przekonania a podstawy znaczenia* Davidson pisze: „Interpretujemy czyjeś zachowanie językowe wtedy, gdy mówimy co znaczą słowa wypowiedziane przez tę osobę w danej sytuacji” (1992/1974b, s. 118). Zwrot „znaczą wypowiedziane słowa”, to tyle co „mówiący powiedział” – pod warunkiem, że mówił w sposób dosłowny, szczerzy, niemetaforyczny i bez pomyłek:

Wyrażeniem „mówi, że” posługuję się tu w taki sposób, że dana osoba mówi (w określonej sytuacji), iż śnieg pada, zawsze i tylko wtedy, gdy używa słów, które (w tej sytuacji) znaczą, że śnieg pada. Tak więc, ktoś może powiedzieć, że śnieg pada, podczas gdy *on sam* nie ma przy tym na myśli, ani nie twierdzi, że śnieg pada (*ibidem*, przyp. 1)⁵⁵.

⁵⁵ To, co mówiący miał na myśli, wypowiadając dane słowa, zostało omówione przez Davidsona w *A Nice Derangement of Epitaphs* (1986). W artykule tym Davidson omawia problem pomyłek i przejęczyń w sytuacji, gdy interpretator zna już idiolekt osoby interpretowanej (tj. dysponuje odpowiednią teorią prawdy). W takim przypadku interpretacja przebiega dzięki pewnym standardom znaczeniowym dla interpretowanego idiolektu. Standardy znaczeniowe można nazywać konwencjami; nie są to jednak konwencje konstytuujące język, lecz poznane regularności użycia wyrażenń znanego idiolektu, które znajdują odzwierciedlenie w teorii prawdy. Warto wspomnieć, że dla Davidsona, w odróżnieniu od Grice’a i Searle’a, to co

Jak zatem zaczyna się proces interpretacji radykalnej? Wedle Davidsona początkiem interpretacji radykalnej jest postawa uznawania zdania za prawdziwe, a dokładniej rozpoznanie przez interpretatora, że ktoś wypowiadając zdanie zamierza wyrazić prawdę, choć interpretator nie wie jaka to prawda (1992/1973, s. 110). Jest to konsekwencja problemu, który Davidson szczegółowo opisuje w *Przekonaniach a podstawach znaczenia*:

Kwestia tą jest natura świadectw na rzecz adekwatności teorii interpretacji. Dla celów, o które nam chodzi, świadectwa te muszą dać się opisać w terminach niesemantycznych i nielingwistycznych; muszą to być świadectwa, których posiadanie nie zakłada posiadania teorii, na rzecz której mają świadczyć (1992/1974b, s. 122).

Nie mogą to być zatem ani konkretne przekonania, ani znaczenia, gdyż uznanie zdania za prawdziwe odbywa się częściowo na podstawie znaczenia, a częściowo na podstawie przekonania. Aby zatem rozpocząć interpretację radykalną:

(...) musimy w jakiś sposób osiągnąć równocześnie teorię przekonania i teorię znaczenia. Jak jest to możliwe? (...) W miejsce wypowiadania wyrażen wezmę pod uwagę pewną postawę wobec wyrażen, postawę, która może towarzyszyć lub nie towarzyszyć rzeczywistym wypowiedziom. Mam na myśli postawę, zrelatywizowaną do czasu, uznawania za prawdziwe (s. 123).

Aby jednak interpretacja mogła faktycznie oprzeć się na tym założeniu, należy dodatkowo założyć Zasadę życzliwości, w formie zbliżonej do maksymy jakości Gricea (zob. 1975), która domaga się, aby rozmówcy mówili prawdę (czyli wypowiadali zdania zgodne z ich przekonaniami), a zarazem stwierdza, że w przeważającej ilości przypadków rozmówcy mówią prawdę. Przy tym założeniu (którego Davidson *explicite* nie wymienia) oraz założeniu, że mamy do czynienia z tym samym środowiskiem zewnętrznym można budować teorię prawdy, która maksymalizuje zgodę między mówiącym a interpretatorem (por. 1992/1973, s. 112). Zgoda ta dotyczy jedynie wypowiedzi wymienionych wcześniej typów zdań, gdyż w wielu przypadkach zgoda nie jest wymagana. Mówiący oraz interpretator nie muszą się zgadzać w kwestiach faktów z przeszłości, ocen, wierzeń itd. Również nie muszą się zawsze zgadzać w przypadku zdań, które określiłem bardzo ogólnikowo jako zdania mające charakter powszechnie akceptowalnych prawd, gdyż poszcze-

mówiący miał na myśli, nie jest domeną komunikowania metaforycznego, lecz dosłownego. Znaczenie dosłowne nie jest jednak konstytuowane przez konwencje, lecz „odkrywane” w procesie interpretacji. Do sprawy tej powrócę w następnym paragrafie.

gólne społeczności mogą się pod tym względem różnić. Można jednak rozsądnie przyjąć, że powszechna zgoda ma miejsce w przypadku prawd tzw. fizyki popularnej, np. „Kamienie są twarde”, „Woda jest mokra” itd., lub prawd biologii popularnej, np. „Wróble latają”, „Ryby żyją w wodzie” itp.

Proces budowania teorii prawdy przypomina pod wieloma względami proces budowania podręcznika przekładu, jednak w wielu szczegółach od niego odbiega – przede wszystkim, ze względu na rekurencyjność teorii prawdy, dla której zachodzi konwencja T: „(...) najpierw szukamy najlepszej drogi dopasowania naszej logiki do nowego języka w tym zakresie, abyśmy otrzymali teorię spełniającą konwencję T” (*ibidem*). Nie jest to jednak – jak to było u Quine’a – przede wszystkim kolejne rozpoznawanie spójników logicznych, lecz:

Może to oznaczać wczytanie struktury rachunku kwantyfikatorów pierwszego rzędu (z identycznością) w ten język, nie zaś wskazanie, jedna po drugiej, stałych logicznych, lecz traktując tę logikę jako siatkę, którą należy jednym posunięciem dopasować do języka. Świadczeniami są tutaj klasy zdań zawsze uznawane za prawdziwe lub zawsze uznawane za fałszywe przez prawie wszystkich, prawie w każdym czasie (potencjalne prawa logiczne), oraz reguły wnioskowania (*ibidem*).

W kroku tym, który pozwala ustalić formę logiczną zdań interpretowanego idiolektu, „rozpoznaje się predykaty, terminy indywiduowe, kwantyfikatory, spójniki i identyczność” (*ibidem*).

Owo rozpoznanie formy logicznej odbywa się całościowo, czy jak pisze Davidson – jednym rzutem. Jest to spowodowane nie tylko holistycznym i rekurencyjnym charakterem teorii prawdy, ale holistycznym charakterem sieci przekonań. Jak pamiętamy z rozdziału drugiego, Davidson odrzucił charakterystyczne dla empirystów utożsamienie porządku uczenia się języka (w tym przypadku dochodzenia do pierwszej teorii prawdy dla interpretowanego idiolektu) z porządkiem logicznym czy epistemologicznym (w tym przypadku weryfikowania istniejącej teorii prawdy i budowania teorii alternatywnej). Dlatego też wymienionych powyżej etapów dochodzenia do teorii prawdy nie należy traktować jako opisu tworzenia tej teorii. Zdaniem Davidsona owa pierwsza teoria prawdy „wyłania się” w ogólnym zarysie, a następnie ulega kolejnym uzupełnieniom i modyfikacjom.

Davidson uważał, że zarówno przekonania mówiącego, jak i przekonania interpretatora tworzą do pewnego stopnia koherentne sieci, a zatem obaj

muszą posiadać pewną „namiastkę logiki”⁵⁶, która przejawia się m.in. nieakceptowaniem oczywistych sprzeczności. Owa struktura logiczna układów przekonań znajduje również swoje odbicie w praktyce używania języka, gdyż mówienie – o czym była przed chwilą mowa – jest zazwyczaj mówieniem prawdy, czyli mówieniem zgodnie z własnymi przekonaniem. Stosując nieco inną terminologię, mówienie polega w dużej części na aktualizowaniu własnych przekonań:

(...) wzajemna zależność przekonania i znaczenia może być uwidoczniła w ten sposób: użytkownik języka uznaje dane zdanie za prawdziwe ze względu na to, co to zdanie znaczy (w jego języku) i ze względu na to jakie ma przekonania (s. 110).

W *Ku jednolitej teorii znaczenia i działania* stwierdza, że interpretacja jest możliwa jedynie „pod warunkiem ustalenia pewnej struktury” między znaczeniami oraz między postawami propozycjonalnymi (przekonaniami) (1992/1980, s. 150). Związki między znaczeniami wyrażen języka przedmiotowego są konstytuowane przez teorię prawdy, lecz „po stronie przekonań nie jest równie jasne, jaką strukturę należy założyć” (*ibidem*):

(...) jednakże nie można zaakceptować zbyt dużych i oczywistych odchylen od racjonalności, jeśli przypisywanie przekonań ma być zrozumiałe. (...) Przede wszystkim musimy rzutować na przekonania owego człowieka naszą własną logikę (s. 150–151).

Pierwszy, wstępny krok interpretacji radykalnej polega zatem na rzutowaniu na wypowiedzenia i na przekonania rozmówcy pewnej struktury logicznej, co jest równoznaczne z przypisaniem mu pewnej dozy racjonalności. Dopiero wtedy możliwy jest drugi krok, który:

(...) skupia się na zdaniach z wyrażeniami wskazującymi [zdaniach okazjonalnych]; zdania te są czasem prawdziwe, a czasem fałszywe, stosownie do wykrytych w świecie zmian (s. 113).

W kroku tym pojawia się możliwość interpretowania, a przynajmniej pojawiają się ograniczenia na możliwe interpretacje poszczególnych predykatów. Jest tak dzięki temu, że znamy już, przynajmniej w zarysie, strukturę logiczną interpretowanego języka. Zauważmy, że oba etapy obejmują zdania, co do

⁵⁶ Davidson był przekonany, że pod powierzchnią form języka kryje się jego struktura logiczna, odpowiadająca logice predykatów pierwszego rzędu z identycznością. Struktura ta rzutowana jest na język przez logiczną strukturę układów (sieci) przekonań.

których prawdziwości lub fałszywości mówiący i interpretator mają zazwyczaj tę samą opinię.

Ostatni etap obejmuje zdania, co do których przekonania mówiącego oraz interpretatora mogą się różnić lub „których wartość logiczna nie zależy w sposób systematyczny od zmian w otoczeniu” (*ibidem*). W zamieszczonym na tej samej stronie przypisie Davidson porównuje krótko swój opis interpretacji radykalnej do opisu radykalnego przekładu z drugiego rozdziału *Słowa i rzeczy*, zwracając uwagę na pewne semantyczne ograniczenia, które „wymuszają na interpretowanym języku strukturę kwantyfikatorową”. Przez owe ograniczenia należy rozumieć wyuczalność języka, której konsekwencją jest rekurencyjność teorii języka. Ponieważ w rezultacie procesu interpretacji powstaje teorii prawdy, rekurencyjność tej teorii, która była zaprojektowana dla języków opartych na języku rachunku predykatów, faktycznie wymusza w sposób jednoznaczny strukturę kwantyfikatorową w języku, dla którego teoria prawdy jest budowana. W artykule *Przekonania a podstawy znaczenia* pisze: „[...] jeśli Konwencja T jest spełniona, to zapewniona zostaje jedyność struktury kwantyfikatorowej” (1992/1974b, s. 138).

Davidson zwraca również uwagę, że w jego opisie nie ma miejsca dla pojęcia znaczenia bodźcowego, gdyż „jego miejsce zajmuje odwołanie się do obiektywnych cech świata” (1992/1973, s. 113, przyp. 13). Jest to wyraźne odwołanie się do różnicy w poglądach między nim a Quinem co do rodzaju świadectw empirycznych, które powinny być uwzględniane w teorii języka. Jak pamiętamy Quine był zwolennikiem tzw. teorii proksymalnej, wedle której należy mówić nie o realnie istniejącym świecie, lecz o działających na nas bodźcach. Davidson był z kolei zwolennikiem tzw. teorii dystalnej, wedle której to właśnie realnie istniejący świat jest tym, co jest przedmiotem ludzkiego poznania⁵⁷.

Metoda interpretowania, którą naszkicował Davidson, „pracuje” przy założeniu Zasady życzliwości, która w jednej ze swych licznych wersji domaga się przyznania racji „(...) – kiedy to tylko możliwe – rodzimym użytkownikom języka” (s. 113). Davidson podaje typowe i wielokrotnie powtarzane uzasadnienie Zasady życzliwości, które odwołuje się do faktu, że można mówić o zgodzie lub niezgodzie przekonań jedynie, gdy istnieje „rozległa zgoda” (*ibidem*). Owa rozległa zgoda stanowi również warunek możliwości interpretacji radykalnej:

⁵⁷ Davidson podjął dyskusję z teorią proksymalną Quine’a w *Meaning, Truth, and Evidence* (1990), starając się jednocześnie zbagatelizować różnicę między obiema teoriami.

Zastosowana do języka, zasada ta głosi; że im więcej zdań zgodnie przyjmujemy lub odrzucamy (przy pośrednictwie lub bez pośrednictwa interpretacji), tym lepiej rozumiemy resztę, niezależnie od tego, czy się co do nich zgadzamy (*ibidem*).

Przywołana tu powszechna zgoda da się wywieść z triangulacyjnego modelu uczenia się języka, który wskazuje, że wyrażona w języku zgoda na temat świata stanowi warunek nabycia pierwszego języka (idiolektu). Nie oznacza to, że uczeń musi zawsze i w każdej kwestii zgadzać się ze swym nauczycielem, ale bez „rozległej” zgody nie ma najmniejszej szansy na nauczenie się języka. Słyszac, że pewne obiekty w naszym otoczeniu nauczyciel nazywa w określony sposób, uczeń początkowo zgadza się ze swoim nauczycielem. A zatem źródłem Zasady życzliwości (we wszystkich jej wariantach i sformułowaniach) jest ostatecznie wspólne środowisko zewnętrzne, czyli obiektywnie istniejący świat oraz triangulacyjny model uczenia się języka. Oczywiście zarówno proces uczenia się języka, jak i proces interpretacji nie odbywa się bez pomyłek. Jednak o pomyłce można mówić jedynie wtedy, gdy w przeważającej ilości przypadków ma się rację.

Davidson utrzymywał, że bez języka nie może być przekonań, a zatem Zasada życzliwości wyprowadzona z faktu istnienia obiektywnego i dostępnego naszemu poznaniu świata, stanowi faktycznie nie tylko warunek interpretacji, lecz również warunek wszelkiego myślenia⁵⁸:

Nie powinno się sądzić, że metodologiczna zasada, by interpretować [wypowiedzenia drugiego człowieka] w sposób maksymalizujący zgodę, opiera się na życzliwym założeniu o ludzkiej inteligencji, które może okazać się fałszywe. Jeżeli nie możemy znaleźć takiego sposobu zinterpretowania wypowiedzi i innych zachowań pewnej istoty żywej, aby wyłonił się zbiór przekonań, które są – według naszych własnych kryteriów – w większości spójne i prawdziwe, to nie mamy powodu uważać tej istoty za racjonalną, mającą przekonania czy mówiącą cokolwiek (s. 113–114).

⁵⁸ W wielu artykułach Davidson podkreślał, że warunkiem myślenia jest posiadanie języka (por. 1974a i 1982). Pogląd ten, który związany jest z Davidsonowską filozofią psychologii, da się również uzasadnić na gruncie triangulacji. Warunkiem wytworzenia się sieci przekonań jest interakcja z drugim człowiekiem oraz ze wspólnym im środowiskiem. W konsekwencji warunkiem istnienia przekonań oraz myślenia jest historia przyczynowa owej interakcji. Początkowo historia ta obejmuje proces nauki języka, a następnie interakcje polegające na interpretowaniu innych ludzi, które również odbywają się przy oddziaływaniu przyczynowym środowiska. Bez historii przyczynowej nie ma zatem myślenia, a w konsekwencji – jak twierdzi Davidson – nawet najlepszy komputer nie może myśleć, chociaż nie wyklucza on możliwości nabycia tej zdolności, jeśli będzie miała miejsce historia interakcji komputera z ludźmi oraz ze środowiskiem (por. 1990a).

Na zakończenie należy wyjaśnić dokładniej niż to czyniłem wcześniej, w jakim sensie znajomość warunków, w jakich użytkownicy języka uznają zdania za prawdziwe wraz z posiadaniem „zadowolającej teorii” pozwala na zinterpretowanie każdego zdania (s. 115). Sama teoria bowiem podaje tylko warunki prawdziwości wyrażone w T-zdaniach. Trzeba zatem pokazać, że narzucenie na teorię wymienionych wcześniej ograniczeń, w szczególności Zasady życzliwości, pozwala na interpretację. Żadne T-zdanie z osobna nie pozwala na zinterpretowanie zdania po lewej stronie równoważności. Aby mogło stanowić interpretację owego zdania, należy uwzględnić proces tworzenia całej teorii prawdy. Zasada życzliwości mówi nam, że w większości przypadków to, co mówiący powiedział, jest tym, co mówiący miał na myśli.

Samo określenie „co mówiący miał na myśli” jest moją propozycją tłumaczenia pochodzącego od Grice’a terminu *speaker’s meaning*. Wydaje się, że wyrażenie „co mówiący miał na myśli” dobrze oddaje istotę interpretacji zdań, gdyż odwołuje się zarówno do znaczeń, jak i przekonań, ale nie odwołuje się do odrzucanych przez Davidsona konwencji, które miałyby konstituować język i znaczenia wyrażen. Co więcej pojęcie to pozwala w pewnych przypadkach na odróżnienie tego co wypowiedziane słowa znaczą od przekonań mówiącego. Jest tak w przypadku pomyłek i przejęzyczeń, o których będzie mowa w paragrafie 8. Jednak nawet jeśli mówiący mówi to, co ma na myśli, to T-zdanie nie musi podawać przekładu w wąskim tego słowa znaczeniu, o czym świadczy przypadek zdań okazjonalnych. W przypadku owych zdań widać szczególnie wyraźnie, że interpretacja musi uwzględnić kontekst wypowiedzenia i związane z nim świadectwa empiryczne, gdyż jedynie w określonym kontekście można przypisać mówiącemu przekonanie, zaś przypisanie przekonań wymaga spełnienia przez teorię wymienionych wcześniej kryteriów formalno-logicznych. Dlatego też właściwie utworzona teoria prawdy nie może prowadzić do T-zdań podających jedynie przekład zdań okazjonalnych:

Zdanie „*This is snow*” jest prawdziwe wtw, gdy to jest śnieg.

W konsekwencji teoria, która pozwala na interpretowanie słów drugiego człowieka, tj. w większości przypadków na przypisanie mu tego, co ma na myśli wypowiadając dane słowa, musi spełniać zarówno kryteria formalne, jak i empiryczne (s. 117). Aby teoria ta spełniała owe kryteria, przy jej tworzeniu musi obowiązywać Zasada życzliwości, którą można wywieść z owych kryteriów⁵⁹. Sam termin „zasada” ma wyraźny wydźwięk normatywny, ale

⁵⁹ Widać to najwyraźniej, odwołując się do sformułowania Zasady życzliwości w postaci Zasady korespondencji (kryteria empiryczne) i Zasady koherencji (kryteria formalno-logiczne).

„Zasada życzliwości” jest faktycznie i zazwyczaj bezwiednie przestrzegana najogólniejszą strategią interpretacyjną.

Oczywiście ktoś może podnieść zarzut, że nie każde wypowiedzenie rozmówcy faktycznie wyraża jego przekonanie. Zdarzają się bowiem kłamstwa, przypadki udawania (np. aktora na scenie), żarty, wypowiedzenia niedosłownie (metafora lub ironia) oraz pomyłki. Aby jednak móc przypisać dosłowne znaczenia takim wypowiedzeniom, tj. wiedzieć co mówiący faktycznie powiedział, interpretator musi lub musiał wcześniej – zgodnie z Zasadą życzliwości – poprawnie zinterpretować przeważającą ilość wypowiedzeń, tj. uchwycić nie tylko co mówiący powiedział, ale również co miał na myśli, wypowiadając te słowa.

W tym miejscu konieczne jest pewne wyjaśnienie, aby nie pomieszać stanowiska Davidsona z poglądami Grice’a, tym bardziej, że użyłem występującego w tłumaczeniach prac Grice’a terminu „co mówiący miał na myśli”. Interpretowanie wypowiedzeń odbywa się w ramach holistycznej teorii prawdy, przy wynikającym z Zasady życzliwości założeniu, że wypowiedzenia zdań w większości przypadków aktualizują przekonania przypisywane interpretowanemu, czyli w przypadku, gdy to, co powiedział mówiący, pokrywa się z tym, co miał na myśli. Teoria ta ma charakter rekurencyjny, co zapewnia zachodzenie związków znaczeniowych między wypowiedzeniami. Zinterpretowane wyrażenie może być użyte do innych celów, a wówczas następuje oddzielenie tego co mówiący powiedział, od tego co miał na myśli. Najlepszym przykładem jest – o czym już była mowa – użycie metaforycznych wyrażen. Zdaniem Davidsona metafory nie interpretuje się jednak w ramach teorii prawdy, a zatem w ścisłym tego słowa znaczeniu nie interpretuje się jej w ogóle, lecz wykorzystuje się jedynie zinterpretowane wcześniej zdanie, aby osiągnąć pewien efekt przyczynowy. W szczególności efekt ten może polegać na zasugerowaniu słuchaczowi pewnego przekonania, które żywi mówiący. Z kolei wedle Grice’a słuchacz dochodzi do przypisania mówiącemu przekonania w pewien sposób wyrażonego przez metaforę, czyli tego, co wypowiadający metaforę miał na myśli, poprzez implikowanie konwersacyjne, czyli poprzez pewnego rodzaju rozumowanie. Wedle Davidsona efekt oddziaływania metafory, nie jest wynikiem jakiegokolwiek rozumowania ani interpretacji, lecz jedynie skutkiem dosłownego jej zinterpretowania w danym kontekście. Różnica jest tu bardzo istotna, gdyż Davidson odróżniał starannie relację przyczynową, zachodzącą między zdarzeniami, od relacji logicznych, zachodzących między zdaniami.

8. INTERPRETACJA W PRAKTYCE

Proces interpretacji realizuje się na wiele sposobów. Do tej pory omówiliśmy dwa z nich: ściśle związany z pojęciem triangulacji proces uczenia się pierwszego języka oraz interpretację radykalną. W pierwszym przypadku interpretacyjny aspekt procesu uczenia się języka polega na przypisywaniu rozmówcy wiedzy o świecie wyrażanej w języku owego rozmówcy, z czym związane jest przypisywanie znaczeń jego słowom i poznawanie w ten sposób tego języka. W przypadku interpretacji radykalnej Davidson pokazuje warunki, które umożliwiają zinterpretowanie przez osobę, która zna już jakiś język, języka całkowicie jej nieznanego. Wbrew deklaracjom poczynionym na początku rozdziału, celowo użyłem terminu „język” nie zaś „idiolekt”, aby zaznaczyć, że w obu przypadkach użycie terminu „język” ma sens, pod warunkiem, że rozumie się go jako klasę podobnych idiolektów, w których obowiązują pewne standardy interpretacyjne. Celem nauki pierwszego języka jest bowiem opanowanie idiolektu możliwie najbardziej zbliżonego do należącego do pewnego języka idiolektu nauczyciela (lub nauczycieli), zaś celem interpretacji radykalnej jest zinterpretowanie pewnej klasy idiolektów, które składają się na nieznaną nam język.

Oparte na standardach interpretacyjnych pojęcie języka okazuje się jednak bezużyteczne w przypadku codziennej interpretacji osób, o których mówimy, że posługują się tym samym językiem co my. Problem ten podjął Davidson w artykule *A Nice Derangements of Epitaphs* (1986)⁶⁰. Na wstępie wyjaśnienia wymaga sam tytuł artykułu, którego dosłowny przekład na język polski mógłby brzmieć „Miłe szaleństwo epitafiów”. W tytule tym Davidson nawiązuje do tzw. malapropizmów⁶¹, czyli pewnego rodzaju przekręceń i przejęzyczeń, z którymi spotykamy się na co dzień, i o ile w ogóle zwracamy na nie uwagę, to najczęściej nie mamy wątpliwości co mówiący miał na myśli. W przytoczonym tytule termin *derangement* zastąpił przez pomyłkę *arrangement*, czyli „ułożenie”, zaś termin *epitaphs* zastąpił *epitets*, czyli „epitety” lub „określenia”. W odpowiednim kontekście niemal każdy czytelnik jest nazwyczaj w stanie zinterpretować wyrażenie występujące w tytule jako „Miła aranżacja określeń”.

Z malapropizmami spotykamy się bardzo często i nie stanowią one zazwyczaj dla nas bariery w komunikacji językowej, gdyż zazwyczaj jesteśmy w stanie odkryć intencje mówiącego. Malapropizmy nie obejmują przypadków

⁶⁰ Artykuł ten omawiany jest dokładnie w (Nowakowski 1977).

⁶¹ Termin „malapropizm” pochodzi od nazwiska bohaterki napisanej w 1775 roku komedii autorstwa Richarda Sheridan’a zatytułowanej *The Rivals*. Pani Malaprop często zamieniała wyrazy na inne, lecz zazwyczaj podobnie brzmiące, co wywoływało efekt komiczny.

celowego wprowadzania w błąd, lecz stanowią zazwyczaj skutek ewidentnego przejęzyczenia, niedbałego sposobu wypowiadania się lub powierzchowności wykształcenia. Popętnienie malapropizmu może wprowadzić ośmieszyć mówiącego, lub wprowadzić słuchacza w przejściowe zakłopotanie, lecz w trakcie dłuższej rozmowy nie ulega zazwyczaj wątpliwości, o co chodziło mówiącemu. Oznacza, że słowa mówiącego zostały faktycznie zinterpretowane. Malapropizmy zdarzają się wprowadzić bardzo często, ale rzadko następuje ich nagromadzenie w jednej wypowiedzi. Davidson rozpoczyna swoje rozważania od przytoczenia takiego właśnie osobliwego przypadku:

Zamiast przykładać, że Ace mówi wprost, słuchacz musi uważać na niespodziankowe *tet a tet*, albo umarł w rękawiczkach. W okrągły sposób będzie on manewrował szukając idealnego wyrażenia dopóki nie wybierze frazy dla danej sytuacji, uderzając prosto w sedno. Wyborowy rozmówca może próbować zatrzeć się z nim w walcu na słowa. W dążeniu do osiągnięcia ekscesu, wielokrotnie przeżyłem mózg, aby osiągnąć spekulacyjny sukces, ale Ace spijał śmietankę, którą sam sobie naważył. Od czasu do czasu Ace w żarliwy sposób monopolował rozmowę błyskotliwą uwagą, zbyt słoną aby ją wspominać. Pobicie go w jego własnej grze jest czarnym krukiem i nigdy mi się to nie udało. Pozostaje tylko zatkać uszy i zapaść pod wodę, bo on zawsze wygrywa z ręką w nocniku (2005/1986, s. 89)⁶².

Przekład tego, co autor mógł mieć na myśli, może wyglądać mniej więcej tak:

Raczej nie zakładając, że Ace mówi wprost, słuchacz musi uważać na nagle słowne utarczki albo umarł w butach. W okrężny sposób, będzie manewrował szukając idealnego wyrażenia, dopóki nie wybierze frazy odpowiedniej dla danej sytuacji, trafiając w sedno. Wytrawny rozmówca może próbować wdać się z nim z nim w walkę na słowa. W dążeniu do wejścia na szczyt powodzenia, wielokrotnie wysilałem rozum, aby osiągnąć spektakularny sukces, ale Ace zawsze był górą. Od czasu do czasu Ace w żartobliwy sposób monopolizował rozmowę błyskotliwą uwagą, zbyt frywolną, aby ją przytaczać. Bardzo rzadko się zdarza, aby ktoś go pokonał w jego własnej grze; w każdym razie mi się to

⁶² Nie jest łatwo tłumaczyć malapropizmy z jednego języka na drugi, gdyż bardzo często zamienione słowa są podobne pod względem kształtu fizycznego w jednym języku, zaś ich odpowiedniki w drugim języku nie są podobne – w podanym przykładzie słowo *pinnacle* (szczyt) zamienione zostało na *pinochle* (rodzaj gry w karty). Propozycja obu tłumaczeń przykładu cytowanego przez Davidsona została sformułowana dzięki pomysłowości uczestników zajęć translatorskich „Wybrane teksty z filozofii języka”, które prowadziłem dla studentów filozofii II stopnia na Uniwersytecie Łódzkim w semestrze letnim roku 2017/18. W tłumaczeniu oraz jego interpretacji staraliśmy się oddać kolokwialny język cytatu.

niegdy nie udało. Pozostaje zamknąć oczy i zapaść się pod ziemię, a on i tak zawsze wygrywa z palcem w nosie⁶³.

W przypadku słów o podobnym kształcie i zbliżonym znaczeniu (np. „adoptować” i „adaptować”) często nie zdajemy sobie sprawy, że mamy do czynienia z malapropizmem i w zależności od kontekstu interpretujemy go w sposób niezgodny z zamiarami mówiącego bez jakiegokolwiek szkody dla komunikacji językowej. Za przykład mogą nam posłużyć słowa „adaptować” i „adoptować” w stosunku do pomysłów, idei oraz teorii, które w wielu kontakstach mogą być używane wymiennie.

Poza malapropizmami Davidson odwołuje się do wprowadzonego przez Keitha Donnellana rozróżnienia na referencyjne i atrybutywne użycie deskrypcji określonych (por. Donnellan 1966). W języku angielskim deskrypcjami określonymi są wyrażenia poprzedzone przedimkiem określonym *the* (*the present president of Poland*)⁶⁴ lub dopełniacza imion własnych (*John's wife*). Teoria deskrypcji, którą zaproponował Bertrand Russell w artykule *On Denoting* (1905) jest teorią atrybutywnego użycia deskrypcji. Polega ono na tym, że deskrypcja oznacza jeden i tylko jeden obiekt, który spełnia opis podany w deskrypcji⁶⁵. Na przykład deskrypcja „obecna królowa Anglii”

⁶³ W podanej parafrazie zdecydowałem się, wraz z uczestnikami translatorium, na zachowanie kolokwialnego stylu cytatu. Należy jednak podkreślić, że fragment ten jest naszą rekonstrukcją tego, co mówiący miał na myśli, gdyż sam Davidson takiej parafrazy nie podaje. Co więcej malapropizmów nie da się tłumaczyć słowo w słowo, gdyż przekład powinien uwzględnić podobieństwo do pomyłonego wyrażenia.

⁶⁴ Aby odróżnić deskrypcję od nazwy ogólnej, w języku polskim musimy odwołać się do kontekstu i odpowiednio zinterpretować wypowiedzenie. Przypadek języka polskiego, w którym nie ma odpowiednika rodzajników określonych, doskonale ilustruje stanowisko prezentowane przez Davidsona, że wszelka interpretacja (nawet w ramach tego samego języka) polega nie tylko na przypisaniu wypowiedzeniom znaczeń, ale również na przypisaniu mówiącym przekonani.

⁶⁵ Faktycznie Russell eliminował deskrypcje ze struktury logicznej języka, gdyż słynne zdanie „Obecny król Francji jest łysy” w postaci kanonicznej było koniunkcją trzech zdań:

Istnieje przynajmniej jeden obecny król Francji.

Istnieje co najwyżej jeden obecny król Francji.

Każdy obecny król Francji jest łysy.

W strukturze logicznej zdania nie ma zatem miejsca na deskrypcje określone. O użyciu atrybutywnym deskrypcji jest zatem sens mówić jedynie na poziomie struktury powierzchniowej języka lub po wprowadzeniu operatora deskrypcji, który został użyty przez Davidsona w jednej z wersji Argumentu procy.

oznacza Elżbietę II, gdyż jest jedynym istniejącym obiektem, który w momencie pisania tego tekstu był królową Anglii⁶⁶.

Donnellan zwrócił uwagę, że bardzo często używamy deskrypcji, aby odnieść się do jakiegoś indywiduum, chociaż nie spełnia ono opisu podanego w deskrypcji. Co więcej, mówiący często wie, lub przynajmniej przypuszcza, że obiekt nie spełnia podanego w deskrypcji opisu. Jest to referencyjne użycie deskrypcji, w którym traktowana jest ona jako narzędzie do wyróżnienia indywiduów w otoczeniu. W oryginalnym przykładzie podanym przez Donnellana niejaki Jones zostaje oskarżony o zamordowanie Smitha. Ponieważ Jones sprawia wrażenie osoby chorej psychicznej, wypowiadamy zdanie „Morderca Smitha jest szaleńcem” (*Smith's murderer is insane*). Donnellan twierdzi, że nawet w przypadku, gdy Jones nie zamordował Smitha, deskrypcja *Smith's murderer* spełnia swoją funkcję, gdyż odnosi się do Jonesa. W konsekwencji przytoczone zdanie głosi coś prawdziwego o Jonesie, chociaż Jones nie spełnia opisu podanego w użytej w zdaniu deskrypcji. Jednak traktując deskrypcję w sposób atrybutywny, przytoczone zdanie w ogóle nie jest o Jonesie. Jeśli zatem Smitha zabił ktoś inny, wówczas zdanie to głosi, że zupełnie inny człowiek, ten który zabił Smitha, jest szaleńcem. Jeśli Smitha nikt nie zamordował (wyobraźmy sobie, że sam Smith upozorował własną śmierć, aby ukryć się przed wierzycielami), to zgodnie z teorią Russella zdanie to jest po prostu fałszywe⁶⁷. Problem podniesiony przez Davidsona polega na tym, że istnieje wyraźna rozbieżność między sposobem rozumienia podanego zdania, a tym, co na ten temat głosi teoria Russella.

Podobnie jak w przypadku malapropizmów, zazwyczaj nie mamy większych problemów z właściwym zinterpretowaniem deskrypcji. Jeśli na przyjęciu powiemy coś człowiekowi, który popija coś ze szklanki z napisem *Martini*, używając deskrypcji „Człowiek, który pije *Martini*”, to nie ma znaczenia czy w jego szklance jest ten, czy też inny napój. Oczywiście bardzo często deskrypcja użyta jest zarówno referencyjnie, jak i atrybutywnie⁶⁸ – tak by było, gdyby wspomniany człowiek był jedyną na przyjęciu osobą, która w tym momencie popija *Martini*. Czasami jednak wyraźnie odgadujemy, kiedy mówiący używa deskrypcji referencyjnie, a kiedy atrybutywnie, jak w przypadku zdania „Szkoda, że jej mąż jest jej mężem” (Lycan 2000, s. 29). W standardowym rozumieniu pierwsze użycie deskrypcji „jej mąż” ma charakter referencyjny (odnosi się do określonego indywiduum), zaś drugie

⁶⁶ Ciekawą dyskusję na temat teorii deskrypcji oraz rozróżnienia dokonanego przez Donnellana można znaleźć w (Lycan 2000).

⁶⁷ Bo fałszywe jest zdanie „Istnieje przynajmniej jeden morderca Smitha”.

⁶⁸ Przykładem użycia jedynie atrybutywnego jest podany przez Russella przykład „obecny król Francji”.

użycie – charakter atrybutywny (tj. odnosi się do jedynego indywiduum spełniającego określony opis). Dlatego też pierwsze użycie „jej mąż” można zastąpić nazwą własną tej osoby, zaś drugiego nie. W niektórych przypadkach różnica w zamierzonym użyciu deskrypcji może prowadzić do chwilowego nieporozumienia, jak w przytoczonym przez Lycana dialogu, w którym jest mowa o przyjaciółce jednego z rozmówców:

(...) inną sprawą jest to, że ona nie jest moją żoną.

To prawda. Mało kto jest Kitty.

Głupcze, to nie Kitty jest tym, kim ona nie jest. To kim ona nie jest, to jest moja żona. A to jest zupełnie coś innego (...) (s. 27).

Tego typu nieporozumienia mogą być – co wyraźnie widać z przebiegu cytowanej rozmowy – łatwo eliminowane.

Podobnie jak w przypadku malapropizmów, z punktu widzenia intuicyjnego sprawa zdaje się być niezwykle prosta: w zależności od okoliczności interpretujemy deskrypcje zazwyczaj w sposób zgodny z intencjami mówiącego (tj. albo w sensie atrybutywnym, albo w sensie referencyjnym). Zupełnie inny problem stanowi zbudowanie teorii, która w sposób adekwatny opisze ten fakt. W przypadku semantyki Davidsona chodzi o dostosowanie jego teorii interpretacji do przypadków malapropizmów oraz różnych sposobów użycia deskrypcji.

Aby zarysować problem, Davidson odwołuje się do MacKay’a krytyki rozróżnienia na referencyjne i atrybutywne użycia deskrypcji (1968) oraz do odpowiedzi udzielonej przez Donnellana na tę krytykę (1968). MacKay zarzuca Donnellanowi, że rozróżnienie na użycie referencyjne i atrybutywne prowadzi do teorii znaczenia w stylu *Humpty Dumpty*. Bohater *Po drugiej stronie zwierciadła* twierdził, że wypowiadając jakieś słowo on sam decyduje w jakim znaczeniu zostało ono użyte. Na przykład – jak powiedział *Humpty Dumpty* do Alicji – użyte przed chwilą przez niego słowo „pochwała” (*glory*) znaczyło tyle co „argument nie do odparcia” (*a nice knockdown argument*) (por. Carrol 2005, s. 170). Faktycznie podana przez Donnellana charakterystyka użycia referencyjnego może nasuwać podejrzenie, że mówiący stosuje teorię znaczenia w stylu *Humpty Dumpty*, gdyż dany obiekt może być oznaczany za pomocą różnych deskrypcji, mimo że nie posiada własności podanych w owych deskrypcjach. W odpowiedzi na ten zarzut Donnellan pisze:

Gdybym miał zakończyć odpowiedź MacKay’owi zdaniem „pochwalam cię”, byłbym winny arogancji oraz – niewątpliwie – zbyt wysokiej oceny tego, co powiedziałem. Jednak biorąc pod uwagę całe tło nie sądzę, aby można mnie było oskarżyć o to, że powiedziałem coś niezrozumiałego. Czyż nie byłbym zro-

zumiany i czy mówiąc „pochwalam cię” nie miałbym na myśli „argument nie do odparcia” (1968, s. 213)⁶⁹.

Davidson zgadza się z MacKay'em, że nie można zmieniać znaczeń słów zgodnie z własnym widzimisie. Można to robić, gdy istnieją podstawy, aby wierzyć, że interpretator ma wskazówki (*clues*) dla nowej interpretacji (2005/1986, s. 98). Wskazówki te zawarte są zazwyczaj w szeroko rozumianym kontekście rozmowy, który znajduje swoje odbicie w przekonaniach rozmówców. W przypadku malapropizmów wskazówką interpretacyjną dla słuchacza jest nie tylko kontekst, ale podobieństwo wypowiedzianego wyrażenia do wyrażenia, które zdaniem słuchacza powinno być przez mówiącego wypowiedziane.

W jaki sposób Davidson adaptuje swą teorię interpretacji codziennych konwersacji⁷⁰? Jak pamiętamy, Tarski zbudował swą semantyczną definicję prawdy dla języków, w których sens wyrażen wyznaczony jest jednoznacznie przez ich kształt fizyczny. Założenie to nie mogło być przez Davidsona utrzymane, choćby dlatego, że adaptował definicję prawdy do wypowiedzi zdań okazjonalnych. W takim przypadku takie same co do fizycznego kształtu wypowiedzenia mogą mieć zupełnie inne warunki prawdziwości. W przypadku malapropizmów oraz różnych sposobów użycia deskrypcji sprawa się jeszcze bardziej komplikuje. Dlatego też Davidson odchodzi od wcześniejszego wymogu, że kompetencja użytkownika języka musi być reprezentowana przez teorię prawdy. Teoria prawdy staje się jedynie jednym z możliwych modeli kompetencji językowej. Inne możliwe modele muszą jednak spełniać warunki narzucone w *Theories of Meaning* na teorie języków wyuczalnych:

W każdym momencie wymiany zdań, interpretator musi posiadać coś, co upieram się nazywać teorią. (Czynię tak gdyż opis kompetencji językowej interpretatora musi być rekurencyjny) (s. 100).

Kompetencja językowa jest pojęciem, które wprowadził i rozpropagował Noam Chomsky. Pojęcie to było od czasów Chomsky'ego powszechnie akceptowane przez lingwistów oraz filozofów języka. Kompetencją w tym rozumieniu miała być wiedza, którą muszą posiadać mówiący oraz interpretator jako warunek możliwości komunikowania się. Wyrazem owej kompetencji mało być znaczenie dosłowne, czy też lepiej – znaczenie standardowe wyrażen

⁶⁹ Cytuję za (Davidson 2005/1986, s. 97).

⁷⁰ Malapropizmy oraz różne sposoby użycia deskrypcji, to jedynie niektóre z problemów, z którymi musimy się zmierzyć w naszych codziennych rozmowach. Spośród pozostałych Davidson zwraca przede wszystkim uwagę na ustawiczne wzbogacanie języka o nowe nazwy własne.

języka. Oczywiście w świetle dotychczasowych rozważań widać, że znaczenia poszczególnych wypowiedzi mogą odbiegać od tak rozumianych znaczeń dosłownych. Jednak przystępując do procesu interpretacji czynimy pewne założenie co do znaczeń słów mówiącego, przypisując im coś, co można określić terminem „znaczenia pierwsze” (*first meaning*). Problem, który postawił Davidson, można wyrazić pytaniem: czy jakaś rekurencyjna teoria, np. teoria prawdy, może odpowiadać tak rozumianej kompetencji. W szczególności, czy kompetencji tej może odpowiadać teoria prawdy wyrażająca znaczenia pierwsze. W tym przypadku Davidson nawiązuje wyraźnie do Grice’a, który opisał proces komunikacji językowej w kategoriach intencji komunikacyjnych⁷¹. Wedle Grice’a mówiący może komunikować to, co ma na myśli (*speakers meanig*) nie wprost (por. Grice 1975). Aby słuchacz zinterpretował odpowiednio jego słowa (tj. zgodnie z intencją), musi znać znaczenie dosłowne (u Davidsona znaczenie pierwsze) oraz kontekst wypowiedzenia. Dopiero wtedy może implikować konwersacyjnie, co mówiący ma na myśli⁷².

W *A Nice Derangement of Epitaphs* Davidson zdecydowanie krytykuje to podejście, również dlatego, że jego dotychczasowe prace mogłyby sugerować, że on sam był jego zwolennikiem. Zdaniem Davidsona krytykowany pogląd da się wyrazić w trzech punktach:

- (1) *Pierwsze znaczenie jest systematyczne* [*systematic*]. Kompetentny użytkownik języka może interpretować wypowiedzenia – swoje własne lub innych – na podstawie semantycznych własności ich części lub słów występujących w wypowiedzeniach oraz na podstawie struktury wypowiedzenia. Aby to było możliwe, musi zachodzić systematyczny związek między znaczeniami wypowiedzi.
- (2) *Pierwsze znaczenia są współdzielone*. Aby mówiący i interpretator mogli się skutecznie i regularnie komunikować, muszą współdzielić metodę interpretacji, opisaną w punkcie (1).
- (3) *Pierwsze znaczenia są oparte na wyuczonych konwencjach lub regularnościach*. Systematyczna wiedza lub kompetencja jest opanowywana zanim interpretacja ma miejsce i ma charakter konwencjonalny (s. 93).

W dalszej części artykułu Davidson pokazuje, że tak zrekonstruowane, tradycyjne i – jak utrzymuje – powszechnie uznawane podejście do problemu kompetencji językowej nie jest w stanie wyjaśnić fenomenu komunikacji językowej.

⁷¹ Klasycznym artykułem, w którym Grice analizuje znaczenie w kategoriach intencji komunikacyjnych, jest *Meaning* (1957). Ważniejsze prace Grice’a zostały zebrane w (1989).

⁷² Teoria Grice’a zapoczątkowała tzw. teorię komunikowania metaforycznego rozwiniętą przez Searle’a (1979) oraz Martinicha (1984). Pamiętać należy, że Davidson był autorem zupełnie innej teorii metafory (teoria przyczynowa).

Mogło by się zdawać, że to właśnie teoria prawdy jest jednym z modeli owej kompetencji, którą Davidson charakteryzuje również jako „maszynę”, która z dowolnego wypowiedzenia zdania oraz pewnych „parametrów” związanych z okolicznościami wypowiedzenia „produkuje” interpretację. Jednak aby zrozumieć jak przebiega interpretacja, nie ma potrzeby wnikać w szczegóły owej teorii:

Nie ma jednak powodu zajmować się szczegółami teorii, która ma stanowić adekwatny model zdolności [językowych] interpretatora. W tej dyskusji liczy się tylko to, że teoria powinna posiadać skończoną bazę oraz być rekurencyjna. Co do tych cech zgadza się większość filozofów oraz lingwistów (s. 95).

W związku z tym cytatem natrafiamy na pewną trudność. Wiąże się ona z niejasnym rozumieniem terminu „kompetencja językowa”. Czy ma być to pewna stała (choć nabyta) dyspozycja do interpretowania słów drugiego człowieka, czy też zdolność, która rozwija się w trakcie interakcji językowych z innymi ludźmi? W rekonstrukcji krytykowanych poglądów Davidson odwołuje się do pierwszego rozumienia kompetencji językowej. Byłaby to zatem stała dyspozycja, która aktualizuje się poprzez użycie słów w ich pierwszym znaczeniu. Podejście to zakłada istnienie języka jako abstrakcyjnego tworu konstytuowanego przez konwencje. Użycie słów byłoby opisane przez sformułowaną dla niego teorię prawdy, a w konsekwencji istniałaby teoria prawdy dla języka polskiego, angielskiego itd. Oczywiście chodzi tu o przypadek wyidealizowany, gdyż kompetencje językowe poszczególnych użytkowników mogą się nieznacznie od siebie różnić. Jednak dla krytykowanego poglądu istotne jest to, że warunkiem skuteczności komunikacji językowej jest istnienie owego idealnego wzorca kompetencji oraz – co jest z tym związane – istnienie języka rozumianego jako idealnego tworu ze znaczeniami dosłownymi przypisanymi jego wyrażeniom na mocy konwencji.

Nie kwestionując istnienia ogólnej i właściwej ludziom zdolności do interpretowania wypowiedzeń, kompetencja językowa może być jednak rozumiana w nieco inny sposób. Można ją bowiem rozumieć jako dynamiczną i zmieniającą się dyspozycję do interpretowania wypowiedzeń innych osób. Tak rozumiana kompetencja jest indywidualną dyspozycją, którą rozwijamy podczas interakcji z innymi ludźmi⁷³. Nie ma nawet potrzeby, aby wprowadzać wyidealizowany wzorec owej kompetencji. Jak się zdaje, w taki właśnie sposób rozumiał kompetencję językową Davidson, a przynajmniej podany

⁷³ Nabywanie owej kompetencji zaczyna się podczas nauki pierwszego języka w procesie, który Davidson nazwał triangulacją.

przez niego opis procesu interpretowania zakłada takie rozumienie kompetencji, chociaż sam nie używa terminu „kompetencja językowa” w tym sensie⁷⁴.

Na początku tego rozdziału zasugerowałem (czyniąc wielkie uproszczenie), że przed przystąpieniem do interpretowania słów innych ludzi władających tym samym językiem, interpretator zakłada trywialną teorię prawdy, tj. teorię, z której wynikają trywialne T-zdania. Uproszczenie polegało między innymi na tym, że nie uwzględniłem zdań okazjonalnych, gdyż teoria, którą stosujemy do zdań okazjonalnych, nie może być trywialna w tym sensie. Następnie w trakcie rozmowy weryfikujemy teorię w obliczu dostępnych świadectw i wprowadzamy w niej stosowne modyfikacje. Pomijając wspomniane uproszczenie, trywialna teoria prawdy może być rozumiana jako model indywidualnej kompetencji językowej, o której *de facto* pisał Davidson. Okazuje się jednak, że pisząc o trywialnej teorii prawdy poczyniłem jeszcze inne uproszczenie. Nie jest bowiem tak, że przystępując do rozmowy zakładamy wstępnie, że nasz rozmówca posługuje się dokładnie takim samym idiolektem co my. Często posiadamy pewną wstępną wiedzę na temat zwyczajów językowych naszych rozmówców. W konsekwencji „wstępna” teoria prawdy dla jednego rozmówcy może się różnić od teorii prawdy dla drugiego.

Przedstawiając własne rozwiązanie, Davidson dokonuje rozróżnienia między tzw. teorią uprzednią (*prior theory*), która ma zastąpić teorię pierwszą, a tzw. teorią bieżącą (*passing theory*). Teoria uprzednia reprezentuje nabyte do tej pory kompetencje językowe interpretatora w sensie, który podałem powyżej:

Zakładam, że teoria interpretatora została dostosowana do świadectw, które były dotychczas dostępne: znajomości charakteru, stroju i płci mówiącego oraz wszystkiego co mogło być do tej pory rezultatem obserwacji jego zachowań językowych oraz pozajęzykowych (s. 100).

Jak widać, kompetencje reprezentowane przez teorię uprzednią obejmują wiedzę na temat sposobu wypowiedzania się przez konkretnego interlokutora w zależności od jego poglądów, przynależności społecznej, wykształcenia oraz doświadczenia interpretatora wyniesionego z dotychczas odbytych z nim rozmów. A zatem teoria uprzednia nie reprezentuje całej kompetencji języko-

⁷⁴ Nieco wcześniej stwierdziłem, że krytykowana przed Davidsona koncepcja daje się wywieść od Chomsky’ego. Na krytykę tę można również spojrzeć z punktu widzenia ogólnej postawy filozoficznej obu filozofów. Chomsky był racjonalistą, zaś Davidson empirystą odrzucającym istnienie wszelkich bytów idealnych. Oczywiście Davidson – podobnie jak Quine – zdawał sobie sprawę, że przyjęcie pewnych bytów można traktować jako założenia teorii. W tym jednak przypadku starał się pokazać, że teoria komunikacji językowej opisana za pomocą powyższych punktów (1)–(3) faktycznie nie wyjaśnia tego fenomenu w sposób satysfakcjonujący.

wej interpretatora nawet w przyjętym tu sensie, lecz jedynie tę jej część, która dotyczy konkretnej osoby. Oznacza to, że ten sam interpretator może dysponować wieloma różnymi teoriami uprzednimi, które są stosowane do różnych osób. Oczywiście żadna teoria uprzednia z pewnością nie jest teorią trywialną w podanym przeze mnie sensie⁷⁵.

Podczas rozmowy interpretator zmienia zazwyczaj teorię, dostosowując ją do okoliczności wypowiedzeń, przejęzyczeń, malapropizmów oraz różnych sposobów użycia deskrypcji:

Kiedy interlokutor wypowie pewną kwestię, interpretator zmienia teorię, wprowadzając hipotezy o nowych nazwach, zmieniając interpretację predykatów i rewiduje dawniejsze interpretacje w obliczu nowych świadectw (*ibidem*).

Powstała w ten sposób teoria nosi nazwę teorii bieżącej.

Oczywiście zmiana teorii nie musi być rezultatem błędów lub malapropizmów. Niekiedy mówiący chce być zrozumiany w określony sposób i stosuje różne środki, aby interpretator zmienił teorię. Rozmawiając, pragniemy bowiem realizować określone zamiary. Dlatego też mówiący zazwyczaj spodziewa się, jaki kształt ma teoria przednia interpretatora i pragnie nadać określony kształt teorii bieżącej. W konsekwencji zarówno teoria uprzednia, jak i teoria bieżąca mogą być rozpatrywane zarówno z perspektywy mówiącego, jak i interpretatora:

Dla słuchacza teoria uprzednia wyraża to, jak jest przygotowany do interpretowania wypowiedzenia mówiącego, zaś teoria bieżąca opisuje to, jak faktycznie je interpretuje. Dla mówiącego teoria uprzednia wyraża to, w co wierzy na temat przygotowania się interpretatora do rozumienia jego słów, zaś teoria bieżąca opisuje to, w jaki sposób chciałby być zrozumiany (s. 101).

W przypadku użytkowników tego samego języka obie teorie nie różnią się znacznie z perspektywy mówiącego i interpretatora. To, co popularnie nazywamy konwencjami językowymi, jest dla Davidsona jedynie wspólną wiedzą mówiącego i interpretatora na temat teorii uprzedniej oraz bieżącej (s. 102), która składa się na wspomniane przeze mnie wcześniej standardy interpretacyjne usprawniające proces interpretacji. Z punktu widzenia tak przedstawionej teorii interpretacji nie ma zatem sensu mówić o języku konstytuowanym przez konwencje, a w konsekwencji nie ma sensu mówić o kompetencji językowej w tradycyjnym sensie:

⁷⁵ Co najwyżej można powiedzieć, że w przypadku całkowicie nieznanego użytkownika tego samego języka, teoria uprzednia zbliża się do teorii trywialnej.

Myślę, że poważnie potraktowane rozróżnienie na teorię przednią i bieżącą podważa powszechnie uznawane pojęcia kompetencji językowej oraz komunikacji językowej (*ibidem*).

Nie ma sensu również mówić o znaczeniach dosłownych (*conventional meanings*) w sensie Grice'a jako różnych od tego, co mówiący miał na myśli. Jeżeli zinterpretujemy jakieś wypowiedzenie, to przypiszemy mu tym samym znaczenie dosłowne. W przypadku malapropizmu byłoby to niekonwencjonalne znaczenie dosłowne⁷⁶.

Nieco wcześniej zaproponowałem niestandardowe, lecz zgodne z duchem poglądów Davidsona, rozumienie kompetencji językowej jako zależnej od kontekstu i zmieniającej się dyspozycji do rozumienia konkretnych osób. Dla każdej z tych osób jest ona reprezentowana przez odpowiadającą danej osobie teorię uprzednią. W trakcie rozmowy z daną osobą tak rozumiana kompetencja może ulec zmianie, co znajduje wyraz w reprezentującej ją teorii bieżącej. Nie oznacza to jednak, że teoria bieżąca, którą formułujemy pod koniec rozmowy, staje się teorią uprzednią dla następnej rozmowy. Co najwyżej można powiedzieć, że formułowana podczas rozmowy teoria bieżąca w pewien sposób wpływa na modyfikację kolejnej teorii uprzedniej. Jest tak, gdyż teoria bieżąca pozwala zinterpretować pojedyncze pomyłki, przejęzyczenia lub malapropizmy. Ponadto „znajomość teorii bieżącej wyraża jedynie wiedzę jak interpretować określone wypowiedzenia w określonych okolicznościach” (*ibidem*). Często spodziewamy się bowiem, że popełnione przez mówiącego błędy nie zostaną więcej popełnione. Modyfikacja kolejnej teorii uprzedniej może mieć miejsce jedynie wtedy, gdy teoria bieżąca opisze pewne zauważone przez interpretatora regularności w wypowiedzeniach rozmówcy. Może to regularnie powtarzany malapropizm lub pewna maniera mówiącego.

Teorie uprzednie prawie nigdy nie są dzielone przez mówiącego i interpretatora (s. 103). Wzajemne zrozumienie interlokutorów może mieć miejsce jedynie na poziomie teorii bieżących, które zbliżają się do siebie podczas rozmowy. Jednak żadna z teorii bieżących nie jest w stanie opisać tego, co tradycyjnie nazywamy językiem oraz kompetencją językową (s. 104). Oczywiście chodzi tutaj o kompetencję rozumianą w sensie punktów (1)–(3), które Davidson proponuje po prostu odrzucić. Davidson uznawał istnienie pewnej zdolności językowej (*linguistic ability*), która umożliwia nam komunikację językową. Ową zdolność, polegającą na dyspozycji do wzajemnego komuni-

⁷⁶ Podejście Grice'a prowadzi do zatarcia różnicy między użyciem metaforycznym a użyciem niezgodnym ze standardami interpretacyjnymi. Davidson wprowadza jasne kryterium rozróżnienia – użycie metaforyczne zaczyna się tam, gdzie kończy się możliwość interpretacji. Z tego powodu metafora nie ma, poza znaczeniem dosłownym, żadnego specjalnego znaczenia.

kowania się mówiącego i interpretatora, pozwoliłem sobie wcześniej określić jako kompetencję językową w nietradycyjnym znaczeniu. Tak rozumianej zdolności nie da się jednak scharakteryzować w tradycyjnych kategoriach kompetencji językowej:

(...) gdyż nie znaleźliśmy żadnego wyuczanego rdzenia spójnego działania, żadnych współdzielonej gramatyki lub reguł, żadnej podręcznej maszyny interpretacyjnej nastawionej na produkowanie znaczenia dowolnego wypowiedzenia. (...) Jeśli jednak tak jest, to należy sobie uświadomić, że nie tylko porzuciliśmy tradycyjne pojęcie języka, ale zatarliśmy różnicę między znajomością języka, a umiejętnością radzenia sobie w świecie (s. 107)⁷⁷.

Wypowiadając coś, lub interpretując wypowiedzenia innych, po prostu działamy, i jak w każdym działaniu możemy osiągnąć zamierzony efekt lub go nie osiągnąć. Efekty, zamierzone lub niezamierzone, są bowiem skutkami naszych działań, nie zaś ich konsekwencjami logicznymi. W szczególności nie istnieją prawa, które głoszą, że wypowiedzenie z zamiarem, aby słuchacz zinterpretował je w określony sposób, jest warunkiem koniecznym i wystarczającym, aby faktycznie wypowiedzenie zostało w taki sposób zinterpretowane⁷⁸. Dlatego też Davidson jest zdania, że:

Nie istnieją żadne reguły, które pozwalają na dochodzenie do teorii bieżących, żadne reguły w ścisłym tego słowa znaczeniu, w przeciwieństwie do ogólnych maksym lub metodologicznych uogólnień (*ibidem*).

Oznacza to, że doświadczony i wyszkolony interpretator (np. zawodowy tłumacz) może sprawniej i bardziej efektywnie interpretować wypowiedzenia mówiącego – jednak nawet on nie jest nieomylny i bezbłędny. W konsekwencji interpretacja zawsze obarczona jest pewnego rodzaju niezdeterminowaniem, które będzie przedmiotem kolejnego, ostatniego już paragrafu.

⁷⁷ W *Ku jednolitej teorii znaczenia i działania* (1980), Davidson wskazuje na to, że zachowania językowe stanowią szczególny przypadek ludzkich działań i postuluje stworzenie jednolitej teorii znaczenia i działania opartej na postawie „ktoś woli, aby prawdziwe było raczej to zdanie, niż tamto” (1992/1980, s. 153).

⁷⁸ Uwaga ta wiąże się z uznawanym przez Davidsona poglądem, że ścisłe prawa przyczynowe mogą być formułowane jedynie w języku fizykalnym i nie da się ich formułować w języku psychologicznym, czyli odwołującym się do opisu postaw propozycyjalnych. Zainteresowanego czytelnika odsyłam do prac z zakresu filozofii działania oraz filozofii psychologii, zebranych głównie w (2001 i 1992).

9. NIEZDETERMINOWANIE INTERPRETACJI

Na końcu artykułu *Interpretacja radykalna* Donald Davidson zwraca uwagę, że ograniczenia, które nakładane są na proces interpretacji, są do pewnego stopnia elastyczne. W konsekwencji wypowiedzenia drugiego człowieka mogą być interpretowane na więcej niż jeden sposób. Każdy z tych sposobów, na który składają się teoria prawdy oraz przyporządkowanie mówiącemu przekonania, dostarcza „poprawnej interpretacji dla każdej potencjalnej wypowiedzi”. Zawsze możemy bowiem – dysponując wszelkimi świadectwami – „dokonać wymiany przekonania, jakie przypisujemy użytkownikowi języka, i interpretacji, jaką nadajemy jego słowom” (1992/1973, s. 117).

Na pierwszy rzut oka wydawać się może, że niezdedeterminowanie, o którym mowa, jest rezultatem tego, że w proces interpretacji uwikłane są dwa parametry – postawy propozycjonalne oraz znaczenia. Skoro bowiem proces interpretacji polega na jednoczesnym konstruowaniu teorii prawdy i przypisywaniu przekonania, mamy do czynienia z procesem analogicznym do rozwiązywania równania z dwoma niewiadomymi, które ma nieskończenie wiele rozwiązań. Ograniczeniem narzuconym na możliwe interpretacje jest – jak twierdzi Davidson – elastyczna Zasada życzliwości, ale nawet ona nie redukuje interpretacji do jednej. Powstaje w ten sposób niezdedeterminowanie, które nie powoduje jednak jakiegokolwiek błędu, gdyż każda poprawnie sformułowana teoria prawdy dostarcza bowiem „poprawnej interpretacji dla każdej potencjalnej wypowiedzi” (*ibidem*). Analogia do równania ma jednak zasadniczą wadę – interpretacja podlega pewnym ograniczeniom. Jest bowiem procesem opartym na przypisywaniu postaw na podstawie świadectw empirycznych oraz budowaniu teorii prawdy, która również jest teorią empiryczną. Dlatego też należy wyjaśnić, jak możliwe jest niezdedeterminowanie pomimo istnienia owych ograniczeń.

O ile w *Interpretacji radykalnej* Davidson odwołał się *explicite* do przekonania i podawanych przez teorię prawdy znaczeń, to w *Three Varieties of Knowledge* przyczyny niezdedeterminowania interpretacji upatruje w świadectwach i znaczeniach:

Ogół świadectw dostępnych interpretatorowi nie determinuje jednej teorii prawdy dla [języka] osoby interpretowanej. Nie jest tak dlatego, że mamy zawsze dostęp jedynie do skończonej próbki świadectw, zaś teoria ma nieskończenie wiele konsekwencji, lecz dlatego, że wszystkie możliwe świadectwa nie ograniczają liczby teorii do jednej (2001a/1991, s. 214).

Między obu podejściami nie ma jednak sprzeczności, gdyż źródłem przekonania są właśnie świadectwa. Należy również zauważyć, co jest w tym przypadku

niezwykle istotne, że z różnych teorii prawdy, o których mowa w cytacie, muszą wynikać równokształtne T-zdania. W przeciwnym przypadku można by było bez problemu jedną z nich sfalsyfikować.

Wobec niezwykle bogactwa dostępnych świadectw, na podstawie których interpretator przypisuje interpretowanemu przekonania oraz inne postawy propozycjonalne niezbędne przy budowaniu teorii prawdy, nie da się wytyczyć granicy między przekonaniami empirycznymi, które bezpośrednio zależą od zdarzeń zachodzących w świecie, a przekonaniami zależnymi od znaczeń słów, które występują w ich aktualizacji. Każde bowiem przekonanie zależy zarówno od świata, jak i znaczeń: „[...] linia graniczna między prawdą empiryczną a prawdą zależną od znaczenia nie da się jasno zdefiniować na gruncie behawioralnym” (s. 215). Cytat ten stanowi nawiązanie do przeprowadzonej przez Quine’a w *Dwóch dogmatach empiryzmu* argumentacji na rzecz braku ostrej granicy między zdaniami analitycznymi (prawdziwymi na podstawie znaczeń) oraz syntetycznymi (prawdziwymi na podstawie faktów).

W jaki sposób należy rozumieć niezdeterminowanie interpretacji? Niech dwie osoby interpretują słowa mówiącego, dysponując jednocześnie tymi samymi świadectwami. Podczas interpretacji mogą dojść do dwóch odmiennych teorii prawdy, które podają znaczenia słów mówiącego. W tym miejscu można by podnieść zarzut, że tak rozumiane niezdeterminowanie jest jedynie błędem, który można stosunkowo łatwo usunąć. Na czym jednak miałyby polegać jego usuwanie? Porównywanie znaczeń nie wchodzi w grę, gdyż Davidson odrzucił zreifikowane znaczenia rozumiane jako byty, które można by w jakikolwiek sposób porównać. Pozostaje nam zatem porównanie całych teorii. Przedsięwzięcie to jednak skazane jest na niepowodzenie, gdyż teoria prawdy nie jest czymś, co się formułuje *explicite*, lecz stanowi jedynie reprezentację pewnego „psychologicznego mechanizmu”, który służy do komunikowania się.

Teorię prawdy można jedynie testować za pomocą T-zdań, odwołując się do dostępnych świadectw. W takim przypadku można by zrewidować jedną z dwóch teorii prawdy, tj. zmienić jakiś jej aksjomat. Takie przypadki nie dotyczą jednak niezdeterminowania w sensie Davidson, które ma być z założenia nieegzemplifikowalne, a w konsekwencji bardziej nieuchwytnie niż Quine’a teza o niezdeterminowaniu przekładu. Pewnej wskazówki dostarcza sam Davidson, wskazując, że niezdeterminowanie dotyczy nie tylko znaczeń, ale również ludzkich myśli: „Akceptuję tutaj tezę Quine’a o niezdeterminowaniu przekładu i rozszerzam ją na interpretację myśli w ogóle. Analogia do pomiaru pochodzi ode mnie” (*ibidem*, przyp. 3).

Zanim przejdziemy do omówienia analogii pomiaru, porównajmy pewien aspekt niezdeterminowania przekładu z niezdeterminowaniem interpreta-

cji. W przypadku Quine'a teza o niezdeterminowaniu głosi, że dwa konkurencyjne podręczniki przekładu nie są zdeterminowane przez ogół dostępnych bodźców lub – równoważnie – przez eksperymentalną bazę przekładu. W Konsekwencji dwa podręczniki przekładu mogą być całkowicie niezgodne między sobą, czy też niewspółmierne (podobnie jak konkurencyjne teorie empiryczne opisujące „ten sam fragment rzeczywistości”) lecz mimo to zgodne z ogółem bodźców. Przeprowadzona przez Davidsona krytyka schematu pojęciowego polegała m.in. na zakwestionowaniu sensowności takich określeń jak „niezgodność” czy też „niewspółmierność”, a w konsekwencji na odrzuceniu pojęcia schematu pojęciowego. Dlatego też „stopień niezdeterminowania” interpretacji będzie – zdaniem Davidson – mniejszy niż w przypadku niezdeterminowania przekładu (1992/1974b, s. 138).

Niezdeterminowanie interpretacji musi mieć zupełnie inne źródło niż niezdeterminowanie przekładu. Mówiacy i interpretator stykają się z tą samą rzeczywistością, która w sposób przyczynowy oddziałuje na ich przekonania. Przekonania oraz inne postawy, które ostatecznie można – jak twierdził Davidson – zredukować do przekonań, aktualizują się (przy wymienianych już założeniach) za pomocą wypowiedzi zdań o określonym znaczeniu. Zarówno przekonania, jak i znaczenia tworzą wzajemnie powiązane (holistyczne) całości. W przypadku przekonań Davidson posługuje się nie do końca jasną metaforą sieci, zaś znaczeniom strukturę nadaje teoria prawdy. Proces interpretacji polega na odtworzeniu obu całości. Jakimi świadectwami dysponuje interpretator? Przede wszystkim ma on dostęp do tej samej rzeczywistości, co osoba interpretowana. Jednak dostęp ten jest zapośredniczony w jego własnych przekonaniach, które mogą się aktualizować w (szczerych) wypowiedzeniach zdań o specyficznych dla niego znaczeniach. Może również obserwować zachowania językowe interpretowanego i na tej podstawie przypisywać mu zarówno przekonania, jak i znaczenia wypowiedzianym przez niego zdaniom. Przypisywanie znaczeń odbywa się za pomocą teorii prawdy zbudowanej dla idiolektu osoby interpretowanej, zaś przypisywanie przekonań może być opisane za pomocą zdań w idiolekcie interpretatora, o których interpretator sądzi, że aktualizują przekonania interpretowanego.

W przypadku niezdeterminowania interpretacji nie chodzi bynajmniej o sytuację, gdy dysponujemy jedynie ograniczonym zasobem świadectw. Taka sytuacja faktycznie zawsze ma miejsce, gdyż nigdy nie dysponujemy ogółem wszystkich możliwych świadectw. Davidsonowi chodzi jednak o przypadek wyidealizowany – nawet hipotetyczny wszechwiedzący (tj. dysponujący ogółem świadectw) interpretator nie jest w stanie dotrzeć do jedynej właściwej interpretacji, gdyż takowej po prostu nie ma. Interpretator występuje zawsze z perspektywy trzecioosobowej i nie ma bezpośredniego dostępu do

postaw propozycyjalnych interpretowanego. Innymi słowy, interpretator nigdy nie będzie dysponował wiedzą subiektywną na temat jego przekonań, lecz jedynie wiedzą intersubiektywną. Znajduje to wyraz w tym, że przypisanie postaw interpretowanemu odbywa się w języku interpretatora.

Interpretator może odtworzyć przekonania interpretowanego i przypisać znaczenia jego słowom jedynie opierając się na Zasadzie życzliwości (we wszystkich z jej wielu sformułowań). Zasada ta jednak jest bardzo elastyczna i pozostawia wiele możliwości interpretacyjnych. Dlatego też dwóch interpretatorów może dojść do różnych interpretacji, które będą całkowicie zgodne z wszystkimi dostępnymi świadectwami. Różnicy tej jednak wykryć się nie da, gdyż jedyna droga do wykrycia różnicy między dwoma konkurencyjnymi interpretacjami wiedzie przez dostępne świadectwa. O ile niezdeterminowanie przekładu przejawia się – zdaniem Quine’a – całkowitą niewspółmiernością podręczników przekładu (wyrażoną również w kategoriach permutacji zdań), to niezdeterminowanie interpretacji jest nieuniknioną konsekwencją epistemologicznych założeń Davidsona. Takie niezdeterminowanie jest jednak całkowicie niegroźne dla procesu interpretacji.

Sytuacja wydaje się zatem paradoksalna. Postuluje się bowiem istnienie niezdeterminowania, ale zaprzecza się jakiegokolwiek możliwości jego egzemplifikacji. Intuicyjnie rzecz biorąc, niezdeterminowanie interpretacji może polegać na tym, że dwaj interpretatorzy przypisują tym samym wypowiedziom interpretowanego różne znaczenia, a tym samym przypisują mu odmienne przekonania, aktualizowane za pomocą tych wypowiedzi. Różnicy tej jednak nie da się w żaden sposób wyartykułować. Pewne intuicje związane z niezdeterminowaniem oddaje – zdaniem Davidsona – analogia do procedury pomiaru. W *Ku jednolitej teorii znaczenia i działania* Davidson pisze:

Powstające w ten sposób niezdeterminowanie interpretacji nie jest w tym ujęciu bardziej znaczące czy kłopotliwe, niż fakt, że ciężar można mierzyć albo w gramach, albo w uncjach (1992/1980, s. 150).

Bardzo podobną analogię Davidson przywołuje w *Reality Without Reference* (1977), odwołując się do procedury pomiaru temperatury, która polega na przyporządkowaniu obiektom liczb wyrażających ich temperaturę. Procedura ta oparta jest na teorii, która nakłada formalne ograniczenia na owo przyporządkowanie przy równoczesnym uwzględnieniu zachodzących w otoczeniu jakościowych zjawisk (1984/1977, s. 224). W przypadku teorii, na której oparta jest skala Celsjusza, zjawiskami tymi są zamarzanie wody, jej wrzenie (w określonych warunkach) oraz wysokość słupka rtęci w rurce umieszczonej w wodzie zamarzającej oraz wrzącej. Ograniczenie formalne

polega w tym przypadku na przyporządkowaniu zamarzającej wodzie oraz znajdującemu się w tych samych warunkach słupkowi rtęci wartości 0, przyporządkowanie wrzącej wodzie oraz znajdującemu się w tych samych warunkach słupkowi rtęci wartości 100 oraz uznanie, że jednej setnej różnicy wysokości słupka rtęci odpowiada różnica 1 stopnia temperatury. Temperaturę można mierzyć nie tylko w skali Celsjusza, również w skali Fahrenheita, przy czym istnieje bardzo prosta liniowa transformacja jednej skali na drugą (s. 225)⁷⁹. Podobnie jak przekład i interpretacja, pomiar temperatury w określonej skali ma charakter holistyczny. Innymi słowy podanie wartości temperatury ciała w określonej skali zawiera *implicite* odniesienie do teorii, która kryje się za ową skalą, nawet jeśli owej teorii nie znamy. W ramach jednej teorii określony stan wody w jeziorze pozwala przypisać mu wartość 20 stopni Celsjusza, zaś w ramach drugiej 68 stopni Fahrenheita. Kontakt termometru z wodą będącą w pewnym obiektywnie zachodzącym stanie fizycznym jest przyczyną zarówno naszego przekonania, że temperatura wody w jeziorze ma 20°C, jak i (w przypadku użycia innego termometru) przekonania, że temperatura wynosi 68°F. Jednak zdarzenie to nie uzasadnia tych przekonań, gdyż uzasadnienia należy szukać w odpowiednich teoriach pomiaru. Wedle Davidsona wszelkie przekonania są jedynie skutkami zdarzeń w świecie, lecz to, które zdania aktualizują owe przekonania, można uzasadnić na gruncie teorii prawdy dla idiolektu, do którego owe zdania należą.

Podobnie jak w przypadku analogii ciężaru, można mieć poważne wątpliwości co do trafności również tej analogii, przede wszystkim dlatego, że dwie osoby posługujące się dwoma skalami pomiaru temperatury doskonale zdają sobie z tego sprawę i potrafią odpowiednio przeliczyć wartości drugiej skali na własną. A zatem analogia upada, gdyż niezdeterminowanie interpretacji nie umożliwia podobnego „przeliczenia”⁸⁰. Spróbujmy jednak doszukać się w tej analogii jakiegoś elementu pozytywnego, tym bardziej, że Davidson przykładał do owej analogii dużą wagę.

Niech odpowiednikiem wyniku pomiaru w pewnej skali będzie T-zdanie wynikające z teorii prawdy zbudowanej dla idiolektu obejmującego wypowiedziane zdanie. Wyniki pomiarów, podobnie bowiem jak T-zdania – służą do weryfikowania odpowiednich teorii. Wyobraźmy sobie, że dwóch ludzi mierzy

⁷⁹ Teoria, na której oparte jest przyporządkowanie ciałom temperatury w stopniach Fahrenheita, jest nieco bardziej skomplikowana. Dość powiedzieć, że temperatura zamarzania wody wynosi w tej skali 32 stopnie, zaś wrzenia wody 212 stopni. Transformacji skali Celsjusza na Fahrenheita dokonuje się wg wzoru: $T_F = 32 + 1,8T_C$.

⁸⁰ Podobnie krytyczne stanowisko wobec tej analogii reprezentują Lepore i Ludwig (por. 2005, s. 243–247).

temperaturę wody w jeziorze, przy czym jeden z nich czyni to za pomocą termometru w stopniach Celsjusza, a drugi za pomocą termometru w stopniach Fahrenheita. Oczywiście na obu termometrach mamy oznaczenia liczbowe temperatur, które to oznaczenia rozumiane są przez nich tak samo. Jeżeli zatem na jednym termometrze odczyt wyniesie 20, a na drugim 68, jedyną możliwością uzgodnienia obu rezultatów (poza założeniem uszkodzenia jednego z termometrów) jest uświadomienie sobie, że oba pomiary odbyły się wg innych skal i ewentualne wyszukanie zależności między wartościami obu pomiarów. W tym momencie analogia między pomiarem a interpretacją ulega załamaniu. W przypadku pomiaru za pomocą takich termometrów wartości liczbowe pełnią funkcję zależnych od teorii pomiaru indeksów dla odczytów tego samego „stanu cieplnego”. W przypadku, o który w naszym przypadku chodzi, tj. dla dwóch interpretacji zgodnych z ogółem świadectw, dla wypowiedzenia zdania *p* odpowiednie T-zdania będą miały identyczny kształt i nie są indeksowane w sposób zależny od teorii, z której wynikają. Na przykład, dla wypowiedzenia zdania *Snow is white* otrzymamy w ramach takich teorii prawdy równokształtne T-zdania postaci: „Zdanie *Snow is white* jest prawdziwe wtw, gdy śnieg jest biały”, które w żaden sposób nie może w obecności jakiegokolwiek świadectwa wskazać właściwej teorii, z której wynika.

Spróbujmy zatem zmodyfikować nieco powyższy przykład⁸¹ i pozbawmy podziałki obu termometrów oznaczeń liczbowych. Ponieważ różnica wysokości słupka rtęci odpowiadająca 1°C na skali pierwszego termometru odpowiada 1,8°F na skali drugiego termometru, można wyobrazić sobie podziałkę, w której kropki odpowiadają wartościom w skali Celsjusza podzielonym przez 10, zaś odległości między kropkami podzielone są na 18 równych odcinków. Dodatkowo na podziałce tej można zaznaczyć (np. stosownym ideogramem) punkt zamarzania, punkt wrzenia wody, punkt zamarzania mieszaniny wody i chlorku amonu (0°F) oraz średnią temperaturę ludzkiego ciała (96°F)⁸². W tym przypadku nie ma znaczenia, czy skala termometru powstała w sposób opisany przez Celsjusza, czy też w sposób opisany przez Fahrenheita, gdyż odczyty dokonywane za pomocą obu termometrów niczym się nie różnią. Analogię można przeprowadzić nieco dalej, zakładając, że dwaj amatorzy pomiarów, z których jeden jest zwolennikiem skali Celsjusza, a drugi skali Fahrenheita, komunikują między sobą wyniki pomiarów posługując się sformułowaniami postaci: „temperatura wynosi dwie kropki i jedna kreskę powyżej punktu zamarzania wody”. Jeżeli pozbedziemy się wartości liczbowych, to można powiedzieć, że zarówno dla zwolennika skali Celsjusza, jak

⁸¹ Modyfikacja ta pochodzi ode mnie.

⁸² Te parametry posłużyły bowiem do stworzenia skali Fahrenheita.

i zwolennika skali Fahrenheita, jedna i ta sama rzecz ma po prostu tę samą temperaturę.

Podany tu przykład miał na celu nie tylko „poprawienie” podanej przez Davidsona analogii z teorii pomiaru, ale również wyjaśnienie pewnego stwierdzenia, które można znaleźć na końcu *Reality Without Reference*:

Sugeruję, że w ten sam sposób [co w przypadku skali Fahrenheita i Celsjusza] znaczenie jest tym, co niezmiennie w różnych, lecz akceptowalnych teoriach prawdy. Znaczenie (interpretacja) zdania jest dane przez przyporządkowanie mu semantycznego położenia w układzie zdań tworzących język. Różne teorie prawdy mogą przyporządkowywać różne warunki prawdziwości temu samemu zdaniu (co stanowi semantyczną analogię do Quine’a tezy o niezdeteminowaniu przekładu), podczas gdy są zgodne (lub wystarczająco zgodne) co do roli zdań w języku (*ibidem*).

W przytoczonym cytacie konieczne jest wyjaśnienie dwóch kwestii. Przede wszystkim znaczenie zdania nie jest wzajemnie jednoznacznie związane z konkretną teorią prawdy, lecz raczej z całą rodziną poprawnie zbudowanych teorii. Zachodzi tu ścisła analogia do przytoczonych powyżej rozważań dotyczących teorii pomiaru – bez względu na to, czy temperaturę jakiegoś obiektu mierzymy w stopniach Celsjusza czy Fahrenheita, powiedzenie, że jest to ta sama temperatura, posiada pewien sens. Znaczenie, podobnie jak temperatura, nie jest rzeczą, lecz pewnym teoretycznym konstruktem, o którym można mówić jedynie na gruncie teorii. Zjawiska termiczne są przyczyną różnych reakcji, np. cofnięcia ręki w przypadku dotknięcia płomienia świecy. Na bodźce termiczne reagują podobnie ludzie i zwierzęta, lecz to ludzie wytworzyli pojęcie temperatury w procesie eksplikacji potocznych pojęć ciepła i zimna. Nawet wobec braku teorii *sensu stricte* możemy przynajmniej powiedzieć, że dwie rzeczy są tak samo ciepłe lub jedna jest cieplejsza od drugiej⁸³. Podobnie jest w przypadku znaczeń. Na wypowiedzenia zdań potrafią reagować zwierzęta. Lecz jedynie ludzie są w stanie przypisać zdaniom znaczenia, budując teorie prawdy. Jeżeli dwie teorie zbudowane są prawidłowo, to nie ma podstaw, aby odróżniać znaczenie, które przypisują zdaniu.

Druga sprawa, która stwarza o wiele większy problem wymagający wyjaśnienia, dotyczy warunków prawdziwości. Jak pamiętamy, warunki prawdziwości zdań języka przedmiotowego są podawane przez T-zdania wynikające z teorii prawdy. W przytoczonym cytacie sugeruje się, że dla tego samego

⁸³ Można to potraktować jako prototeorię temperatury, która jest zawarta *implicite* w języku naturalnym za sprawą terminów tworzących skalę porządkową: „zimny”, „letni”, „ciepły”, „gorący”, „zimniejszy”, „cieplejszy”, „gorętszy”.

zdania z dwóch prawidłowo zbudowanych teorii prawdy mogą wynikać różne T-zdania. Przyjęcie tego za dobrą monetę ratuje oczywiście oryginalne sformułowanie analogii z teorii pomiaru, ale stwarza problemy w świetle tego, co zostało powiedziane do tej pory. Jeżeli bowiem jakimś zdaniu odpowiadają dwa różne warunki prawdziwości (tj. T-zdania) na gruncie dwóch różnych teorii prawdy, to istnieje podstawa, aby zweryfikować obie teorie pod tym względem, uwzględniając wszelkie dostępne świadectwa. Można jednak rozumieć przytoczony cytat w nieco inny sposób. Przez warunki prawdziwości, poza T-zdaniami, można bowiem uznać również przekonania przypisywane interpretowanemu przez dwóch interpretatorów oraz przekonania owych interpretatorów. Jest to oczywiście inne rozumienie warunków prawdziwości, ale uzasadnione tym, że to właśnie przekonania pozwalają – poza teorią prawdy – na przypisanie zdaniom wartości logicznych. Dwie teorie tego samego idiolektu sformułowane w sposób prawidłowy przez dwóch interpretatorów mogą być identyczne, tj. mogą wynikać z nich dokładnie te same co do kształtu T-zdania. Są zatem jak pozbawione wartości liczbowych dwa termometry, które mogą służyć w zależności od wiedzy i przekonań użytkowników, do pomiaru zarówno w skali Celsjusza, jak i Fahrenheita.

Warto w tym momencie przywołać pewien argument, który podałem na początku tego rozdziału. Budowana przeze mnie przed laty trywialna teoria prawdy dla mojego ówczesnego idiolektu oraz budowana przeze mnie obecnie trywialna teoria prawdy dla mojego obecnego idiolektu o tym samym zasobie słów co przed laty, będą identyczne co do kształtu, chociaż znaczenia słów mogły ulec zmianie. Teoria prawdy bowiem podaje jedynie „semantyczne położenie zdania w układzie zdań tworzących język”. W tym sensie teoria prawdy nie jest – jak twierdził Davidson w cytacie na początku rozdziału – teorią znaczenia, bo stanowi jedynie warunek konieczny podania znaczeń języka. Warunek konieczny i wystarczający stanowią łącznie teoria prawdy oraz sieć postaw interpretatora (wiedza subiektywna) i zrekonstruowana sieć postaw interpretowanego (wiedza intersubiektywna). Ponieważ przypisywanie postaw propozycyjalnych osobie interpretowanej odbywa się w procesie kontruwowania teorii prawdy, stanowi ono historię przyczynową powstania teorii prawdy. A zatem dwie teorie prawdy mogą się nie różnić co do kształtu fizycznego, lecz różnić się historią swego powstania, podobnie jak dwa identyczne co do kształtu termometry, z których pierwszy ma mierzyć, według zamysłu swego twórcy, temperaturę w skali Celsjusza, zaś drugi w skali Fahrenheita.

Davidson przykładął dużą wagę do historii przyczynowej interakcji z otoczeniem i innymi ludźmi, traktując ją jako warunek myślenia⁸⁴. W przypadku każdego z nas niepowtarzalną historią przyczynową jest po prostu triangulacja. Dlatego też dwóch interpretatorów – założmy dla uproszczenia, że posługują się identycznymi idiolektami – może zbudować dla interpretowanego idiolektu dwie równokształtne teorie prawdy. Jednak interpretacje te nie będą identyczne, a ów brak identyczności jest właśnie niezdeterminowaniem. Co więcej, niezdeterminowania tego nie da się w żaden sposób wykryć na gruncie dostępnych świadectw empirycznych. Chociaż zdaję sobie sprawę z kontrowersyjności tego rozwiązania, niezdeterminowanie stanowi tutaj warunek posiadania indywidualnych i niepowtarzalnych przekonań oraz innych postaw propozycyjalnych.

Na zakończenie warto podać jeszcze dwie analogie, na które powołuje się Davidson przy okazji niezdeterminowania interpretacji. Jak pamiętamy proces interpretacji musi zatem rozwikłać dwa, wzajemnie powiązane zagadnienia: aby poznać znaczenia słów idiolektu drugiego człowieka, należy przypisać mu przekonania, zaś aby przypisać przekonania drugiemu człowiekowi, należy poznać znaczenia jego słów. Jak pamiętamy z paragrafu poświęconego interpretacji radykalnej, Davidson starał się przerwać owo błędne koło, szukając postaw propozycyjalnych, których przypisanie nie wymaga uprzedniego przypisania znaczeń. Aby rozwiązać ten problem Davidson próbował również – jak się okazało nieskutecznie – zastosować aparaturę teorii decyzji w wersji F. Ramseya. Miała ona pomóc określeniu na podstawie obserwowalnych zachowań, np. dokonywanych wyborów lub oznak preferencji, jakie przekonania przypisze interpretator osobie interpretowanej i jakie znaczenia przypisze jej słowom. Interpretacja słów drugiego człowieka może być zatem traktowana jako proces analogiczny do podejmowania decyzji, np. w grze, w której musimy się zdecydować czy zadowolimy się bardzo prawdopodobną, lecz niewielką wygraną, czy też zaryzykować wygraną bardzo wysoką, lecz mało prawdopodobną. W takim przypadku grający dokonuje wyboru na podstawie dwóch czynników: wartości, jaką przywiązuje do określonego wyniku oraz subiektywnego prawdopodobieństwa uzyskania tego wyniku. Jednym z zadań teorii decyzji jest wyznaczenie obu czynników dla konkretnego wyboru gracza, który stanowi odpowiednik świadectwa empirycznego.

⁸⁴ W *Knowing One's Own Mind* (1987) Davidson podaje eksperyment myślowy, w którym na skutek uderzenia pioruna powstaje dokładna kopia Donalda Davidsona, która w dodatku zachowuje się dokładnie jak jej pierwowzór. Zdaniem Davidsona, ów Bagnolud (*Swampman*) może zacząć myśleć dopiero po pewnym czasie, w którym będą miały miejsce interakcje z otoczeniem i innymi ludźmi, które składają się na historię przyczynową.

Teorią decyzji Davidson zajmował się już we wczesnym okresie swej działalności, zanim poświęcił się filozofii języka⁸⁵, i powracał do niej wielokrotnie w esejach z teorii interpretacji, filozofii umysłu oraz filozofii prawdy. Niestety nigdy natomiast nie udało mu się w pełni, i w sposób dla siebie satysfakcjonujący, implementować teorii decyzji do teorii języka i umysłu. Nie przedstawiając tutaj dokładnie aparatury teorii decyzji, warto zauważyć, że temu samemu wyborowi w grze można przypisać różne układy preferowanych wartości i subiektywnych prawdopodobieństw. Teoria decyzji nie wyznacza zatem procedury w pełni determinującej wybór⁸⁶, podobnie jak w przypadku interpretacji rozumianej jako jednoczesne przypisanie przekonań i znaczeń⁸⁷.

Druga analogia dotyczy wspomnianej już poprzednio próby stworzenia jednolitej teorii znaczenia i działania. W (1980) Davidson zarysował projekt szerszego ujęcia wypowiedzeń rozumianych jako ludzkich działań⁸⁸. Przypomnijmy, że wypowiedzenie zdania jest nie tylko zdarzeniem, lecz również działaniem, gdyż posiada powód (*reason*) (por. 2001/1963, s. 3). Na powód składają się tzw. przednastawienie (*pro attitude*) do działania pewnego typu, tj. chęć, uznanie za słuszne lub wartościowe, poczucie obowiązku itp., oraz przekonanie, że faktycznie podjęte działanie (np. wypowiedzenie pewnego zdania) jest działaniem tego typu.

Jak twierdzi Davidson powód działania jest również jego przyczyną. W przypadku zastosowania teorii działania do analizy komunikacji językowej pojawia się jeszcze większy problem niezdeterminowania niż w przypadku interpretacji radykalnej, w której Davidson był skłonny redukować wszystkie postawy propozycyjalne do przekonań. Pojawiają się mianowicie trzy całkowicie niezależne parametry: znaczenie, przekonanie i pragnienie (*desire*) (ta ostatnia etykieta oznacza wszystkie postawy wchodzące w skład przednastawienia) (1992/1980, s. 143). Aby jednak przypisać mówiącemu powód wypowiedzenia, tj. wyrazić odpowiednie pragnienie oraz przekonanie, należy przypisać znaczenia wypowiedzianym słowom; aby z kolei przypisać znaczenia wypowiedzianym słowom należy znać powód, dla którego zostało ono wypowiedziane. Podobnie jak w poprzednich przypadkach mamy tutaj do czynienia z procedurą całkowicie niezdeterminowaną. W konsekwencji dwie różne osoby mogą przypisać komuś zupełnie odmienne powody tego samego działania i nie da się rozstrzygnąć przez odwołanie się do jakichkolwiek

⁸⁵ Chodzi tu o kilka wczesnych publikacji, które Davidson napisał we współpracy z Patrickiem Suppesem.

⁸⁶ Warto zauważyć, że gdyby tak było, to nie dałoby się obronić tezy o wolności człowieka.

⁸⁷ W świetle wcześniejszej analizy chodzi tu o znaczenia podawane przez teorię prawdy.

⁸⁸ Bardziej adekwatny zdaje się tytuł oryginalny *Toward Unified Theory of Thought, Meaning and Action*, gdyż wskazuje na nieobecne w tytule polskim postawy propozycyjalne.

świadczeń, który z owych powodów jest tym właściwym. Zakładają identyczność znaczeń słów, za pomocą których oba konkurencyjne powody zostały opisane, niezdeterminowanie to – w odróżnieniu od niezdeterminowania Interpretacji – przynajmniej daje się wyartykułować.

Davidson sformułował podstawy semantyki w połowie lat 60. Jednak budując przez lata spójny system filozoficzny zmuszony był do dopasowywania zaproponowanej semantyki do całości systemu. Dlatego też proponowana przez niego semantyka stała się coraz bardziej uwikłana w rozważania epistemiczne, psychologiczne, a nawet metafizyczne. Owo uwikłanie było szczególnie widoczne w ostatnim rozdziale niniejszej pracy, a zwłaszcza w paragrafie poświęconym niezdeterminowaniu. Opracowanie to nie wyczerpuje oczywiście problemów związanych z semantyką Davidsona. Pozostały do dokładnego omówienia jej zastosowania: analiza parataktyczna, analiza zdań opisujących działania (*action sentences*) oraz przyczynowa teoria metafory. Teoria metafory wykracza wprawdzie poza semantykę, ale ją zakłada i jest nad nią nadbudowana. Na dokładne opracowanie zasługują również te poglądy Davidsona z zakresu innych dyscyplin filozoficznych, które wpływały pośrednio na jego filozofię. Dlatego też mam nadzieję, że monografia ta stanowi początek większego cyklu.

BIBLIOGRAFIA

- Ajdukiewicz K. (1931) *O znaczeniu wyrażeń*, „Księga Pamiątkowa Polskiego Towarzystwa Filozoficznego we Lwowie”, s. 31–77. Przedruk [w:] Ajdukiewicz 1960, s. 102–136.
- Ajdukiewicz K. (1934) *Sprache und Sinn*, „Erkenntnis” IV, s. 100–138. Tłumaczenie polskie: *Język i znaczenie* (tł. F. Zeidler), [w:] Ajdukiewicz 1960, s. 145–174.
- Ajdukiewicz K. (1934a) *Das Weltbild und die Begriffsapparatur*, „Erkenntnis” IV, s. 259–287. Tłumaczenie polskie: *Obraz świata i aparatura pojęciowa* (tł. F. Zeidler), [w:] Ajdukiewicz 1960, s. 175–195.
- Ajdukiewicz K. (1960) *Język i poznanie*, t. 1, PWN.
- Aristoteles (1983) *Metafizyka* (tł. K. Leśniak), PWN.
- Aristoteles (1990) *Hermeneutyka*, [w:] Aristoteles, *Dziela wszystkie*, t. 1 (tł. K. Leśniak), PWN.
- Aristoteles (1990a) *Kategorie*, [w:] Aristoteles, *Dziela wszystkie*, t. 1 (tł. K. Leśniak), PWN.
- Austin J. L. (1962) *How to Do Things with Words*, Clarendon Press. Tłumaczenie polskie: *Jak działać słowami?* (tł. B. Chwedeńczuk), [w:] J. L. Austin, *Mówienie i poznawanie*, PWN, 1993, s. 543–713.
- Barwise J. i J. Perry (1981) *Semantic Inocence and Uncompromising Situations*, [w:] T. E. Uerling, Jr. and H. K. Wettstein (eds.), *The Foundations of Analytic Philosophy*, „Midwest Studies in Philosophy” 6, University of Minnesota Press, s. 387–403.
- Bell D. (1990) *Husserl*, Routledge.
- Belnap N. D. (1962) *Tonk, Plonk and Plink*, „Analysis” 22, s. 130–134.
- Carnap R. (1947) *Meaning and Necessity*, University of Chicago Press. Wydanie II rozszerzone 1956.
- Carroll L. (2005) *Alicja w krainie czarów. Alicja po drugiej stronie zwierciadła* (tł. A. Marianowicz i H. Baltyn), Nasza Księgarnia.
- Chomsky N. (1959) *A review of B.F. Skinner's 'Verbal Behavior'*, „Language” 35, s. 26–58.
- Davidson D. (1963) *The Method of Extension and Intension*, [w:] A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*, Open Court.
- Davidson D. (1963a) *Actions, Reason, and Causes*, „The Journal of Philosophy” 60, s. 685–701. Przedruk [w:] Davidson 2001, s. 3–19.
- Davidson D. (1965) *Theories of Meaning and Learnable Languages*, „Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science”, North Holland Publishing Co., s. 383–394. Przedruk [w:] Davidson 1984, s. 3–15.
- Davidson D. (1967) *Truth and Meaning*, „Synthèse” 17, s. 304–323. Tłumaczenie polskie: *Prawda i znaczenie* (tł. J. Gryz), [w:] Davidson 1992, s. 3–32.
- Davidson D. (1967a) *The Logical Form of Action Sentences*, [w:] N. Rescher (ed.), *The Logic of Decision and Action*, University of Pittsburg Press. Przedruk [w:] Davidson 1980, s. 105–122.

- Davidson D. (1968) *On Saying That*, „Synthèse” 19, s. 130–146. Przedruk [w:] Davidson 1984, s. 93–108.
- Davidson D. (1973) *Radical Interpretation*, „Dialectica” 27, s. 313–328. Tłumaczenie polskie: *Interpretacja radykalna* (tł. P. Józefowicz), [w:] Davidson 1992, s. 95–117.
- Davidson D. (1970) *Semantics for Natural Language*, [w:] *Linguaggi nella Società e nella Tecnica*, Edizioni di Comunità. Tłumaczenie polskie: *Semantyka dla języków naturalnych* (tł. M. Witkowski), [w:] Davidson 1992, s. 60–75.
- Davidson D. (1974) *On the Very Idea of a Conceptual Scheme*, „Proceedings and Adresses of the American Philosophical Association” 47, s. 5–20. Tłumaczenie polskie: *O pojęciu schematu pojęciowego* (tł. J. Gryz), [w:] B. Stanosz (red.), *Empiryzm Współczesny*, Wydawnictwo Uniwersytetu Warszawskiego, 1991, s. 280–297.
- Davidson D. (1974a) *Psychology as Philosophy*, [w:] S. C. Brown (ed.) *Philosophy of Psychology*, Macmillan Press. Tłumaczenie polskie: *Psychologia jako filozofia* (tł. C. Cieśliński), [w:] Davidson 1992, s. 217–233.
- Davidson D. (1974b) *Belief and Basis of Meaning*, „Synthèse” 27, s. 309–323. Tłumaczenie polskie: *Przekonania a podstawy znaczenia* (tł. B. Stanosz), [w:] Davidson 1992, s. 118–140.
- Davidson D. (1975) *Thought and Talk*, [w:] S. Guttenplan (ed.), *Mind and Language*, Oxford University Press, s. 7–23. Przedruk [w:] Davidson 1984, s. 155–170. Przekład polski: *Myślenie i mowa* (tł. B. Stanosz), [w:] B. Stanosz (red.), *Język w świetle nauki*, Warszawa 1980, s. 340–362.
- Davidson D. (1976) *Reply to Foster*, [w:] G. Evans and J. McDowell (eds.), *Truth and Meaning: Essays in Semantics*, Oxford University Press. Przedruk [w:] Davidson 2001, s. 171–179.
- Davidson D. (1977) *Reality Without Reference*, „Dialectica” 31, s. 247–253. Przedruk [w:] Davidson 1984, s. 215–225.
- Davidson D. (1978) *What Metaphors Mean*, „Critical Inquiry” 5, s. 31–47. Przedruk [w:] Davidson 1984, s. 245–264.
- Davidson D. (1979) *Quotation*, „Theory of Decision” 11, s. 27–40. Przedruk [w:] Davidson 1984, s. 79–92.
- Davidson D. (1979a) *Moods and performance*, [w:] A. Margalit (ed.), *Meaning and Use*, Reidel. Przedruk [w:] Davidson 1984, s. 109–121.
- Davidson D. (1980) *Toward Unified Theory of Thought, Meaning and Action*, „Gratzer Philosophische Studien” 11, s. 1–12. Przedruk pod tytułem *A Unified Theory of Thought, Meaning and Action*, [w:] Davidson 2004, s. 151–166. Tłumaczenie polskie: *Ku jednolitej teorii znaczenia i działania* (tł. B. Stanosz), [w:] Davidson 1992, s. 141–159.
- Davidson D. (1982) *Rational Animals*, „Dialectica” 36, s. 317–327. Tłumaczenie polskie: *Zwierzęta racjonalne* (tł. C. Cieśliński), [w:] Davidson 1992, s. 234–250.
- Davidson D. (1984) *Inquiries into Truth and Interpretation*, Clarendon Press.
- Davidson D. (1984a) *First Person Authority*, „Dialectica” 38, s. 101–111. Przedruk [w:] Davidson 2001a, s. 3–14.
- Davidson D. (1985) *Adverbs of Action*, [w:] B. Vermazan and M. Hintikka (eds.), *Essays on Davidson: Actions and Events*, Oxford University Press. Przedruk [w:] Davidson 2001, s. 293–304.
- Davidson D. (1986) *A Nice Derangements of Epitaphs*, [w:] E. Grandy and R. Warner (eds.), *Philosophical Grounds of Rationality*, s. 156–174. Przedruk [w:] Davidson 2005, s. 86–107.

- Davidson D. (1987) *Knowing One's Own Mind*, „Proceedings and Addresses of the American Philosophical Association”, s. 441–458. Przedruk [w:] Davidson 2001a, s. 15–38.
- Davidson D. (1989) *What is Present to the Mind?*, „Grazer Philosophische Studien” 36. Przedruk [w:] Davidson 2001a, s. 53–67.
- Davidson D. (1990) *Meaning, Truth, and Evidence*, [w:] R. Barret and G. Gibson (eds.), *Perspectives on Quine*, Blackwell. Przedruk [w:] Davidson 2005, s. 47–62.
- Davidson D. (1990a) *Turing's Test*, [w:] W. H. Newton-Smith and K. V. Wilkes (eds.), *Modelling of Mind*, Oxford University Press, s. 1–11. Przedruk [w:] Davidson 2004, s. 77–86.
- Davidson D. (1991) *Three Varieties of Knowledge*, [w:] A. Phillips Griffiths (ed.), „A. J. Ayer Memorial Essays: Royal Institute of Philosophy Supplement” 30, Cambridge University Press. Przedruk [w:] Davidson 2001a, s. 205–220.
- Davidson D. (1992) *Eseje o prawdzie, języku i umyśle*, PWN.
- Davidson D. (2001) *Essays on Actions and Events*, Clarendon Press. Wydanie II uzupełnione.
- Davidson D. (2001a) *Subjective, Intersubjective, Objective*, Clarendon Press.
- Davidson D. (2004) *Problems of Rationality*, Clarendon Press.
- Davidson D. (2005) *Truth, Language, and History*, Clarendon Press.
- Donnellan K. (1966) *Reference and Definite Descriptions*, „The Philosophical Review” 75, s. 281–304.
- Donnellan K. (1968) *Putting Hampty Dumpty Together Again*, „The Philosophical Review” 77, s. 203–215.
- Field H. (1972) *Tarski's Theory of Truth*, „The Journal of Philosophy” 69, s. 347–375.
- Foster J. A. (1976) *Meaning and Truth Theory*, [w:] G. Evans and J. McDowell (eds.), *Truth and Meaning. Essays in Semantics*, Oxford University Press, s. 1–32.
- Frege G. (1892) *Über Sinn und Bedeutung*, „Zeitschrift für Philosophie und Philosophische Kritik”, s. 25–50. Tłumaczenie polskie: *Sens i znaczenie*, [w:] Frege 1977, s. 60–88.
- Frege G. (1918) *Der Gedanke. Eine logische Untersuchung*, „Beiträge zur Philosophie des deutschen Idealismus” R. 1, s. 58–77. Tłumaczenie polskie: *Myśl. Studium logiczne*, [w:] Frege 1977, s. 101–129.
- Frege G. (1977) *Pisma semantyczne* (tł. B. Wolniewicz), PWN.
- Grice H. P. (1957) *Meaning*, „Philosophical Review” 66, s. 377–388.
- Grice H. P. (1975) *Logic and Conversation*, [w:] P. Cole and J. Morgan (eds.), *Syntax and Semantics*, vol. 3, Academic Press. Tłumaczenie polskie: *Logika a konwersacja* (tł. B. Stanosz), [w:] B. Stanosz (red.), *Język w świetle nauki*, Warszawa 1980, s. 91–114.
- Grice H. P. (1989) *Studies in the Way of Words*, Harvard University Press.
- Hahn L. D. (ed.) (1999) *The Philosophy of Donald Davidson*, The Library of Living Philosophers vol. XXVII, Open Court.
- Ingarden R. (1960) *O dziele literackim*, PWN.
- Ingarden R. (1961) *Spór o istnienie świata*, PWN.
- Kotarbiński T. (1922) *O istocie doświadczenia wewnętrznego*, „Przegląd Filozoficzny” 2. Przedruk [w:] Kotarbiński 1958a, s. 40–59.
- Kotarbiński T. (1935) *Zasadnicze myśli pansomatyzmu*, „Przegląd Filozoficzny” 5. Przedruk [w:] Kotarbiński 1958a, s. 117–136.
- Kotarbiński T. (1958) *Wybór pism*, t. 2, PWN.
- Kotarbiński T. (1958a) *Zdania niepsychologiczne*, [w:] Kotarbiński 1958, s. 156–160.
- Kotarbiński T. (1958b) *Próba redukcji poznania psychologicznego do ekstraspekcji*, „Studia Filozoficzne” 5(8), s. 138–141.

- Kripke S. (1980) *Naming and Necessity*, Blackwell. Wydanie drugie zmienione. Tłumaczenie polskie: *Nazywanie i konieczność* (tł. B. Chwedeńczuk), Wydawnictwo PAX, 1988.
- Lepore E. (ed.) (1986) *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson*, Basil Blackwell.
- Lepore E. i K. Ludwig (2005) *Donald Davidson. Meaning, Truth, Language, and Reality*, Clarendon Press.
- Lepore E. i K. Ludwig (2007) *Donald Davidson's Truth-Theoretic Semantics*, Clarendon Press.
- Lepore E. i K. Ludwig (eds.), (2013) *A Companion to Donald Davidson*, Blackwell Companions to Philosophy, Blackwell.
- Ludwig K. (ed.) (2003) *Donald Davidson*, Cambridge University Press.
- Lycan W. G. (2000) *Philosophy of Language. A Contemporary Introduction*, Routledge.
- Lynch M. P. (2001) (ed.) *The Nature of Truth. Classic and Contemporary Perspectives*, The MIT Press.
- Maciaszek J. (2003) *Znaki logiczne*, „Acta Universitatis Lodzensis. Folia Philosophica” 15, Wydawnictwo Uniwersytetu Łódzkiego.
- Maciaszek J. (2008) *Znaczenie, prawda, przekonania. Problematyka znaczenia w filozofii języka*, Wydawnictwo Uniwersytetu Łódzkiego.
- Maciaszek J. (2013) *Holistyczna teoria znaczenia Kazimierza Ajdukiewicza. Próba krytycznej analizy*, „Przegląd Filozoficzny” 88, s. 225–284.
- Maciaszek J. (2014) *Poppera i Davidsona krytyka mitu schematu pojęciowego*, „Przegląd Filozoficzny. Nowa Seria” 92, s. 209–230.
- Maciaszek J. (2015) *Analiza parataktyczna wypowiedzeń zdań performatywnych*, [w:] P. Stalmaszczyk (red.), *Studia z metodologii i filozofii językoznawstwa 3. Od zdań do aktów mowy – rozważania lingwistyczne i filozoficzne*, s. 133–156. Wydawnictwo PRIMUM VERBUM.
- Maciaszek J. (2015a) *Składalność znaczeń w teorii dyrektywnej*, „Przegląd Filozoficzny. Nowa Seria” 93, s. 27–45.
- Maciaszek J. (2016) *Kilka uwag krytycznych o Tadeusza Kotarbińskiego analizie zdań psychologicznych*, „Studia z Filozofii Polskiej” 11, s. 23–37.
- Maciaszek J. (2016a) *Donald Davidsona przyczynowa teoria metafory*, „Studia Semiotyczne” XXX, s. 121–144.
- Maciaszek J. (2017) *Tadeusza Kotarbińskiego analiza zdań psychologicznych i Donald Davidsona analiza parataktyczna*, [w:] D. Łukaszewicz i R. Mordarski (red.), *Racjonalność w myśleniu i działaniu. Filozofia Tadeusza Kotarbińskiego*, Wydawnictwo Uniwersytetu Kazimierza Wielkiego, s. 63–79.
- MacKay A. (1968) *Mr Donnellan and Humpty Dumpty on Referring*, „The Philosophical Review” 77, s. 197–202.
- Martinich A. P. (1984) *A Theory of Metaphor*, „Journal of Literary Semantics” 13, s. 35–56.
- Matthews R. J. (1986) *Learnability of Semantic Theory*, [w:] Lepore 1986, s. 49–58.
- Mill J. S. (1843) *System of Logic Ratiocinative and Inductive*. Tłumaczenie polskie: *System logiki dedukcyjnej i indukcyjnej* (tł. Cz. Znamierowski), PWN, 1962.
- Modrak D. (2001) *Aristotle's Theory of Language and Meaning*, Cambridge University Press.
- Modrak D. (2006) *Philosophy of Language*, [w:] M. L. Gill i P. Pellegrin (eds.), *A Companion to Ancient Philosophy*, Blackwell, s. 640–663.
- Nowaczyk A. (2003) *Co naprawdę powiedział Tarski o prawdzie w roku 1933?*, [w:] J. J. Jadacki (red.), *Alfred Tarski: dedukcja i semantyka*, Wydawnictwo Semper.

- Nowaczyk A. (2006) *Polawianie sensu w filozoficznej głębi*, Wydawnictwo Uniwersytetu Łódzkiego.
- Nowakowski A. (1997) *Semantyka Donalda Davidsona*, Wydawnictwo UMCS.
- Partee, B. H. (1984) *Compositionality*, [w:] F. Landman and F. Veltman (eds.), *Varieties of Formal Semantics*, s. 281–311, Foris.
- Pelc, J. (red.) (1967) *Logika i język. Studia z semiotyki logicznej* (tł. J. Pelc), PWN.
- Popper K. (1994) *The Myth of the Framework*, [w:] *The Myth of the Framework. In Defence of Science and Rationality*, Routledge. Tłumaczenie polskie: *Mit schematu pojęciowego: w obronie nauki i racjonalności* (tł. B. Chwedeńczuk), Książka i Wiedza, 1997.
- Prior A. N. (1960) *The Runabout Inference-Ticket*, „Analysis” 21, s. 38–39.
- Putnam H. (1975) *The Meaning of ‘Meaning’*, [w:] K. Gunderson (ed.), *Language, Mind, and Knowledge*, „Minnesota Studies in the Philosophy of Science” t. 7, University of Minnesota Press, s. 358–398. Tłumaczenie polskie: *Znaczenie wyrazu znaczenie* (tł. A. Grobler), [w:] H. Putnam, *Wiele twarzy realizmu i inne eseje*, Wydawnictwo Naukowe, 1998, s. 325–429.
- Quine W. V. O. (1940) *Mathematical Logic*, Harvard University Press. Tłumaczenie polskie: *Logika matematyczna* (tł. Leon Koj), PWN, 1974.
- Quine W. V. O. (1951) *Two Dogmas of Empiricism*, „The Philosophical Review” 60, s. 20–43. Przedruk zmienionej wersji artykułu [w:] W. V. O. Quine *From the Logical Point of View*, Harvard University Press, 1961. Tłumaczenie polskie: *Dwa dogmaty empiryzmu* (tł. B. Stanosz), [w:] W. V. O. Quine, *Z punktu widzenia logiki*, Wydawnictwo Spacja, 2000.
- Quine W. V. O. (1960) *Word and Object*, The Massachusetts Institute of Technology. Tłumaczenie polskie: *Słowo i rzecz* (tł. B. Stanosz), Wydawnictwo Spacja, 1999.
- Quine W. V. O. (1990) *Three Indeterminacies*, [w:] R. B. Barrett and R. F. Gibson (eds.), *Perspectives on Quine*, Blackwell, s. 1–16. Tłumaczenie polskie: *Trzy niezdeterminowania* (tł. B. Stanosz), [w:] B. Stanosz (red.), *Filozofia języka*, Wydawnictwo Spacja, 1993, s. 121–139.
- Ramsey F. (1990) *The Nature of Truth*, „Episteme” 16, s. 6–16.
- Rawling P. (2003) *Radical Interpretation*, [w:] Ludwig 2003, s. 85–112.
- Russell B. (1905) *On Denoting*, „Mind” 14, s. 479–493. Tłumaczenie polskie: *Denotowanie*, [w:] Pelc 1967, s. 253–275.
- Scheffler I. (1954) *An Inscriptural Approach to Indirect Quotation*, „Analysis” 10, s. 83–90.
- Scheffler I. (1997) *Symbolic worlds. Art, science, language, ritual*, Cambridge University Press.
- Searle J. R. (1979) *Metaphor*, [w:] *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge University Press, s. 76–116.
- Shramko Y. i H. Wansing (2009) *The Slingshot Argument and Sentential Identity*, „Studia Logica” 91, s. 429–455.
- Skinner B. F. (1957) *Verbal Behavior*, Appleton Crofts.
- Soames S. (1984) *What Is a Theory of Truth?*, „The Journal of Philosophy” 81, s. 411–429.
- Tarski A. (1930/1931) *O pojęciu prawdy w odniesieniu do sformalizowanych nauk dedukcyjnych*, „Ruch Filozoficzny” 12, s. 210–211. Przedruk [w:] Tarski 1995 s. 3–8.
- Tarski A. (1932) *Der Wahrheitsbegriff in den Sprachen der deduktiven Disziplinen*, „Der Akademie der Wissenschaften in Wien, Mathematisch-naturwissenschaftlichen Klasse, Akademischer Anzeiger” t. 69, s. 23–25. Tłumaczenie polskie: *Pojęcie prawdy w językach nauk dedukcyjnych* (tł. J. Zygmunt), [w:] Tarski 1995, s. 9–12.

- Tarski A. (1933) *Pojęcie prawdy w językach nauk dedukcyjnych*, „Prace Towarzystwa Naukowego Warszawskiego, Wydział II Nauk Matematyczno-Fizycznych” nr 34, Warszawa. Przedruk [w:] Tarski 1995, s. 13–172.
- Tarski A. (1936) *O ugruntowaniu naukowej semantyki*, „Przegląd Filozoficzny” 39, s. 50–57. Przedruk [w:] Tarski 1995, s. 173–185.
- Tarski A. (1936a) *O pojęciu wynikania logicznego*, „Przegląd Filozoficzny” 39, s. 58–68. Przedruk [w:] Tarski 1995, s. 186–202.
- Tarski A. (1936b) *Głos w dyskusji nad referatem Marii Kokoszyńskiej zatytułowanym „W sprawie względności i bezwzględności prawdy”*, Księga Pamiątkowa Trzeciego Polskiego Zjazdu Filozoficznego, Kraków. Przedruk [w:] Tarski 1995, s. 203–205.
- Tarski A. (1944) *The Semantic Conception of Truth and the Formulation of Semantics*, „Philosophy and Phenomenological Research” 4, s. 341–375. Tłumaczenie polskie: *Semantyczna koncepcja prawdy i podstawy semantyki* (tł. J. Zygmunt), [w:] Tarski 1995, s. 228–282.
- Tarski A. (1995) *Pisma logiczno-filozoficzne. Tom 1. Prawda*, PWN.
- Thomason R. H. (ed.) (1976) *Formal Philosophy. Selected Papers of Richard Montague*, Yale University Press.
- Twardowski K. (1900) *O tak zwanych prawdach względnych*, [w:] *Księga pamiątkowa Uniwersytetu lwowskiego ku uczczeniu pięćsetnej rocznicy fundacji Jagiellońskiej Uniwersytetu krakowskiego*, Lwów. Przedruk [w:] Twardowski *Wybrane pisma filozoficzne*, PWN, 1965, s. 315–336.
- Żegleń U. (red.) (1996) *Dyskusje z Donaldem Davidsonem o prawdzie, języku i umyśle*, Towarzystwo Naukowe KUL.
- Żegleń U. (ed.) (1999) *Donald Davidson. Truth, Meaning, and Knowledge*, Routledge.

INDEKS OSÓB

- Ajdukiewicz K. 7, 25, 88, 115, 116, 192, 207
Arystoteles 5, 6, 9, 27, 28, 76, 77, 80, 82, 83, 113, 116, 117, 120, 171, 181
Barwise J. 131
Belnap N. D. 133
Boas F. 133
Carnap R. 89, 105, 120, 134, 194
Chomsky N. 79–81, 232, 234
Church A. 104, 105, 109, 131
Donnellan K. 229–231
Dummett M. 81, 95
Field H. 20, 41–43, 50–67
Foster J. A. 9, 145, 159–172, 215
Frege G. 29, 70, 100, 101, 104–107, 109, 118–125, 129, 131, 133
Grice H. P. 173, 203, 204, 220, 225, 226, 233, 237
Hahn L. D. 6
Hume D. 76, 82, 114, 180, 181, 187
Husserl E. 117
Ingarden R. 117
Kokoszyńska M. 25, 26
Kotarbiński T. 7, 102, 149
Kripke S. 64, 116, 118, 134
Lepore E. 6, 108, 109, 127, 199
Locke J. 76, 82, 83, 113, 114
Ludwig K. 6, 108, 109, 127, 190
Lynch M. 30
Łukasiewicz J. 133
Matthews R. J. 109
Modrak D. 77, 82
Montague R. 72, 73
Nowaczyk A. 22, 69, 70, 197
Partee B. H. 106, 107
Pawłowski T. 89
Perry J. 131
Platon 6, 120
Protagoras 193
Putnam H. 64, 116, 118, 182
Quine W. V. O. 9, 10, 55, 64, 65, 67, 76, 77, 83–86, 96, 97, 103, 104, 111, 114, 115, 131, 145, 190, 194–196, 198, 199, 201, 208–211, 213–215, 221, 223, 234, 240, 242, 245
Ramsey F. 30, 247
Russell B. 29, 63, 64, 66, 112, 229, 230
Sapir B. 133, 134
Scheffler I. 102–104
Searle J. 219, 233
Skinner B. F. 80
Soames S. 23, 26, 41, 55, 65–69
Spinoza B. 6
Strawson P. 76, 82, 84–88
Tarski A. 4, 10, 11–70, 91, 92, 96–99, 101, 130, 137, 151, 153, 156, 164, 217, 232
Whorf E. 133, 134
Żegleń U. 6

INDEKS POJĘĆ

- Aksjomaty semantycznej definicji prawdy 13, 14, 16, 17, 26, 34, 43, 48–50, 55, 57, 58, 60, 62, 93, 108, 130, 139 140, 141, 164, 174, 179, 207, 208, 213
- Akt mowy 145, 147–150
- Analiza parataktyczna 101–103, 146–150, 162, 167, 177, 168, 185, 188, 249
- Analogia pomiaru 240, 242–246
- Antykonwencjonalizm (Davidsona) 9, 91, 173, 236, 237
- Argument procy 130–136, 168, 180, 229
- Argument trzeciego człowieka 120–125
- Asercja 150
- Atom znaczeniowy 83, 93, 113
- Atrybutywne użycie deskrypcji 229, 231
- Behawioryzm (w semantyce) 80, 81, 188, 199, 209–213, 240
- Byty postulowane 54, 60, 63, 70, 104, 111, 130, 131, 182
- Definicja nominalna 82, 83, 113, 122
- Definicja realna 82
- Definicja redukcyjna 89
- Deskrypcja określona 55, 70, 85, 106, 112, 116, 128, 133, 135, 229, 230–232, 236
- Dyrektywalna teoria znaczenia 7, 29, 78, 88, 115–117, 157, 192, 194, 197
- Działanie 5–7 19, 66, 91, 141, 177, 203, 208, 238, 248, 249
- Eksplikacja 26, 28, 33, 35, 59, 137
- Empirystyczne teorie 9, 76–85, 111, 112, 115, 183, 187
- Fenomenologia 117, 119, 187
- Filozofia dialogu 187
- Fizykalistyczna dopuszczalność 21, 26, 41, 59, 60, 62–67
- Fizykalizm 8, 20, 24, 34, 59, 74
- Hermeneutyka 187, 188
- Hipoteza Sapira-Whorfa 84
- Hipotezy analityczne 199, 209, 213, 214
- Historia przyczynowa 246, 247
- Historyczno-przyczynowa teoria 64, 66, 116–118
- Holizm konformacyjny 83, 84, 213
- Holizm przekonaniowy 186, 202, 205
- Holizm znaczeniowy 69, 86, 128
- Ideacyjna teoria znaczenia 76, 171
- Idiolekt 91, 138–140, 142, 143, 167, 171–175, 178–180, 184, 190, 195, 198, 204–208, 216, 219, 221, 224, 227, 235, 243, 246, 247
- Inskrypcyjna teoria mowy zależnej 102
- Interpretacja 7–10, 59, 91, 115, 156, 159, 162, 166, 167, 169, 171–249
- Interpretacja radykalna 10, 115, 190, 191, 196, 202, 208, 214–226, 227, 247, 248
- Interpretacyjność teorii 152
- Intesja 120, 134, 135
- Kartezjanizm 180, 187
- Klasyczna koncepcja prawdy 12, 26–29, 32, 55, 96
- Koherencyjna koncepcja prawdy 19
- Koło wiedeńskie 20, 25, 60, 64, 83, 114, 195
- Kompetencja językowa 81, 86–88, 151, 152, 183, 205, 206, 209, 232–238
- Konotacja 114, 120, 172

- Konwencja T 15, 17, 19, 24, 33, 136, 138, 151, 154, 156, 157, 159, 163, 173, 217, 218, 221, 223, 225,
Konwencjonalizm (Ajdukiewicza) 116
Korespondencyjna koncepcja prawdy 19, 27–29, 55, 201
Kryterium prawdziwości 15, 19, 28, 148, 155
Maksymy konwersacyjne 204, 220
Malapropizm 227, 228, 230–232, 236, 237
Merytoryczna trafność 12, 14, 15–20
Metafora (przyczynowa teoria) 6, 197, 226, 249
Metodologiczna poprawność 12, 20–22
Mowa zależna 100–105, 112, 118, 120, 129
Możliwy świat 134, 135
M-zdanie 142, 143, 150, 151, 153, 154, 163, 164, 206
Naukowa semantyka 20, 27, 32–35, 53, 59
Nazwy puste 39, 83, 113, 114
Neopozytywizm 20, 26
Neutralność filozoficzna 23, 29, 39, 40, 44, 53–55, 64
Niedefiniowalność prawdy 29, 92, 183, 217
Niekonserwatywność 135
Niezdeterminowanie interpretacji 10, 67, 214, 238–249
Niezdeterminowanie przekładu 67, 84, 196, 199, 209, 214, 245
Okazjonalność 26, 63, 65, 92, 94, 97, 115, 144–150, 163–169, 172, 185, 199, 202, 206, 207, 210–212, 219, 222, 225, 232, 235
Opis stanu 134, 135
Ostensja 85, 86
Paradoks 18, 20, 21
Perspektywa pierwszoosobowa 178–183, 185
Perspektywa trzecioosobowa 174–183, 208, 241
Podręcznik przekładu 115, 161, 196, 199, 208–216, 221, 240, 242
Porządek odkrycia 78, 80, 83, 87, 181
Porządek uzasadniania 78, 80–83, 87, 181, 183
Postawa propozycjonalna 90–92, 104, 105, 118, 145, 146, 149, 150, 172, 175–180, 182, 184–186, 188, 190, 192, 193, 201, 205, 206, 222, 239, 241, 246–248
Powód działania 203, 248
Powszechnik (uniwersale) 77, 82, 113
Pragmatyka 172, 173
Pragmatystyczna koncepcja prawdy 19
Pragmatyzm 209
Prawda obiektywna 190, 191, 193, 199
Prawda w modelu 8, 22, 54, 69–73
Predykat dyspozycyjny 89
Produktywność 81
Przekład radykalny 10, 115, 190, 198, 212, 208–214
Przekonanie 7, 19, 57, 78, 90, 150, 166, 167, 175–183, 185, 186, 188–191, 193, 194, 196, 200–204, 212, 213, 220–226, 232, 239–243, 246, 247
Racjonalność 203, 204, 222, 224
Redukcjonizm 20, 23, 26, 32, 35, 41, 64, 83, 187
Referencyjne użycie deskrypcji 229, 230, 231
Regres w nieskończoność 121–124
Reifikacja znaczenia 112, 115, 117–126, 131, 134–136, 166, 191, 199, 214, 240
Schemat pojęciowy 10, 84, 193–200, 209, 213, 240
Semantyczna definicja prawdy (sformułowanie syntaktyczne) 40–53
Semantyczna definicja prawdy 7–10, 11–72, 81, 90–93, 98–100, 108, 111, 125, 130, 131, 136–143, 144, 155, 161, 171, 200, 232
Semantyczna definicja spełniania 40
Semantyka nieodniesieniowa 10, 70, 101, 106, 108, 109, 110–169
Semantyka odniesieniowa 9, 70, 106–109, 112–120, 122, 125, 126, 128, 180
Semantyka teoriomodelowa 54, 70–73, 101, 106, 123, 125

- Sens wyrażenia (Frege) 104–106, 108, 109, 118, 120, 123, 129, 131.
- Sensowność wyrażenia 25, 60, 82–84, 90, 92, 112–118, 126
- Sieć postaw propozycyjalnych 186, 246
- Składalność 104–109, 215
- Struktura logiczna 93
- Strukturalnoopisowa nazwa 17–19, 143, 154
- S-zdania 33, 34, 41–44, 46, 47, 49
- Świadectwo (empiryczne) 81, 90, 140, 141, 143, 155, 176, 177, 179, 180, 182, 184–189, 193, 198, 199, 201, 216, 217, 220, 236, 239, 240, 242, 244, 246, 247, 249
- Teoria bieżąca 236, 237
- Teoria decyzji 10, 247, 248
- Teoria dystalna 201, 223
- Teoria proksymalna 201, 223
- Teoria uprzednia 235–237, 247
- Teoria znaczeniowa 81, 95
- Testowanie (teorii prawdy) 30, 56, 57, 59, 64, 141, 213
- Teza Duhema-Quine’a 89
- Triangulacja 192, 246
- Warunek prawdziwości 15, 16, 18, 67, 69, 92, 104, 136–145, 148, 154, 155, 157, 158, 163–165, 172, 207, 217, 225, 232, 245, 246
- Weryfikacja teorii prawdy 30, 31, 59, 81, 90, 91, 138, 141, 144, 217, 218
- Weryfikacjonizm 60, 83, 114–118
- Wiedza intersubiektywna 184, 186–189, 192, 193, 246
- Wiedza obiektywna 184–189, 192
- Wiedza subiektywna 181, 184–189, 192, 241, 246
- Wypowiedzenie performatywne 92, 146–149
- Wyrażenia cudzysłowowe 92, 95–100, 103, 109, 146, 149
- Wyrażenie semantycznie proste 93, 94, 96, 102, 104, 105, 107, 108, 119
- Wyczuwalność 75–109, 21, 152, 223
- Zasada koherencji 200–204
- Zasada kooperacji 203–204
- Zasada korespondencji 200–204
- Zasada zastępowania ekstensjonalnego 100, 101, 104, 133
- Zasada życzliwości 10, 141, 156, 160, 188, 190, 200–204, 210, 220, 223–226, 239, 241
- Zdanie analityczne 83, 211, 218, 240
- Zdanie analityczno-bodźcowe 210–221
- Zdanie nieprzejryste poznawczo 202, 211
- Zdanie obserwacyjne 115, 202, 210–212, 218
- Zdanie protokolarne 64, 83, 114, 115, 195
- Zdanie wieczne 210–212
- Ziemia bliźniacza 182
- Znaczenie analityczne 25, 60, 114
- Znaczenie bodźcowe 115, 196, 207, 218
- Znaczenie empiryczne 60, 114
- Związek przyczynowy 181–183, 185, 186, 192, 196, 197, 200–202, 204, 226, 241, 24

PHILOSOPHY OF DONALD DAVIDSON THE BASIS OF SEMANTICS (SUMMARY)

Donald Davidson (1917–2002) was one of the most influential philosophers of the 20th century. His philosophical views are expressed in over 200 papers. They constitute a subtle philosophical system comprising theory of action, philosophy of language, epistemology, metaphysics, philosophy of mind, theory of rationality, and the applications of his philosophy in literary theory and the interpretations of classical philosophers like Plato, Aristotle, Aquinas, Spinoza and Hume. Davidson is an exceptionally demanding author who writes in an elliptical or even cryptic way. So interpreting his philosophy is a real challenge and resembles doing a complicated puzzle where every piece would have to fall in just the right spot to create a completed image. Perhaps some rival solution of particular philosophical problems could seem more attractive, or even better, but it is necessary to have in mind that Davidson's philosophy is intended as a holistic system wherein, e.g. the philosophy of action and epistemology intermingle with the philosophy of language, and confining it at the same time.

Although there is a vast English literature on Donald Davidson's philosophy, there are only few publications in Polish as well as only a few translations of Davidson's papers. The aim of this book – conceived as the first of the series – is to fill, at least partially, the gap, and to present the foundations of Donald Davidson's semantics to the Polish reader. In consequence, the presented monograph comprises the detailed overview of Alfred Tarski's semantic definition of truth, the problem of adequate theory of natural language conceived as learnable language, the theory of meaning, and the theory of interpretation. The next volume should be a monograph dealing with other topics of Davidson's philosophy of language: paratactic analysis, action sentences, causal theory of metaphor, and applications in literary theory. In spite of being original, Davidson's philosophy stems from Western philosophical tradition from Plato and Aristotle to Tarski and Quine. In consequence it is hard to appreciate Davidson's ideas without more general knowledge of

the history of philosophy. Therefore, the presentation of Davidson's thoughts is embedded in the presentation of the relevant philosophical views. This is necessary especially when Davidson only gives hints about other philosophers' positions.

In Chapter 1 *Alfred Tarski's Semantic Definition of Truth*, I give a comprehensive analysis of Tarski's semantic definition of truth or theory of truth. I start with a simple example from (1930/31) spelling out the axioms of the definition and giving a detailed proof of a T-sentence. Then I examine some details of the definition provided by Tarski, i.e. material adequacy and formal correctness. In the former case I focus on convention T, the role of T-sentences and the fact that semantic definition of truth gives truth conditions for sentences of an object language, but it does not give the criteria of being true. Tarski's demand of formal correctness is connected with his general approach to semantics implying the reduction of unclear semantic concepts to clear syntactic concepts (1936). The problem is that Tarski's definition of truth uses semantic concept of satisfaction by sequences of objects. Moreover the definition presupposes the semantic notion of translation from an object language to the meta-language. So one can suspect that, contrary to Tarski's intentions, the definition is not philosophically neutral. The problem of philosophical neutrality of theory of truth is crucial for its application in the philosophy of language as Davidson adopted Tarski's theory of truth as a theory of meaning for natural language. A theory of meaning should ascribe meanings to all meaningful expressions: those describing the world of existing objects and those in the literary fiction. The minor challenge for Davidson's program is Tarski's general skepticism about the possibility of the application of his definition to natural language.

Tarski's claim of the philosophical neutrality is partially derived from classical concept of truth formulated by Aristotle in *Metaphysics*. The semantic definition of truth was intended as an explication of this concept. The classical concept of truth is contrasted with the correspondence theory of truth that gives ontologically involved criteria for sentences to be true. Another problem concerning philosophical neutrality can be formulated as in the question – are semantic concepts in Tarski's definition really semantic or they are just syntactic concepts in "semantic disguise"? This question leads me to the detailed analysis of the semantic definition of satisfaction that constitutes the salient part of Tarski's definition of truth. Having introduced basic concepts based on hints that can be traced in Tarski, I try to formulate semantic definition of satisfaction exclusively in syntactic terms where sequences of objects are substituted by sequences of names. Then the definition of satisfaction is extended to the formulation of semantic

definition of truth exclusively in syntactic terms, but I identify the problem of this approach in case of axioms for quantifiers. The difficulty is similar to the one posed by Tarski in (1936a) where he examined the possibility of defining the notion of logical consequence exclusively in purely syntactic terms.

The failure of the enterprise does not mean that the philosophical neutrality is in danger. It suffices to follow Carnap, Quine and Davidson in treating the existence of objects as an ontological commitment of the theory, and the objects as posits of the theory. To illustrate this strategy I adapt Davidson's example of the "theory" of reference of the definite description "The father of x " from *Truth and Meaning* (1967). As everybody has a father, I formulate the theory of the definite description "The oldest son of x ". In this case it is obvious that references of descriptions are only posits of the theory, as many people do not have sons. Moreover, the theory gives us some partial insight into the meaning of the description. The theory in question does not, in fact, give the meaning of the description for two reasons. First of all, in his theory of truth Tarski presupposes the translation from the object language to the meta-language and, in consequence, the theory of reference of the definite description, "The oldest son of x ", conceived as a part of the theory of truth presupposes the meaning of the description. The second reason is that, according to Davidson, meanings can be only given by the whole theory of truth conceived as a holistic theory of meaning that represents all meaning dependencies between words in its axioms. To achieve this goal Davidson presupposed the concept of truth as basic and undefinable, and only under this assumption the theory of truth allows us the translation of words and sentences onto meta-language, i.e. give us meanings of object language expressions. In the same time the role of T-sentences has been changed. They are no longer partial definitions of truth predicate, but they serve to test the theory of truth conceived as the theory of meaning on the grounds of the available evidence.

At the end of the chapter I compare my interpretation of Tarski's semantic definition of truth with the one given by Hartry Hield in (1972), by Scott Soames in (1984), and I look into the differences between Tarski's conception of truth in a language analyzed in this chapter with his later conception of truth in a model. Though the latter conception has been developed from the former, it is the semantic definition of truth that is usually considered as being of great philosophical interest and importance. The concept of truth in a model gave rise to model semantics of natural language which has been criticized by Davidson in *Truth and Meaning*.

In the second chapter *Theory of Meaning and Learnability*, I focus on Davidson's 1965 paper, *Theories of Meaning and Learnable Languages*. The

paper is written in an extremely elliptical way and is less known than Davidson's seminal paper *Truth and Meaning*. Nevertheless, in *Theories of Meaning* Davidson is justifying and sketching, for the first time, his semantic program comprising theory of meaning and its applications to non-extensional contexts, and putting restrictions on possible theories of meaning of a natural language, i.e. he is formulating a meaning theory in Michael Dummett's terms. The main thesis of the paper is that any theory of a natural language has to take into account its learnability. Davidson starts his argument with the contention that the empiricist theories of meaning have usually wrongly interpreted the demand of learnability identifying it with the description of the process of language acquisition. In consequence, they confuse the epistemological or logical order of justification with the order of discovery involved in the process of language learning. As Davidson puts it, the adequate theory of learnable language does not call for the psychological description of language acquisition. As Davidson does not exemplify his statement, I give some examples of theories of meaning of this sort. First of all I focus on Aristotle's theory of linguistic sign from *De Interpretatione*. The concept of non-existing object like centaur (i.e. the meaning of the empty name "centaur") is introduced by a nominal definition which applies the concepts of existing objects, a human and a horse (i.e. the meanings of "human" and "horse") in *definiens*. The latter concepts have real definitions. The empiricists would maintain that people learn the concept of a centaur *via* the concepts of a human and a horse, but definitions belong to the order of justification and say nothing about the concept formation in the process of language acquisition. The Aristotelian theory of meaning has been adopted by English empiricists, Locke and Hume who explicitly stated that we learn meanings (concepts, ideas) of the empty names, constructing them by operating on simple ideas conceived as meanings of the names of existing things. But as Davidson remarks, there is no evidence that we learn concepts in this way. The conclusion is as follows: the empiricists committed two closely related errors: they took learnability restrictions on the theory of natural language for the description of language acquisition. As they could not formulate the proper psychological description of language learning, they confused the process of concept learning with the justification of the existence of concepts. The operations on concepts justify their existence but do not shed light on the process of their acquisition.

The only empiricist theory described by Davidson is Peter Strawson's critique of Quine's thesis of elimination of singular terms from language. Davidson neither analyzes nor criticizes the thesis itself, but he investigates Strawson's argument against the thesis. Strawson argues that the possible lack of singular terms in language would obstruct the description of its acquisition.

The description of this process involved in a theory of language is, according to Davidson, a typical fault of empiricism. Nevertheless there is a positive aspect of Strawson's theory. In spite of the wrong ideas concerning learnability Strawson accurately recognized the holistic and compositional aspect of natural language that should be taken into account in any adequate theory. In consequence, the proper theory should be recursive, i.e. should give a constructive account of meanings of compound expressions as a function of meanings of a finite number of semantic primitives as words and idioms. A good candidate for the theory of this kind is Tarski's theory of truth. As "learnability" is a dispositional predicate I give a short explanation of the relation between a learnable language and its adequate theory. Though it is an obvious fact that a natural language is learnable, there is no room in the theory either for the predicate "learnable", or the verb "to learn". In consequence, the adequate and illuminating theory of natural language should explain the phenomenon of its learnability in other terms, e.g. the theory should be recursive. Analogically an illuminating physical theory of solubility should not take advantage of terms like "to dissolve", but should explain the phenomenon in terms of molecular structure of soluble objects.

In the second part of the paper Davidson criticizes some theories that even have not a chance to be the theories of a learnable language. In the first place he investigates Tarski's and Quine's theories of quotations. The main sin of those theories is that they treat quotation as the name of an inscription between quotation marks. In consequence, all quotations are semantic primitives and as there are infinitely many of them, the language described by the theories is not learnable. Then he analyzes some theories of indirect speech – Israel Scheffler's inscriptional theory of "said that" phrases, and Quine's solution of belief-sentences from *Word and Object*. Both theories try to get rid of any ontological commitments and, in consequence, have to treat infinitely many expressions as semantic primitives. Therefore, they cannot be the theories of natural language. The last example is Frege-Church theory of sense and reference applied to indirect speech. From the ontological point of view the theory is extremely generous as it presupposes normal references (senses of level 0), normal senses (senses of level 1), meta-senses (senses of level 2) etc. The idea is that in non-extensional context the reference of an expression is its normal sense, and its sense becomes the sense of level 2. As we can iterate non-extensional phrases (e.g. "John thought that Mary believed that Ann doubted that ...") we need, for every expression of a language, an unlimited sequence of senses. As there is no rule of how to construct the sense of level $n+1$ from the sense of level n , every expression of a language represents an infinite sequence of semantic primitives. Therefore, any language described by

the theory cannot be counted as learnable. In the end I examine the Fregean notion of compositionality of meanings and Davidson's notion of learnability of language. The former does not require the finite number of semantic primitives. But the most important difference is that the latter leads to the theory that does not reify meanings. The justification of this fact was given by Davidson in *Truth and Meaning* which I discuss in the next chapter.

I start Chapter 3, *The Non-referential Semantics*, with the distinction of the two types of non-referential semantics. The first type presupposes the existence of, and makes use of, referents of particular expressions as a necessary condition of meaningfulness of all other expressions of a language. The reference, of course, does not need to be a material entity – it could be a mental entity, Platonic idea, a connotation, an impression, or even a stimulus. The latter reifies the meanings, i.e. treats them as special kind of referents. In consequence, every theory of the second type belongs to the first one, but not *vice versa*. In the first group we find Aristotelian theory of sign, Locke's and Hume's psychological theories of meaning, Mill's semantic theories of connotative and non-connotative names, verificationist theory of meaning, Saul Kripke's and Hilary Putnam's causal theories of naming, phenomenological theories of meaning (Edmund Husserl and Roman Ingarden), and Quine's theory of meaning in terms of translation manual. With some hesitation, I also include Kazimierz Ajdukiewicz's directival theory of meaning in the first type. The second type would comprise Frege's theory of sense (in consequence, it must be included also in the first type), the Aristotelian theory of linguistic sign, psychological theories of meaning, Mill's semantic theory of connotative names, and phenomenological theories of meaning. I exclude from this type the verificationist theory of meaning, the causal theory of naming, Quine's theory of meaning in terms of translation manual, and the directival theory of meaning I neglect here Putnam's representation of meaning as a vector, Quine's notion of a stimulus meaning, and representing meanings Ajdukiewicz's matrix of language.

Davidson's arguments from *Truth and Meaning* are aimed at both types of referential theories, therefore, he does not distinguish between a meaning and a referent. The first Davidson's argument which I analyze and develop in relation to the demand of learnability, can be called The Third Man Argument. It resembles Plato's argument from *Parmenides* against his own theory of ideas. Plato's problem can be formulated as follows: two people are humans because they both exemplify the same Idea of man. But then the problem appears – why a particular man exemplifies the Idea of man, or what is the criterion of this relation. In order to solve this problem one has to introduce some higher order idea exemplified by both – a particular man, and the Idea of

man, what leads to an infinite regress. In case of reified meanings or referents the problem is similar. As an example we can adopt, after Davidson, a simple sentence "Theaetetus flies", where the words "Theaetetus" and "flies" are joined by concatenation. If both words had reified meanings, we would face a serious problem how to constitute a reified meaning of the whole sentence from reified meanings of its parts. As concatenation is the syntactic, and not semantic device, in order to solve the problem we have to ascribe to it also a reified meaning, and represent it on the syntactic level by the word "concatenation". Unfortunately this maneuver does not solve the problem because we have now three words and two concatenations in the sentence, and, in consequence, we start the infinite regress. According to Davidson, Frege tried to hide the problem dividing expressions into saturated and unsaturated. I show that the solution based on Frege's proposal leads to a semantics that violates the demand of learnability. In case of our example one has to ascribe to "flies" a function from reified meanings of proper names to reified meanings of all sentences of the form "x flies". In general case if we presuppose the compositionality of reified meanings, then in order to know reified meanings of functors, we have to know the reified meaning of all sentences of the language first.

In *Truth and Meaning* Davidson approximates step by step the semantic definition of truth as a non-referential theory of meaning. He starts with the already cited theory of the description "the father of x". The theory does not identify or presuppose the existence of a referent for any particular expression, but it treats objects (i.e. the referents of the descriptions) instrumentally as the posits of the theory. The next step consists in formulating a similar theory of a predicate where the problem of referents of sentences appears. What is then a referent of a declarative sentence – a fact, an event, or a state of affairs? To solve the problem Davidson employs the famous but very controversial Slingshot Argument concluding that the only possible referent of a declarative sentence is its true value. Though Davidson argues that the Slingshot Argument works for all possible referents or reified meaning, I indicate that he is right only if there is no relation between the referent or reified meaning of a sentence and the truth value. The problem appears if one applies the notion of intension conceived of as the function from possible worlds or state descriptions to logical values, where the Slingshot Argument does not work. Davidson's conclusion is that Tarski's theory of truth can be considered as a theory of meaning that fulfills all earlier demands, and as it neither uses reified meanings nor its formulation contain the word "meaning", it can be metaphorically nicked as "a theory of meaning without meanings".

Since T-sentences, that are the logical consequences of the theory, can be used to its empirical verification, the theory can be considered as an empirical theory. Though, as Davidson argues, no adequate theory of meaning can contain the term “meaning” which is a non-extensional term, we can exchange in T-sentences of the theory of truth accepted as the theory of meaning for a given language, the phrase “(...) is true iff (...)” for “(...) means that (...)”, and obtain M-sentences that give us a translation of sentences of the object language into its meta-language. This trick can be accepted as long as we approve of the theory of truth. Nevertheless if we encounter evidence that compelling to falsify it and, in consequence, to change the theory, we have to do it within the very theory of truth. The problem of translation of sentences from the object language to the meta-language (given by T-sentences or M-sentences) appears when we take into account occasional sentences with indexical devices. In this case the relevant T-sentences do not give us their translations into the meta-language, but rather give their interpretations. In the last part of the chapter I analyze in details the well known John Foster’s critique of Davidson’s theory of meaning (1976) and Davidson’s *Reply to Foster* (1976) where he explains some problems of his theory of interpretation as an extension of theory of meaning.

The last chapter, *The Theory of Interpretation*, focuses on some aspects of Davidsonian theory of interpretation that should not be considered as a new theory that substituted his theory of meaning, but rather as its extension. In a nut-shell: theory of interpretation focuses on the process of arriving at a formulation of theory of meaning, i.e. theory of truth, and on the process of falsifying and changing the theory. From the point of view of Davidsonian theory of interpretation there is no sense in maintaining that there is something like a language conceived as an abstract code constituted by conventions. Instead Davidson uses the notion of individual idiolect that evolves in time. What we call a language (e.g. English or Polish) is just a class of similar idiolects. The similarity in question consists in triviality or nearly-triviality of theory of truth built in one idiolect for another, i.e. the fact that all (or nearly all) T-sentences that are logical consequences of the theory of truth are of the form “The sentence ‘*p*’ is true iff *p*”. In consequence, the conventions (or rather interpretative standards) do not constitute the language, but are just the means to make mutual process of interpreting utterances more efficient.

The process of interpreting consists in two interrelated actions –ascribing meanings to speakers’ utterances (i.e. constructing a theory of truth) and ascribing propositional attitudes, especially beliefs, to the speakers. The two processes are inseparable, and the resultant theory of truth cannot be taken in isolation without the causal history of the process of interpretation. Let

us imagine that one interprets her own text written some years ago (so she treats herself from the third person perspective), and let us assume, for the sake of argument, that though the interpreter did not change her vocabulary, she now ascribes different meanings to some utterances of the same physical shape. The theory of truth constructed in the past is trivial, as well as the theory of truth constructed in the present for her present utterances (but the theory constructed now for her past utterances is not trivial). The conclusion that I draw from the example is as follows – the theory of truth is only a necessary condition of meaning ascription. The conclusion is obvious from the point of view of the first person perspective. The triviality of the definition of truth for the speaker's own utterances does not suffice for her to ascribe meanings to her own words. What constitutes meanings is the fact that she holds true some utterances of sentences of her own idiolect. This leads to Ajdukiewicz's directival theory of meaning where meanings are constituted by beliefs.

The first person perspective is contrasted with the third person perspective in various aspects, which leads Davidson to distinguish three varieties of knowledge that are involved in the process of interpretation and must be always expressible in a linguistic form. Subjective knowledge is the knowledge of my own propositional attitudes (beliefs, desires, preferences etc.) that I know directly from the privileged first person perspective. This knowledge does not demand any evidence. Objective knowledge is the knowledge about the external world existing independently of me, and in this case the empirical evidence is indispensable. Intersubjective knowledge is about propositional attitudes of other people. In this case we have to take into account various evidence: empirical evidence, human behavior, and the words of the speakers. No type of knowledge can be reduced to any other. They are acquired in the process of triangulation that is in fact the process of the first language acquisition. Triangulation makes the interpretation possible, as in this process we acquire the concept of absolute truth which is an indispensable condition of constructing any theory of truth as a theory of meaning.

The concept of absolute truth is the only one of the various conditions of interpretation. Another condition is the rejection of the concept of a conceptual scheme presupposed, sometimes tacitly, by many anthropologists and philosophers, e.g. Carnap and Quine. Though the concept has never been well defined, it has been usually declared as an obstacle in understanding other people. In (1974), Davidson criticizes the concept as unclear and characterized in metaphorical terms and, in consequence, he rejects it. A positive condition of interpretation is the Principle of Charity formulated by Davidson in different papers in various ways. In (1991) Davidson gives the ultimate

formulation of the Principle as a pair consisting of the Principle of Coherence and the Principle of Correspondence. The former states that while interpreting other person's utterances we have to assume her to be a rational agent, and the latter states that the beliefs of the interpreter as well as the beliefs of the interpreted person correspond, in a great part, to the external world. Davidson conceives this correspondence as a causal relation between events and beliefs.

In order to have an insight into the very process of interpretation, Davidson introduces the notion of the radical interpretation analogically to Quine's notion of the radical translation. The interpretation "from the roots" is not so frequent as every day interpretation of the utterances of the users "of the same language" who use their own idiolects. This typical everyday interpretative situation was described by Davidson in (1986) where he focused mainly on malapropisms and two ways of use of definite descriptions – the attributive and the referential. The conclusion is that interpretation is a dynamic process that cannot be adequately described as an actualization of any kind of traditionally conceived linguistic competence. In consequence the traditional concept of linguistic competence as well as language should be rejected.

The last part of the book is devoted to the controversial problem of the indeterminacy of interpretation. As the process of interpretation consists of the ascription of meanings and the ascription of beliefs, the process is undetermined, but this indeterminacy cannot be exemplified, i.e. two interpreters can arrive at different interpretations of the same utterances not being conscious of the fact. Moreover they even cannot tell where the difference is. In order to make the notion of indeterminacy more intelligible Davidson introduced the measurement analogy. This analogy was strongly criticized by many authors mainly because of the fact that if two people measure something in different scales, e.g. the temperature in scales of Celsius and of Fahrenheit, there is always a simple and known formula to transform the obtained values, e.g. from °C into °F and *vice versa*, while nothing like this is possible in case of two interpretations. In order to save the analogy I give an example of the two thermometers without numerical values but with a system of dots and dashes. In consequence one person can have in mind Celsius' scale and another person the Fahrenheit's, they can communicate perfectly unconscious of the fact of measuring in different scales.

SPIS TREŚCI

Wstęp	5
Rozdział 1. Alfreda Tarskiego semantyczna definicja prawdy	11
1. Prosty przykład semantycznej definicji prawdy	12
2. Wymóg merytorycznej trafności	15
3. Metodologiczna poprawność	20
4. Filozoficzne problemy związane z semantyczną definicją prawdy	23
5. Semantyczna definicja prawdy a klasyczna koncepcja prawdy	27
6. Rola T-zdań w semantycznej definicji prawdy	30
7. Program naukowej semantyki	32
8. Wstępna rekonstrukcja semantycznej definicji spełniania	35
9. W stronę semantycznej definicji prawdy w sformułowaniu syntaktycznym	40
10. Uogólnienie definicji spełniania do definicji prawdy	45
11. Problemy semantycznej definicji prawdy w sformułowaniu syntaktycznym	49
12. Semantyczna definicja prawdy w sformułowaniu semantycznym	53
13. Holistyczny aspekt semantycznej definicji prawdy	57
14. Hartry Fielda interpretacja semantycznej definicji prawdy	59
15. Scotta Soamesa interpretacja semantycznej definicji prawdy	65
Dodatek. Pojęcie prawdy w modelu	69
Rozdział 2. Teoria znaczenia a wyuczalność	75
1. Empirystyczne teorie znaczenia i ich krytyka	77
2. Krytyka poglądów Strawsona	84
3. Wyuczalność języka a teoria znaczenia	88
4. Krytyka teorii wyrażeń cudzysłowowych	95
5. Krytyka wybranych teorii mowy zależnej	100
6. Wyuczalność a składalność	105
Rozdział 3. Semantyka nieodniesieniowa	111
1. Przegląd wybranych semantyk odniesieniowych	112
2. W stronę semantyki nieodniesieniowej	119
3. Teoria odniesienia bez odniesień	125
4. Argument procy	130
5. W stronę Alfreda Tarskiego semantycznej definicji prawdy	136
6. Okazjonalność	144
7. Johna Fostera zarzuty wobec teorii prawdy jako teorii znaczenia	150

Rozdział 4. Teoria interpretacji	171
1. Perspektywa trzecioosobowa	174
2. Ontogeneza procesu interpretacji – trzy rodzaje wiedzy i triangulacja	183
3. Warunki możliwości interpretacji – krytyka schematu pojęciowego	193
4. Warunki możliwości interpretacji – Zasada życzliwości	200
5. Holistyczny aspekt interpretacji	204
6. Przekład radykalny	208
7. Interpretacja radykalna	214
8. Interpretacja w praktyce	227
9. Niezdeteminowanie interpretacji	239
Bibliografia	251
Indeks osób	257
Indeks pojęć	259
Summary	263