

MARK SAINSBURY

LOGIKA FILOZOFICZNA*

WPROWADZENIE

Wiedząc, że wszyscy bogaci ludzie są szczęśliwi i Jan jest bogaty, możemy wyciągnąć wniosek, że Jan jest szczęśliwy. Wiedząc, że ziemniaki gotują się od dwudziestu minut, możemy wyciągnąć wniosek, że ziemniaki są już ugotowane.

Mamy tu do czynienia z wyraźnie różnymi rodzajami wnioskowania. Pierwszy, który można nazwać wnioskowaniem *dedukcyjnym*, charakteryzuje się tym, że jeśli to, z czego wyprowadzamy wniosek, jest prawdziwe, to jest zupełnie niemożliwe, aby to, co wyprowadzamy, było fałszywe. Drugi, który możemy nazwać *indukcyjnie mocnym* wnioskowaniem, charakteryzuje się tym, że chociaż to, z czego wyprowadzamy wniosek, dostarcza dobrej racji dla tego, co wyprowadzamy, to racja ta nie jest rozstrzygająca. Na dużych wysokościach woda gotuje się w niższej temperaturze, przez co czas jej gotowania ulega wydłużeniu. Zresztą nawet na normalnych wysokościach zdarzają się bardzo twarde ziemniaki!

Niniejszy esej dotyczy wnioskowania pierwszego rodzaju, czyli wnioskowania, które, przynajmniej w zamierzeniu, jest dedukcyjne. [...] Własnością rozumowania, która nadaje mu wartość i która interesuje logików, jest to, że zachowuje ono prawdziwość czy, inaczej mówiąc, że jest *niezawodne*. (Spra-

Prof. MARK R. SAINSBURY – Department of Philosophy, University of Texas, Austin; e-mail: marksainsbury@mail.utexas.edu

* Niniejszy tekst jest nieznacznie skróconą wersją eseju *Philosophical Logic*, [w:] A. C. Grayling (ed.), *Philosophy: A Guide Through the Subject*, t. 1, Oxford: Oxford University Press 1995, s. 61-122. Przekład za zgodą Autora i Oxford University Press.

wa ta została dokładniej omówiona w SAINSBURY 1991: rozdz. 1.) Zazwyczaj rozumowanie przechodzi od jednego lub kilku twierdzeń, które traktuje się – przynajmniej tymczasowo – jako dane, do pewnego innego twierdzenia. Te wyjściowe twierdzenia nazywa się *przesłankami*, a twierdzenia końcowe – *wnioskiem*. Zbiór twierdzeń złożony z przesłanek i wniosku nazywa się *wnioskowaniem*. Standardowa definicja niezawodności wnioskowania przedstawia się następująco:

Wnioskowanie jest niezawodne zawsze i tylko, gdy jest niemożliwe, aby wszystkie przesłanki były prawdziwe, a wniosek fałszywy.

Aby ustalić, czy wnioskowanie jest niezawodne, trzeba zbadać odniesienie przedmiotowe. Weźmy przykład. Czy wnioskowanie przechodzące od przesłanki „Napoleon był dynamiczny” do wniosku „Ktoś był dynamiczny” jest niezawodne? Inaczej mówiąc, czy jest niemożliwe, aby przesłanka była prawdziwa, a wniosek fałszywy? Skłonni jesteśmy odpowiedzieć twierdząco. Czy w takim razie w podobny sposób należy ocenić wnioskowanie przechodzące od przesłanki „Nie istnieje największa liczba pierwsza” do wniosku „Coś nie istnieje”? W tym wypadku powinniśmy poczuć pewien opór przed odpowiedzią twierdzącą. Skoro przesłanka jest prawdziwa, a niezawodne wnioskowanie prowadzi zawsze od prawdziwej przesłanki do prawdziwego wniosku, to twierdząca odpowiedź na to pytanie zmusiłaby nas do uznania, że coś nie istnieje. W najlepszym razie wydaje się to dziwne, a w najgorszym – szalone. Jednym z celów tego eseju jest doprowadzenie czytelnika do wyrobienia sobie własnego zdania w tego rodzaju kwestiach.

Pierwsza część eseju dotyczy najbardziej, jak się wydaje, oczywistych przykładów słów, których zadaniem jest odnoszenie się do przedmiotów, na przykład nazw „Napoleon” i „Francja”. W tej części przedstawię teorie Johna Stuarta Milla, Gottloba Fregego, Bertranda Russella i Saula Kripkego.

Tematem drugiej części eseju są deskrypcje określone, czyli wyrażenia o postaci „(jedyne) to-a-to”, na przykład „pierwszy cesarz Francji”. Wprawdzie na pierwszy rzut oka wydaje się, że są to wyrażenia, których zadaniem jest odnoszenie się do przedmiotów, ale „teoria deskrypcji” Bertranda Russella prowadzi do odmiennego wniosku. Zapoznamy się z tą teorią, z argumentami Russella na jej rzecz oraz z zarzutami, jakie wobec niej wysunięto. (Główne z nich są omówione w SAINSBURY 1979: rozdz.4, warto jednak również poznać oryginalne źródła tych zarzutów.)

Trzecia część eseju analizuje problem dotyczący istnienia. Jeśli mówimy, że coś nie istnieje, to – jak się wydaje – zakładamy w ten sposób, że to

istnieje. Przedstawię znaną teorię na ten temat pochodzącą od Russella oraz nowsze ujęcie, którego autorem jest Gareth Evans.

Część czwarta dotyczy twierdzeń o identyczności, a zwłaszcza tego, czy te z nich, które są prawdziwe, są prawdziwe zawsze i w sposób konieczny. Rozważania na ten temat skierują nas ku pojęciu konieczności, które zostanie dokładniej omówione w części piątej na podstawie tekstów Alana Gibbarda, Saula Kripkego i Davida Lewisa.

Część szósta w sposób dość pobieżny omawia pojęcie prawdy. [...] Próbowałem nakreślić mapę niektórych głównych sporów, nie wchodząc w szczegóły. Jeśli któryś z wątków wzbudzi zainteresowanie czytelnika, konieczne będzie sięgnięcie po jedno z sugerowanych na końcu tego eseju opracowań.

Do najważniejszych powodów, dla których warto studiować filozofię, należy możliwość nauczenia się, jak formułować własne poglądy i jak ich bronić. Własny pogląd nie musi być poglądem, którego nikt wcześniej nie głosił. W logice trudno w ogóle spotkać nowe poglądy. Własny pogląd to taki, do którego jest się przekonany i którego potrafi się bronić. W trakcie lektury czytelnik zetknie się z wieloma konkurencyjnymi poglądami i będzie musiał zdecydować, które z nich uznać za własne.

Niekiedy, zwłaszcza na początku studiów, trudno jest się zdecydować, która z konkurencyjnych teorii jest najlepsza. Ważne jednak, aby podejmować przynajmniej „mniejsze” decyzje: czy to wnioskowanie ostatecznie ustala lub obala daną teorię? Czy danego zarzutu można uniknąć odpowiednio modyfikując tę teorię? Czy nie można by wyróżnić przynajmniej dwóch wersji koncepcji, że jest tak-a-tak? Aby odkryć wartość studiowania filozofii, trzeba sobie stawiać takie pytania i na nie odpowiadać.

1. NAZWY

1.1. *Mill*

Takie nazwy jak „Napoleon” i „Francja” wyglądają na wzorcowe przypadki wyrażen, których celem jest odnoszenie się do przedmiotów w świecie. Najprostsza teoria głosi, że na tym można poprzestać: „Napoleon” odnosi się do Napoleona, a „Francja” do Francji. Mniej więcej taką teorię zaproponował John Stuart Mill.

Mill wprowadził rozróżnienie między nazwami ogólnymi, takimi jak „biały” (występującymi np. w zdaniu „Ta róża jest biała”), a nazwami jedno-

stkowymi, takimi jak „Dartmouth” i „obecny premier Wielkiej Brytanii”. Nazwa jednostkowa ma się odnosić do dokładnie jednego przedmiotu (denotować ją), podczas gdy nazwa ogólna może się odnosić – i na ogół rzeczywiście się odnosi – do wielu przedmiotów (denotuje je). Niektóre nazwy, takie jak „biały” i „obecny premier Wielkiej Brytanii”, są *konotatywne*: konotują one własność i denotują wszystkie i tylko te przedmioty, które posiadają tę własność. Toteż „biały” denotuje wszystkie białe przedmioty i denotuje każdą z nich na mocy posiadania przez nią własności, którą „biały” konotuje, mianowicie bieli.

Nazwa własna, taka jak na przykład „Dartmouth”, to według Milla nazwa jednostkowa, która nie konotuje,. Co prawda mogliśmy mieć powód, aby nadać czemuś taką, a nie inną nazwę. Jakiegoś człowieka możemy nazwać „Janem”, ponieważ takie było imię jego ojca, a pewne miasto nazwać „Dartmouth”, ponieważ leży ono u ujścia rzeki Dart.

Ale nie jest częścią znaczenia słowa „Jan” to, że ojciec osoby tak nazwanej nosił to samo imię; ani nawet nie jest częścią znaczenia słowa „Dartmouth” to, że miasto to jest położone u ujścia rzeki Dart. Gdyby piasek zasypał ujście tej rzeki albo gdyby trzęsienie ziemi zmieniło jej bieg i odsunęło ją na pewną odległość od tego miasta, to nazwa jego nie zmieniłaby się z konieczności rzeczy [...] Imiona własne są przywiązane do przedmiotów samych przez się i nie są zależne od tego, czy trwa jakaś cecha danego przedmiotu.

(MILL 1843: 26 [przekład polski – t. I, s. 52-53])

Mill wysuwa dwa zarzuty przeciwko sugestii, że nazwy własne konotują. Po pierwsze, ewentualne konotowane własności nie określają denotacji nazwy. Po drugie, nazwę własną można zrozumieć, nawet nie znając kandydatów na jej konotację. (Który z naszych dwóch przykładów, „Jan” i „Dartmouth”, najlepiej pasuje do której z tych uwag?)

Szczegółowe uwagi Milla wydają się trafne: zamierzone konotacje nazw „Jan” i „Dartmouth” są niezadowolające. Aby jednak przejść od nich do wniosku, że nazwy własne nie konotują, Mill musi dokonać dwóch uogólnień. Po pierwsze, musi wykazać, że nie istnieją inni, lepsi kandydaci na konotację nazw „Jan” i „Dartmouth”. (Czy przychodzą ci na myśl jacyś kandydaci?) Po drugie, musi następnie wykazać, że jest tak w wypadku wszystkich nazw własnych, a nie tylko w wybranych przez niego przykładach.

Można by podsunąć Millowi następujący ogólniejszy argument: dwoje ludzi, którzy używają jakiejś nazwy własnej, może wiązać z nią całkiem inne informacje, tak że nie istnieje żaden wspólny im element, który byłby kono-

tacją tej nazwy. Ta uwaga jest trafna tylko pod warunkiem, że znajomość konotacji musi być wspólna osobom, które skutecznie się porozumiewają. (Czy Mill jest do tego zobowiązany?)

Teoria Milla staje przed dwoma dość poważnymi trudnościami: musi ona uporać się z takimi nazwami jak „Zeus” i „Święty Mikołaj”, które nie odnoszą się niczego, oraz z „łamigłówką Fregego”.

Nazwa, która nie odnosi się do niczego, nie ma denotacji. W myśl ujęcia Milla nie ma ona również konotacji. A zatem, w myśl jego teorii, nazwa bez denotacji nie posiada żadnych własności semantycznych. To zaś oznacza, że powinna być niezrozumiała jako część języka. Tymczasem wydaje się, że nazwy bez nosicieli są w sposób całkiem zrozumiały używane w naszym języku.

Łamigłówka Fregego dotyczy identyczności i doprowadziła ona do opracowania jego teorii sensu (*Sinn*) i odniesienia przedmiotowego (*Bedeutung*).

1.2. Frege

Frege twierdził, że nie ma gwarancji, iż dwóch nazw tego samego przedmiotu można używać zamiennie; gdyby jedyną istotną własnością nazwy było jej odniesienie, to mielibyśmy taką gwarancję. Przypuśćmy, że „Hesperos” i „Fosforos” odnoszą się do tego samego ciała niebieskiego. Historia jest taka, że nazwę „Hesperos” nadano planecie Wenus ze względu na to, iż była ona widoczna wieczorem, natomiast nazwę „Fosforos” nadano tej samej planecie ze względu na to, że była ona widoczna rankiem (w innym punkcie nieba i w innej porze roku). Starożytni astronomowie nie zdawali sobie sprawy, że jest to jedno i to samo ciało niebieskie (planeta, a nie gwiazda), które ukazuje się w różny sposób.

Porównajmy teraz twierdzenia:

(1) Hesperos jest to Hesperos.

(2) Hesperos jest to Fosforos.

Wydaje się, że mówią one coś innego. (1) jest banalne, natomiast (2) mogłoby wyrażać cenną nową informację: oba twierdzenia „w oczywisty sposób różnią się pod względem wartości poznawczej”. A jednak różnią się one tylko tym, że nazwa „Fosforos” zastąpiła jedno z wystąpień nazwy „Hesperos”. Frege wyciągnął z tego wniosek, że dwóch nazw tego samego przedmiotu nie można używać zamiennie, i dlatego sposób funkcjonowania nazwy nie sprowadza się wyłącznie do posiadania przez nią odniesienia.

Frege nazwał tę dodatkową własność nazwy jej sensem (*Sinn*). W swoim artykule „Sens i odniesienie” rozpoczyna on od określenia sensu nazwy jako

„sposobu, w jaki przedmiot jest nam dany”. Następnie sugeruje, że „sens nazwy chwyta każdy, kto zna wystarczająco język” (FREGE 1892a: 158 [przekład polski – s. 62]). Pojawiają się kolejne tezy: sens określa odniesienie (to jest wyrażenia, które zgadzają się pod względem sensu, muszą się również zgadzać pod względem odniesienia), ponadto o całych zdaniach również można powiedzieć, że mają sens i odniesienie: sensem całego zdania jest myśl, a jego odniesieniem – wartość logiczna, w zależności od przypadku: Prawda lub Fałsz.

Argument Fregego o „Hesperosie” i „Fosforosie” pokazuje, że funkcja nazwy nie sprowadza się wyłącznie do posiadania przez nią odniesienia, pozostawia jednak otwartą sprawę, czy ta dodatkowa własność jest Fregeowskim sensem, czy może raczej jest ona czysto syntaktyczna albo nie jest czymś, co zmienia się w zależności od osoby, i z tego względu nie może być stabilną publiczną własnością języka. [Zwraca na to uwagę SALMON (1986).] Należy zatem zgodzić się, że argument Fregego uzasadnia słabszą tezę, iż nazwa to nie tylko jej odniesienie. Jednakże celem artykułu Fregego jako uzasadnienie mocniejszej tezy, że tym czymś dodatkowym jest sens.

Frege informuje nas, że przyjmował wcześniej (1879) pogląd, który teraz odrzuca. W myśl wcześniejszej teorii Fregego zdanie identycznościowe, takie jak „Hesperos jest to Fosforos”, można potraktować jako mówiące, że dwie nazwy: „Hesperos” i „Fosforos” odnoszą się do tego samego przedmiotu. Zgodnie z tym „poglądem metajęzykowym”, identyczność jest relacją między nazwami („rozważania dotyczyłyby samych znaków”). Wydaje się, że ten pogląd dopuszcza różnicę pod względem wartości poznawczej między „Hesperos jest to Hesperos” a „Fosforos jest to Fosforos”, ponieważ być może jest banalne, że „Hesperos” i „Hesperos” odnoszą się do tego samego obiektu, natomiast to, że „Hesperos” i „Fosforos” odnoszą się do tego samego obiektu, jest bardziej interesujące. Ten pogląd metajęzykowy jest jednak nie do przyjęcia. Pociągałby on bowiem twierdzenie, że w zdaniach identycznościowych nie mówimy faktycznie o świecie, lecz o języku, a zatem nie mogłby wyjaśnić naszej zdolności do używania takich zdań do wyrażenia wiedzy o świecie („Zdanie $a = b$ dotyczyłoby zatem nie przedmiotu, lecz jedynie naszej symboliki, nie wyrażając właściwie żadnej wiedzy”). Ponadto istotne fakty dotyczące nazw są „arbitralne”: jest po prostu kwestią konwencji, że używamy nazw „Hesperos” i „Fosforos”, aby odnosić się do tego samego przedmiotu, nie jest natomiast kwestią konwencji, że Hesperos jest Fosforesem.

A zatem, w myśl mojej interpretacji, Frege uznaje swój wcześniejszy pogląd za pogląd metajęzykowy i odrzuca go teraz z podanych powodów: czyni

twierdzenia identycznościowe przygodnymi, arbitralnymi i dotyczącymi nie świata, lecz spraw, o których rozstrzyga konwencja językowa.

Według Fregego nazwa, która odnosi się do czegoś (ma *Bedeutung*), posiada również sens. Czy każda nazwa, która posiada sens, ma również *Bedeutung*? Odpowiedź Fregego wydaje się jasna: „Wyrażenie „szereg najwolniej zbieżny” ma sens, ale można dowieść, że nie ma odniesienia” (FREGE 1892a: 159 [por. przekład polski – s. 63]). Ta interpretacja została zakwestionowana przez Garetha Evansa (1982), który argumentuje, że chociaż Frege definitywnie dopuszczał sens bez odniesienia w kontekstach fikcyjnych, to nie powinniśmy wyciągać z tego wniosku, że sens bez odniesienia jest rzeczywiście możliwy. Mówiąc, że nazwa fikcyjna, taka jak „Odyszeusz”, nie posiada odniesienia, mówimy, że nie posiada ona żadnego *rzeczywistego* odniesienia; bez wątplenia jednak ta nazwa miała posiadać odniesienie w danej historii. Można by jednak również wysunąć sugestię, że ta nazwa nie posiada żadnego *rzeczywistego* sensu: raczej, opowiadając historię, udajemy, że nazwa ma sens. A zatem zgodnie z poglądem, który Evans przypisuje Fregemu, sens i odniesienie są nierozłączne: w wypadku, gdy mamy do czynienia z fikcyjnym czy udawanym sensem, mamy równocześnie do czynienia z fikcyjnym czy udawanym odniesieniem; tam, gdzie mamy rzeczywisty sens, mamy rzeczywiste odniesienie.

Jest to niewątpliwie ciekawy pogląd i być może to właśnie powinien był powiedzieć Frege. Jednakże, w najlepszym razie, nie mówił tego za każdym razem, jak tego dowodzi cytat o najwolniej zbieżnym szeregu.

Widzimy teraz, jak teoria Fregego podchodzi do dwóch zagadek, które zadaliśmy w związku z Millem. Frege posłużył się rozróżnieniem między sensem a odniesieniem do rozwiązania zagadki identyczności: prawda lub fałszywość jest ustalana przez odniesienie, natomiast wartość poznawcza – przez sens. W myśl proponowanej tu standardowej interpretacji, możliwość sensu bez odniesienia wyjaśnia, jak nazwy nieposiadające denotacji mogą mimo to odgrywać rolę semantyczną.

Ktoś mógłby oskarżyć Fregego o traktowanie sensu jako tego czegoś, co jest potrzebne do wyjaśnienia zagadki identyczności; w takim wypadku jego wyjaśnienie byłoby nieadekwatne, ponieważ sprowadzałoby się do innego nazwania problemu. Aby wyjaśnienie Fregego było zadowalające, musi on podać taki opis sensu, który wykracza poza scharakteryzowanie go jako tego czegoś, co wyjaśnia zagadkę.

Frege bez wątplenia podaje niezależne własności sensu. Jak wspomniałem, rozpoczyna od powiedzenia, że sens nazwy jest sposobem, w jaki przedstawia

ona swoje odniesienie, a następnie mówi, że sens nazwy jest chwytny przez każdego, kto rozumie język. Jednakże te własności prowadzą do niepokojącego napięcia. Ten sam przedmiot może się różnie przedstawiać różnym ludziom; intuicyjnie jednak to, co trzeba uchwycić, aby zrozumieć nazwę, jest tym samym dla wszystkich, którzy ją rozumieją. A zatem sposób przedstawiania bywa idiosynkratyczny, podczas gdy wydaje się, że to, co musi być uchwycone, nie powinno takie być.

Frege dostrzega to w artykule „Sens i odniesienie” (Porównaj również bardziej zaawansowane rozważania w FREGE 1918):

Co do sensu prawdziwych imion własnych – jak „Arystoteles” – zdania mogą się co prawda różnić. Można np. przyjmować tu za sens – uczeń Platona i nauczyciel Aleksandra Wielkiego. Kto tak czyni, ten łączy ze zdaniem „Arystoteles pochodził ze Stagiry” inny sens niż ktoś, kto przyjmuje tu za sens: nauczyciel Aleksandra Wielkiego pochodzący ze Stagiry.

[przekład polski – s. 62]

Następnie Frege mówi, że te zmiany można tolerować dopóki nie prowadzą one do zmiany odniesienia, jednakże powinno się je usunąć z doskonałego języka. Tolerancyjna pierwsza część tej uwagi pociąga albo to, że pewne nazwy, takie jak „Arystoteles”, nie mają jednoznacznego użycia, albo to, że w ich wypadku rozumienia nie należy identyfikować z ujmowaniem sensu.

Sądzę że Frege nigdy nie rozwiązał tego konfliktu w sposób zadowalający. W tej chwili najważniejszą rzeczą jest zdanie sobie sprawy z jego istnienia; spróbujcie sformułować go tak jasno, jak tylko potraficie.

1.3. *Russell*

Teoria nazw Russella ma szansę rozwiązać wspomniany konflikt. Aby ją przedstawić, muszę najpierw poczynić kilka uwag wstępnych.

Russellowi wydawało się oczywiste, że nazwa własna powinna się zachowywać w sposób opisany przez Milla. (O ile wiem, w tym kontekście nie wymienił on Milla z nazwiska, wiemy jednak, że czytał jego pracę.) Powiada on: „nazwa własna jako taka jest jedynie środkiem wskazania przedmiotu” (RUSSELL 1918-1919: 245) i z entuzjazmem przystępuje do rozwiązania dwóch „problemów”, na które, jak widzieliśmy, natrafia ten pogląd: istnienia i identyczności.

„Gdyby [„Romulus”] rzeczywiście był nazwą własną, to problem istnienia w ogóle nie mógłby powstać, ponieważ nazwa musi coś nazywać, albo nie jest

nazwą” (1918-1919: 243). Russell dochodzi do wniosku, że wbrew pozorom „Romulus” nie jest nazwą własną, ponieważ powstaje problem, czy Romulus istniał (jest on postacią jedynie legendarną). Mamy więc nowy rodzaj teorii: nazwy własne dzielą się na co najmniej dwie kategorie i „Romulus” nie należy do tych, które funkcjonują w sposób opisany przez Milla.

Z podobną strukturą mamy do czynienia w wypadku identyczności: „Jeśli jeden przedmiot ma dwie nazwy własne, to stwierdza się dokładnie to samo używając którejkolwiek z nich”. Stąd gdyby „Hesperos” i „Fosforos” rzeczywiście były nazwami własnymi, to stwierdzilibyśmy dokładnie to samo mówiąc „Hesperos to Hesperos” i „Hesperos to Fosforos”. Dałoby nam to nową klasę wyrażen wykluczonych z kategorii nazw własnych czy raczej z kategorii „nazw w wąskim sensie logicznym, czyli słów, których znaczeniem jest konkretny przedmiot” (1918-1919: 201) – czyli, jak będę mówił, zgodnie z utartym zwyczajem, „logicznych nazw własnych”.

Ów test wykluczy z kategorii logicznych nazw własnych prawie wszystkie wyrażenia uznawane przez nas za nazwy własne. Nawet gdyby coś miało w danej chwili tylko jedną nazwę, moglibyśmy z łatwością wprowadzić drugą i tak to zaaranżować, by powstała łamigłówka Fregego. Wydaje się, że taka próbka społecznej inżynierii nie powinna zmieniać pierwotnego charakteru nazwy. A zatem nawet przed wprowadzeniem konkurencyjnej nazwy dana nazwa nie była logiczną nazwą własną. Wydaje się, że idąc w tym kierunku, trzeba by powiedzieć, iż logiczne nazwy własne w ogóle nie istnieją.

Russell zbliża się to tego wniosku: „Do jedynych słów, których używa się jako nazw w sensie logicznym, należą takie słowa jak «to» i «tamto»” (1918-1919: 201). Rozważenie tych słów prowadzi do zupełnie nowego zbioru zagadnień. Powróćmy jednak do naszego wyjściowego tematu, czyli takich słów jak „Arystoteles”, w których przypadku poglądy Fregego prowadzą do pewnego konfliktu. Jak twierdziłem, Russell może nas od tego konfliktu uwolnić.

Aby się o tym przekonać, musimy pominąć logiczne nazwy własne „to” i „tamto” i zająć się Russellową teorią „potocznych nazw własnych”. (I tym razem, zgodnie z utartym zwyczajem, używam określenia „potoczne”, aby odróżnić te nazwy od logicznych nazw własnych.) Wokół teorii Russella powstał pewien mit. W tym paragrafie przedstawię prawdziwą historię, sygnalizując od czasu do czasu naturę tego mitu; w następnym paragrafie przedstawię sam mit.

„Zazwyczaj słowa pospolite, a nawet imiona własne, są w rzeczywistości deskrypcjami” (RUSSELL 1912: 29 [przekład polski – s. 62.]). Russell nie

mógł tego rozumieć dosłownie. Deskrypcja, to jest deskrypcja określona, jest zwrotem o postaci „(jedyne) to-a-to”; nie ulega jednak wątpliwości, że na przykład „Arystoteles” nie jest zwrotem o takiej postaci (pierwszą literą tego wyrażenia jest „a”, a nie „j”, itd.). Zgodnie z mitem Russell utrzymuje, że nazwy są synonimiczne z deskrypcjami określonymi. Ale już następne zdanie w tekście każe powątpiewać, czy taki był zamysł Russella: „To znaczy, że myśl w umyśle osoby używającej imienia własnego można w ogólnym przypadku poprawnie wyrazić *explicite* jedynie wtedy, gdy zastąpimy imię własne przez deskrypcję” [przekład polski – s. 62.] „Znaczenie” nazwy własnej powinno być czymś publicznym i wspólnym wszystkim jej użytkownikom. Na podstawie przytoczonego zdania nie można jeszcze powiedzieć, czy Russell uważa, że „myśl w umyśle” osoby używającej nazwy będzie obdarzona analogicznymi własnościami: czy wszyscy użytkownicy nazwy będą mieć myśli, które są takie same w istotnym względzie. Wątpliwości zostają jednak rozwiane w kolejnych zdaniach: „Co więcej, deskrypcja niezbędna do wyrażenia tej myśli będzie różna dla różnych ludzi lub dla tego samego człowieka w różnym czasie. Jedyną niezmienną rzeczą (jeśli imienia użyto prawidłowo) jest ów przedmiot, do którego się owo imię stosuje” (1912: 29-30 [przekład polski – s. 62.]).

A zatem w myśl teorii Russella nazwa własna jest związana z różnymi deskrypcjami, potrzebnymi do opisanego *explicite* myśli rozmaitych użytkowników tej nazwy. To zróżnicowanie pokazuje, że – według Russella – deskrypcje nie mogą podawać publicznego znaczenia nazwy, ponieważ znaczenie powinno być wspólne w ramach wspólnoty językowej.

Obiecałem, że teoria Russella pomoże nam rozwiązać konflikt, który odkryliśmy u Fregego. Polegał ono na tym, że związane z nazwą deskrypcje, pozwalające wyjaśnić różnice w wartości poznawczej i inne własności użycia, wydają się idiosynkratyczne, podczas gdy sens czy znaczenie nazwy powinno być takie samo dla użytkowników języka. Rola, jaką Russell przypisuje deskrypcjom, jest całkowicie idiosynkratyczna. Oddaje ona sprawiedliwość co najwyżej jednej stronie rzeczonego konfliktu.

Niemniej jednak Russell w błyskotliwy sposób pokazuje, jak powiązać stały aspekt nazw z ich aspektem zmiennym. Spotkaliśmy już pierwszą wskazówkę: „Jedyną niezmienną rzeczą (jeśli imienia użyto prawidłowo) jest ów przedmiot, do którego się owo imię stosuje”.

Można powiedzieć, że nazwa posiada publiczne odniesienie w ramach jakiejś wspólnoty, o ile istnieje coś, co spełnia wszystkie deskrypcje, jakie wiążą z nią jej użytkownicy (a przynajmniej większość z nich). Przyjmując tę

koncepcję, możemy powiedzieć, że ludzie skutecznie się komunikują używając danej nazwy, o ile należą oni do wspólnoty, w której ta nazwa ma publiczne odniesienie. Możemy też powiedzieć, że osoba, która wypowiada zdanie o postaci „*n* jest tym-a-tym” dla pewnej nazwy *n*, mówi prawdę zawsze i tylko, gdy *o* jest tym-a-tym, gdzie *o* jest publicznym odniesieniem *n*.

Przypuśćmy, na przykład, że wypowiadam słowa „Bismarck był wytrawnym dyplomata”. Zgodnie ze sposobem, w jaki używam nazwy własnej „Bismarck”, myśl w moim umyśle najlepiej opisać jako myśl, że *pierwszy kanclerz Niemiec* był wytrawnym dyplomata. Zdaję sobie jednak sprawę, że wy możecie wiązać tę nazwę z inną deskrypcją, powiedzmy z deskrypcją „najpotężniejszy człowiek w Europie”. Gdy zatem wypowiadam moje zdanie, to nie zmierzam do wywołania w was dokładnie tej samej myśli; staram się raczej wywołać w was myśl o tym, że Bismarck (bez względu na sposób, w jaki o nim myślicie) był wytrawnym dyplomata. Mój zamiar byłby spełniony, gdybyście uświadomili sobie, że próbowałem powiedzieć o najpotężniejszym człowieku w Europie, iż był wytrawnym dyplomata. (Myślę, że tak należy zinterpretować dość zawiły wywód Russella od s. 31 (w RUSSELL 1912 [przekład polski – s. 63-65].)

Russell rozwiązuje więc konflikt, z którym zetknęliśmy się u Fregego, następująco: idiosynkratyczne elementy pozostają tym, czym są, i nie wymaga się, aby odgrywały one jakąś samodzielną rolę semantyczną; zostają one jednak powiązane z publicznym odniesieniem, które jest wystarczająco stabilne i niezmiennie, aby stanowić znaczenie nazwy własnej.

Chociaż pogląd Russella oddaje chyba sprawiedliwość zagadce Fregego (o identyczności i wartości poznawczej), ale znacznie słabiej radzi sobie z zagadnieniami istnienia i nazw nie posiadających nosiciela.

W wypadku pierwszej z tych kwestii Russell może wyjaśnić różnicę między „Hesperos jest to Hesperos” a „Hesperos jest to Fosforos”, odwołując się do odmiennych deskrypcji wiązanych z tymi dwoma nazwami. Wyjaśnienie nie musi przebiegać na poziomie prawdy lub odniesienia, lecz może być po prostu takie, że te zdania są z reguły powiązane z odmiennymi myślami. (W istocie Russell może dopuścić taki sposób rozumienia zdania „Paderewski jest Paderewskim”, przy którym ma ono – tak jak powinno – wartość poznawczą. Por. KRIPKE 1979.)

W ujęciu Russella publiczne odniesienie jest czymś stabilnym, a nazwie bez nosiciela brak takiego odniesienia. To zaś oznacza, że nie można zagwarantować, iż dwie różne wypowiedzi „Mojżesz nie istniał” mówią to samo. Nie można zagwarantować, że istnieje tu wspólny przedmiot, ponieważ nie

można zagwarantować, że istnieje tu publiczne odniesienie (nie można zagwarantować, że ktokolwiek spełnia związane z nazwą „Mojżesz” deskrypcje, takie jak „jako jedyny wyprowadził Izraelitów z Egiptu”). Nie można również zagwarantować, że istnieje wspólna deskrypcja, ponieważ deskrypcje mogą być różne u różnych mówiących. A zatem taka wypowiedź może być pozbawiona publicznych własności semantycznych.

Zarzut ten nie jest ostateczny, ponieważ nie jest oczywiste, że w takich przypadkach musi istnieć opis semantyczny. Moglibyśmy powiedzieć, że jeśli związane z daną nazwą deskrypcje nie wiążą jej ze światem, to jej użycie jest wadliwe (o ile nie jest zamierzone jako fikcyjne, co prowadzi do innych zagadnień.)

Frege, w myśl przyjętej tu interpretacji, może po prostu zaliczyć nazwy bez nosiciela do przypadków sensu bez odniesienia; zmusza to jednak do określenia pozytywnej natury sensu. Sens ma być sposobem myślenia o czymś, ale – jak widzieliśmy – nie jest łatwo pogrupować sposoby myślenia wystarczająco „grubo”, aby zapewnić dokładnie jeden sposób dla każdej z pozoru jednoznacznej nazwy, takiej jak „Mojżesz”. Choć więc Russell boryka się z pewnymi trudnościami, konkurencyjne rozwiązanie Fregego nie ma tu oczywistej przewagi.

1.4. *Kripke*

Saul Kripke sformułował ważne i szeroko akceptowane twierdzenie: nazwy są „sztywnymi desygnatorami”, które desygnują to samo bez względu na możliwe zmiany okoliczności. Sugeruje on, że wynika z tego, iż pogląd, który nazywa deskryptywną teorią nazw i który przypisuje Fregemu i Russellowi, jest błędny. Omówię najpierw pozytywną tezę Kripkego, a następnie zajmę się jego krytyką pod adresem Fregego i Russella.

W czasach, gdy Kripke głosił swoje wykłady „Nazywanie a konieczność”, wielu filozofów pod wpływem Quine’a podejrzliwie traktowało pojęcia modalne (konieczności, możliwości itd.). Quine nazwał je stworami ciemności. W swojej wczesniej pracy Kripke proponuje formalną semantykę światów możliwych dla języka modalnego (zob. część 5 poniżej). Jednym z wielu efektów jego wykładów było odnowienie zainteresowania modalnością.

W szczególności pojęciem modalnym jest pojęcie sztywnego desygnatora. Deskrypcja „pierwszy cesarz Francji” desygnuje Napoleona, ale sprawy mogły się potoczyć inaczej. Gdyby nigdy nie wybuchła Rewolucja Francuska, mogłoby nie być żadnego francuskiego cesarza albo mógłby nim być ktoś

inny niż Napoleon. W pierwszym wypadku deskrypcja „pierwszy cesarz Francji” nie desygnowałaby nikogo; w drugim – desygnowałaby inną osobę. Deskrypcja nie jest więc sztywna, ponieważ to, co desygnuje, może się zmieniać w zależności od możliwych zmian okoliczności, nawet jeśli znaczenie wyrażenia pozostaje takie samo.

„Napoleon”, dla odmiany, wydaje się sztywnym desygnatorem. Moglibyśmy opisać sytuację, w której Napoleon istnieje, choć nie jest desygnowany przez wyrażenie „pierwszy cesarz Francji”; wydaje się jednak, że nie jesteśmy w stanie opisać sytuacji, w której Napoleon istnieje, ale nie jest desygnowany przez nazwę „Napoleon”. Możemy sobie wyobrazić sytuację, w której nosiłby on inne imię, ale nie o to tutaj chodzi: musimy sobie wyobrazić, że używamy naszego wyrażenia „Napoleon” tak, jak go używamy, ale ze względu na okoliczności nie desygnuje ono Napoleona. Możemy więc użyć jego imienia, aby opisać możliwą sytuację, w której nikt nie nosi tego imienia: mogłoby się zdarzyć, że Napoleona nazywano Janem Michałem i że nikt w rodzinie Bonaparte nie nosił imienia Napoleon.

To zjawisko, jak się twierdzi, dotyczy wszystkich nazw i można je opisać w sposób ogólny (por. PEACOCKE 1973). Termin jednostkowy (*s*), jest sztywnym desygnatorem w wypowiedzi „...*s*...” zawsze i tylko, gdy dla pewnego przedmiotu *o*, w sposób konieczny ta wypowiedź jest prawdziwa zawsze i tylko, gdy *o* spełnia „...*s*...”. Przyjrzymy się bliżej tej definicji, wyjaśniając jej szczegóły techniczne i pokazując, że w jej świetle „pierwszy cesarz Francji” nie jest sztywnym desygnatorem, natomiast „Napoleon” jest.

Powiedzmy, że mówię teraz: „Pierwszy cesarz Francji urodził się na Korsyce”. Aby ustalić, czy termin jednostkowy (mogący pełnić funkcję podmiotu) „pierwszy cesarz Francji” jest sztywnym desygnatorem, musimy ustalić, czy istnieje przedmiot, od którego zależy (mówiąc nieco metaforycznie) warunek prawdziwości tego zdania. Wypowiedź ta jest prawdziwa, ponieważ Napoleon urodził się na Korsyce. Jeśli zatem istnieje przedmiot, od którego zależy odpowiedni warunek prawdziwości, to musi nim być Napoleon. Toteż Napoleon musiałby spełniać warunek (nałożony na zmienną *o* w definicji), że w sposób konieczny wypowiedź ta jest prawdziwa zawsze i tylko, gdy Napoleon spełnia „*s* urodził się na Korsyce”. Nie spełnia on jednak tego warunku, ponieważ gdyby nie został cesarzem – co niewątpliwie by się stało, gdyby poległ w bitwie pod Mondovi – prawdziwość lub fałszywość tej wypowiedzi nie miałyby z nim żadnego związku. A zatem „pierwszy cesarz Francji” nie jest sztywnym desygnatorem. (Czy powinniśmy uznać tę wypowiedź za fałszywą, gdyby Francja nigdy nie miała cesarza?)

Rozważmy z kolei wypowiedź „Napoleon urodził się na Korsyce”. Czy istnieje przedmiot, od którego zależy warunek prawdziwości tego zdania? Wydaje się, że tak – jest nim Napoleon. W sposób konieczny, wypowiedź ta jest prawdziwa zawsze i tylko, gdy Napoleon urodził się na Korsyce. Prawdziwość tej wypowiedzi nie mogłaby pod żadnym względem zależeć od kogokolwiek innego.

Stwierdzenie, że nazwa jest sztywnym desygnatorem, pozostawia wiele niedopowiedzeń i wymaga dalszych wyjaśnień. Musimy wyjaśnić łamigłówkę Fregego dotyczącą identyczności. Musimy wyjaśnić kwestię nazw nieposiadających nosiciela: nawet jeśli są one – banalnie – sztywnymi desygnatorami, to nie mogą od pozostałych nazw różnić się tylko syntaktycznie. Ponadto Kripke pisze w taki sposób, jak gdyby twierdzenie, że nazwy są sztywnymi desygnatorami, było złą wiadomością dla takich myślicieli jak Frege i Russell. Dlaczego? Krótka odpowiedź Kripkego wydaje się następująca: Frege i Russell sądzą, że w wypadku każdej nazwy własnej istnieje pewna deskrypcja podająca jej znaczenie, a przynajmniej ustalająca jej odniesienie, oraz że ta deskrypcja jest niesztywnym desygnatorem. Jednakże znaczenia sztywnego desygnatora nie można nigdy utożsamić ze znaczeniem niesztywnego desygnatora. A zatem teoria Russella i Fregego jest błędna.

Problem z tym argumentem polega na tym, że jest wątpliwe, czy ktokolwiek w ogóle przyjmuje atakowany przez Kripkego pogląd. Jak widzieliśmy, Frege wspomina o deskrypcjach wyjaśniając sens nazwy „Arystoteles”, ale nie mówi nigdzie, że sens nazwy można określić jedynie przez deskrypcję. Być może jedynym sposobem wyrażenia sensu nazwy jest użycie samej tej nazwy. To prawda, że nazwa „Hesperos” desygnuje Fosforosa, aby jednak powiedzieć, co ona desygnuje w taki sposób, który ujawniłby jej sens, powinniśmy powiedzieć, że desygnuje Hesperosa. Jeśli posłużymy się tą regułą interpretując wypowiedzi użytkowników (tj. jeśli przyjmiemy, że używają oni nazwy „Hesperos” do desygnowania Hesperosa), uzyskamy prawidłowe rezultaty, lepsze niż gdybyśmy próbowali traktować ich jako używających „Hesperosa” do desygnowania Fosforosa. Różnice uwidoczniłyby się, na przykład, w wypadku łamigłówki Fregego. Osoby, która wypowiedziała słowa „Hesperos jest to Fosforos”, nie należy traktować jako mówiącej, że Hesperus jest Hesperusem. (Mam tu na myśli wersję stanowiska Fregego sformułowaną przez McDowella (1977).)

Gdyby to ujęcie okazało się poprawne, to Fregeowski sens można by uzależnić od deskrypcji, podważając w ten sposób argument Kripkego.

Argument ten nie trafia również w stanowisko Russella. Główna idea Russella była taka, że deskrypcje są powiązane z nazwami w taki sposób, iż

można ich używać, aby wyjaśnić dokładniej, co ma na myśli podmiot, gdy używa jakiejś nazwy. Nie jest to relacja synonimiczności. Podany przez Russella bardzo krótki opis sposobu, w jaki nazwy wnoszą swój wkład w warunki prawdziwości, sugeruje, że traktuje je on w istocie jako sztywne desygnatory. Warunki prawdziwości wypowiedzi „Bismarck był wytrawnym dyplomata” zależą od publicznego odniesienia nazwy „Bismarck”, którym jest sam Bismarck. Zgodnie ze sposobem, w jaki funkcjonuje obecnie nasz język, „Bismarck” będzie się odnosił do Bismarcka nawet w sytuacji, gdy nie jest on pierwszym kanclerzem. A zatem Russell mógłby się zgodzić z tym, że nawet gdyby Francuzi wygrali wojnę francusko-pruską, tak iż żadne Imperium Niemieckie formalnie by nie powstało i nie byłoby pierwszego kanclerza, Bismarck nadal byłby wytrawnym dyplomata.

Na podstawie wykładu Kripkego można by pomyśleć, że zarzutem wobec tezy, że każda nazwa własna jest synonimiczna z pewną deskrypcją określoną jest to, iż deskrypcje określone nie są sztywnymi desygnatorami. Starałem się do tej pory wykazać, że tej tezy nie można przypisać ani Fregemu, ani Russellowi. Pogląd Russella ma nawet pewne cechy, które powinny mu zjednać przychyłność Kripkego: Russell bowiem traktuje w istocie nazwy własne jak sztywne desygnatory. Jednakże Kripke wysuwa jeszcze inne zarzuty przeciwko teoriom, które – tak jak teoria Russella – wykorzystują deskrypcje na pewnym etapie charakteryzowania nazw. Rozważę dwa z nich.

„Człowiek z ulicy [...] może [...] używać nazwy «Feynman». Zapytany, powie: No tak, to jest fizyk lub ktoś taki. Może on nie myśleć, że wyznacza to kogokolwiek jednoznacznie (*picks out uniquely*). Myślę jednak, że posługuje się nazwą «Feynman» jako nazwą Feynmana” (KRIPKE 1972: 292 [przekład polski – s. 82]). W myśl tego zarzutu niektórzy użytkownicy mogą nie wiązać z nazwą żadnej deskrypcji i wobec tego nie mogą mieć przypisywanych im przez Russella myśli.

Russell zaprzecza w odpowiedzi, że takie przypadki są możliwe. Wysuwa on także określoną propozycję dotyczącą ustalania deskrypcji, propozycję, która stosuje się przypuszczalnie do wielu przypadków:

Aby ustalić, co kryje się w moim umyśle, gdy wydaję sąd dotyczący Juliusza Cezara, musimy zastąpić to imię deskrypcją złożoną z pewnych rzeczy, które o nim wiem. (Zadanie to może często spełniać deskrypcja „człowiek imieniem *Juliusz Cezar*”. Nawet jeśli nic więcej o nim nie pamiętam, jest oczywiste, że gdy o nim wspominam, pamiętam jego imię.)

(RUSSELL 1911: 160)

Można, jak sądzę, sensownie powiedzieć, że gdyby niedouczona osoba opisana przez Kripkego nie mogła nawet podać takiej deskrypcji jak „osoba (czy słynny fizyk) nazywana «Feynemanem»”, to nie można by uznać, że rozumie ona tę nazwę.

Kripke sprzeciwia się temu posunięciu, zarzucając mu błędne koło. Z pewnością błędem błędnego koła obarczone byłoby powiedzenie, że każdy użytkownik języka rozumie nazwę dzięki temu, że chwytą taką deskrypcję, ale ani Russell, ani nikt inny nie musi tak twierdzić.

Kolejny zarzut Kripkego jest poważniejszy. Dla wielu z nas Gödel to osoba, która jako pierwsza udowodniła niezupełność arytmetyki; jest to zatem po prostu fakt psychologiczny, że jego imię dla wielu z nas wiąże się z tą właśnie deskrypcją. Kripke każe nam sobie wyobrazić, że to Schmidt jako pierwszy udowodnił niezupełność arytmetyki, a Gödel jedynie przedstawił dowód jako własny. Zakładając, że nazwę „Gödel” łączymy z deskrypcją „osoba, która udowodniła niezupełność arytmetyki”, wypowiedź „Gödel urodził się w Wiedniu” byłaby prawdziwa zawsze i tylko, gdy Schmidt urodził się w Wiedniu. A jednak, intuicyjnie, odnieśliśmy się do Gödela, a nie do Schmidta i to, co powiedzieliśmy, jest prawdą tylko wtedy, jeśli Gödel urodził się w Wiedniu; miejsce urodzenia Schmidta nie ma tu nic do rzeczy (por. KRIPKE 1972: 294 [przekład polski – s. 85].)

Ten zarzut dotyka trudnych kwestii. Russell dopuszcza możliwość poczynienia ciekawych rozróżnień w obrębie użytkowników danej nazwy. Sugeruje, na przykład, że Bismarck mógłby używać nazwy „Bismarck” jako nazwy samego siebie, bez korzystania z deskrypcji, podczas gdy my nie znajdujemy się w tak dogodnym położeniu. Nie jest to bardzo odległe od rozróżnienia wprowadzonego przez Garetha Evansa między „wytwórcami” i „konsumentami” w praktyce używania nazwy (por. EVANS 1982: rozdz. 11). Wytwórcy mają kontakt z nazywanym przedmiotem, konsumenci nie mają takiego kontaktu, a zatem w swym życiu muszą polegać na wytwórcach. W wypadku imion osób dawno zmarłych wszyscy jesteśmy konsumentami; w wypadku imion współczesnych sław niektórzy z nas są wytwórcami, jednak większość to konsumenci. W wypadku imion współczesnych anonimowych przechodniów istnieją wytwórcy, ale niewielu konsumentów.

Z punktu widzenia tego rozróżnienia wymóg, aby konsumenci kierowali się w swoim życiu życiem wytwórców, nie jest niczym nierozsądnym. Wracając do podanego przez Kripkego przypadku Gödla-Schmidta: nas, czyli konsumentów w praktyce używania imienia „Gödel”, można by opisać jako uznających *dominację* deskrypcji „osoba nazywana «Gödlem»” nad wszyst-

kimi innymi deskrypcjami łączonymi z tą nazwą w następującym sensie: w wypadku rozbieżności odniesienia wspomnianą przed chwilą deskrypcję należy uznać za wskazującą poprawne odniesienie. Bycie *nazywanym* „Gödem” zostałyby tu wyjaśnione w kategoriach rozróżnienia między konsumentami a wytwórcami: praktyki wytwórców miałyby pierwszeństwo przed praktykami konsumentów w określaniu, kto jest nazywany tak-a-tak.

To ujęcie wykracza znacznie poza to, co można znaleźć u Russella, i wymagałoby dopracowania, zanim można by je poddać poważnej ocenie. Pokazuje ono jednak, że pomimo niezrozumienia Russella w niektórych kwestiach Kripke rzeczywiście stawia celny i istotny zarzut pod adresem teorii Russella. Tym niemniej jest to zarzut, który ktoś, kto przyjmuje podstawowe stanowisko Russella, w myśl którego nazwy wiążą się z deskrypcjami, ale nie są ich znaczeniami, może być w stanie uchylić. (Zob. też: EVANS 1982: rozdz. 1, 2, 11; MCCULLOCH 1989.)

2. DESKRYPCJE

2.1. *Teoria Russella*

Russell uważał, że chociaż na pozór funkcją deskrypcji określonych, czyli zwrotów o postaci „(jedyne) to-a-to” jest, tak jak w wypadku nazw własnych, odnoszenie się do przedmiotu, ich rzeczywista funkcja jest zupełnie inna: są one zwrotami kwantyfikatorowymi, a nie wyrażeniami odnoszącymi się. Zdanie „Pierwszy cesarz Francji urodził się na Korsyce” wcale nie funkcjonuje tak jak zdanie „Napoleon urodził się na Korsyce”. Ze względu na funkcję przypomina ono bardziej zdanie „Żaden cesarz nie urodził się na Korsyce”. Jest to rodzaj generalizacji.

Dokładniej: zdanie o formie „(Jedyne) to-a-to jest F ” (dla dowolnego predykatu „ F ”) należy analizować następująco:

- (1) Istnieje przynajmniej jedno to-a-to.
- (2) Istnieje co najwyżej jedno to-a-to.
- (3) Cokolwiek jest tym-a-tym, jest F .

Co ważne, w tej analizie nie pojawia się w ogóle słowo „jedyne” (*the*). Jest ona równoważna bardziej zwięzłemu sformułowaniu:

Istnieje dokładnie jedno to-a-to i ono jest F . [Zastępując „to-a-to” przez „ G ”, możemy je przedstawić symbolicznie: $\exists x(Gx \ \& \ \forall y (Gy \rightarrow y = x) \ \& \ Fx)$].

Russell uważał, że „istnieje” nie jest predykatem (z racji, które zostaną omówione w części 3). Z tego względu zdanie o postaci „(Jedyny) to-a-to istnieje” wymaga odmiennej analizy, która przedstawia się następująco:

- (1) Istnieje przynajmniej jedno to-a-to.
- (2) Istnieje co najwyżej jedno to-a-to.

Krótko mówiąc:

Istnieje dokładnie jedno to-a-to.

Zdanie „(Jedyny) obecny król Francji jest łysy” jest prawdziwe zawsze i tylko, gdy obecnie istnieje dokładnie jeden król Francji i jest on łysy. Zdanie to jest fałszywe, ponieważ nie istnieje obecnie król Francji. Zdanie „(Jedyny) mieszkaniec Londynu jest zamożny” jest prawdziwe zawsze i tylko, gdy istnieje dokładnie jeden mieszkaniec Londynu i ta osoba jest zamożna. Zdanie to jest fałszywe, ponieważ mieszkańców Londynu jest więcej. [Te wnioski zależą od zrównania „fałszywe” z „nieprawdziwe”. Być może w niektórych wypadkach, na przykład nieostrości, nie jest ono uprawnione.] Zdanie „(Jedyna) złota góra nie istnieje” jest prawdziwe zawsze i tylko, gdy nie jest tak, że istnieje dokładnie jeden przedmiot, która jest zarazem górą i jest ze złota. Zdanie to jest prawdziwe, ponieważ nie istnieje nawet jeden taki przedmiot. Russell przytoczył te przykłady, aby pokazać, iż jego teoria prowadzi do intuicyjnie poprawnych wyników.

O ile wiem, Russell podał tylko jeden bezpośredni argument na rzecz swej teorii. (Inne argumenty są pośrednie – pokazują one albo że jego teoria rozwiązuje pewne łamigłówki, albo że inne teorie wklęają się w trudności. Omówimy niektóre z tych argumentów w paragrafie 2.3). Ten bezpośredni argument jest prosty: Russell twierdzi, że jedyna różnica między słowami „jakieś” (*a*) a „jedyne” (*the*) jest taka, że drugie z nich implikuje jedyność [RUSSELL 1919: 176]. Ponieważ słowo „jakieś” należy rozumieć jako kwantyfikator egzystencjalny, tak że „Spotkałem jakiegoś człowieka” jest analizowane jako „Istnieje coś ludzkiego, co spotkałem”, słowo „jedyne” musi być kwantyfikatorem egzystencjalnym z dodaną jedynością. I takie właśnie rozumienie proponuje teoria Russella.

Argument ten nie przekona głównych przeciwników, uznają oni bowiem deskrypcje określone za autentyczne wyrażenia odnoszące się i dlatego nie zgodziliby się, że jedyną różnicą między słowami „jakieś” a „jedyne” jest to, że drugie z nich implikuje jedyność.

2.2. Porównanie nazw i deskrypcji

Pierwszą rzeczą, z której należy sobie zdać sprawę w przypadku deskrypcji określonej jest to, że nie jest ona nazwą.

(RUSSELL 1918-1919: 244)

W tym kontekście Russell rozumie przez „nazwę” logiczną nazwę własną, czyli inaczej mówiąc „to” lub „tamto”. Chodzi mu tu o podkreślenie, że deskrypcje nie są wyrażeniami odnoszącymi się. W tym celu używa on następujących argumentów.

2.2.1. Rozumienie

Jeśli rozumiecie język polski, to zrozumielibyście znaczenie zwrotu „autor *Waverleya*”, nawet jeśli nigdy wcześniej tego zwrotu nie słyszeliście, natomiast nie zrozumielibyście znaczenia nazwy „Scott”, gdybyście nigdy wcześniej jej nie słyszeli, ponieważ znać znaczenie nazwy to wiedzieć, do kogo się ona stosuje.

(RUSSELL 1918-1919: 245)

Uwaga Russella jest, jak sędzę, trafna w wypadku przynajmniej niektórych deskrypcji. Nazwy są semantycznie proste: ich rozumienie nie jest wypadkową rozumienia ich części, ponieważ części te nie mają udziału w ich znaczeniu. (W słowie „Sokrates” występują wprawdzie litery „sok”, ale nie wpływają one na znaczenie tego słowa.) Natomiast deskrypcje są – z definicji – semantycznie złożone i przynajmniej w niektórych przypadkach (takich jak „największa liczba pierwsza”) nie ulega wątpliwości, że rozumiemy je wyłącznie na mocy rozumienia ich składowych słów. Znajomość denotacji (jeśli takowa istnieje) nie jest tu wcale konieczna.

2.2.2. Umowa

Nazwę można przypisać w sposób umowny; żadna umowa nie może jednak sprawić, aby deskrypcja „autor *Waverleya*” denotowała Scotta. Jest raczej tak, że Scott musi sięgnąć po pióro i napisać tę powieść bez niczyjej pomocy (RUSSELL 1918-1919: 245). Ta uwaga również wydaje się trafna i stanowi pochodną semantycznej prostoty nazw, która różni je od semantycznie złożonych deskrypcji.

2.2.3. Identyczność

Jeśli jeden przedmiot ma dwie nazwy, wypowiadasz dokładnie to samo stwierdzenie używając którejkolwiek z nich.

(RUSSELL 1918-1919: 245)

Jednak twierdzenie „Scott jest *c*”, gdzie *c* jest dowolną nazwą Scotta, nie jest tym samym twierdzeniem, co „Scott jest autorem *Waverleya*”, ponieważ pierwsze, w przeciwieństwie do drugiego, jest tautologią pozbawioną wartości informacyjnej. A zatem „autor *Waverleya*” nie jest nazwą.

Russell zakłada, że nazwy, które nazywają to samo, znaczą to samo, podczas gdy argument Fregego można by uznać za podważający to założenie. Podstawa wniosku Russella jest więc w najlepszym razie sporna.

2.2.4. Istnienie

Można otrzymać sensowne sądy zaprzeczając istnieniu „(jedynego) tego-a-tego” [...] [np.:] „Nie istnieje największa skończona liczba”.

(RUSSELL 1918-1919: 248)

Ponieważ jednak nazwa musi coś nazywać, każdy taki sąd byłby fałszywy i w rzeczywistości moglibyśmy powiedzieć, że jest on fałszywy właśnie na podstawie rozumienia zdania. Dowodzi to niewątpliwie, że przynajmniej niektóre deskrypcje różnią się od nazw, jeśli nazwa jest czymś, co musi się do czegoś odnosić.

2.3. „Łamigłówki” Russella

Teorię logiczną można wypróbować, badając, w jakiej mierze radzi sobie ona z łamigłóvkami; i zdrowy to sposób postępowania, rozmyślając nad logiką zadawać sobie do rozwiązania możliwie jak najwięcej łamigłówek, te bowiem służą podobnemu celowi, jak eksperymenty w fizyce.

(RUSSELL 1905: 47 [przekład polski – s. 262])

Taka metoda nie przynosi ostatecznych wyników, ponieważ nie wyklucza możliwości, że jakaś inna teoria również jest w stanie rozwiązać te łamigłówki. Podobnie jednak jak w nauce mamy dobrą rację, aby przyjąć teorię, która pomyślnie zdaje testy – przynajmniej dopóki nie jest nam znana konkurencyjna teoria, które również je zdaje.

2.3.1. Prawo identyczności

Jeśli a jest identyczne z b , to cokolwiek jest prawdą o jednym z nich, jest też prawdą o drugim oraz każde z tych dwóch można podstawić na miejsce pozostałego w dowolnym sądzie, nie zmieniając przez to prawdziwości ani fałszywości owego sądu. Otóż Jerzy IV pragnął się dowiedzieć, czy Scott jest autorem *Waverleya*; i faktycznie Scott był autorem *Waverleya*. Możemy zatem podstawić *Scott* na miejsce *autor Waverleya* i tym samym dowieść, że Jerzy IV pragnął się dowiedzieć, czy Scott jest Scottem. A jednak zainteresowanie prawem identyczności trudno przypisać pierwszemu dżentelmenowi Europy.

(RUSSELL 1905: 47-8 [przekład polski – s. 262])

Gdy Russell mówi o zastępowaniu jednego przedmiotu (w odróżnieniu od słowa) przez drugi, należy to rozumieć dosłownie. Russell rozumie sąd jako sekwencję przedmiotów, znaczeń słów występujących w zdaniu wyrażającym ten sąd. Zdaniu „Scott jest człowiekiem” odpowiada sąd obejmujący rzeczywiste przedmioty: Scotta i człowieczeństwo. Taka koncepcja sądu uniemożliwia Russellowi odróżnienie sądu wyrażonego przez zdanie „Hesperos jest to Hesperos” od sądu wyrażonego przez zdanie „Hesperos jest to Fosforos”, przy założeniu, że „Hesperos” i „Fosforos” to logiczne nazwy własne.

Przesłanka o zastępowalności przedmiotów identycznych jest niekiedy nazywana zasadą Leibniza o nieodróżnialności przedmiotów identycznych. Zasada ta wymaga precyzyjnego sformułowania. Jeśli, na przykład, rację ma Frege, to niepoprawne byłoby następujące sformułowanie: Jeśli „ $a = b$ ” jest prawdziwe, to „... a ...” ma taką samą wartość logiczną, jak „... b ...” (Kropki reprezentują dowolny sposób utworzenia zdania, w którym zostaje użyta jedna z tych nazw.) Mimo że zdanie „Hesperos jest to Fosforos” jest prawdziwe, to według Fregego następujące zdania, w których „Fosforos” zastępuje „Hesperosa” różnią się wartością logiczną: „Starożytni wierzyli, że Hesperos jest to Hesperos”; „Starożytni wierzyli, że Hesperos jest to Fosforos”. (Czy ta sama uwaga przemawiałaby również przeciwko podanemu przez Russella sformułowaniu zasady Leibniza w kategoriach zastępowania przedmiotów, a nie wyrażen?)

Russellowskie rozwiązanie tej łamigłówki jest takie, że zdanie „Scott był autorem *Waverleya*” nie wyraża sądu identycznościowego i wobec tego nie można tu prawomocnie odwołać się do zasady Leibniza. Sąd ten nie wyraża identyczności, ponieważ – jak ująłby to Russell – deskrypcja określona nie przedstawia składnika sądu, czyli – inaczej mówiąc – jej rola nie polega na odnoszeniu się. Ujawnia to następująca analiza tego sądu: Istnieje dokładnie

jeden autor *Waverleya* i Scott jest tą osobą. Nie jest to identyczność, lecz generalizacja egzystencjalna (mówi ona, że istnieje coś jedyne, spełniającego określony warunek).

Choć Russell ma rację, że zdania z deskrypcjami określonymi nie są, w myśl jego analizy, zdaniami identycznościowymi, pewne zastrzeżenia budzi sposób, w jaki sformułowana jest ta łamigłówka. Istnieje bowiem rzeczywisty problem, unaoczniony przez Fregego, czy identyczność tak czy inaczej umożliwia zastępowalność w kontekstach zależnych (*indirect contexts*), generowanych przez takie wyrażenia jak „Starożytni wierzyli, że...”.

2.3.2. Prawo wyłączzonego środka

Na mocy prawa wyłączzonego środka musi być prawdą albo „A jest B”, albo „A nie jest B”. A zatem prawdą musi być albo: „Obecny król Francji jest łysy”, albo: „Obecny król Francji nie jest łysy”. A jednak gdybyśmy wyliczyli rzeczy łyse, następnie zaś rzeczy niełyse, to na żadnej z tych list nie znaleźlibyśmy (jedynego) obecnego króla Francji.

(RUSSELL 1905: 48 [por. przekład polski – s. 262])

Prawo wyłączzonego środka, w ogólnej postaci, stwierdza, że prawdziwe jest każde podstawienie „*p* lub nie-*p*”, gdzie „nie-*p*” jest negacją „*p*”. Tymczasem, powiada Russell, skłonni jesteśmy uznać zarówno zdanie „Obecny król Francji jest łysy”, jak i zdanie „Obecny król Francji nie jest łysy” za fałszywe, tak że ich alternatywa jest fałszywa, a równocześnie uznać drugie z tych zdań za negację pierwszego.

Rozwiązanie Russella odwołuje się do rozróżnienia *zakresu*. [W logice formalnej zakres łącznika zdaniowego to najkrótsza poprawnie zbudowana formuła, w której on występuje.] („Zakres” jest nowszym określeniem; sam Russell mówi o pierwotnym i wtórnym wystąpieniu.) Musimy odróżnić, na przykład, zdanie „Niektórzy ludzie nie są szczęśliwi” od zdania „Nie jest tak, że niektórzy ludzie są szczęśliwi”. Drugie, w przeciwieństwie do pierwszego, wyklucza, że istnieją jacyś szczęśliwi ludzie. Różnica ta nie wynika, jak się wydaje, ze znaczenia słów wchodzących w skład tych zdań, lecz raczej z pozycji, jaką względem innych wyrażen zajmuje znak negacji („nie” lub „nie jest tak, że”). W pierwszym zdaniu „nie” rządzi jedynie słowem „szczęśliwi” (tj. obejmuje swoim zakresem tylko słowo „szczęśliwi”); w drugim zadaniu rządzi ono całym zdaniem (tj. obejmuje swoim zakresem całe zdanie). Jeśli do zdania „*p*” dodajemy znak negacji, to mówimy, że wynikiem jest negacja zdania „*p*” tylko wówczas, gdy ten znak obejmuje swoim zakresem całe „*p*”.

W świetle tego rozróżnienia, zauważa Russell, zdanie „Obecny król Francji nie jest łysy” jest wieloznaczne. Najbardziej naturalne jest przypisanie partykule „nie” stosunkowo wąskiego zakresu, obejmującego „łysy”, ale nie obejmującego innych części zdania. W takim rozumieniu zdanie to nie jest negacją zdania „Obecny król Francji jest łysy”, a zatem nie powstaje problem z prawem wyłączonego środka.

Alternatywnie można uznać, że „nie” rządzi całym zdaniem, mniej wieloznacznie wyrażonym jako „Nie jest tak, że obecny król Francji jest łysy”. To zdanie rzeczywiście jest negacją wyjściowego zdania; jednak, zgodnie z teorią Russella, jest ono prawdziwe (ponieważ prawdziwie głosi, że nie istnieje nic jedyne, co obecnie jest królem Francji i jest łyse.) A zatem alternatywa „Obecny król Francji jest łysy albo nie jest tak, że obecny król Francji jest łysy” okazuje się prawdziwa i znowu nie powstaje problem z prawem wyłączonego środka.

Moim zdaniem to rozumowanie Russella jest bez zarzutu, a rozróżnienie zakresów ma ważne znaczenie dla opisu funkcjonowania języka.

2.3.3. Istnienie

Jakże jednak jakiś nie-byt może być podmiotem sądu logicznego?

(RUSSELL 1905: 48 [przekład polski – s. 263])

Jeśli nie ma różnicy między A i B , to zdanie „Nie istnieje różnica między A i B ” powinno być prawdziwe. Wydaje się jednak, że orzeka ono coś (nieistnienie) o tym, co – z założenia – jest niebytem, a przecież niebyt nie może być podmiotem sądu.

Rozwiązanie polega na uznaniu, zgodnie z teorią deskrypcji, że zwrot „różnica między A i B ” nie odnosi się do czegoś nieistniejącego, lecz jest zwrotem kwantyfikatorowym. Odpowiednia analiza wygląda więc następująco: „Nie ma jedynej (*unique*) różnicy między A i B ”. To zdanie nie sprawia wrażenia, że podmiotem jest niebyt.

Teoria Russella rzeczywiście rozwiązuje ten problem. W części 3 zastanowimy się, czy można podać jakieś konkurencyjne rozwiązania.

2.4. Zarzuty

Nicią przewodnią krytyk pod adresem teorii Russella jest twierdzenie, że przynajmniej w niektórych przypadkach ich użycia, deskrypcje określone są wyrażeniami odnoszącymi się do przedmiotów. Ten ogólny zarzut przybiera różne formy szczegółowe.

2.4.1. Brak jedyności

Przypuśćmy, że w pokoju z dwójgiem drzwi zamykacie jedno z nich. Słyszając ich skrzypienie mówię: „(Jedynie) drzwi skrzypią”. Jeśli potraktujemy analizę Russella poważnie, to moje słowa są prawdą jedynie wówczas, gdy istnieją tylko jedno drzwi. Faktycznie jednak na świecie istnieją ich miliardy – nawet w pokoju, o którym mowa, jest ich dwoje. Wydaje się zatem, że w myśl analizy Russella powiedziałem coś fałszywego, a to kłóci się z naszymi intuicjami.

Każde ujęcie kwantyfikatorów musi uwzględniać fakt, że na ich dziedzinę może mieć wpływ kontekst. Gdy podczas wykładu wołam: „Nie ma już kredy”, to moje słowa są prawdziwe, jeśli kreda skończyła się w sali wykładowej, choćby nawet było jej pod dostatkiem na klifach w Dover. Russell powinien powiedzieć, że przypadek dwójga drzwi nie różni się w zasadzie od tego przypadku. Dziedzinę kwantyfikatora w „Istnieją dokładnie jedno drzwi i te drzwi skrzypią” należy rozumieć jako ograniczoną przez kontekst.

Na czym jednak polega to ograniczenie? Nie jest to, z założenia, ograniczenie do drzwi w pokoju, w którym wypowiadam swoje słowa. Być może należy powiedzieć, że chodzi o ograniczenie do drzwi, które są istotne czy wyróżnione z punktu widzenia toczącej się rozmowy. Takich drzwi może być jednak więcej niż jedno. Mógłbym, na przykład, opowiadać wam o problemach związanych z posiadaniem pokoju z dwójgiem drzwi, wyróżniając w ten sposób oboje z nich. Tym, co czyni wchodzące w grę drzwi szczególnie istotnymi z punktu widzenia toczącej się rozmowy, jest bez wątpienia fakt, że je zamknęliście; dość trudno jednak opisać tę własność w sposób ogólny. Być może powinniśmy powiedzieć: jest to ograniczenie do tych drzwi, które są najbardziej istotne z punktu widzenia rozmowy i związanych z nią czynności rozmówców.

Nie rozwiewa to do końca podejrzenia, że niekiedy mówimy: „(Jedynie) (*the*) drzwi...”, gdy równie dobrze moglibyśmy powiedzieć: „Te (*that*) drzwi...”. A ten drugi zwrot – pomijając jego semantyczną złożoność – przypomina wzorcowe przypadki Russellowskich wyrażen odnoszących się, czyli logiczne nazwy własne „to” i „tamto” (*that*). Pozostaje więc pytanie, czy używając w taki sposób „(jedynie)”, gdy „te” byłoby równie dobre, jeśli nie lepsze, właściwie posługujemy się językiem. Z drugiej strony należy podkreślić, że tego rodzaju argument ma ograniczony zasięg: z oczywistych względów nie moglibyśmy użyć „ta” zamiast „jedyna” w deskrypcji takiej jak „(jedyna) największa liczba pierwsza”.

2.4.2. Brak istnienia

Peter F. Strawson (1950a) zwrócił uwagę, że jeśli ktoś podchodzi do nas i mówi: „Obecny król Francji jest łysy”, to nie należy się spodziewać, że odpowiemy: „To nieprawda”. Jednakże w myśl analizy Russella wypowiedziane zdanie jest fałszywe. Strawson uważa, że takie sytuacje można lepiej wyjaśnić mówiąc, że dana osoba w ogóle niczego nie stwierdza; nie mówi nic prawdziwego ani fałszywego. Wypowiada zdanie sensowne, które w danych okolicznościach jest wadliwe (występującej w nim deskrypcji określonej brak odniesienia), co przekreśla możliwość użycia go do powiedzenia czegokolwiek, zgodnie z prawą lub nie. Właśnie dlatego nie jesteśmy skłonni odpowiedzieć: „To nieprawda”.

I tym razem spotykamy się z ideą, że deskrypcje określone są wyrażeniami odnoszącymi się. Jeśli bowiem zadaniem deskrypcji określonej jest odniesienie, to gdy nie spełnia ona tego zadania, nie przyczynia się też do określenia warunku prawdziwości wypowiedzi, w której występuje. Tylko wypowiedź, która *posiada warunek* prawdziwości lub fałszywości, może *być* prawdziwa lub fałszywa.

Pierwsza nasuwająca się uwaga jest taka, że omawiane ujęcie ma zastosowanie tylko w niektórych wypadkach. Nie chcemy, aby zdanie „(Jedyna) największa liczba pierwsza nie istnieje” okazało się nieprawdziwe przez to, że „(jedyna) największa liczba pierwsza” nie posiada odniesienia.

Druga uwaga jest taka, że wniosek Strawsona nie wynika ze wskazanego przezeń zjawiska. Być może na wypowiedź „Obecny król Francji jest łysy” nie odpowiadamy mówiąc „To nieprawda”, ponieważ w ten sposób wywołalibyśmy u słuchaczy mylne wrażenie, że zdanie to jest fałszywe w zwykłym sensie: zazwyczaj bowiem zdanie o formie „(Jedynę) to-a-to jest *F*” jest fałszywe dlatego, że choć istnieje dokładnie jedno to-a-to, to nie jest ono *F*.

Drugą z tych uwag można rozwinąć na podstawie pewnej sugestii Kripkego (1977). Nawet gdyby ludzie zaczęli wprost używać analizy Russella, a więc wypowiadać zdania w rodzaju „Istnieje obecnie dokładnie jeden król Francji i jest on łysy”, to nadal mógłby pozostać opór przed odpowiedzią „To nieprawda” – z powodów wymienionych w poprzednim akapicie. A zatem zjawisko to nie stanowi poważnego wyzwania dla analizy Russella.

2.4.3. Atrybutywne i referencyjne użycie deskrypcji

Keith Donnellan (1966) zasugerował, że istnieją dwa różne sposoby użycia deskrypcji: atrybutywne i referencyjne. Twierdzi on, że analiza Russella

jest trafna tylko w wypadku użycia atrybutywnego, natomiast analiza Strawsona – tylko w wypadku użycia referencyjnego.

Rozróżnienie to można zilustrować następującym przykładem. Gdy na widok okaleczonego ciała Smitha mówimy: „(Jedyny) morderca Smitha jest szalony”, to używamy deskrypcji „(jedyny) morderca Smitha” atrybutywnie. Nie wiemy, kto zamordował Smitha, ale chcemy powiedzieć, że ktokolwiek to uczynił, był szalony. Jeśli natomiast Jones został aresztowany za morderstwo Smitha i będąc w sądzie wypowiadamy dokładnie to samo zdanie, to w sposób naturalny zostanie ono potraktowane jako odnoszące się do Jonesa. Według Donnellana zdanie to będzie prawdziwe zawsze i tylko, gdy Jones jest szalony, bez względu na to, czy Jones rzeczywiście zamordował Smitha.

Opisane zjawisko jest z pewnością ciekawe, ale można powątpiewać, czy jego znaczenie jest takie, że Donnellan może je użyć przeciwko Russellowi. Rozważmy następujący, podany przez Donnellana przykład użycia referencyjnego. W czasie przyjęcia mówię: „(Jedyny) człowiek pijący martini jest pijany”. Mam na myśli kogoś, kto pije wodę z kieliszka na martini podczas przyjęcia, na którym nikt nie pije martini; zmylony kształtem kieliszka, błędnie przyjmuję, że zawiera ona martini. W myśl poglądu Donnellana to, co mówię, jest prawdziwe pod warunkiem, że ów człowiek rzeczywiście jest pijany. Wydaje się jednak, że takie postawienie sprawy nie zostawia miejsca na uwzględnienie faktu, że popełniłem pewnego rodzaju błąd.

Obrońca Russella dokonałby wyraźnego rozróżnienia między tym, co ktoś ściśle i dosłownie mówi, a tym, co ma na myśli. Nie ulega wątpliwości, że moim zamiarem było powiedzenie czegoś o człowieku pijącym wodę z kieliszka na martini; nie jest natomiast oczywiste, czy udało mi się spełnić ten zamiar. Zwolennik Russella uznałby, że nie: to, co ściśle i dosłownie powiedziałem, było fałszywe, ponieważ nikt w danej chwili nie pije martini, chociaż moim zamiarem było powiedzenie czegoś o człowieku pijącym wodę ze szklanki na martini i chciałem, aby to, co mówię, było prawdziwe zawsze i tylko, gdy ów człowiek jest pijany.

Omawiany kontrast wziął pod lupę Paul Grice (1961). Jeden z podanych przezeń przykładów jest szczególnie wymowny. Kolega pyta mnie, co sądzę o kimś jako o filozofie. Obaj wiemy, że mam wszelkie dane po temu, by wydać dobrze poinformowany sąd w tej sprawie. Odpowiadam: „Ma piękny charakter pisma”. Przypuszczalnie chcę przez to powiedzieć, że kiepski z niego filozof, ale z pewnością nie mówię tego ściśle i dosłownie.

Russell może umieścić referencyjne użycie deskrypcji określonych na poziomie „pragmatycznym”, na którym opisuje się sposób użycia języka, za-

miast na poziomie semantycznym, na którym opisuje się warunki prawdziwości wypowiedzi. Ponieważ teoria deskrypcji jest teorią semantyczną, można ją dostosować do opisanych przez Donnellana różnic pragmatycznych.

Nie jest również jasne, czy Donnellan słusznie twierdzi, że ujęcie Russella jest poprawne w wypadku użycia atrybutywnego. Niekiedy Donnellan zdaje się utożsamiać użycie atrybutywne z użyciem niereferencyjnym. Czasami podejmuje jednak również próby podania bardziej pozytywnej charakterystyki takiego użycia, proponując (choć nie całkiem wprost) warunki prawdziwości dla takich sposobów użycia deskrypcji, które nie są Russellowskie. I tak pisze on: „Osoba mówiąca, która używa deskrypcji określonej atrybutywnie stwierdza coś o kimś (czymś), kto (co) jest taki-a-taki (takie-a-takie)” (1966: 285). Sugeruje to, wbrew Russellowi, że zadaniem deskrypcji używanej atrybutywnie jest odnoszenie się. Ponadto Donnellan sugeruje, że jeśli nic nie jest *F*, to mówiąc „(Jedynę) *F* jest *G*”, o niczym nie powiedziano, że jest *G* (1966: 287), a w istocie, że w ogóle nic nie powiedziano (1966: 288). W myśl ujęcia Russella sugestia ta jest nietrafna. Wypowiadając „(Jedynę) *F* jest *G*”, mówi się to samo, bez względu na to, czy jest jakieś *F*, czy nie. Jeśli nie ma żadnego *F*, to mówi się coś fałszywego.

2.4.4. Użycie odwołujące się do bytu?

U podstaw większość wspomnianych sugestii leży przekonanie, że deskrypcje określone przynajmniej niekiedy funkcjonują jako wyrażenia odnoszące się: istnieje przedmiot, który mają one wskazywać. Christopher Peacocke (1973) proponuje nazywać takie użycie deskrypcji „odwołującym się do bytu” (*entity invoking*). Nasuwa się pytanie, czy ma on rację, twierdząc, że jest takie użycie.

Zjawiska związane z brakiem jedyności i brakiem istnienia mogą skłaniać do odpowiedzi twierdzącej, podobnie jak Donnellanowskie pojęcie użycia referencyjnego. Widzieliśmy jednak, że nie przemawiają one rozstrzygająco przeciwko Russellowi. Peacocke (1973:117) sugeruje, że istnieją niezależne racje za przyjęciem użycia zakładającego byt.

Jednym z testów takiego użycia jest możliwość zamiany „jedynę” (*the*) i „to” (*that*), czyli „jeśli w wypowiedzi «(Jedynę) *F* jest *G*» to, co jest ściśle i dosłownie powiedziane, można równie dobrze powiedzieć mówiąc «To *F* jest *G*»” (PEACOCKE 1973: 117). Przyjmuje się tu, że złożone wyrażenia wskazujące o formie „to *F*” bez wątpienia odwołują się do bytu (dla naszych obecnych celów możemy się z tym zgodzić), a zatem omawiana zamienność dowodziłaby, że również deskrypcja odwołuje się do bytu.

Peacocke podaje następujący przykład:

Jeśli odwiedzając wczoraj kasyno w Monte Carlo, widzieliśmy, jak ktoś rozbił kasę, i tego samego dnia widzieliśmy również, jak ktoś rozbił kasę w Nicei, i jeśli obaj o tym wiemy, to wówczas deskrypcja „(Jedyny) człowiek, który rozbił wczoraj kasę w Monte Carlo”, tak jak występuje ona w wypowiedzianym *dziś* zdaniu „(Jedyny) człowiek, który rozbił wczoraj kasę w Monte Carlo, miał dziurawe buty” [...] odwołuje się do bytu – zarówno intuicyjnie, jak i na mocy naszego kryterium.

(PEACOCKE 1973: 117)

Można powątpiewać, czy przykład ten rozstrzyga sprawę. Po pierwsze, kryterium polegające na gotowości do powiedzenia „to” zamiast „jedyne” nie musi ustalać równoważności. Mogę być gotów powiedzieć „Zmarł z powodu choroby wieńcowej” zamiast „Zmarł z powodu zawału serca”, lecz te twierdzenia nie są równoważne. Po drugie, wydaje się, że ktoś, kto przypadkowo usłyszał tę wypowiedź, a nie posiada wspólnej nam wiedzy, od której, według Peacocke’a, zależy odwołanie się do bytu, może, intuicyjnie, zrozumieć, co się mówi równie dobrze jak uczestnicy rozmowy.

Uważam, że nie istnieje rozstrzygający argument za potraktowaniem deskrypcji określonych inaczej, niż robi to Russell. [Pogląd ten rozwija i broni w szczegółach Neale (1992).]

3. ISTNIENIE

3.1. *Pogląd Russella*

Klasyczny problem istnienia dotyczy tego, jak gramatycznie pojedyncze przeczące zdania egzystencjalne mogą być prawdziwe. Przykładami takich zdań są „Wulkan nie istnieje” (chodzi nie o boga mitologii greckiej, ale o planetę postulowaną dla wyjaśnienia orbity Merkurego; okazało się, że taka planeta nie istnieje) i „Złota góra nie istnieje”. Problem polega na tym, że wydaje się, iż aby takie zdania były sensowne, gramatycznie pojedyncze wyrażenie („Wulkan”, „złota góra”) powinno się do czegoś odnosić; jeśli jednak wyrażenie do czegoś się odnosi, to przypuszczalnie jest to coś, co istnieje, a zatem takie zdanie jest fałszywe.

Jedna z odpowiedzi na tę trudność polega na przyjęciu, że są przedmioty, które nie istnieją. Możemy wówczas powiedzieć, że „Wulkan” i „złota góra” odnoszą się do takich nieistniejących przedmiotów, a zdania mówią o nich prawdziwie, że nie istnieją. Tę odpowiedź łączy się z Meinongiem (1904).

Russell podał dwa argumenty przeciwko tej odpowiedzi. Jedna jest taka, że sprzeciwia się ona „silnemu poczuciu rzeczywistości”, które jest „nieodzwone do podania poprawnej analizy sądów o jednoróżcach, złotych górach, kwadratowych kołach i innych takich pseudo-przedmiotach” (1919: 170). Wydaje się, że ta odpowiedź to jedynie wyraz pewnej preferencji.

Drugi argument wygląda na silniejszy, ponieważ głosi, że pogląd Meinonga wikła się w sprzeczność, z tego względu, że kwadratowe koło jest zarazem okrągłe i nieokrągłe oraz że istniejący obecny król Francji zarazem istnieje i nie istnieje. Ogólna zasada, do której odwołuje się ten argument, mówi, że prawdziwe musi być orzeczenie o takim-a-takim przedmiocie (*the so-and-so*), że jest on taki-a-taki (*a so-and-so*).

Każdy przyzna, że istniejące przedmioty nie mogą być sprzeczne, nie jest jednak jasne, dlaczego trzeba przyjąć, że dotyczy to również przedmiotów nieistniejących. Wewnętrzna sprzeczność kwadratowego koła można by uznać za dowód jego nieistnienia.

Russell mówi, że jeśli można podać teorię, która nie wikła się w te sprzeczności, to należy ją preferować (RUSSELL 1905: 45). Bez wątplenia musimy rozstrzygnąć, czy względy logiczne zmuszają nas do przyjęcia stanowiska Meinonga. Russell pokazuje, że istnieje stanowisko konkurencyjne.

Widzieliśmy już wcześniej, jak Russell radzi sobie z sytuacją, w której gramatycznie pojedyncze wyrażenie w przeczącym zdaniu egzystencjalnym jest deskrypcją (np. „Złota góra nie istnieje”). Zaprzecza on, że to wyrażenie rzeczywiście jest pojedyncze i że jego funkcją jest odnoszenie się, traktując je w zamian jako kwantyfikator egzystencjalny. Przypomnijmy, że analiza jest następująca: Nie jest tak, że istnieje dokładnie jeden przedmiot, który jest zarazem złoty i jest górą. Przy takiej analizie nie powstaje w ogóle problem odniesienia do przedmiotu nieistniejącego. Prawdziwość tego zdanie można równie łatwo zrozumieć jak prawdziwość zdania „Nie ma jednoróżców”.

Gdy gramatycznie pojedynczym wyrażeniem jest nazwa (np. „Wulkan”), to standardowa interpretacja Russella przyjmuje, że traktuje on ją jako synonim pewnej deskrypcji określonej, w ten sposób sprowadzając ów przypadek do poprzedniego. Nie zgadza się to z moją interpretacją Russella, w myśl której musi on dopuścić możliwość, że zdanie „Wulkan nie istnieje” nie ma publicznego znaczenia i wobec tego nie można podać jego ogólnej analizy. Każdy użytkownik języka będzie mógł używać tych słów do formułowania sądów, ale nie ma gwarancji, że istnieje jeden sąd, który wszyscy sformułują. Stąd jedna osoba mogłaby użyć tych słów do wydania sądu, że nie istnieje jedna jedyna planeta leżąca między Merkurym a Słońcem, a ktoś

inny mógłby ich użyć do zaprzeczenia, że promienie z jakiejś jednej planety powodują raka (zakładając, że popularność zyskała teoria, zgodnie z którą promienie z Wulkanu powodują raka).

W wypadku niektórych nazw bez nosiciela istnieje jakaś deskrypcja, którą osoba musi znać, aby zrozumieć tę nazwę. Na przykład: ktoś, kto nie wie, że „Święty Mikołaj” jest brodatym, prowadzącym sanie Lapończykiem, który przynosi świąteczne prezenty, nie zrozumie nazwy „Święty Mikołaj”. Analiza Russella może sprowadzić te przypadki do tych, w których gramatycznie pojedyncze wyrażenie jest deskrypcją określoną.

Russell wyraża po części swoje stanowisko mówiąc, że istnienie nie jest predykatem orzekanym o indywiduach. Rozumie przez to, że nie możemy uznać lub zaprzeczyć istnienia żadnego konkretnego przedmiotu – w przeciwnym razie otrzymalibyśmy tautologię (w wypadku uznania) lub sprzeczność (w wypadku zaprzeczenia: odnieść się do czegoś, jednocześnie zaprzeczając jego istnieniu to sprzeczność). Stawia on sprawę następująco: „Powiedzieć, że one [aktualne przedmioty, które są w świecie] istnieją to ściśle biorąc nonsens, ale powiedzieć, że one nie istnieją, to również ściśle biorąc nonsens” (RUSSELL 1918-1919: 233). „Istnieje” jest predykatem orzekanym nie o indywiduach, lecz o funkcjach zdaniowych. Przez funkcję zdaniową Russell rozumie wyrażenie w rodzaju „ x jest jednorożcem”. Powiedzieć, że jednorożce istnieją, to powiedzieć, że ta funkcja zdaniowa jest prawdziwa o przynajmniej jednym przedmiocie; powiedzieć, że jednorożce nie istnieją, to powiedzieć, że nie jest ona prawdziwa o niczym. Russellową analizę zdań w rodzaju „Złota góra nie istnieje” można wyrazić podobnie: „ x jest złote i jest górą i cokolwiek jest złote i jest górą jest identyczne z x ” nie jest prawdziwe o niczym.

Pewne racje przemawiają jednak za tym, że dopuszczamy, iż „istnieje” może być uznawane lub zaprzeczane o indywiduach. Ocenimy je w następnej części (3.2.)

3.2. Czy „istnieje” jest predykatem orzekanym o indywiduach?

Pomijając pogląd Meinonga, jeśli „istnieje” jest predykatem orzekanym o indywiduach, to jest to predykat prawdziwy o nich wszystkich. W idei tej nie ma nic zasadniczo błędnego: upodabnia się tu „istnieje” do takich predykatów jak „jest identyczne z samym sobą”.

W preferowanym przez Russella języku formalnym istnieje zwrot odpowiedni do wyrażenia istnienia indywiduów. Powiedzenie, że x istnieje, jest równoważne powiedzeniu, że jest coś, czym x jest. Można to sformalizować następująco:

$$\exists y(y = x)$$

Jest to prawdą o wszystkim (a odpowiednia formuła jest twierdzeniem), a zaprzeczenie tego jest zawsze fałszywe. Wyrażenie to jest jednak zbudowane całkowicie poprawnie. Russell nie powinien więc mówić, że próba uznania lub zaprzeczenia istnienia indywiduum prowadzi do nonsensu. Sytuacja jest raczej taka, że uznanie będzie zawsze udane, a zaprzeczenie – nieudane. Nie ma powodów logicznych, dla których predykat ten nie miałby być orzekany o indywiduach.

Z łatwością można zmodyfikować „istnieje” w taki sposób, że uzyska się predykat, który nie jest prawdziwy o wszystkim. Można, na przykład, używać „obecnie istnieje” tylko w odniesieniu do tych przedmiotów, które istnieją w chwili wypowiedzi, tak że predykat ten jest fałszywy w wypadku przedmiotów, które już lub jeszcze nie istnieją. Nie powinno to budzić kontrowersji; rzeczywisty problem dotyczy tego, czy możliwe jest wolne od wszelkich takich modyfikacji pojęcie *istnienia*, które stosuje się do indywiduów. Na korzyść takiego pojęcia można by przytoczyć nasuwające się wyjaśnienie semantyki, na przykład, „obecnie istnieje”: „obecnie” służy do ograniczenia *predykatu* „istnieje” do pewnego podzbioru wszystkich przedmiotów istniejących, podobnie jak „szczęśliwy” w „szczęśliwy człowiek” ogranicza zbiór wszystkich ludzi do pewnego podzbioru.

Można by się skłaniać do przypuszczenia, że niezmodyfikowane „istnieje”, jako orzekane o indywiduach, nie odgrywałoby w naszym języku żadnej roli, ponieważ nie pozwalałoby nam nigdy powiedzieć nic interesującego. Przypuszczenie to może być zbyt pochopne, ponieważ uznanie istnienia za predykat orzekany o indywiduach może się okazać konieczne do zrozumienia bardziej skomplikowanych konstrukcji. Potencjalny przykład pojawił się pod koniec poprzedniego akapitu. Inny podaje Moore (1936: 143-144), który słusznie twierdził, że zdanie „To mogłoby nie istnieć” jest zrozumiałe i, przynajmniej potencjalnie, dyskusyjne. (Jeśli „to” zostaje użyte, aby odnieść się do osoby, to przypuszczalnie to, co zostaje powiedziane, jest prawdziwe; jeśli do liczby, to przypuszczalnie fałszywe; mogą również istnieć przypadki sporne.) Najprostsze ujęcie „To mogłoby nie istnieć” przebiega następująco:

Mogłoby być tak, że: to nie istnieje.

Trudno byłoby nie uznać, że to, co następuje po dwukropku jest zaprzeczeniem orzeczenia istnienia o indywiduum.

Kolejną główną racją za tezą idei, że „istnieje” jest predykatem orzekanym o indywiduach, jest to, iż w przeciwnym razie musimy powiedzieć

coś niezadowolającego o przeczących zdaniach egzystencjalnych, w których występują nazwy. Musimy albo powiedzieć, wraz z mitycznym Russellem, że są one synonimami deskrypcji, i wpaść w kłopoty próbując znaleźć jakieś deskrypcje wspólne wszystkim użytkownikom języka; albo powiedzieć, wraz z prawdziwym Russellem, że nie ma gwarancji, iż takie zdania mają w ogóle znaczenie w języku publicznym. Ponieważ żadna z tych opcji nie jest w pełni zadowolająca (choć drugą trudno rozstrzygająco obalić), warto sprawdzić, to czy nie istnieje jakieś alternatywne rozwiązanie.

3.3. *Pogląd Evansa*

Jednym z takich alternatywnych rozwiązań jest pogląd metajęzykowy. Zgodnie z analizą językową takie zdanie jak „Wulkan nie istnieje” mówi, że nazwa nie ma nosiciela. Problem z owym poglądem polega na tym, że nie nakłada on żadnych warunków na rozumienie tego rodzaju zdań. Intuicyjnie nie rozumiemy zdania „Wulkan nie istnieje”, jeśli nie znamy pewnych faktów astronomicznych. Pogląd metajęzykowy uniemożliwia wyjaśnienie tego warunku, ponieważ rozumienie wymaga jedynie wiedzy o tym, o której nazwie się mówi, że nie posiada ona nosiciela.

O ile wiem, jedyną poważną próbę wprowadzenia w życie idei, że „istnieje” jest predykatem orzekanym o indywiduach, podjął Gareth Evans. Sednem jego propozycji jest to, że pojedyncze przeczące prawdy egzystencjalne wykorzystują fikcję po to, aby obnażyć ją jako fikcję. Ktoś, kto używa zdania w ten sposób

nie jest jak ktoś, kto chcąc zapobiec temu, aby widownię teatralną poniosło, wskazuje na scenę wołając: „Spójrzcie, ci ludzie są tylko aktorami i nie ma tutaj rusztowania ani budynku – to są tylko rekwizyty”. Raczej jest on jak ktoś, kto wskazuje na scenę i woła: „Spójrzcie, Zuzanna i tamten złodziej to tylko postaci w sztuce, a to rusztowanie i ten budynek są tylko rekwizytami”. Widownia musi być gotowa uczestniczyć – lub przynajmniej być gotowa uczestniczyć – w tym udawaniu, aby zrozumieć, co on mówi.

(EVANS 1982: 369)

Kontrast polega na tym, że – w wypadku, który interesuje Evansa – wykorzystujemy konwencje fikcji i traktujemy osoby na scenie jako postaci, a nie jako aktorów, tło natomiast jako domy, a nie zwykłą scenografię, tym samym zaś do pewnego stopnia pozostajemy wewnątrz fikcji w samym akcie przedstawiania jej jako tylko fikcji.

Evans uważa, że aby zastosować tę ideę do przypadku językowego, musimy przyjąć, przynajmniej milcząco, operator „rzeczywiście”, dla którego podaje precyzyjne warunki prawdziwości. Kiedy mówimy – fałszywie – „Hamlet istnieje”, to powinniśmy rozumieć to jako wypowiedź o formie logicznej „Hamlet *rzeczywiście* istnieje”. Podobnie „Wulkan nie istnieje” ma formę logiczną „Wulkan *rzeczywiście* nie istnieje”.

Aby wyjaśnić warunki prawdziwości dla „rzeczywiście”, musimy najpierw wprowadzić ideę jedynie fikcyjnych warunków prawdziwości. I tak możemy ocenić zdanie „Hamlet był stanowczy” jako prawdziwe lub fałszywe w pewnej fikcji, *prawdziwe* lub *fałszywe*, jak ujmuje to Evans, ponieważ fikcja nakłada na zdanie fikcyjne warunki prawdziwości (tj. *warunki prawdziwości*). Z grubsza zdanie „Hamlet był stanowczy” jest *prawdziwe* zawsze i tylko, gdy według pewnej fikcji był on stanowczy, a *fałszywe* zawsze i tylko, gdy według tej fikcji nie był on stanowczy.

„Rzeczywiście *p*” (dla dowolnego zdania „*p*”) jest prawdziwe zawsze i tylko, gdy:

- (1) „*p*” posiada *warunki prawdziwości*.
- (2) „*p*” posiada warunki prawdziwości.
- (3) „*p*” jest prawdziwe (a nie jedynie *prawdziwe*).

Pierwszy warunek wymaga, aby „*p*” miało użycie w pewnego rodzaju fikcji. Drugi warunek wymaga, aby to zdanie miało również właściwe użycie nie-fikcyjne. Rzadko się zdarza, by spełnione były oba te warunki. Jedynymi przypadkami, dla zdań zawierających nazwy, są te, w których jakaś fikcja została osnuta wokół prawdziwej osoby. I tak historia El Cida rzekomo nawiązywała do prawdziwej osoby, a zatem takie zdanie jak „El Cid był odważny” posiada zarówno *warunki prawdziwości*, nadane mu przez tę historię, jak i warunki prawdziwości, związane ze zwykłym użyciem nie-fikcyjnym. A zatem zdanie „El Cid był *rzeczywiście* odważny” jest prawdziwe zawsze i tylko, gdy składnik „El Cid był odważny” posiada zarówno *warunki prawdziwości*, jak i warunki prawdziwości oraz jest prawdziwy; z tego względu jest ono (jak miemam) prawdziwe.

Zdanie „Hamlet *rzeczywiście* istnieje” (bardziej idiomatycznie: „istniał”) jest prawdziwe tylko, gdy zdanie „Hamlet istnieje” posiada warunki prawdziwości. Ponieważ ich nie posiada – nie posiada użycia poza fikcją – „Hamlet *rzeczywiście* istnieje” jest fałszywe, a „Hamlet *rzeczywiście* nie istnieje” jest prawdziwe. Wszystko jest tak jak należy.

Istnieją prawdy o formie „ x rzeczywiście istnieje”. Evans sugeruje, że jeśli widzicie małego zielonego ludzika, ale uważacie, iż ulegacie przywidzeniu, a ja to wszystko wiem, to mogę prawdziwie powiedzieć „Ten mały zielony ludzik rzeczywiście istnieje”. „Ten mały zielony ludzik istnieje” posiada *warunki prawdziwości*, generowane przez wasze przywidzenie oraz warunki prawdziwości generowane przez fakt, że istnieje mały zielony ludzik będący odniesieniem zwrotu „ten mały zielony ludzik” oraz przez fakt, że zdanie to jest prawdziwe. A zatem zdanie „Ten mały zielony ludzik rzeczywiście istnieje” jest prawdziwe, podobnie jak zdanie „El Cid rzeczywiście istniał”. Te wyniki również wydają się intuicyjnie właściwe.

Jest to atrakcyjna teoria, dzięki której można sobie radzić z przypadkami fikcyjnymi, nie jest jednakże oczywiste, jak zastosować ją do przypadków takich jak „Wulkan nie istnieje”, w których nie ma wprost żadnej fikcji. Astronomowie, którzy wprowadzili tę nazwę, traktowali ją poważnie; nie opowiadali niestworzonych historii. Być może moglibyśmy ukuć pojęcie „nieświadomej fikcji”, z którą mamy do czynienia wówczas, gdy językowa aktywność ludzi rozmija się z rzeczywistością. Przypomina ona fikcję, z tym wyjątkiem, że nikt nie miał zamiaru stwarzać fikcji. Moglibyśmy wtedy uznać, że astronomowie, którzy wprowadzili nazwę „Wulkan”, uczestniczyli w nieświadomej fikcji, tak że zdanie „Wulkan istnieje” posiada *warunki prawdziwości*. Jednakże zdanie to nie posiada warunków prawdziwości, ponieważ nie posiada autentycznego użycia (to jest: nie posiada użycia, w którym występuje odniesienie do czegoś), a jedynie użycia wewnątrz nieświadomej fikcji. To wystarczy do zagwarantowania, że zdanie „Wulkan rzeczywiście istnieje” jest fałszywe, a zatem zdanie „Wulkan rzeczywiście nie istnieje” jest prawdziwe. (Zob. też STRAWSON 1967.)

4. IDENTYCZNOŚĆ

Identyczność prowadzi do różnych pytań. Z jednym z najważniejszych zetknęliśmy się już, omawiając wcześniej łamigłówkę Fregego: jak wyjaśnić oczywistą różnicę między prawdą o formie $a = a$ a prawdą o formie $a = b$? W tej części rozważymy kolejne pytanie: czy identyczność jest konieczna czy przygodna, a dokładniej – czy jakieś identyczności są przygodne?

Sąd jest przygodny, czy przygodnie prawdziwy, jeśli jest prawdziwy, ale mógłby nie być prawdziwy. Jedynie przygodnie jest prawdą, że w tej chwili na moim biurku znajdują się dwa długopisy; mogłoby być inaczej. Sąd jest

konieczny czy koniecznie prawdziwy, jeśli jest prawdziwy i nie mógłby być fałszywy. Standardowy przykład stanowią prawdy matematyczne: nie mogłoby być inaczej niż tak, że $5 + 7 = 12$.

Za Kripkem (1972) musimy rozróżnić między koniecznością sądu a jego poznawalnością *a priori*. Sąd jest poznawalny *a priori*, jeśli może być poznany bez odwołania się do doświadczenia, na przykład, dzięki czystemu rozumowi. Jest sprawą otwartą, czy wszystko, co jest poznawalne *a priori*, jest konieczne oraz czy wszystko, co jest konieczne, jest poznawalne *a priori*. Sam Kripke argumentował przeciwko obu tym sugestiom. W szczególności twierdzi on, że identyczności są konieczne, ale nie wszystkie są poznawalne *a priori*. Na przykład jest konieczne, że Hesperos jest to Fosforos; nie mogłoby być inaczej. Nie jest to jednak poznawalne *a priori* – w przeciwnym razie starożytni astronomowie bez wątpienia by o tym wiedzieli.

Rozważając, czy jest konieczne, że Hesperos jest to Fosforos, moglibyśmy spróbować wyobrazić sobie sytuację, w której ta identyczność nie zachodzi. Możemy sobie przedstawić sytuację, w której jedna planeta jako pierwsza pojawia się wieczorem, a inna planeta jako ostatnia znika rankiem; w ten sposób nie przedstawiamy sobie jednak jeszcze sytuacji, w której Hesperos jest różny od Fosforosa. Może to być po prostu sytuacja, w której Hesperos (= Fosforos) posiada inną orbitę, tak że nie pojawia się rankiem; albo taka sytuacja, w której Hesperos (= Fosforos) posiada jeszcze inną orbitę, tak że nie pojawia się wieczorem. Twierdzenie, że Hesperos jest różny od Fosforosa, wydaje się teraz arbitralne; krytycy powiedzą, że nie można w ten sposób opisać niczego, co sobie wyobraziliśmy.

Na poparcie poglądu, że identyczność jest konieczna, można przytoczyć następujący argument formalny:

- (1) Zasada Leibniza: jeśli x jest identyczne z y , to cokolwiek jest prawdą o x , jest prawdą również o y .
- (2) Konieczność identyczności z samym sobą: wszystko jest w sposób konieczny identyczne z samym sobą.
- (3) Zastosowanie zasady Leibniza do własności bycia w sposób konieczny identycznym z x : x w sposób konieczny posiada własność bycia identycznym z samym sobą (na mocy (2)), a zatem jeśli x jest identyczne z y , to y w sposób konieczny posiada własność bycia identycznym z x (na mocy (1)). Inaczej mówiąc, jeśli x jest identyczne z y , to x jest w sposób konieczny identyczne z y .

Argument ten można też wyrazić symbolicznie:

- (1) $\forall x \forall y (x = y \rightarrow (Fx \leftrightarrow Fy))$.
 (2) $\forall x \Box (x = x)$.
 (3) $\forall x \forall y (x = y \rightarrow \Box (x = x) \leftrightarrow \Box (x = y))$.
 Zatem: $\forall x \forall y (x = y \rightarrow \Box (x = y))$.

Wiersz (3) wersji symbolicznej wynika z (1) po zastąpieniu „ $F...$ ” w (1) przez „ $\Box (x = ...)$ ”; wniosek zostaje wyprowadzony z (2) i (3). Stwierdza on, że prawdziwa identyczność jest prawdziwa w sposób konieczny. (Formalną wersję tego argumentu przypisuje się na ogół Barcan Marcus (1947). Argument ten podają, na przykład, Kripke (1971) i Wiggins (1980a: 109-110).)

Najbardziej wątpliwą częścią tego argumentu jest twierdzenie, że „ $\Box (x = ...)$ ” jest właściwym podstawieniem „ $F...$ ” w zasadzie Leibniza. Ujmując rzecz mniej formalnie, wątpliwość dotyczy tego, czy własność bycia w sposób konieczny identycznym z x można w sposób prawdziwy orzec o czymkolwiek. Jak zobaczymy, niektórzy przeczą istnieniu takich własności jak bycie z czymś w sposób konieczny identycznym, a zatem zasada Leibniza nie ma żadnego zastosowania. Skąd bierze się to przekonanie?

Wynika ono częściowo z poczucia, że istnieją niezależne argumenty na rzecz twierdzenia, że niektóre identyczności są przygodne. Usuńmy najpierw na bok pewne nieinteresujące przypadki. Bez wątplenia jest czymś przygodnym, że główny poczmistrz Stanów Zjednoczonych był wynalazcą soczewek dwuogniskowych (zob. KRIPKE 1972). Ponadto intuicyjnie zdanie to można by zaklasyfikować jako zdanie identycznościowe. Jak jednak łatwo zauważyć, źródłem przygodności jest tu fakt, że deskrypcje określone mogłyby, znacząc to, co znaczą, odnosić się do innych przedmiotów; w terminologii Kripkego są one niesztynnymi desygnatorami. Przypadek ten nie przemawia więc w żaden sposób za przygodnością samej relacji identyczności.

Trzeba rozróżnić między twierdzeniem, że zdania o identyczności są przygodne, a twierdzeniem, że sama relacja identyczności jest przygodna. Aby relacja była przygodna, muszą istnieć przedmioty, między którymi zachodzi ona jedynie przygodnie. Aby relacja była konieczna, musi być tak, że jeśli ta relacja zachodzi między jakimiś przedmiotami, to zachodzi ona między nimi w sposób konieczny. (Jest pewną osobliwością, że zgodnie z gramatyką języka polskiego trzeba użyć liczby mnogiej („przedmiotami”), gdy faktycznie chodzi o jeden przedmiot i w związku z tym liczba pojedyncza byłaby bardziej na miejscu.) Widać to wyraźnie w formalnej wersji wniosku omawianego tu argumentu ((1)-(3) powyżej). Można spójnie powiedzieć, że istnieją przygodne zdania o identyczności, mimo że sama relacja jest konieczna.

Można zatem powiedzieć, że zdanie „Główny poczmistrz Stanów Zjednoczonych był wynalazcą soczewek dwuogniskowych” jest przygodne i jest zdaniem identycznościowym, ale że jeśli weźmiemy pod uwagę przedmiot x , do którego faktycznie odnosi się deskrypcja „pierwszy poczmistrz Stanów Zjednoczonych”, i przedmiot y , do którego odnosi się deskrypcja „wynalazca soczewek dwuogniskowych”, to jest konieczne, że x jest identyczne z y . (Czy Russell mógłby dopuścić, że „Główny poczmistrz Stanów Zjednoczonych był wynalazcą soczewek dwuogniskowych” jest zdaniem identycznościowym?)

Pora teraz omówić poważne wyzwanie wobec konieczności relacji identyczności (zob. GIBBARD 1975). Rozważmy posągi, które są wykonane z kawałków gliny. Zazwyczaj taki kawałek zaczyna istnieć wcześniej niż posąg, który tworzy: dopiero później nadaje mu się kształt posągu. Dość często kawałek przestaje istnieć wcześniej niż posąg: jeśli posąg zostaje wyszczerbiony, mamy już inny kawałek, choć nadal ten sam posąg. A zatem zazwyczaj posągi różnią się od kawałków gliny, z których są wykonane, ponieważ mają inne historie życia.

Rozważmy teraz przypadek, w którym posąg zostaje uformowany w dokładnie tym samym momencie, w którym zaczyna istnieć określony kawałek gliny. Na przykład dwie ukształtowane połówki posągu zostają ze sobą połączone, tworząc w ten sposób równocześnie nowy kawałek i nowy posąg. Nazwijmy ten kawałek „Kawałkiem₁”, a posąg „Goliatem”. Przypuśćmy, że istnienie Goliata kończy się rozbiciem go na drobne kawałki. Zdarzenie to niszczy równocześnie Kawałek₁. Tak więc Goliat i Kawałek₁ zaczynają istnieć w dokładnie tej samej chwili i kończą swoje istnienie w dokładnie tej samej chwili. Gibbard sugeruje, że powinniśmy uznać następujące twierdzenie:

(A) Goliat jest identyczny z Kawałkiem₁.

Gdybyśmy uznali, że Goliat i Kawałek₁ się różnią, to stanęlibyśmy przed koniecznością rozwiązania najprzeróżniejszych łamigłówek. Jak dwie różne rzeczy mogą równocześnie zajmować to samo miejsce? Jeśli Goliat waży 2 funty i Kawałek₁ waży tyle samo, a mimo to są różnymi przedmiotami, to dlaczego nie ważą łącznie 4 funtów? Jeśli się różnią, to wydaje się, że nie można poprawnie przejść od przesłanki, że zniszczyliśmy Goliata, do wniosku, że zniszczyliśmy Kawałek₁.

Równocześnie wiemy, że mogłoby być inaczej. Kawałek₁ mógłby zostać zniszczony, podczas gdy Goliat by przetrwał. A zatem musimy uznać zarówno (A), jak i

(B) Goliat mógłby nie być identyczny z Kawałkiem₁.

Jeśli (A) i (B) są prawdziwe, to dają nam one łącznie przypadek przygodnej identyczności.

Gdyby się z tym zgodzić, to uzyskalibyśmy rację dla zaprzeczenia, że indywidua posiadają własności modalne, takie jak bycie w sposób konieczny identycznym z x . Takiego właśnie stanowiska broni Gibbard i podaje dość drobiazgową rekonstrukcję pozornie modalnych własności indywiduów w innych kategoriach.

Sądzę jednak, że nie został podany dobry argument za przyjęciem (A). O ile wcześniej nie dokonaliśmy jakiegoś szalbierstwa z zasadą Leibniza, to możemy udowodnić, że (A) jest fałszywe: Goliat posiada własność, której brak Kawałkowi₁, mianowicie może przetrwać utratę części swojej materii. A jak mówi zasada Leibniza, jeśli x posiada własność, której brak y , to x jest różny od y .

Poza tym ze wspomnianymi łamigłówkami musimy się uporać niezależnie od tego, czy uznajemy (A), czy nie, a zatem nie dostarczają nam one dodatkowej racji uznania (A). Rozważmy przypadek, w którym, jak przyznałby Gibbard, posąg i kawałek są różne, na przykład dlatego, że kawałek zaczyna istnieć wcześniej niż posąg. Nawet wówczas jest okres, w którym dwie rzeczy zajmują równocześnie to samo miejsce i w którym każda z nich posiada jakiś ciężar, a ich łączny ciężar jest taki sam, jak ciężar każdej z tych rzeczy z osobna i w którym nie da się logicznie wywnioskować zniszczenia jednej z nich ze zniszczenia drugiej.

Dlatego, według mnie, argumenty Gibbarda nie obalają konieczności relacji identyczności.

5. KONIECZNOŚĆ

5.1. Pudełka i diamenty

Posługiwaliśmy się już wcześniej pojęciem konieczności. Pytanie, które podejmiemy w tej części, jest następujące: jaka jest logiczna forma (*logical form*) zdań modalnych języka polskiego? Modalne zdanie języka polskiego to takie zdanie, które zawiera jakieś wyrażenie modalne (wyrażające możliwość, przygodność lub konieczność). Oto przykłady:

- (A) Jest konieczne, że $5 + 7 = 12$.
- (B) Hesperos jest w sposób konieczny identyczny z Fosforesem.
- (C) Musisz zapewnić odpowiednie zabezpieczenie finansowe dla swoich dzieci.
- (D) Musisz wyjść teraz, albo spóźnisz się na samolot.

- (E) Jeśli naciśniesz koniec sztywnego drążka, swobodnie zrównoważonego pośrodku, drugi koniec musi się unieść.
 (F) To po prostu musi być błąd.

Nie ulega wątpliwości, że w tych przykładach mamy do czynienia z różnymi standardami czy kryteriami konieczności. Można uważać, że (A) odwołuje się do konieczności logicznej, (B) do konieczności metafizycznej, (C) do konieczności moralnej, (D) do konieczności roztropnościowej, (E) do konieczności przyrody, a (F) do konieczności epistemicznej. W tej części zajmujemy się koniecznością w stosunkowo wymagającym sensie, obejmującym co najwyżej przykłady (A) i (B).

Zadając pytanie dotyczące formy logicznej zdań modalnych, pytamy w istocie o to, jak te zdania funkcjonują i jaka jest ich logika. Sugestia, która zostanie rozważona w tej części, jest taka, że właściwym wyrazem modalności jest jednoargumentowy operator zdaniowy: wyrażenie, które, podobnie jak „nie”, tworzy zdanie ze zdania. Operator dla konieczności zapisuje się jako „ \square ” i nazywa „pudełkiem”. Operator dla możliwości zapisuje się jako „ \diamond ” i nazywa „diamentem”¹. A zatem „jest konieczne, że p ” zapisuje się jako „ $\square p$ ”, a „jest możliwe, że p ” jako „ $\diamond p$ ”. Moglibyśmy przyjąć jeden z tych operatorów – diament lub pudełko – jako pierwotny i wprowadzić drugi za pomocą definicji. Na przykład:

„ $\diamond p$ ” zostaje zdefiniowane jako „nie- \square nie- p ”.

Inaczej mówiąc, możemy zdefiniować to, że coś jest możliwe, jako to, że nie jest niemożliwe – tj. jako to, że nie jest konieczne, iż to coś nie zachodzi.

Istnieją tu ciekawe różnice zasięgu. Na przykład, musimy odróżnić:

- (G) \square każdy matematyk jest matematykiem.
 (H) Każdy matematyk jest \square matematykiem.

(G) mówi zgodnie z prawdą, że jest konieczne, iż matematycy są matematykami. (H) mówi fałszywie, o każdym matematyku, że w sposób konieczny jest matematykiem. Jest to fałszywe, ponieważ ludzie mogą wybierać swój zawód.

Inna para zdań, którą należy odróżnić, jest następująca:

- (I) \square coś jest człowiekiem.
 (J) Coś jest \square człowiekiem.

¹ Określenia „pudełko” i „diament” funkcjonują wyłącznie w angielskiej terminologii logicznej (przyp. tłum.)

Pierwsze jest fałszywe: niewielkie zakłócenia w wartościach niektórych parametrów Wielkiego Wybuchu doprowadziłyby do tego, że na żadnej planecie nie pojawiłoby się życie. Drugie jest najprawdopodobniej prawdziwe, ponieważ chyba wszystko, co jest człowiekiem, musi nim być.

Jaka jest tu odpowiednia logika? Z pewnością zechcemy przyjąć, że

(K) $\Box p$ pociąga p ,

czyli, że wszystko, co w sposób konieczny jest takie-a-takie, jest takie-a-takie. Zachcemy również przyjąć następującą regułę wnioskowania:

(L) $Z \Box p$ i \Box (jeśli p , to q) wynika q .

Najprawdopodobniej powinniśmy również przyjąć

(M) $\Box p$ pociąga $\Box\Box p$.

Jest to charakterystyczny aksjomat logiki nazywanej, za C. I. Lewisem, S4 (LEWIS i LANGFORD 1932). Mówi on, że rzeczy nie są konieczne w sposób przypadkowy: to, co jest konieczne, musi być konieczne. Istnieją filozoficzne argumenty przeciwko tej tezie. Na przykład: być może to, co jest konieczne, nie jest obiektywną cechą świata, lecz zostaje określone przez nasze zdolności i zainteresowania. Gdyby mogły one być inne, niż są, wówczas to, co faktycznie jest konieczne, nie musiałyby takie być.

Teza charakterystyczna dla logicznego systemu S5 jest następująca:

(N) p pociąga $\Box\Diamond p$.

Mówi ona, że to, co aktualnie jest takie-a-takie, jest, z konieczności, możliwe takie-a-takie. W konsekwencji, jeśli coś jest w sposób możliwy konieczne, to jest takie aktualnie. Ma to zaskakujące zastosowanie do pewnej wersji dowodu ontologicznego. Argument ten można przedstawić następująco:

(O) Pojęcie Boga jako istoty koniecznej jest pojmowalne.

(P) Wszystko, co istnieje z konieczności, istnieje.

(Q) Zatem Bóg istnieje.

Traktując (P) jako punkt wyjścia, moglibyśmy zinterpretować (O) jako mówiące, że jest możliwe, iż Bóg koniecznie istnieje: $\Box\Diamond$ Bóg istnieje. W S5 pociąga to, że Bóg istnieje (bez potrzeby odwoływania się do innych przesłanek). Można z tego wyciągnąć morał, że nie powinniśmy zbyt pośpiesznie formalizować (O) w kategoriach pudełek i diamentów albo że logika S5 jest mocniejsza niż ta, którą posługujemy się w naszych rozumowaniach.

Jak sformułować semantyczną teorię dla języka pudełek i diamentów? W grę wchodzi dwie możliwości. Jedna jest taka, że powinniśmy postępować „homofonicznie” i przy podawaniu warunków prawdziwości posłużyć się ponownie pojęciem konieczności, wyrażanym przez operator zdaniowy. Weźmy dla porównania negację. Jesteśmy przyzwyczajeni mówić, że „nie- p ” jest prawdziwe zawsze i tylko, gdy „ p ” nie jest prawdziwe. Analogicznie odpowiedni semantyczny aksjomat dla pudełka mógłby być następujący: „ $\Box p$ ” jest prawdziwe (w określonym języku) zawsze i tylko, gdy „ p ” jest w sposób konieczny prawdziwe (w tym języku).

Żywo dyskutowanym rozwiązaniem konkurencyjnym jest podanie semantyki w kategoriach światów możliwych. Intuicyjnie powiedzieć, że jest konieczne, iż p , to powiedzieć, że p zachodzi w każdej możliwej sytuacji czy w każdym możliwym świecie. Moglibyśmy więc zaproponować aksjomat: „ $\Box p$ ” jest prawdziwe zawsze i tylko, gdy dla każdego możliwego świata w , p jest prawdziwe w w . Nasuwa się natychmiast pytanie, czy zobowiązuje nas to do uznania istnienia możliwych światów. Graeme Forbes (1985) argumentował, że nie, ale nie zdołał mnie przekonać (zob. SAINSBURY 1991: 284-2855). Jeśli posługiwanie się semantyką światów możliwych dla języka pudełek i diamentów zakłada ontologię światów, to równie dobrze można by podać formy logiczne wprost w języku światów możliwych, co omówimy w trzech następnych paragrafach (5.2-5.4).

5.2. Kwantyfikatory (1): Dodatkowe miejsce na argument

Podstawowa idea kryjąca się za podejściami omawianymi w tym i dwóch kolejnych paragrafach jest taka, że modalność należy reprezentować w formie logicznej jako kwantyfikację po światach możliwych. Konieczność jest reprezentowana przez kwantyfikację ogólną (prawdziwość we wszystkich możliwych światach), a możliwość – przez kwantyfikację egzystencjalną (prawdziwość w niektórych możliwych światach).

Samo zastąpienie pudełka przez zwrot „w każdym możliwym świecie” nie gwarantuje zachowania sensu. Można, na przykład, powątpiewać, czy zdanie „W każdym możliwym świecie, p ” jest sensowne. Jeśli jest to skrót dla „ $\langle p \rangle$ jest prawdziwe w każdym możliwym świecie”, to zmusza nas to do przyjęcia poglądu metajęzykowego, co jest niepożądane. (Tak jak w wypadku innych teorii metajęzykowych, nie moglibyśmy wyjaśnić, dlaczego rozumienie pierwotnego twierdzenia „W sposób konieczny p ” wymaga rozumienia „ p ”.)

Jednym ze sposobów na wybrnięcie z tej sytuacji jest uznanie, że każdy predykat posiada o jedno miejsce argumentowe więcej niż się początkowo wydaje. I tak zdanie „Jest konieczne, że Sokrates jest człowiekiem” zo- stałoby sformalizowane jako:

Dla każdego możliwego świata w Sokrates jest człowiekiem- w - w .

Normalnie myślimy o wyrażeniu „jest człowiekiem” jako o predykanie jedno- argumentowym: dodając do niego jedną nazwę, możemy utworzyć zdanie. Zgodnie z obecną sugestią jest to predykat dwuargumentowy: potrzebujemy dwóch nazw lub innych terminów jednostkowych (np. zmiennej), aby utwo- rzyć zdanie. Okazuje się więc, że wyrażenie „jest człowiekiem” przypomina potocznie rozumiane „kochać”, które z kolei nie jest dwu-, lecz trójargumen- towe i posiada – poza dwoma miejscami na [imiona] zakochanych – dodatko- we miejsce na nazwę świata lub na zmienną przebiegającą światy.

David Lewis (1986) zasugerował, że jest to niezadowalające, ponieważ nie oddaje sprawiedliwości przypadłościowym własnościom wewnętrznym (*accidental intrinsic properties*). Własność przedmiotu jest wewnętrzna, jeśli jej posiadanie przez ten przedmiot nie jest uzależnione od żadnego innego przedmiotu. Własność jest przypadłościowa, jeśli jest przygodna: przedmiot ją posiada, ale mógłby jej nie posiadać. W wypadku takiej własności mu- simy dokonać relatywizacji do różnych światów. W myśl branej teraz pod uwagę propozycji logiczna forma zdania „Ta strona jest biała, lecz mogłaby nie być” jest następująca:

Ta strona jest biała- w - w^* , lecz nie jest biała- w - w ,

gdzie „ w^* ” reprezentuje aktualny świat, a „ w ” jakiś inny świat. Intuicyjnie, *biały* jest własnością wewnętrzną – posiadanie jej przez pewien przedmiot nie powinno zależeć od żadnego innego przedmiotu. Jednak powyższa forma logiczna przedstawia *biały* nie jako własność wewnętrzną, lecz jako włas- ność relacyjną, zachodzącą między przedmiotem a światem. Kłóci się to z naszym rozumieniem natury takich własności. Jeśli więc istnieje inny sposób podania ich formy logicznej, należy go preferować.

5.3. Kwantyfikator (2): „ $\forall w$ ”

David Lewis (1986) zaproponował inny sposób użycia kwantyfikatorów do podania logicznej formy zdań modalnych. (Ten paragraf omawia jeden aspekt teorii Lewisa; kolejny aspekt, pojęcie odpowiednika (*counterpart*), zostanie omówiony w następnym paragrafie – 5.4).

W zdaniu takim jak

W Australii wszystkie piwa są dobre

zwrot „w Australii” ogranicza zbiór wartości zmiennej związanej przez kwantyfikator w tym, co po nim następuje. W tym kontekście, nie wszystkie piwa są istotne dla prawdziwości tego zdania, a jedynie piwa australijskie. Lewis sugeruje, że może nam to posłużyć jako model dla formalizowania zdań modalnych. Kluczowe wyrażenia mają formę „w w”, gdzie w jest nazwą możliwego świata lub zmienną przebiegającą możliwe światy. Zadaniem takiego wyrażenia jest ograniczenie wartości zmiennych związanych przez kwantyfikator do określonego świata. I tak

Dla pewnego świata w, w w (wszyscy politycy są uczciwi)

jest prawdziwe zawsze i tylko, gdy istnieje jakiś świat, w którym wszyscy politycy są uczciwi. W myśl omawianej propozycji taką formę logiczną ma zdanie „Mogłoby być tak, że wszyscy politycy są uczciwi”.

W jaki sposób wybrać między tego rodzaju podejściem a – na przykład – językiem pudełek i diamentów? Są tutaj dwa kluczowe argumenty. Po pierwsze, Lewis zasugerował, że pewnych rzeczy, które chcemy powiedzieć, nie można wyrazić w języku pudełek i diamentów, a można je wyrazić w języku kwantyfikatorów wiążących zmienne przebiegające światy możliwe. Jeśli Lewis ma rację, to z pewnością jest to bardzo istotna zaleta podejścia kwantyfikatorowego.

Po drugie, metafizyka światów możliwych jest dla wielu osób nie do przyjęcia. Traktowanie światów możliwych jako autentycznie istniejących bytów to według nich czyste szaleństwo. Z tego względu wiele osób próbuje opisywać światy możliwe w kategoriach aktualnych bytów, na przykład traktując je jako teoriomnogościowe konstrukcje z bytów aktualnych. Lewis nazwał takie teorie „ersacyzmem”. Problem w tym, że jeśli ersacyzm się nie sprawdzi, to pozostaje nam dziwaczny wniosek, że światy możliwe inne niż nasz własny rzeczywiście istnieją. Niemal wszyscy, nie wyłączając Lewisa, uznają to za wadę; pytanie tylko, czy wady tej nie są w stanie zrekompenzować jakieś zalety, a zwłaszcza większa ekspresyjna moc języka z kwantyfikatorami wiążącymi zmienne przebiegające światy możliwe.

Odłóżmy wyłaniające się w tym miejscu zagadnienia metafizyczne do części 5.5. Na razie podamy dwa przykłady mające rzekomo ilustrować większą ekspresyjną moc podejścia kwantyfikatorowego.

Jeśli założymy, że podejście kwantyfikatorowe dysponuje nazwą dla świata

aktualnego, powiedzmy „ w^* ”, to z pewnością pozwala nam to powiedzieć rzeczy, których nie da się powiedzieć w języku pudełek i diamentów. Rozważmy polskie zdanie

Wszystko, co jest czerwone, mogłoby być błyszczące,

które rozumiem jako równoważne zdaniu „Jest możliwe, że wszystko, co jest aktualnie czerwone, mogłoby być również błyszczące” (zob. DAVIES 1981: 220-221). Z łatwością można to sformalizować w kwantyfikatorowym języku „ w w^* ” następująco:

Dla pewnego świata w i dla wszystkich x , jeśli w w^* (x jest czerwone), to w w (x jest błyszczące),

gdzie „ w^* ” to nazwa naszego aktualnego świata. Wydaje się natomiast, że nie można podać adekwatnej formalizacji tego zdania w języku pudełek i diamentów.

Dla każdego x , jeśli x jest czerwone, to $\diamond x$ jest błyszczące

nie nadaje się, ponieważ wymaga jedynie, aby każda czerwony przedmiot z osobna mógł być błyszczący, podczas gdy w wyjściowym sformułowaniu mowa o możliwości, w której wszystkie czerwone przedmioty jednocześnie są błyszczące.

\diamond dla wszystkich x , jeśli x jest czerwone, to x jest błyszczące

również nie zdaje egzaminu, ponieważ byłoby prawdziwe, gdyby wszystkie czerwone przedmioty w pewnym możliwym świecie były błyszczące, podczas gdy w wyjściowym sformułowaniu mowa o czerwonych przedmiotach w aktualnym świecie.

Język pudełek i diamentów można jednak rozszerzyć w taki sposób, aby usunąć ten defekt. Musimy dodać jakieś wyrażenie na określenie aktualności, odpowiadające „ w^* ”, ale zgodne z podejściem operatorowym. Nadaje się do tego w sam raz polskie słowo „aktualnie” lub jakiś jego skrót. Pozwoli nam to poprawnie sformalizować problematyczne zdanie następująco:

\diamond dla każdego x , jeśli x jest aktualnie czerwone, to x jest błyszczące.

(Jak moglibyśmy użyć „aktualnie”, aby przekształcić nieszytywne deskrypcje określone w szytywne?)

Kolejny problem dotyczący ekspresyjnej adekwatności, który omówię, również posłużył Lewisowi jako racja tego, aby przedkładać kwantyfikatory

nad pudełka i diamenty w formalizacji modalności. Rozważa on tezę o superweniencji tego, co mentalne, względem tego, co fizyczne:

nie mogłoby być różnicy mentalnej między dwojgiem ludzi, gdyby nie było między nimi pewnej różnicy fizycznej [...] Odczytując „mogłoby” jako diament, otrzymujemy następującą tezę: nie istnieje świat [...], w którym dwoje ludzi różni się mentalnie, choć nie ma jakiegóż różnicy fizycznej [...] między nimi. Nie jest to jednak w pełni poprawne. Niepotrzebnie ograniczyliśmy naszą uwagę do fizycznych różnic między dwojgiem ludzi w tym samym świecie, pomijając w ten sposób te różnice zewnętrzne, które powstają jedynie między ludźmi w różnych światach.

(LEWIS 1986: 16)

W myśl teorii Lewisa ogólne strukturalne własności czasoprzestrzeni są jednakowe w całym świecie, tak więc osoba, której czasoprzestrzeń jest Riemannowska, musi znajdować się w innym świecie niż osoba, której czasoprzestrzeń odpowiada opisowi Łobaczewskiego. Skoro jednak różnią się oni mentalnie, to dlaczego nie uznać różnicy w naturze ich czasoprzestrzeni za tę różnicę fizyczną, która uzasadnia superweniencję?

Jest coś dialektycznie niezadowolającego w tym argumentacie. Poszukujemy języka form logicznych, który pozwoli nam wyrazić wszystko, co możemy wyrazić w potocznej polszczyźnie. Jednakże myśl, którą Lewis chce wyrazić – że ludzie mogą się różnić jedynie pod względem zewnętrznych własności fizycznych, jeśli znajdują się w różnych możliwych światach – nie jest sformułowana w potocznej polszczyźnie, lecz w języku właściwym teorii Lewisa. Nie jest więc oczywiste, czy jest to udany kontrprzykład wobec ekspresyjnej adekwatności pudełek i diamentów w potocznej polszczyźnie.

Potrzebne jest sformułowane w potocznej polszczyźnie określenie tego, co zostało pominięte przez formalizację tezy o superweniencji w kategoriach pudełek i diamentów. Formalizację tę można by rozwinąć w taki mniej więcej sposób:

\square dla każdego x , \square dla każdego y , \square (jeśli x i y różnią się pod pewnym względem mentalnym, to różnią się pod pewnym względem fizycznym).

Trudno mi dostrzec, dlaczego ta formalizacja nie wyraża adekwatnie rozważanego twierdzenia o superweniencji.

Istnieje wiele subtelnych zagadnień związanych z próbami porównania ekspresyjnych środków językiem pudełek i diamentów z jednej strony oraz języka kwantyfikującego po światach z drugiej. Celem tego paragrafu było je-

dynie zasygnalizowanie głównych kierunków dyskusji. (Zob. też SAINSBURY 1991: rozdz. 5.8 oraz podane tam odsyłacze do LEWIS 1986 i HAZEN 1976).

5.4. Kwantyfikatory (3) Teoria odpowiednika

Lewis łączy kwantyfikatorowe podejście do modalności z innym charakterystycznym, ale teoretycznie niezależnym, poglądem. Utrzymuje on, że nic nie istnieje w więcej niż jednym świecie. Na przykład rozważając, czy zdanie

Sokrates jest w sposób konieczny człowiekiem

jest prawdziwe, nie możemy pytać, czy „Sokrates jest człowiekiem” jest prawdziwe w każdym możliwym świecie, w którym Sokrates istnieje, ponieważ istnieje on wyłącznie w świecie aktualnym. Musimy natomiast spytać, czy każdy odpowiednik Sokratesa jest człowiekiem; i właśnie taka będzie logiczna forma omawianego zdania. Odpowiednikiem Sokratesa w danym świecie jest przedmiot, którym w tym świecie jest najbardziej do niego podobny; jak można by powiedzieć: przedmiot „maksymalnie podobny” do Sokratesa. Tak więc omawiane zdanie jest prawdziwe zawsze i tylko, gdy w każdym świecie cokolwiek jest maksymalnie podobne w tym świecie do Sokratesa z naszego świata, jest człowiekiem.

Przyjęcie przez Lewisa teorii odpowiednika wynika z przekonania, że identyczność międzyświatowa jest „niespójna” (*inconsistent*) (LEWIS 1986: 199; kluczowe rozważania są zawarte na s. 198-209). Zdaniem Lewisa za sprzeczne należy bowiem uznać przypuszczenie, że coś istnieje w dwóch światach posiadając w nich niespójne własności. Hubert Humphrey ma pięć palców u lewej ręki, ale mógłby mieć ich sześć. Jeśli fakt, że mógłby on posiadać sześć palców u lewej ręki przedstawimy jako należenie przez niego do pewnego alternatywnego świata, w którym ma on sześć palców, to otrzymamy:

Hubert Humphrey ma pięć palców u lewej ręki (w tym świecie) i Hubert Humphrey ma sześć palców u lewej ręki (w pewnym innym świecie).

Celem wtrącenia w nawiasach jest, oczywiście, próba uniknięcia grożącej sprzeczności, ale Lewis argumentuje, że próba ta kończy się niepowodzeniem, ponieważ żaden ze standardowych sposobów, w jakie modyfikatory mogą pozwolić na uniknięcie sprzeczności nie stosuje się do tego przypadku.

Dana wieża może być okrągła na trzecim piętrze i kwadratowa na czwartym. Jednakże niespójne predykaty stosują się tu do różnych części wieży,

podczas gdy w omawianym przypadku mówi się, że cały Humphrey należy do tego świata i zarazem do innego.

Humphrey może pozostawać w różnych relacjach do różnych indywiduów, będąc ojcem jednego, a synem drugiego. To jednak również nie stanowi dobrego modelu omawianego przypadku, ponieważ mowa jest nie o jego własnościach relacyjnych, lecz o własnościach wewnętrznych. Posiadanie pięciu palców jest własnością osoby jako takiej; nie jest to kwestia jego relacji do innych rzeczy. Nie możemy więc unikać tej sprzeczności, mówiąc, że chodzi tu o predykaty *pozostaje w relacji posiadania pięciu palców do tego świata* oraz *pozostaje w relacji posiadania sześciu palców do innego świata*.

Przestrzenny przykład wieży można z pożytkiem porównać z przypadkiem czasowym, powiedzmy tym:

Ten sklep jest otwarty (w dni pracujące) i nie jest otwarty (w dni wolne od pracy).

Lewis sądzi, że jedynym sposobem zrozumienia mechanizmu, który wyjaśnia wchodzącą tu w grę spójność jest potraktowanie przedmiotów trwających w czasie jako złożonych z części czasowych. Spójność zostaje osiągnięta dzięki temu, że uznajemy, iż początkowy fragment powyższego zdania mówi o niektórych z tych części (tych związanych z dniami pracującymi), że są one otwarte, a końcowy fragment mówi o innych częściach (tych związanych z dniami wolnymi), że są one zamknięte. Mechanizm ten nie może jednak wyjaśnić spójności przykładu Humphreya, ponieważ zakładamy, że cały Humphrey jest obecny w każdym z dwóch światów.

Lewis uważa, że problem ten nie dotyczy każdej teorii. Nie powstaje on w wypadku podejścia operatorowego, ponieważ jego zwolennik musi po prostu założyć spójność

p i \diamond nie- p .

Nawet w ramach teorii możliwych światów problem ten nie powstaje w wypadku ersacyzmu, który musi dopuścić, że z dwóch przedstawień jedno może reprezentować, że p , a drugie, że nie- p . Powstaje on natomiast w wypadku systemu Lewisa ze względu na jego realizm w kwestii światów możliwych. Realizm ów wymaga, aby rozumiał on problematyczne zdanie jako należące do tej samej grupy, co przypadki przestrzenne i czasowe. Choć większość ludzi uznałaby, że spójność w przypadkach przestrzennych (takich jak wieża) można wyjaśnić w kategoriach części przestrzennych, podobne wyjaśnienie w przypadkach czasowych jest bardziej kontrowersyjne. Gdyby można podać dobre wyjaśnienie spójności tych przypadków bez odwoływania się do

części czasowych, to mogłoby one objąć również przypadki modalne i można by w ten sposób zerwać związek między modalnym realizmem a niespójnością założenia, że niektóre indywidua znajdują się w różnych możliwych światach i mają w nich różne własności wewnętrzne. (Takie rozumienie czasowych modyfikacji sugeruje Lowe (1987)).

Niezależnie od tego, czy teoria odpowiednika znajduje adekwatne uzasadnienie w problemach dotyczących identyczności międzyświatowej, trzeba się zastanowić, czy podaje ona adekwatne formy logiczne zdań modalnych. Zaczniemy od niedobrego argumentu za tym, że tak nie jest (por. KRIPKE 1972: 344; PLANTINGA 1974: 115-16). Kiedy, na przykład, mówimy, że Sokrates mógłby być głupi, to chcemy przypisać pewną własność, własność bycia w sposób możliwy głupim, samemu Sokratesowi, a nie jego odpowiednikowi, choćby najbardziej podobnemu. Lewis odpowiada, zupełnie słusznie, że zarzut ten jest chybiony, ponieważ w teorii odpowiednika twierdzenie o Sokratesie pozostaje twierdzeniem o Sokratesie: jest to twierdzenie, że Sokrates posiada głupiego odpowiednika (LEWIS 1986).

A oto ciekawszy zarzut, ujawniający pewne własności teorii odpowiednika. Alvin Plantinga (1974: 109-110) wysuwa zarzut, że, intuicyjnie, Sokrates mógłby być bardzo podobny do aktualnego Ksenofonta, a Ksenofont mógłby być bardzo podobny do aktualnego Sokratesa. W terminologii Lewisa oznacza to, że powinien istnieć świat, w którym odpowiednik Sokratesa jest dokładnie taki jak aktualny Ksenofont, a odpowiednik Ksenofonta jest dokładnie taki jak aktualny Sokrates. To jednak wydaje się niemożliwe, ponieważ odpowiednik Sokratesa musi być bardziej podobny do aktualnego Sokratesa niż do kogokolwiek innego, podczas gdy w opisanym świecie to odpowiednik Ksenofonta najbardziej przypomina aktualnego Sokratesa.

W odpowiedzi Lewis odwołuje się do tezy, że istnieje wiele różnych względów podobieństwa i często o wyborze jednego z nich będzie decydował kontekst. W rozważanym wypadku własności jakościowe nie są istotne przy określaniu relacji odpowiednika właśnie dlatego, że zmiana własności jakościowych jest tym, co sobie przedstawiamy. Tak więc w tym wypadku o relacji bycia odpowiednikiem będą decydować takie czynniki jak pochodzenie (to, kim byli rodzice). Odpowiednik Sokratesa, oceniany na podstawie tego standardu podobieństwa, będzie kimś, kto ma tych samych rodziców co aktualny Sokrates; a to pozostawia wiele miejsca na różnice między aktualnym Sokratesem a jego odpowiednikiem pod nieskończenie wieloma względami jakościowymi, jak bycie brodatym i niebycie brodatym lub bycie mądrym i niebycie mądrym itd.; a zatem pozostawia miejsce na to, aby od-

powiednik Sokratesa był jakościowo (choć nie ze względu na rodziców) podobny do Ksenofonta.

Lewis wyraźnie zaznacza, że teoria odpowiednika nie uzasadnia konieczności identyczności (w jej rozumieniu przyjętym w części 4). Powodem jest to, że dany przedmiot może mieć więcej niż jeden odpowiednik w jakimś świecie. Przypuśćmy, że x posiada dwa odpowiedniki w w , x_1 i x_2 . Załóżmy następnie, że x jest identyczny z y w aktualnym świecie, tak że y również posiada dwóch odpowiedników w w . Jeśli warunkiem możliwości bycia różnym dla x i y jest to, że istnieje taki świat, w którym pewien odpowiednik x jest różny od pewnego odpowiednika y , to warunek ten jest spełniony przez to, iż x_1 jest różny od x_2 . Lewis (1986: rozdz. 4.5) podaje również argument, który ma dowieść, że nie kłóci się to z naszymi intuicjami.

Główną przeszkodą w przyjęciu teorii odpowiednika jest jej metafizyka: w wydaniu Lewisa, teoria ta wymaga wiary w rzeczywiste istnienie dużej liczby przyczynowo odizolowanych światów możliwych. Następny paragraf (5.5) określa ramy, w których można przedyskutować niektóre z tych zagadnień metafizycznych.

5.5. *Nieco metafizyki*

W tej części prześledzimy powiązania między problemami formy logicznej a problemami metafizycznymi.

Pogląd zdroworoządkowy przyjmuje, jak sądzę, następujące tezy dotyczące modalności. (Nie sugeruję, że jest to racja za uznaniem ich za prawdziwe.)

- (A) *Realizm*: niektóre zdania modalne, np. „Jest konieczne, że wszystko, co jest okrągłe, jest okrągłe” są obiektywnie prawdziwe; prawdziwe ze względu na to, jak się rzeczy mają, niezależnie od naszych umysłów i zdolności poznawczych.
- (B) *Aktualizm*: wszystko jest aktualne, nie istnieją przedmioty nieaktualne.

Poza tym pogląd zdroworoządkowy wyraża modalność bez wyraźnego kwantyfikowania po światach możliwych. Formy powierzchniowe są tu dość podobne do tych proponowanych przez podejście odwołujące się do pudełek i diamentów. Na pierwszy rzut oka współgra to dobrze z tezą (B): ponieważ, na pierwszy rzut oka, wszystkie światy możliwe poza światem aktualnym są przedmiotami nieaktualnymi, zwolennik aktualizmu sprzeciwi się kwantyfikowaniu po światach możliwych. Podobnie, skoro zwolennik aktualizmu unika możliwych światów na poziomie form logicznych, używając w zamian

pudełek i diamentów, to czymś niewłaściwym wyda mu się wprowadzenie ich do semantyki przy użyciu jakiegoś aksjomatu dla \square sformułowanego w kategoriach wszystkich możliwych światów.

To powiązanie poglądu zdroworozsądkowego z tezą (B) istnieje jednak tylko, jeśli traktuje się światy możliwe naiwnie – jako w większości istniejące, ale nie-aktualne byty. Widzieliśmy, że istnieje rozwiązanie konkurencyjne, ersacyzm. Teoria ta głosi, że to, co nazywamy nieaktualnym światem możliwym, jest przedmiotem aktualnym, ale takim, który w pewien sposób reprezentuje nieaktualny stan rzeczy. Świat możliwy jest jak obraz: obraz istnieje aktualnie, mimo że może on przedstawiać rzeczy w sposób odbiegający od aktualnego.

Zwolennik ersacyzmu może więc połączyć tezę (B) z takimi analizami formalnymi modalności, które kwantyfikują po światach możliwych. Nie istnieje bezpośredni związek między problemem formy logicznej a problemem metafizycznym. Mogą natomiast istnieć mniej bezpośrednie powiązania. Na przykład Lewis argumentował, że ersacyzm nie może oddać sprawiedliwości modalności, również z powodów czysto logicznych.

Sam Lewis określa siebie mianem „modalnego realisty”: odrzuca aktualizm i twierdzi, że jego światy możliwe są (z wyjątkiem świata aktualnego) nieaktualnymi przedmiotami zamieszkiwanymi przez nieaktualnych mieszkańców. Przyznaje on, że jest to stanowisko, które należy przyjąć tylko w ostateczności, ponieważ kłóci się ono ze zdrowym rozsądkiem. Niektóre racje przemawiające za nieaktualizmem dotyczą szczegółów technicznych. Nie możemy, na przykład, zadowolić się pudełkami i diamentami z powodu ich ekspresyjnej nieadekwatności i nie możemy zaakceptować ersacyzmu, ponieważ nie oddaje on sprawiedliwości naszym intuicjom modalnym. Istnieją jednak poza tym racje innego rodzaju.

Mówi się często, że scharakteryzowanie zwrotu „jest konieczne, że” przy użyciu „wszystkich możliwych światów” jest niedobre, ponieważ charakterystyka ta wykorzystuje pojęcie *możliwości*, a już wcześniej wiedzieliśmy, że pojęcia *konieczności* można definiować przez odwołanie się do pojęcia *możliwości* i odwrotnie. Uwaga ta, przynajmniej w obecnej postaci, opiera się na nieporozumieniu. Charakterystyki możliwości w kategoriach światów możliwych nie trzeba wcale traktować jako analizy. Być może jest ona po prostu sugestią dotyczącą tego, jakie rodzaje formy logicznej najlepiej oddają sprawiedliwość naszym idiomom modalnym.

Analiza modalności jest jednak czymś pożądanym i cechą teorii Lewisa, na którą chciałbym zwrócić uwagę, jest to, że dostarcza ona bodaj jedynej liczącej się analizy tego rodzaju. (Istnieją inne charakterystyki modalności, jednak wiele z nich (jak choćby projektywizm, zob. BLACKBURN 1986), re-

zygnują z podania takiej analizy. Konkurencyjną analizę podaje Armstrong (1989). Dzieje się tak dzięki temu, że Lewis definiuje możliwy świat w kategoriach nie-modalnych. Najogólniej mówiąc, idea jest taka, że świat możliwy składa się z czasoprzestrzennie powiązanych przedmiotów, gdzie dowolny przedmiot pozostający w dowolnej relacji czasoprzestrzennej do dowolnego innego przedmiotu w jakimś świecie, również należy do tego świata i nie należy do żadnego innego. Nie wspomina się tu ani słowem o modalności. Jeśli można podać redukcyjną analizę *możliwego świata* oraz logiczne formy polskich zdań modalnych przy użyciu kwantyfikatorów wiążących zmienne przebiegające możliwe światy, to można tym samym podać pełną redukcyjną charakterystykę modalności.

6. PRAWDA

6.1. Schemat Tarskiego i antynomia kłamcy

Zdanie jest prawdziwe zawsze i tylko, gdy jest tak, jak ono mówi; fałszywe zaś zawsze i tylko, gdy nie jest tak, jak ono mówi. Ta myśl wydaje się podstawowa dla naszego rozumienia prawdy. Od czasu Alfreda Tarskiego sugeruje się zwykle, że ową myśl można – przynajmniej w przybliżeniu – wyrazić w postaci następującego schematu:

„ p ” jest prawdziwe zawsze i tylko, gdy p

Schemat ten ma przedstawiać „odczyszławiający” charakter prawdy: powiedzenie o zdaniu, że jest ono prawdziwe (w tym celu musimy się odnieść do tego zdania, na przykład cytując je), jest równoważne wypowiedzeniu samego tego zdania (tutaj używamy go bez cudzysłowu).

Przebieg naszej dyskusji na temat prawdy wyznaczą dwa pytania dotyczące schematu Tarskiego: (1) Czy jest on poprawny? i (2) Czy zawiera on wszystko, co należy powiedzieć o prawdzie?

W tym paragrafie wskażę dwie podważalne (a nawet, jak zasugeruję, podważone) racje, aby sądzić, że ów schemat jest niepoprawny. W kolejnych paragrafach rozważę, jak ewentualnie należałoby go uzupełnić, aby uzyskać adekwatną teorię prawdy.

Zasada dwuwartościowości głosi, że wszystko, do czego można sensownie zastosować prawdę lub fałsz, jest albo prawdziwe, albo fałszywe. Nieostrość może stanowić rację dla odrzucenia tej zasady. Rozważmy zdanie „To jest czerwone”. W wypadku niektórych przedmiotów jest ono prawdziwe,

a w wypadku innych – fałszywe, a zatem jest ono czymś, do czego można sensownie zastosować prawdę lub fałsz. Istnieją jednak pewne przedmioty – przypadki graniczne – co do których nie jesteśmy skłonni powiedzieć, ani że to zdanie jest o nich prawdziwe, ani że jest o nich fałszywe. Ową niechęć można by wyjaśnić tym, że to zdanie nie jest ani prawdziwe, ani fałszywe. (Czy przychodzi ci na myśl jakieś inne wyjaśnienie?) Oznaczałoby to odrzucenie zasady dwuwartościowości.

To odrzucenie byłoby niespójne z przyjęciem schematu Tarskiego (tak jak był on rozumiany do tej pory). Jeśli bowiem „ σ ” jest zdaniem, które nie jest ani prawdziwe, ani fałszywe, to

„ σ ” jest prawdziwe

jest fałszywe, podczas gdy (na mocy hipotezy) samo σ takie nie jest. A zatem podstawienie schematu Tarskiego

„ σ ” jest prawdziwe zawsze i tylko, gdy σ

posiada jeden składnik fałszywy i drugi taki, który nie jest ani prawdziwy, ani fałszywy. Taką równoważność trudno byłoby uznać za prawdziwą. Stanowi ona więc wyjątek od schematu Tarskiego.

Jedną z odpowiedzi byłoby powstrzymanie się od odrzucenia zasady dwuwartościowości. (Należałoby rozważyć po kolei wszystkie możliwe racje za jej odrzuceniem.) Jednak w naszych obecnych rozważaniach moglibyśmy udzielić jeszcze innej odpowiedzi: moglibyśmy zastrzec, że przyjęcie tego schematu sprowadza się wyłącznie do przekonania, że nie posiada on fałszywego podstawienia, gdzie fałszywe podstawienie jest rozumiane jako takie, w którym jeden składnik równoważności jest prawdziwy, a drugi fałszywy.

Inną racją dla uznania schematu Tarskiego za niepoprawny jest to, iż gra on rolę w zaskakującym argumencie związanym z antynomią kłamcy. Rozważmy następujące zdanie K („Kłamca”):

K nie jest prawdziwe.

K próbuje powiedzieć o sobie, że nie jest prawdziwe.

W schemacie Tarskiego wolno nam zastąpić schematyczną literę „ p ” dowolnym zdaniem, do którego możemy sensownie zastosować pojęcie prawdy (patrzac na lewy człon schematu) lub które może się sensownie pojawić jako składnik równoważności (patrzac na prawy człon schematu). Nie oczekiwaliśmy więc, że schemat Tarskiego będzie spełniony, gdy zastąpimy „ p ” zdaniem rozkazującym (np. „Strzelaj!”) lub pozbawionym sensu ciągiem słów. Jednakże K wydaje się być zdaniem, któremu można sensownie przypisać

prawdziwość i które może być składnikiem równoważności. Wydaje się zatem, że dopuszczalne powinno być zastąpienie w naszym schemacie „*p*” przez *K*, tak aby uzyskać:

„*K* nie jest prawdziwe” jest prawdziwe zawsze i tylko, gdy nie jest prawdziwe.

Biorąc pod uwagę sposób, w jaki wprowadziliśmy *K*, wiemy że

$K = \text{„}K \text{ nie jest prawdziwe”}$.

Na mocy zasady Leibniza, która głosi, że to, co identyczne, można zastępować z zachowaniem prawdziwości, możemy zastąpić pierwsze wystąpienie „*K* nie jest prawdziwe” w schemacie przez „*K*”, uzyskując w ten sposób:

K jest prawdziwe zawsze i tylko, gdy *K* nie jest prawdziwe.

A to jest sprzeczność.

Jedno z rozwiązań polega na zanegowaniu poprawności, a przynajmniej uniwersalnego zastosowania, schematu Tarskiego. Zgodnie z tym rozwiązaniem uznaje się, że *K* jest zdaniem, do którego można sensownie zastosować pojęcie prawdy lub które może wystąpić jako człon równoważności, ale zaprzecza się, że ten schemat stosuje się do każdego zdania, które spełnia te warunki. Tak więc antynomia kłamcy stanowi powód do odrzucenia schematu Tarskiego.

Istnieją jednak jeszcze inne odpowiedzi na ową antynomię, pozwalające zachować schemat Tarskiego. Jedną z nich, stosunkowo mało popularną, jest zaprzeczenie, że w takich kontekstach obowiązuje zasada Leibniza (zob. SKYRMS 1982). Znacznie większą popularnością cieszy się pogląd, że – z tych czy innych względów – *K* nie jest wyrażeniem, do którego ten schemat ma zastosowanie. Ta odpowiedź wymaga wyjaśnienia, dlaczego, wbrew pozorom, *K* jest zdaniem, do którego nie stosuje się sensownie pojęcie prawdy, oraz zdaniem, które nie może w sposób właściwy wystąpić jako człon równoważności. (Dlaczego trzeba wyjaśnić obie te własności *K*? Dlaczego nie wystarczyłoby wyjaśnienie jednej z nich?) Jedną z sugestii idących w tym kierunku jest taka, że coś jest nie w porządku z pojawiającym się w *K* samoodniesieniem. Mogłoby się wydawać, że możliwe powinno być określenie wszystkich własności zdania, które są konieczne do jego zrozumienia, bez konieczności presuponowania samego tego zdania. Wydaje się jednak, że w wypadku *K* nie da się tego zrobić. Określając zdanie *K*, musimy, już w pierwszym słowie *K*, odnieść się do tego właśnie zdania, które próbujemy

określić. Taką sugestię wysunął jako pierwszy Russell (1908) w swojej „zasadzie błędnego koła”, a później podjęto kilka prób doprecyzowania tej ogólnej idei (na przykład: KRIPKE 1975; MARTIN 1984: wprowadzenie)

Odpowiedź samego Tarskiego trudno zaklasyfikować w nakreślonych tutaj ramach. Uważał on, że antynomia kłamcy dowodzi, iż nasze potoczne pojęcie prawdy jest z natury paradoksalne. Zadaniem filozofii (i matematyki) jest skonstruowanie możliwie najwierniejszego, unikającego tej antynomii, zastępnika tego potocznego pojęcia. Tarski uważał, że podstawowym problem stanowi założenie, że może istnieć pojęcie prawdy, które jest prawomocnie stosowalne do zdania zawierającego samo to pojęcie. Aby temu zapobiec, zaproponował on hierarchiczne uporządkowanie języków. Podstawowy język, powiedzmy L_0 , nie zawierałby w ogóle pojęcia prawdy. Następny język, L_1 , zawierałby każde zdanie L_0 wraz z predykatem prawdy, nazwijmy go „prawdziwe₁”, zdefiniowanym wyłącznie dla zdań L_0 . Kluczową sprawą jest to, że predykatu „prawdziwe₁” nie można stosować do zdania zawierającego ten predykat; to jest: nie można go stosować do żadnego zdania L_1 , o ile to zdanie nie należy do L_0 . Hierarchia ta idzie w górę. Prawda dla zdań L_1 jest wyrażalna za pomocą predykatu „prawdziwe₂”, który należy do L_2 itd.

Zdania K nie można wyrazić w żadnym języku należącym do hierarchii Tarskiego. Przypuśćmy (na zasadzie *reductio ad absurdum*), że K należy do pewnego języka L_n . Wówczas predykat prawdy, który ono zawiera, jeśli ma być stosowalny do samego K, musi należeć do pewnego języka wyższego rzędu w hierarchii, minimalnie do L_{n+1} . W pełnym sformułowaniu zdanie K miałoby więc następującą postać:

K nie jest prawdziwe_{n+1}.

Ale wówczas K nie może należeć do L_n . *QED*.

Pogląd Tarskiego był taki, że każde *spójne* (*coherent*) wystąpienie słowa „prawdziwe” można zastąpić przez predykat prawdy z indeksem: „prawdziwe₁”, „prawdziwe₂” itp. Według Tarskiego, na tyle na ile schemat jest zrozumiały, na tyle też jest poprawny. Jeśli wydaje się, że prowadzi on do antynomii, to tylko dlatego, że nie została zachowana odpowiednia hierarchia, a w takim wypadku schemat od początku nie był zrozumiały. W ujęciu Tarskiego omawiany schemat nie jest pomyślany jako twierdzenie o naszym potocznym pojęciu prawdy, lecz raczej jako kryterium spójności pojęcia prawdy.

Tarski uważał, że aby osiągnąć spójność, musimy odejść od potocznego języka i jego z natury paradoksalnego predykatu prawdy. Czy jest jakiś mniej

radykalny wariant tej teorii? Czy można twierdzić, że coś w rodzaju hierarchii Tarskiego jest już *implicite* obecne w naszym języku, ratując go w ten sposób przed niespójnością? Beznadziejne wydaje się przypuszczenie, że istnieją takie z góry ustalone poziomy, których wymaga hierarchia Tarskiego. Warto jednak zastanowić się, czy taka hierarchia nie wyłania się w naszym języku okazjonalnie – to jest: czy nie pojawia się na różne sposoby w zależności od kontekstu. (Zob. BURGE 1979.)

Z punktu widzenia naszych obecnych celów najważniejsze są dwa wnioski: po pierwsze, pojęcie prawdy jest uwikłane w głębokie paradoksy i adekwatne ujęcie tego pojęcia powinno uwzględnić ten fakt; po drugie, rozwiązanie tych paradoksów nie wymaga, jak się wydaje, odrzucenia schematu Tarskiego. W następnym paragrafie zakładam, że schemat ten stanowił część każdego ujęcia prawdy.

6.2. Wzbogacenie schematu Tarskiego

Dysponując schematem Tarskiego, co *jeszcze* trzeba powiedzieć o prawdzie? Jedną możliwą odpowiedzią jest taka: nic (lub prawie nic). W pierwszych dwóch częściach zostaną podane pewne warianty tej odpowiedzi. W pozostałych częściach zostaną zaproponowane inne uzupełnienia.

6.2.1. Minimalizm

Czy moglibyśmy posłużyć się schematem Tarskiego do *zdefiniowania* prawdy? Pogląd, który możemy luźno nazwać „minimalizmem”, jest taki, że schemat ten mówi wszystko, co należy wiedzieć: „prawdziwe” znaczy cokolwiek to wyrażenie musi znaczyć, aby we właściwy sposób wypełnić lukę w następującym schemacie:

„*p*” jest ____ zawsze i tylko, gdy *p*.

Pierwszą wątpliwość, jaka się nasuwa wobec tego poglądu, dotyczy jego wyrażonej kołowości, ponieważ „wypełnienie w sposób właściwy luki” w powyższym schemacie to wypełnienie jej w taki sposób, że jakiegokolwiek zdanie (odpowiedniego rodzaju) wstawi się za „*p*”, podstawienie to będzie *prawdziwe*. Wątpliwość tę można częściowo rozwiązać, przechodząc do mniej formalnej wersji stanowiska minimalistycznego. Przypuśćmy, że ktoś mówi językiem, który nie zawiera predykatu „prawdziwe” ani jego odpowiednika. Moglibyśmy zapoznać go z tym predykatem w następujący sposób: zawsze,

gdy wypowiada on coś twierdząco, na przykład „*p*”, moglibyśmy mówić: równie dobrze mógłbyś powiedzieć „*«p» jest prawdziwe*”. Posługujemy się tu pojęciem równoważności; nie posługujemy się jednak – przynajmniej nie wprost – pojęciem prawdy. Można by uznać, że posługujemy się podstawieniami schematu, aby wyjaśnić prawdę, nie zakładając z góry samej prawdy.

Przejdę teraz do trzech poważniejszych wątpliwości: (1) czy coś innego mogłoby wypełnić lukę; (2) jak określić, które zdania nadają się do bycia prawdziwymi; i (3) czy powiedziano wystarczająco wiele, aby wyjaśnić pojęcie prawdy, a nie jedynie określić jego zakres.

1. Oto lista możliwych kandydatów: wyrażień, które, jak można utrzymywać, nadają się do zastąpienia w tej luce słowa „prawdziwe”, mimo że mają inne znaczenie:

 poznawalne,
 odpowiadające faktom,
 korzystne do uznania,
 w dłuższej perspektywie czasu przyjęte przez wszystkich bezstronnych badaczy.

Minimaliści muszą rozważyć kolejno wszystkie te możliwości. Mają oni do dyspozycji dwie strategie: jedną jest zakwestionowanie różnicy znaczeń, drugą – zaprzeczenie, że dany kandydat nadaje się do zastąpienia słowa „prawdziwe” w omawianym schemacie. Pierwsza strategia mogłaby poskutkować w wypadku zwrotu „odpowiadające faktom”. Druga mogłaby się sprawdzić w wypadku wyrażenia „w dłuższej perspektywie czasu przyjęte przez wszystkich bezstronnych badaczy”, na przykład dlatego, że możemy sobie wyobrazić zdania, które rozumiemy i które są prawdziwe, lecz których prawdziwości nie ustalą nigdy żadne badania.

2. Odpowiedź na pierwsze pytanie nie zajdzie daleko, jeśli nie poruszy się drugiego pytania, o to, jak należy określić odpowiednie podstawienia „*p*” w schemacie. Widzieliśmy, że odpadają zdania nonsensowne oraz zdania w trybie rozkazującym. Jednakże w dziejach filozofii twierdzono, o bardzo różnych zdaniach, że wbrew wszelkim pozorom nie można do nich sensownie stosować pojęcia prawdy. Słynnym przykładem jest dyskurs moralny. Twierdzono, że do zdań w rodzaju „Torturowanie niewinnych jest złem” nie można sensownie stosować prawdy ani fałszu. (Takie ponoć było stanowisko Davida Hume’a, 1738: ks. 3). Wersja stanowiska, które byłoby w ogólnym zarysie Hume’owskie, mogłaby wykorzystać następujące przesłanki: (1) zdanie może wyrażać przekonanie zawsze i tylko, gdy stosuje się do niego sensownie pojęcie prawdy; (2) motywacja każdego działania musi obejmować

coś, co nie jest przekonaniem (mianowicie pragnienie); (3) pogląd moralny, wraz z przekonaniem, jest wystarczający jako motywacja działania. Wynika z tego, że pogląd moralny nie jest przekonaniem, a zatem że jego wyraz nie czymś, do czego sensownie stosuje się prawda lub fałsz.) W tak zwanej „emotywistycznej” szkole etyki upodobniono je do wyrażania odczuć, takich jak „Tfu!”, „Fe!” (dla tortur) i „Ach!”, „Hurra!” (dla wyrażania aprobaty). (Zob. np. AYER 1936: rozdz. 6.)

pozytywna charakterystyka mogłaby głosić, że zdanie jest odpowiednim podstawieniem „p” w schemacie Tarskiego, tylko jeśli reprezentuje niezależną rzeczywistość, której może ono odpowiadać lub nie. Pytanie o to, czy zdanie nadaje się do bycia prawdziwym, będzie istotnym pytaniem metafizycznym, którego nie można zbyć odwołując się do tak powierzchownych własności zdania jak jego składnia czy jak nasza nieskażona filozofią skłonność do używania w odniesieniu do tego zdania słów „prawdziwe” i „fałszywe”.

Poprawność tego poglądu podważyłaby minimalizm w kwestii prawdy. Albowiem gdyby nawet można było wyjaśnić, za pośrednictwem schematu Tarskiego, na czym polega prawdziwość zdania nadającego się do bycia prawdziwym, wyjaśnienie, na czym polega nadawanie się do bycia prawdziwym, wymagałoby zaawansowanej filozofii. Nie możemy używać tego schematu bez wyraźnego określenia, do których zdań należy go stosować, a to może nas uwikłać w kontrowersyjną metafizykę.

3. Czy zwolennik minimalizmu w wystarczającym stopniu ukazał nam rolę prawdy w życiu i języku? Oto dwie racje za odpowiedzią przeczącą.

(a) Mogłoby się wydawać, że zwolennik minimalizmu określił co najwyższej, dla każdego zdania, na czym polega jego prawdziwość. Pomiął natomiast doniosłość bycia prawdziwym dla zdania – nie określił, jakie są następstwa prawdziwości dla przekonania, twierdzenia czy wniosku. Na podstawie samego schematu Tarskiego nie moglibyśmy chyba orzec, że następujące stwierdzenia są banałami (*platitudes*): powinniśmy dążyć do posiadania prawdziwych przekonań; osoba mówiąca prawdziwie mówi, jak się rzeczy mają; stwierdzić to przedstawić jako prawdę; prawdy nie mogą pozostawać w konflikcie itd. Dopóki nie uchwyci się znaczenia pojęcia prawdy dla bycia przekonaniem, mówienia i stwierdzania, nie zrozumie się jego w pełni.

(b) Ponieważ różne własności mogą się stosować do tych samych rzeczy (np. własność posiadania serca i własność posiadania nerek), możemy wiedzieć, jak używać słowa „prawdziwe” (stosujcie go tylko do tych zdań, które skłonni jesteście szczerze uznać), nie wiedząc zarazem, jaką własność to słowo wyraża, a zatem nie rozumiejąc go w pełni.

W odpowiedzi na ten drugi zarzut zwolennik minimalizmu mógłby wprowadzić pojęcie *minimalnego* (*the least*) znaczenia: słowo „prawdziwe” ma najmniejsze ze znaczeń, umożliwiającego weryfikację schematu Tarskiego. Nawet gdyby uchyliło to drugi zarzut, pierwszy nadal nie doczekał się odpowiedzi. Musimy ustalić (w dwóch następnych paragrafach), jak wzbogacić schemat Tarskiego tak, aby lepiej zrozumieć prawdę. Najpierw jednak rozważę pogląd, którego celem jest skierowanie nas w inną stronę.

6.2.2. Redundancja

Namysł nad schematem Tarskiego mógłby sugerować, że predykat „prawdziwe” jest zbędny, i dałoby się go bez straty wyeliminować: zamiast orzekać prawdę o zdaniu, moglibyśmy po prostu stwierdzić to zdanie. Jest to wersja tak zwanej „redundancyjnej teorii” prawdy.

Teoria ta jest jawnie niepoprawna w podanej tu formie, ponieważ możemy orzekać prawdę nawet w sytuacji, gdy nie jesteśmy w stanie stwierdzić zdania, o której ją orzekamy. Wiedząc, na przykład, że ktoś jest stuprocentowo uczciwy i dobrze poinformowany, ale nie wiedząc, co powiedział podczas wystąpienia w radio, moglibyśmy powiedzieć: „Wszystko, co on powiedział w radio, było prawdziwe”. Ponieważ nie wiemy, co ten ktoś powiedział, nie możemy użyć jego zdań; potrzebujemy predykatu „prawdziwe”.

Standardowa odpowiedź obrońców teorii redundancyjnej głosi, że nie trzeba wcale używać predykatu „prawdziwe”, ponieważ to, co chcemy powiedzieć, można wyrazić następująco:

Dla każdego p , jeśli wypowiedział on p podczas wystąpienia w radio, to p .

Jak należy rozumieć tę kwantyfikację? Być może zmienna „ p ” przebiega zdania; inaczej mówiąc, zajmuje ona takie miejsce, jakie mogłaby zajmować nazwa zdania. (Porównajmy: w „dla każdego x , jeśli x jest człowiekiem, to x jest śmiertelny” zmienna przebiega przedmioty. A zatem można by ją zastąpić nazwą jakiegoś przedmiotu, powiedzmy nazwą „Sokrates”.) Nada to sens wystąpieniu „ p ” w „On wypowiedział p ”, ponieważ faktycznie stwierdzamy tu (w ogólny sposób) relację między osobą mówiącą a zdaniem, do którego się odnosimy. Nie nada to jednak sensu ostatniemu wystąpieniu tej zmiennej, zajmowane przez nią miejsce musi bowiem zająć zdanie, a nie nazwa zdania. (Czy sytuacja polepszyłaby się, gdybyśmy zastąpili „wypowiedział p ” przez „powiedział, że p ” i zinterpretowali kwantyfikację jako wiążącą zmienne przebiegające sądy – przedmioty, które zdania stwierdzają?)

Zwolennik teorii redundancyjnej zaproponuje, byśmy interpretowali kwantyfikatory jako *podstawieniowy*. (Zob. SAINSBURY 1991: rozdz. 4.18.) Wyjaśniamy to mówiąc, że przedstawiona formuła jest równoważna twierdzeniu, iż niezależnie od tego, jakim zdaniem zastąpimy „*p*” w

Jeśli on wypowiedział *p* podczas wystąpienia w radio, to *p*,
rezultat jest prawdziwy. Może to nadać sens kwantyfikacji (choć nawet to jest sporne), ale z pewnością nie służy to celom redundancyjnej teorii prawdy, jako że kwantyfikacja zostaje wyjaśniona w sposób, który istotnie korzysta z rzekomo redundantnego predykatu „prawdziwe”.

6.2.3. Korespondencja

Zdanie jest prawdziwe tylko wtedy, gdy odpowiada faktom. Na pierwszy rzut oka wydawałoby się, że twierdzenie to jest bezsporne. Pozostaje jednak pytanie, czy mówi ono coś istotnego o prawdzie; czy służy ono jako dobre uzupełnienie schematu Tarskiego. „Korespondencyjna teoria” prawdy udziela na to pytanie twierdzącej odpowiedzi.

Wspomniane twierdzenie stało się przedmiotem słynnego sporu między Johnem Austinem a Peterem Strawsonem. Strawson zaprzeczył, że teorii korespondencyjnej udało się wyjaśnić prawdę jako relację między zdaniem a czymś innym, mianowicie faktami. Wymagałby to bowiem podania takiego ujęcia faktów, które jest niezależne od pojęcia prawdy; jednak „fakty są tym, co twierdzenia (jeśli są prawdziwe) stwierdzają [...] Gdyby usunąć twierdzenia ze świata, to znikłyby również fakty; sam świat nie stałby się jednak przez to ani trochę uboższy” (STRAWSON 1950b: 38-39).

Formułując szczegóły swej teorii prawdy, Austin (1950) podał teoretyczne ujęcie tego, na czym miałyby polegać korespondencja. (Stąd Strawson był chyba trochę niesprawiedliwy, twierdząc, że Austin nie określił natury faktów w sposób nieodwołujący się do prawdy. Zob. AUSTIN 1961. Austin przykłada dużą wagę do rozróżnienia między zdaniem a twierdzeniem, tu jednak pomijam to rozróżnienie.) Twierdził, że zdanie jest powiązane z dwoma odmiennymi rodzajami rzeczy w świecie, przez dwa różne zbiory konwencji. Konwencje opisu łączą je ze stanem rzeczy – to jest czymś, co może, ale nie musi zachodzić. Konwencje wskazywania łączą je z określoną sytuacją. A zatem (by posłużyć się przykładem Austina) wypowiedzenie zdania „Kot jest na macie” jest powiązane przez konwencje opisu ze stanem rzeczy polegającym na byciu kota na macie, a przez konwencje wskazywania –

z określoną sytuacją w pewnej części świata. (Austin, co podkreśla Strawson, wyraża się dość niejasno na temat tego, o jaką konkretną sytuację chodzi.) Wypowiedzenie tego zdania jest prawdziwe tylko wtedy, gdy ta określona sytuacja jest tego samego rodzaju, co ten stan rzeczy (to jest egzemplifikuje go). Przez wiele lat od czasu ogłoszenia tej teorii uważano powszechnie, że została ona obalona przez Strawsona. (Strawson postawił zarzut, że Austin nadał prawdzie charakter czysto konwencjonalny, ale wydaje się, że jest to nieporozumienie.) Jednak ostatnimi czasy teoria Austina ożyła na nowo za sprawą Barwise'a i Etchemendy'ego (1987), którzy twierdzą, że pozwala ona uniknąć antynomii kłamcy.

Jak dotąd zakładaliśmy, że problemem poglądu traktującego prawdę jako korespondencję jest jego banalność. Musimy teraz osłabić to założenie: uważanie prawdy za korespondencję traktuje się często jako oznakę najróżniejszych plag, w rodzaju empiryzmu, realizmu i atomizmu. Jeśli jest to zasadne, to powinniśmy przestać myśleć o doktrynie korespondencji jako o banale. Jej poprawność stałaby się zależna od poprawności empiryzmu, realizmu i atomizmu.

Jeśli obrońcy teorii korespondencyjnej twierdzą, że wiedzieć to być biernym odbiorcą skutków działania zewnętrznej rzeczywistości – faktów – to są oni zobowiązani do (nieco komiksowej wersji) empiryzmu. Ale ten pogląd na wiedzę jest czymś dodatkowym; korespondencyjna teoria prawdy go nie pociąga. Jej zwolennik może się skłaniać ku racjonalistycznej epistemologii.

Jeśli obrońcy teorii korespondencyjnej przyjmują realistyczny pogląd na temat faktów – istnieją one niezależnie od umysłu i mogą wykraczać poza nasze zdolności poznawcze – to są oni zobowiązani do realizmu. Ale ten pogląd jest czymś dodatkowym; korespondencyjna teoria prawdy go nie pociąga. Nawet jeśli traktuje się fakty jako w jakiś sposób zależne od ludzkich zdolności poznawczych, to prawda nadal mogłaby polegać na korespondencji z nimi.

Jeśli obrońcy teorii korespondencyjnej traktują korespondencję jako rodzaj izomorfizmu, w którym jakaś część świata koresponduje z każdą znaczącą częścią prawdziwego zdania, to są oni zobowiązani do jakiejś formy atomizmu (np. RUSSELL 1918-1919). Jednakże korespondencja jako taka nie wymaga tak dokładnego odwzorowania.

To prawda, że ci, którzy uznają pogląd korespondencyjny za atrakcyjny i odkrywczy, mogą być skłonni do przyjęcia empiryzmu, realizmu i atomizmu (osobno lub w połączeniu). Historia uczy, że taka tendencja rzeczywiście istnieje. Ściśle jednak biorąc, teoria korespondencyjna w podanym tu sformułowaniu (zdanie jest prawdziwe zawsze i tylko, gdy odpowiada ono fak-

tom) nie pociąga tych dodatkowych poglądów. Z tego względu nie trafiają w nią argumenty wymierzone w te poglądy.

Mimo wszystko można ją uznać za banał. Dodając ją do schematu Tarskiego, nie wzbogacamy w istotny sposób naszego ujęcia prawdy.

6.2.4. Konwergencja

Zarozumiałstwem byłoby utrzymywać, że, w wypadku każdego aspektu rzeczywistości, posiadamy lub moglibyśmy uzyskać środki pojęciowe konieczne do jego adekwatnego opisu. Zarozumiałstwem byłoby więc utrzymywać, że każdy aspekt rzeczywistości jest poznawalny. Znacznie jednak mniejszym zarozumiałstwem byłoby wysunięcie sugestii – jak niektórzy sądzą trafnej – że moglibyśmy w zasadzie poznać wszystkie te aspekty rzeczywistości, które możemy ująć pojęciowo. Oznaczałoby to, że moglibyśmy w zasadzie poznać prawdziwość lub fałszywość każdego sensownego zdania.

Tę linię rozumowania, obecnie określaną na ogół jako „semantyczny antyrealizm”, rozwinął Michael Dummett (w wielu pismach, szczególnie lubię 1973c i wprowadzenie do 1978). Pogląd przeciwny, czyli (semantyczny) realizm, głosi, że możemy sformułować sensowne zdania, które są prawdziwe lub fałszywe, lecz o których prawdziwości lub fałszywości możemy się nigdy nie dowiedzieć. W myśl tego poglądu prawda, jak to się niekiedy ujmuje, „przekracza świadectwo”.

Przy wszechstronnym ujęciu zagadnienia prawdy należałoby poświęcić nieco miejsca tej kwestii (wiąże się ono z innym poruszonym wcześniej zagadnieniem. Realista ma powód, aby przyjąć zasadę dwuwartościowości, Dummett uznał brak takiego powodu za cechę rozpoznawczą antyrealizmu; zob. np. DUMMETT 1973c i 1978: przedmowa.), przejdę jednak teraz do innych, choć powiązanych pytań: Czy istnieje tendencja, aby najlepsza opinia zbliżała się do prawdy? Czy moglibyśmy nawet zdefiniować prawdę jako to, ku czemu najlepsza opinia się zbliża? Twierdząca odpowiedź na przynajmniej pierwsze z tych pytań byłaby atrakcyjna dla semantycznego antyrealisty. Jeśli zdanie jest prawdziwe, to jest poznawalne, a zatem zastosowanie odpowiednich metod badania powinno ostatecznie doprowadzić do jego odkrycia; a zatem konwergencja.

Za ideą konwergencji kryje się jeszcze inna myśl. Prawda jest ściśle powiązana z obiektywnością. Jeśli jesteśmy realistami, obiektywność prawdy zostaje zapewniona przez obiektywność przedmiotu dyskursu, czegoś, co istnieje niezależnie od tego wszystkiego, co wiąże się z nami, w tym zwłaszcza od

naszych metod badawczych. Jeśli jednak nie jesteśmy realistami, to pozostajemy odcięci od tego źródła obiektywności. Co moglibyśmy powiedzieć, aby zapewnić, że mimo wszystko zajmujemy się autentycznymi faktami, a nie tylko igraszkami wyobraźni? Konwergencja to jedna z możliwych odpowiedzi.

Zbyt proste byłoby przypuszczenie, że istnieje nieunikniona tendencja do konwergencji. Główna myśl jest taka, że to samo świadectwo na „wejściu” będzie mieć tendencję do prowadzenia nas do tych samych poglądów na „wyjściu”. Nie oczekiwaliśmy konwergencji między ludźmi, którzy mają do czynienia z całkowicie różnymi świadectwami. A zatem, po poprawce, moglibyśmy powiedzieć, że brak konwergencji wymaga wyjaśnienia w kategoriach odmiennych danych wyjściowych lub defektów mechanizmu poznawczego. Po tej poprawce uzyskujemy pogląd, który mogliby równie dobrze zaakceptować realisci. Szkopuł w tym, że w tym rozumieniu konwergencja nie wydaje się raczej odpowiednią podstawą do *zdefiniowania* prawdy, nawet z punktu widzenia semantycznego antyrealisty. Ujęcie to w sposób istotny wykorzystuje pojęcie mechanizmu *poznawczego*. Bez tego zastrzeżenia z pewnością nie mielibyśmy pojęcia nadającego się do zdefiniowania prawdy, ponieważ nawet rozbieżności w kwestiach, w których nie pojawia się problem prawdy, takich jak reakcje uczuciowe, również muszą być wyjaśnialne w kategoriach różnicy w danych wyjściowych lub różnicy mechanizmu poznawczego. Jednakże po wprowadzeniu tego zastrzeżenia, mamy obowiązek wyjaśnić, czym jest ów *poznawczy* mechanizm, a trudno pojąć, jak można by wywiązać się z tego zadania nie odwołując się do pojęcia prawdy. Mechanizm poznawczy to po prostu mechanizm, który dostarcza takich danych wyjściowych, do których stosuje się pojęcie prawdy (a wobec tego również pojęcie przekonania, uzasadnienia itd.).

Być może sama próba *zdefiniowania* prawdy jest chybiona. Powinniśmy poprzestać na ujęciu, które – wsparte schematem Tarskiego – określa właściwą relację prawdy do twierdzenia, przekonania, uzasadnienia (*warrant*), interpretacji i obiektywności. (Zob. też: HAACK 1978; WIGGINS 1980b).

KOLEJNE TEMATY

1. Nie zostały dokładnie omówione (a niektóre nie zostały nawet wspomniane) rozmaite tematy dotyczące konieczności, na przykład:

(a) Rozróżnienie między koniecznością *de re* a koniecznością *de dicto*. Jedno z pytań brzmi: „Czy to rozróżnienie można adekwatnie wyjaśnić w kategoriach zasięgu?”. SAINSBURY 1991 (rozdz. 5.4) wprowadzi was w ten temat.

(b) Zarzuty wobec samej idei konieczności, zwłaszcza konieczności *de re*. Chodzi tu przede wszystkim o argumenty Quine'a i o argument pochodzący od Fregego. SAINSBURY 1991 (rozdz. 5.5 i 5.6) wprowadzi was w ten temat.

(c) Nierzeczywiste okresy warunkowe, w tym słynne analizy takich okresów w kategoriach światów możliwych. Zapoznanie się z nimi dałoby wam pojęcie o użyteczności teorii światów możliwych. SAINSBURY 1991 (rozdz. 5.2) może służyć jako wprowadzenie, ale nie można pominąć pięknego artykułu LEWIS 1973.

(d) Językowa analiza konieczności. Stary, ale niedościgniony opis tej teorii (która od jakiegoś czasu wyszła z mody) znajduje się w PAP 1958.

2. Istnieje grupa tematów ściśle związanych z dziedziną logiki. Czym jest stała logiczna? Pytanie brzmi: co sprawia, że wyrażenie („i”, „nie”, „wszystkie”) jest słowem logicznym czy stałą logiczną? Czy \square również powinno się uznać za stałą logiczną? A co z symbolem (ϵ) dla teoriomnogościowego bycia elementem? Istnieje tu ścisły związek z pytaniami dotyczącymi tego, co sprawia, że jakaś prawda jest prawdą logiczną. SAINSBURY 1991 (rozdz. 6.5) zawiera wstępny przegląd niektórych głównych stanowisk.

3. Przedmiotem filozoficznego namysłu są również kluczowe pojęcia logiki: niezawodność (*validity*) i wynikanie (*entailment*). Czy można podać doskonalsze pojęcie niezawodności niż to w kategoriach zachowywania prawdy? Motywem pytania jest fakt, że wnioskowanie z niespójnymi przesłankami są logicznie poprawne w świetle potocznego standardu zachowywania prawdy, mają one jednak małą wartość poznawczą. Analogiczne zagadnienia dotyczą wynikania. To zagadnienie zostało podjęte w sławnych artykułach Geacha (1958, 1970).

4. Epistemologia logiki, a zwłaszcza uprawomocnienie dedukcji. Jak można uprawomocnić to, co uważa się za prawa logiki? Zob. DUMMETT 1973b oraz 1991; w wypadku drugiej z tych prac warto posłużyć się zawartym w indeksie hasłem „uprawomocnienie praw logiki”.

*Z języka angielskiego przełożyli
Marcin Iwanicki i Joanna Odrowąż-Sypniewska*

Dr MARCIN IWANICKI – Katedra Historii Filozofii Nowożytnej i Współczesnej na Wydziale Filozofii Katolickiego Uniwersytetu Lubelskiego Jana Pawła II; adres do korespondencji: Al. Racławickie 14, 20-950 Lublin; e-mail: miwanick@nd.edu

Dr hab. JOANNA ODROWĄŻ-SYPNIEWSKA – Zakład Semiotyki Logicznej, Instytut Filozofii na Wydziale Filozofii i Socjologii Uniwersytetu Warszawskiego; adres do korespondencji – e-mail: j.odrowaz@uw.edu.pl

BIBLIOGRAFIA

- AMSTRONG D.: *A Combinatorial Theory of Possibility*, New York 1989.
- AUSTIN J. L.: *Truth*, „Proceedings of the Aristotelian Society“ 24(1950), s. 111-128 [*Prawda*, przeł. B. Chwedeńczuk, [w:] TENŻE, *Mówienie i poznawanie. Rozprawy i wykłady filozoficzne*, Warszawa: PWN 1993, s. 156-178].
- *Unfair to Facts*, [w:] *Philosophical Papers*, Oxford 1961 [*Niewiernie wobec faktów*, przeł. B. Chwedeńczuk, [w:] TENŻE, *Mówienie i poznawanie. Rozprawy i wykłady filozoficzne*, Warszawa: PWN 1993, s. 206-233].
- AYER A. J.: *Language, Truth and Logic*, London 1936.
- BARCAN MARCUS R.: *The Identity of Individuals in a Strict Functional Calculus of Second Order*, „Journal of Symbolic Logic” 12 (1947), s. 12-15.
- BARWISE J., ETCHEMENDY J.: *The Liar: An Essay in Truth and Circularity*, New York 1987.
- BLACKBURN S.: *Morals and Modals*, [w:] G. MACDONALD, C. WRIGHT (red.), *Fact, Science and Morality*, Oxford 1986.
- BURGE T.: *Semantical Paradox*, „Journal of Philosophy” 76 (1979), s. 169-198.
- CHELLAS B.: *Modal Logic: An Introduction*, Cambridge 1980.
- DAVIES M.: *Meaning, Quantification, Necessity: Themes in Philosophical Logic*, London 1980.
- DONNELLAN K.: *Reference and Definite Descriptions*, „Philosophical Review” 77 (1966), s. 203-215 [W języku polskim dostępny jest inny tekst tego autora: *Nazwy własne i deskrypcje identyfikujące*, przeł. T. Szubka, [w:] B. STANOSZ (red.), *Filozofia języka*, Warszawa: Fundacja ALETHEIA-Wydawnictwo SPACJA 1993, s. 203-225].
- DUMMETT M.: *Frege: Philosophy of Language*, London 1973.
- *The Justification of Deduction*, „Proceedings of the British Academy” 59 (1973), s. 201-232.
- *The Philosophical Basis of Intuitionistic Logic*, [w:] H. E. ROSE, J. C. SHEPHERDSON (red.), *Logic Colloquium 1973*, North-Holland 1975. Przedruk w: DUMMETT 1978.
- *Truth and Other Enigmas*, London 1978.
- *The Logical Basis of Metaphysics*, London 1991 [*Logiczna podstawa metafizyki*, przeł. W. Sady, Warszawa: PWN 1998].
- EVANS G.: *The Varieties of Reference*, Oxford 1982 [W języku polskim dostępny jest jeden artykuł tego autora: *Przyczynowa teoria nazw*, przeł. T. Szubka, [w:] B. STANOSZ (red.), *Filozofia języka*, Warszawa: Fundacja ALETHEIA-Wydawnictwo SPACJA 1993, s. 226-245].
- FORBES G.: *The Metaphysics of Modality*, Oxford 1985.
- FREGE G.: *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, 1879. [W języku polskim dostępny jest fragment pierwszego rozdziału tej pracy: *Sąd*, przeł. B. Wolniewicz, [w:] TENŻE, *Pisma semantyczne*, Warszawa: PWN 1977, s. 3-9].
- *Function and Concept*, 1891 [Funkcja i pojęcie, przeł. B. Wolniewicz, [w:] TENŻE, *Pisma semantyczne*, Warszawa: PWN 1977, s. 18-44].
- *On Sense and Meaning*, 1892 [*Sens i odniesienie*, przeł. B. Wolniewicz, [w:] TENŻE, *Pisma semantyczne*, Warszawa: PWN 1977, s. 60-88].
- *On Concept and Object*, „Vierteljahrsschrift für wissenschaftliche Philosophie” 16 (1892), s. 192-205 [*Pojęcie i przedmiot*, przeł. B. Wolniewicz, [w:] TENŻE, *Pisma semantyczne*, Warszawa: PWN 1977, s. 45-59].
- *Der Gedanke: eine logische Untersuchung*, „Beiträge zur Philosophie des deutschen Idealismus” 1 (1918). S. 58-77 [*Mysł. Studium logiczne*, przeł. B. Wolniewicz, [w:] TENŻE, *Pisma semantyczne*, Warszawa: PWN 1977, s. 101-129].

- GEACH P.: *Entailment*, „Proceedings of the Aristotelian Society” 32 (1958), s. 157-172.
— *Entailment Again*, „Philosophical Review” 79 (1970), s. 237-239.
- GIBBARD A.: *Contingent Identity*, „Journal of Philosophical Logic” 4 (1975), s. 187-221.
- GRICE H. P.: *The Causal Theory of Perception*, „Proceedings of the Aristotelian Society” 35 (1961), s. 121-154.
— *Logic and Conversation*, [w:] P. COLE, J. L. MORGAN (red.), *Syntax and Semantics*, New York 1975.
- GUTTENPLAN S. D.: *The Languages of Logic*, Oxford 1988.
- HAACK S.: *Philosophy of Logics*, Cambridge 1978. [W języku polskim dostępne są dwa rozdziały tej pracy w przekładzie A. Sierszulskiej: *Logika modalna i Logika wielowartościowa*, [w:] J. WOLEŃSKI (red.), *Filozofia logiki*, Warszawa: Fundacja ALETHEIA-Wydawnictwo SPACJA 1997, s. 183-235].
- HAZEN A.: *Expressive Incompleteness in Modal Logic*, „Journal of Philosophical Logic” 76 (1976), s. 25-46.
- HODGES W.: *Logic*, London 1978.
- HORWICH P.: *Truth*, Oxford 1990.
- HUME D.: *A Treatise of Human Nature*, London 1738 [*Traktat o naturze ludzkiej*, t. 1-2, przeł. Cz. Znamierowski, Warszawa: PWN 1963].
- KRIPKE S.: *Identity and Necessity*, [w:] M. K. MUNITZ (red.), *Identity and Individuation*, New York 1971 [*Identyczność a konieczność*, przeł. T. Szubka, [w:] T. SZUBKA (red.), *Metafizyka w filozofii analitycznej*, Lublin: TN KUL 1995, s. 95-126].
— *Naming and Necessity*, [w:] D. DAVIDSON, G. HARMAN (red.), *Semantics of Natural Language*, Dordrecht 1972 [*Nazywanie a konieczność*, przeł. B. Chwedeńczuk, Warszawa: PAX 1988].
— *Outline of a Theory of Truth*, „Journal of Philosophy” 72 (1975), s. 690-716 [*Zarys pewnej teorii prawdy*, przeł. P. Garbacz, „Kwartalnik Filozoficzny” 29 (2001), s. 97-131].
— *Speaker's Reference and Semantic Reference*, [w:] P. A. FRENCH, T. E. UEHLING, H. K. WETTSTEIN (red.), *Midwest Studies in Philosophy*, Minneapolis 1977.
— *A Puzzle about Belief*, [w:] A. MARGALIT (red.), *Meaning and Use*, Dordrecht 1979.
- LEMMON E. J.: *Beginning Logic*, London 1964.
- LEWIS C. I., LANGFORD C. H.: *Symbolic Logic*, London 1932.
- LEWIS D.: *Counterfactuals*, Oxford 1973. [W języku polskim dostępny jest fragment tej pracy: *Światy możliwe*, przeł. U. Żegleń, [w:] T. SZUBKA (red.), *Metafizyka w filozofii analitycznej*, Lublin: TN KUL 1995, s. 127-135].
— *On the Plurality of Worlds*, Oxford 1986. [W języku polskim dostępny jest wykład Lewisa nawiązujący do tematyki poruszonej w tej pracy: *Możliwości: konkretne światy czy abstrakcyjne obiekty proste?*, przeł. U. Żegleń, [w:] T. SZUBKA (red.), *Metafizyka w filozofii analitycznej*, Lublin: TN KUL 1995, s. 155-176].
- LOWE E. J.: *Lewis on Perdurantism versus Endurantism*, „Analysis” 47 (1987), s. 152-154.
- MCCULLOCH G.: *The Game of the Name: Introducing Logic, Language and Mind*, Oxford 1989.
- MCDOWELL J.: *On the Sense and Reference of a Proper Name*, „Mind” 86 (1977), s. 159-185.
- MARTIN R. L. (red.): *Recent Essays on Truth and the Liar Paradox*, Oxford 1984.
- MEINONG A.: *The Theory of Objects*, [w:] *Untersuchungen zur Gegenstandstheorie und Psychologie*, Leipzig 1904 [*O teorii przedmiotu*, przeł. T. Lubowiecki i J. Mizera, „Principia” 8-9 (1994), s. 171-212].
- MILL J. S.: *A System of Logic*, 1843 [*System logiki*, t. 1-2, przeł. Cz. Znamierowski, Warszawa: PWN 1962].
- MOORE A. W. (red.): *Meaning and Reference*, Oxford: Oxford Readings in Philosophy 1993.

- MOORE G. E.: *Is Existence a Predicate?*, „Proceedings of the Aristotelian Society” 15 (1936), s. 157-188.
- NEALE S.: *Descriptions*, Cambridge, Mass. 1990.
- PAKULUK M.: *The Interpretation of Russell's Elogy Argument*, [w:] A. IRVINE, G. WEDEKIND (red.), *Russell and Analytic Philosophy*, Vancouver 1993.
- PAP A.: *Semantics and Necessary Truth*, New Haven, Conn. 1958.
- PEACOCKE C. A. B.: *Proper Names, Reference and Rigid Designation*, [w:] S. BLACKBURN (red.), *Meaning, Reference, Necessity: New Studies in Semantics*, Cambridge 1973.
- PITCHER G. (red.): *Truth*, Englewood Cliffs, NJ 1964.
- PLANTINGA A.: *The Nature of Necessity*, Oxford 1974.
- RAMSEY F. P.: *Facts and Propositions*, „Proceedings of the Aristotelian Society”. Supplement 7 (1927), s. 153-170. Przedruk w: *The Foundations of Mathematics and Other Logical Essays*, ed. R. B. Braithwaite, London 1931, s. 138-155; *Philosophical Papers*, ed. D. H. Mellor, Cambridge 1990, s. 34-51.
- RUSSELL B.: *On Denoting*, „Mind” 14 (1905), s. 479-493 [*O denotowaniu*, przeł. J. Pelc, [w:] J. PELC (red.), *Logika i język*, Warszawa: PWN 1967, s. 254-275].
- *Mathematical Logic as Based on the Theory of Types*, „American Journal of Mathematics” 30 (1908), s. 222-262.
- *Knowledge by Acquaintance and Knowledge by Description*, „Proceedings of the Aristotelian Society” 11 (1911), s. 108-128.
- *The Problems of Philosophy*, ed. R. C. Marsh, London 1912 [*Problemy filozofii*, przeł. W. Sady, Warszawa: PWN 2004].
- *Lectures on the Philosophy of Logical Atomism*, „Monist” 28 (1918-1919), s. 29.
- *Introduction to Mathematical Philosophy*, London 1919 [*Wstęp do filozofii matematyki*, przeł. Cz. Znamierowski, Warszawa: ALETHEIA 2003].
- *Logic and Knowledge*, London 1956.
- RUSSELL B., WHITEHEAD A. N.: *Principia Mathematica*, Cambridge 1910-1913.
- SAINSBURY R. M.: *Russell*, London 1979.
- *Gareth Evans: The Varieties of Reference*, „Mind” 94 (1985), s. 120-142.
- *Logical Forms: An Introduction to Philosophical Logic*, Oxford 1991.
- *Russell on Names and Communication*, [w:] A. IRVINE, G. WEDEKIND (red.), *Russell and Analytic Philosophy*, Vancouver 1993.
- *Paradoxes*, Cambridge 1995.
- SALMON N.: *Frege's Puzzle*, Cambridge, Mass. 1986.
- SKYRMS B.: *Intensional Aspects of Semantical Self-Reference*, 1982.
- STRAWSON P. F.: *On Referring*, „Mind” 59 (1950a), s. 269-286 [*O odnoszeniu się użycia wyrażień do przedmiotów*, przeł. J. Pelc, [w:] J. PELC (red.), *Logika i język*, Warszawa: PWN 1967, s. 377-413].
- *Truth*, „Proceedings of the Aristotelian Society” 24 (1950b), s. 129-156. [W języku polskim dostępny jest inny artykuł tego autora na temat prawdy: *Prawda*, przeł. J. Pelc, [w:] J. PELC (red.), *Logika i język*, Warszawa: PWN 1967, s. 353-375].
- *Truth: A Reconsideration of Austin's Views*, „Philosophical Quartely” 15 (1965), s. 289-301.
- *Is Existence Never a Predicate?*, „Critica” 1 (1967), s. 5-15.
- TARSKI A.: *The Concept of Truth in Formalized Language*, 1937 [*Pojęcie prawdy w językach nauk dedukcyjnych*, przeł. J. Zygmunt, [w:] TENŻE, *Pisma logiczno-filozoficzne*, t. 1, Warszawa: PWN, s. 13-172].

- *Truth and Proof*, „Scientific American” 194 (1969), s. 63-77 [*Prawda i dowód*, przeł. J. Zygmunt, [w:] TENŹE, *Pisma logiczno-filozoficzne*, t. 1, Warszawa: PWN, s. 292-332].
- WIGGINS D.: *Frege's Problem of the Morning Star and the Evening Star*, [w:] M. SCHIRN (red.), *Studies on Frege*, Stuttgart 1976.
- *Sameness and Substance*, Oxford 1980.
- *What Would Be a Substantial Theory of Truth?*, [w:] Z. VAN STRAATEN (red.), *Philosophical Subjects: Essays Presented to P. F. Strawson*, Oxford 1980.
- WRIGHT C.: *Truth and Objectivity*, Cambridge, Mass. 1992. [W języku polskim dostępny jest artykuł tego autora poświęcony pojęciu prawdy: *Prawda: przegląd tradycyjnego sporu*, przeł. T. Szubka, „Kwartalnik Filozoficzny” 28 (2000), s. 153-198].

Słowa kluczowe: nazwy, deskrypcje, istnienie, konieczność, prawda.

Key words: names, descriptions, existence, necessity, truth.

Information about Author: Prof. MARK R. SAINSBURY – Department of Philosophy, University of Texas at Austin; address for correspondence: 1 University Station, Austin, Texas 78713; e-mail: marksainsbury@mail.utexas.edu

Information about Translators:

MARCIN IWANICKI, Ph.D. – Chair of the History of Modern and Contemporary Philosophy, Faculty of Philosophy, The John Paul II Catholic University of Lublin; address for correspondence: Al. Raławickie 14, PL 20-950 Lublin; e-mail: miwanick@nd.edu

JOANNA ODROWĄŻ-SYPNIEWSKA, Ph.D. – Assistant Professor, Section of Logical Semiotics, Faculty of Philosophy and Sociology, University of Warsaw; address for correspondence – e-mail: j.odrowaz@uw.edu.pl