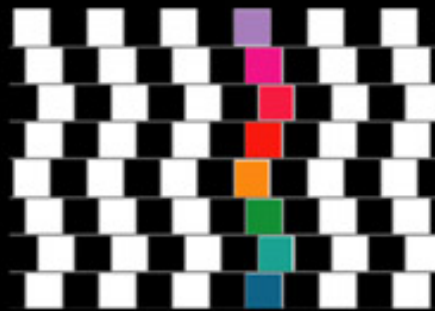


filozofia  
wobec  
nauk  
szczegółowych



zbiór prac  
pod redakcją  
Tadeusza Ciecierskiego,  
Lecha M. Nijakowskiego  
Jakuba Szymanika

# **FILOZOFIA I NAUKI SZCZEGÓŁOWE**

Pod redakcją  
Tadeusza Ciecierskiego  
Lecha M. Nijakowskiego  
Jakuba Szymanika

Warszawa 2002

Recenzenci:

dr Krzysztof Wójtowicz (Filozofia a matematyka)

prof. dr hab. Barbara Stanosz (Filozofia a teoria języka)

dr Bohdan Chwedeńczuk (Filozofia a biologia)

dr hab. Robert Piłat (Filozofia a teoria umysłu)

prof. dr hab. Stanisław Siekierski (Filozofia a nauki społeczne)

Wydawca: Koło Filozoficzne przy MISH

Uniwersytet Warszawski

Kolegium Międzywydziałowych Indywidualnych Studiów Humanistycznych

Nowy Świat 69

00-927 Warszawa

Projekt okładki: Piotr Labenz

Korekta: Barbara Ciecierska

Książka ukazała się dzięki dotacji Rady Konsultacyjnej ds. Studenckiego Ruchu Naukowego przy JM Rektorze Uniwersytetu Warszawskiego oraz dzięki dotacji Fundacji „Instytut Artes Liberales” ze środków przeznaczonych na wspieranie Kolegium MISH UW



ISBN 83 – 915323 – 1 - 3

Spis treści

Wstęp

## **FILOZOFIA A MATEMATYKA**

Leszek Aleksander Kołodziejczyk, O „redukowaniu filozofii do matematyki”. Przykład Tezy Churcha

Piotr Labenz, O pozorności sporu realizm vs formalizm

Krzysztof Wójtowicz, Czy spór realizm vs formalizm jest sporem pozornym?

## **FILOZOFIA A TEORIA JĘZYKA**

Jakub Szymanik, Problemy formalnej teorii języka naturalnego

Tomasz Puczyłowski, Zdania Moore’a w świetle teorii implikatury

Witold Kieraś, Teza o niezdeterminowaniu przekładu jako przykład Quine’owskiej koncepcji lingwistyki

## **FILOZOFIA A BIOLOGIA**

Jakub Jakubowski, Próby przewyciężenia filozoficznego paradoksu Empedoklesa w protobiologii

## **FILOZOFIA A TEORIA UMYŚLU**

Marcin Zajenkowski, Kilka uwag o poznawaniu cudzych stanów psychicznych

Marcin Gokieli, Searle i jego gwoździe

Marcin Miłkowski, Czy istnieje granica między kognitywistyką a filozofią umysłu?

## **FILOZOFIA A NAUKI SPOŁECZNE**

Sebastian Szymański, Czy współcześnie można być etnologicznym ewolucjonistą?

Lech M. Nijakowski, Analogia jako źródło innowacji w naukach społecznych i filozofii

Maciej M. Czapnik, Realizm i konstruktywizm w historii a język historycznego dyskursu

Marta Bucholc, Użyteczność pojęcia „gra językowa” dla nauk społecznych

## Wstęp

Terminy „filozofia” i „nauka” można rozumieć różnie. W najszerszym z możliwych znaczeń filozofia jest to dyscyplina obejmująca wszystkie teoretyczne dociekania człowieka. Jako taka łączy w sobie elementy i cechy każdej z nauk. Natomiast filozofia pojmowana jako dyscyplina zajmująca się np. zagadnieniami etycznymi nie ma nic wspólnego z naukami. W innych jeszcze rozumieniach niektóre zagadnienia filozoficzne wiążą się w ten lub inny sposób z nauką, niektóre zaś nie mają z nauką nic wspólnego. Z drugiej strony termin „nauka” również nie należy do najjaśniejszych – według niektórych obejmuje on tylko dyscypliny nazywane popularnie przyrodniczymi, według innych – również nauki formalne lub humanistykę. Sprawa skomplikuje się jeszcze bardziej, gdy zauważymy, że na uczelniach istnieją instytucje zajmujące się programowo tzw. badaniami interdyscyplinarnymi- np. ogólną teorią umysłu.

Niejasności te nie mają oczywiście nic wspólnego z ogólną kondycją nauki lub filozofii - ta pierwsza miewa się znakomicie, ta druga- dobrze. Nie jest celem ani tego wstępu, ani tej książki podanie definicji tego, czym jest filozofia lub nauka. Przyjmujemy arbitralnie, że zakres nazwy „nauka” występującej w tytule książki jest szeroki i obejmuje obok dociekań przyrodniczych: matematykę, nauki społeczne, kognitywistykę i teorię języka. Jeżeli chodzi o związki tak pojętej nauki z filozofia, to autorzy artykułów zastanawiają się albo nad filozoficznym znaczeniem pewnych rezultatów naukowych, albo nad filozoficzną interpretacją pewnych dziedzin nauki (np. nad ich ontologicznymi zobowiązaniami), albo nad użytecznością pewnych pojęć lub teorii filozoficznych w nauce. Redaktorzy pozostawili autorom swobodę wyboru jednej z tych perspektyw, zdając sobie sprawę z tego, że narzucanie którejkolwiek z nich miałooby niekorzystny wpływ na jakość tekstów. To pobieżne wyjaśnienie tytułu nie może zadowolić Czytelnika, wypada wyrazić więc nadzieje, że wszystkie wątpliwości rozwieje dalsza lektura.

Praca ta nie ukazałaby się, gdyby nie życzliwość wielu osób. Ogromne podziękowania za wnikliwą i życzliwą zarazem krytykę należą się naszym Recenzentom: dr Bohdanowi Chwedeńczukowi, doc. Robertowi Piłatowi, prof. Stanisławowi Siekierskiemu, prof. Barbarze Stanosz oraz dr Krzysztofowi Wójtowiczowi. Bardzo wiele zawdzięcza ta książka naszym

przyjaciołom : Romanowi Chymkowskiemu, który bardzo nam pomógł w redagowaniu działu nauk społecznych oraz Piotrowi Labenzowi, którego zasługą jest graficzna część książki. Obu im bardzo dziękujemy.

Pragniemy podziękować również Radzie Konsultacyjnej ds. Studenckiego ruchu Naukowego przy JM Rektorze Uniwersytetu Warszawskiego oraz Fundacji „Instytutu *Artes Liberales*” za sfinansowanie publikacji.

Redaktorzy

# **FILOZOFIA A MATEMATYKA**

## O „redukowaniu filozofii do matematyki” – Przykład Tezy Churcha

### *1. Wprowadzenie*

Matematyka jest dla filozofii ważna przynajmniej dwojako: jako przedmiot badań i jako narzędzie badawcze.

Matematyka jest przedmiotem badań wyspecjalizowanego działu filozofii, zwanego – jak nietrudno odgadnąć – „filozofią matematyki”. Filozofia matematyki zajmuje się między innymi: naturą obiektów, o których mowa w matematyce (albo tym, czy w matematyce w ogóle jest mowa o jakichś obiektach); charakterem rozumowań matematycznych i dopuszczalnością niektórych stosowanych w matematyce rozumowań; cechami odróżniającymi twierdzenia matematyki od twierdzeń innych nauk. Są to zagadnienia istotne, nie tylko dla filozofii matematyki, ale i dla całej filozofii. Jest tak choćby dlatego, że matematyka to nauka nader specyficzna: z jednej strony, wyraźnie odmienna od pozostałych pod względem metody i treści, z drugiej zaś – stosowana we wszystkich niemal dziedzinach wiedzy, i to w taki sposób, że stopień zaawansowania teoretycznego danej dziedziny można by wręcz mierzyć stopniem jej zmatematyzowania. Właśnie owa specyfika matematyki czyniła ją i czyni tak ważnym przedmiotem rozważań filozoficznych – począwszy od pitagorejczyków i Platona aż po współczesność.

Matematyka jest jednak dla filozofii ważna także jako narzędzie. Filozof stosujący matematykę jako narzędzie działa w przybliżeniu tak: napotykając jakiś problem filozoficzny ( $F$ ), który opiera się próbom rozstrzygnięcia (albo wydaje się niejasno sformułowany), stara się go „zredukować” do pewnego problemu matematycznego ( $M$ ). „Redukcja” taka polegać winna na uzasadnieniu, że  $M$  jest (w jakimś sensie) równoważne  $F$  albo, że niektóre rozstrzygnięcia  $M$  implikują pewne rozstrzygnięcia  $F$ . Oczywiście, skutek „redukcji”, to co w matematyce wiadomo na temat  $M$ , można odnieść również do  $F$ . Nieraz zachodzi także sytuacja poniekąd odwrotna: najpierw pojawiają się pewne wyniki matematyczne, których



„filozoficzna doniosłość” jest łatwo rozpoznawalna, a dopiero potem próbuje się podać ich filozoficzną interpretację, tj. wysnuć z nich filozoficzne wnioski.

Idea wykorzystywania wyników matematyki – i szerzej: wyników nauki – w celu rozjaśnienia czy nawet rozstrzygnięcia problemów filozoficznych jest znana od dawna, a w XX wieku stała się szczególnie popularna. Klasycznym przykładem wprowadzania tej idei w życie mogą być prace Hansa Reichenbacha, zwłaszcza jego *Powstanie filozofii naukowej* (Reichenbach [1954]). Reichenbach, jak wiadomo, uważał, że osiągnięcia nauki dostarczają odpowiedzi na niektóre spośród tradycyjnych pytań filozofii: powstanie geometrii nieeuklidesowych obala kantowską koncepcję przestrzeni, teoria względności – kantowską koncepcję czasu, mechanika kwantowa – zasadę przyczynowości itd. Przykładem z dziejów filozofii polskiej jest tzw. „metoda parafraz” Kazimierza Ajdukiewicza (por. Ajdukiewicz [1934], [1937], [1948]), który m. in. skrytykował tezę idealizmu transcendentalnego, „redukując” ją do pewnej tezy metamatematycznej – jak się okazało, fałszywej.

Pomysł nadawania zagadnieniom filozoficznym interpretacji naukowej jest ciekawy i, jak się wydaje, owocny, budzi wszelako czasem pewne zastrzeżenia. Zastrzeżenia te bywają niekiedy całkiem zasadnicze: przykładowo, zwolennicy fenomenologicznego programu filozofii bezzakołoniowej w ogóle nie akceptują odwoływania się filozofii do wyników badań naukowych<sup>1</sup>. Idee fenomenologów można odrzucić jako utopijne, pozostanie jednak problem następujący: jest jasne, że tezy filozoficzne nie są *sensu stricto* (tj. w sensie konsekwencji logicznej) równoważne twierdzeniom naukowym ani z nich nie wynikają. Skoro tak, to należy jakoś uzasadnić prawomocność ich „redukcji” do twierdzeń naukowych – czyli w szczególności wyjaśnić, na czym polega owa „równoważność” tez filozoficznych i twierdzeń naukowych bądź też- „wynikanie” jednych z drugich.

Jan Woleński zaproponował (por. Woleński [1993], s. 10-17), by mówić, że pewne tezy filozoficzne są (mogą być) „konsekwencjami interpretacyjnymi” twierdzeń naukowych. Zasugerował także, że „konsekwencja interpretacyjna” jest to w istocie „entymematyczna konsekwencja logiczna”, czyli konsekwencja logiczna ze względu na dodatkowe przesłanki. By przywołać przykład podany przez Woleńskiego: filozoficzne zdanie „nasze poznanie jest niepewne” jest logiczną konsekwencją zdania „niesprzeczność formalnej arytmetyki pierwszego rzędu nie jest w niej dowodliwa” (czyli luźno sformułowanego drugiego

---

<sup>1</sup> Fenomenolog, uznający za naukę również filozofię, mówiłby tu zapewne raczej o wynikach „nauk szczegółowych”. My jednak filozofii za naukę nie uważamy, a ponadto pisząc o „wynikach badań naukowych”, mamy na myśli między innymi, a nawet przede wszystkim, wyniki matematyki, która „nauką szczegółową” nie jest; trudno o mniej „szczegółową” naukę niż matematyka.

twierdzenia Gödla) oraz dwu dodatkowych przesłanek: (1) „dowodliwość niesprzeczności naszego poznania w nim samym jest warunkiem koniecznym jego pewności” i (2) „formalna arytmetyka pierwszego rzędu należy do naszego poznania”<sup>2</sup>.

Jeśli ujęcie Woleńskiego jest trafne (przyjmujemy, że jest), to prawomocność „redukcji” tez filozoficznych do twierdzeń naukowych sprowadza się do prawdziwości (słuszności?) owych dodatkowych przesłanek, które są potrzebne do wyprowadzania konsekwencji interpretacyjnych. Rodzą się więc pytania: skąd wziąć takie przesłanki? jak je uzasadniać? czy mogą one w ogóle (a jeśli tak, to które, kiedy?) być w ścisłym słowa znaczeniu prawdziwe?

Zwolennicy filozoficznej interpretacji wyników naukowych byli czasem świadomi wagi tych pytań i próbowali na nie odpowiadać, jednak ich odpowiedzi były nieraz zdawkowe i niezadowolające. Dla przykładu, wspomniany już Ajdukiewicz rozważał (Ajdukiewicz [1934], s. 214) dwie ewentualne metody uzasadniania swoich „parafraz” (czyli w naszym ujęciu: uzasadniania dodatkowych przesłanek potrzebnych do wyprowadzania konsekwencji interpretacyjnych z twierdzeń nauki): pierwszą jest „analiza znaczeniowa wyrazów mowy potocznej”, drugą – „podnoszenie ich [parafraz – L.A.K.] do rzędu postulatów, które nie troszcząc się o znaczenie, jakie wyrazy posiadały w mowie potocznej, nadawałyby im znaczenia w sposób arbitralny”. Jak łatwo zauważyć, stosowanie drugiej metody (rygorystycznie rozumianej, tj. w pełni arbitralnej) zupełnie uniemożliwiłoby „redukcję” zagadnień filozoficznych do naukowych. Metoda pierwsza z kolei ma, jak się wydaje, bardzo ograniczony zakres stosowalności.

Można zatem uznać, że zagadnienie statusu przesłanek potrzebnych do wyprowadzania konsekwencji interpretacyjnych (a co za tym idzie, zagadnienie prawomocności „redukowania” problemów filozoficznych do naukowych) pozostaje otwarte. Warto zagadnienie to rozważyć na przykładzie takich przesłanek, które z jednej strony są

---

<sup>2</sup> Marcin Mostowski zwrócił nam uwagę, że przykład Woleńskiego nie jest w pełni udany. W obecnej postaci rozumowanie jest formalnie poprawne tylko, wtedy gdy przesłankę (1) rozumiemy jako „dowodliwość niesprzeczności każdego fragmentu naszego poznania w obrębie samego tego fragmentu jest koniecznym warunkiem pewności naszego poznania”. W takim jednak ujęciu przesłanka (1) staje się absurdalna, co czyni cały przykład mało interesującym. Należałoby więc jakoś zmodyfikować przesłanki, np. zmienić (1) na „dowodliwość niesprzeczności wszystkich odpowiednio dużych fragmentów naszego poznania w nich samych jest koniecznym warunkiem pewności naszego poznania”, a (2) – na „formalna arytmetyka pierwszego rzędu jest (odpowiednio dużym) fragmentem naszego poznania”.

filozoficznie istotne, z drugiej – mają szczególnie prostą postać. W naszym przekonaniu, do tego rodzaju przesłanek należy tzw. „teza Churcha”.

## 2. Teza Churcha

Intuicyjne pojęcie algorytmu, czyli efektywnej metody, pozwalającej w skończonej liczbie z góry przepisanych kroków rozstrzygnąć dane zagadnienie, było matematykom znane (choć nie pod nazwą „algorytm”) już w starożytności. Świadczy o tym istnienie tzw. algorytmu Euklidesa, służącego do znajdowania największego wspólnego dzielnika dwóch liczb. Z biegiem czasu poszukiwanie algorytmów rozstrzygających rozmaite problemy matematyczne stawało się zadaniem coraz istotniejszym. Warto przypomnieć np. dziesiąty spośród słynnych dwudziestu trzech problemów sformułowanych na kongresie matematyków w Paryżu (1900 r.) przez Dawida Hilberta: należało znaleźć metodę (tj. algorytm), za której pomocą można by w skończonej liczbie kroków rozstrzygnąć, czy dane równanie diofantyczne<sup>3</sup> ma rozwiązanie w liczbach całkowitych.

Z pojęciem algorytmu wiążą się blisko dwa dalsze: pojęcie problemu rozstrzygalnego i pojęcie funkcji obliczalnej. Pojęcia te odnoszą się w zasadzie do – odpowiednio – problemów dotyczących liczb naturalnych (problemy takie interpretujemy jako  $n$ -argumentowe relacje nad zbiorem liczb naturalnych – gdzie  $n$  jest liczbą naturalną<sup>4</sup>) oraz funkcji (być może wieloargumentowych) z liczb naturalnych w liczby naturalne<sup>5</sup>. Problem jest rozstrzygalny, jeśli istnieje algorytm, który go rozstrzyga; funkcja jest obliczalna, jeśli istnieje algorytm, który, mając dane argumenty tej funkcji, oblicza jej wartość dla tych argumentów. Oba te ważne pojęcia pozostawały przez długi czas niedoprecyzowane: podobnie jak pojęciu algorytmu, brak im było ścisłej, matematycznej charakteryzacji.

---

<sup>3</sup> Równanie diofantyczne – to równanie postaci  $f(x_1, \dots, x_n) = 0$ , gdzie  $f(x_1, \dots, x_n)$  jest wielomianem od  $n$  zmiennych o współczynnikach całkowitych.

<sup>4</sup> Zasadę tej „interpretacji” można zobrazować następującymi dwoma przykładami: problem, czy dana liczba naturalna dzieli jakąś inną daną liczbę, interpretujemy jako relację dwuargumentową  $\{(a,b) \in \omega^2 : a \text{ dzieli } b\}$ ; problem, czy dana liczba jest pierwsza – jako relację jednoargumentową  $\{a \in \omega : a \text{ jest liczbą pierwszą}\}$ .

<sup>5</sup> „W zasadzie”, gdyż, jak wiadomo, dzięki odpowiednim technikom kodowania wiele zagadnień dotyczących *prima facie* obiektów innych niż liczby naturalne można traktować jako zagadnienia dotyczące liczb naturalnych.

Sytuacja ta zmieniła się w latach trzydziestych ubiegłego wieku. Narodziło się wówczas pojęcie – tym razem już ściśle zdefiniowane – które miało wkrótce zostać uznane za matematyczną wersję pojęcia obliczalności (i rozstrzygalności): *rekurencyjność*.

Istnieje wiele równoważnych definicji klasy funkcji rekurencyjnych. Definicje te mają bardzo zróżnicowany charakter: jedne odwołują się do abstrakcyjnych maszyn (maszyny Turinga, maszyny RAM), inne – do pewnych systemów sformalizowanych ( $\lambda$ -definiowalność, obliczalność za pomocą systemu równości Herbrand-Gödla), jeszcze inne – do zamkniętości na pewne operacje na funkcjach. Jedną z takich definicji przytaczamy w *Dodatku*. Inne można znaleźć w podręcznikach teorii rekursji<sup>6</sup>.

Stosunkowo szybko zorientowano się, że pojęcie funkcji rekurencyjnej dość dobrze odpowiada intuicjom związanym z pojęciem funkcji obliczalnej (a pojęcie relacji rekurencyjnej, czyli takiej, której funkcja charakterystyczna jest rekurencyjna – intuicjom związanym z pojęciem problemu rozstrzygalnego). W połowie lat trzydziestych Alonzo Church, a po nim (niezależnie) Alan M. Turing i (niezależnie od Turinga, ale już nie od Churcha) Emil Post wysunęli następującą tezę (zwaną w literaturze zazwyczaj „tezą Churcha” lub „tezą Churcha–Turinga”):

(TC) Funkcja jest obliczalna wtedy i tylko, wtedy gdy jest rekurencyjna.

TC ma też wersję dotyczącą rozstrzygalności:

(TC<sub>1</sub>) Relacja (problem) jest rozstrzygalna wtedy i tylko, wtedy gdy jest rekurencyjna<sup>7</sup>.

Od tej pory nie będziemy rozróżniać TC od TC<sub>1</sub>, lecz potraktujemy je jako jedną tezę. Ponadto, od tej pory wszystkie funkcje to dla nas funkcje z liczb naturalnych w liczby naturalne.

---

<sup>6</sup> Np. Odifreddi [1989].

<sup>7</sup> Por. Church [1936], Turing [1936], Post [1936]. Church publicznie zaproponował identyfikację klasy funkcji obliczalnych z klasą funkcji rekurencyjnych po raz pierwszy podczas wypowiedzi dla Amerykańskiego Towarzystwa Matematycznego 19 kwietnia 1935 r. (choć prywatnie proponował ją także przed tą datą; por. Sieg [1997]). Należy podkreślić, że autorzy TC nie byli zgodni co do tego, jaki jest status tej tezy: przykładowo, Church uważał ją za definicję, a Post – za „hipotezę roboczą”, która „wymaga nieustannej weryfikacji” (o problemie statusu TC – patrz niżej).

Znaczenie TC jest fundamentalne: dzięki niej, można z twierdzeń o nierekurencyjności pewnych zbiorów wyciągać wnioski o nierozstrzygalności odpowiadających im problemów, tj. o nieistnieniu algorytmów rozstrzygających te problemy. I tak, już w 1936 r. Church i Turing udowodnili, że (przy założeniu TC) nie ma algorytmu rozstrzygającego, czy dane zdanie klasycznego rachunku predykatów jest twierdzeniem tegoż rachunku. Podobnie w 1970 r. Matijasiewicz wykazał, że (znów: *modulo* TC) nie istnieje algorytm pozwalający stwierdzić, czy dane równanie diofantyczne ma rozwiązanie w liczbach całkowitych<sup>8</sup>. Wynik ten – to, rzecz jasna, rozwiązanie X problemu Hilberta, aczkolwiek jest to rozwiązanie innego typu, niż oczekiwał Hilbert.

Stwierdzenie, że istnieją problemy matematyczne niedostępne żadnym algorytmom, jest już w zasadzie stwierdzeniem filozoficznym. Za pomocą TC można jednak z wyników matematycznych próbować wyciągać jeszcze „bardziej filozoficzne” wnioski, dotyczące np. syntetyczności zdań matematycznych (por. dyskusję takiej propozycji C. Castonguaya w Woleński [1993], s. 157-158, 161). Co więcej, choć wyrażane czasem poglądy, że sama TC „wiąże działanie umysłu z obliczaniem” (Hofstadter [1979], s. xii) albo, że „wspiera mechanistyczną teorię umysłu” (Nelson [1987], s. 581), są raczej bezpodstawne, to TC wraz z pewnymi dodatkowymi założeniami (np. właśnie z tak czy inaczej sformułowaną tezą mechanicyzmu) pozwala wnioskować o zasadniczych ograniczeniach matematycznych możliwości ludzkiego umysłu. Świadczy to, naszym zdaniem, o tym, że TC jest dobrym przykładem uzupełniającej przesłanki używanej do wyprowadzania interpretacyjnych konsekwencji twierdzeń matematycznych. Skoro tak, to można zadać w odniesieniu do niej pytania, które we *Wprowadzeniu* przytoczyliśmy w odniesieniu do wszelkich tego typu przesłanek – tj. pytania o jej status epistemiczny i o możliwości jej uzasadniania. Ponieważ TC jest tezą szczególnie prosto sformułowaną, a ponadto- szczególnie intensywnie analizowaną, o odpowiedzi może w jej przypadku być łatwiej niż w przypadku ogólnym. Mając zaś choćby zarys takich odpowiedzi, można by się było z kolei pokusić o wyciągnięcie z nich ogólniejszych wniosków.

---

<sup>8</sup> Stwierdzenia te należy rozumieć następująco: Church i Turing pokazali, że (jeśli w odpowiedni sposób zakodujemy zdania klasycznego rachunku predykatów jako liczby naturalne) zbiór twierdzeń KRP nie jest zbiorem rekurencyjnym. Matijasiewicz zaś udowodnił twierdzenie, z którego wynika w szczególności, że istnieje wielomian  $p(x_1, \dots, x_n, y_1, \dots, y_k)$  o współczynnikach całkowitych, dla którego zbiór takich n-tek liczb naturalnych  $a_1, \dots, a_n$  (ograniczenie się w tym miejscu do liczb naturalnych nie zmniejsza istotnie ogólności problemu), że istnieją  $y_1, \dots, y_k$  spełniające  $p(a_1, \dots, a_n, y_1, \dots, y_k) = 0$ , nie jest rekurencyjny.

### 3. Status tezy Churcha

W podręcznikach teorii rekursji czy teorii złożoności obliczeniowej można natknąć się na zapewnienia, że TC jest „powszechnie przyjęta” (Immerman [1998], s. 23), „niemal powszechnie przyjęta” (Enderton [1997], s. 533) czy „przyjęta przez niemal wszystkich logików” (Schoenfield [1993], s. 26). Zapewnienia takie skłaniają do namysłu nad przynajmniej dwiema kwestiami. Po pierwsze: skoro wszyscy czy prawie wszyscy przyjmują TC, to jakie argumenty ich do tego skłaniają? Po drugie: co tak naprawdę uznają ci wszyscy, którzy „przyjmują TC”?

Zacznijmy od uwagi *à propos* sprawy pierwszej, tj. argumentów na rzecz TC. Uwaga brzmi: niektóre spośród najczęściej spotykanych w literaturze argumentów są bardzo podejrzane. Oto dwa przykłady.

*Argument pierwszy:* wszystkie znane nam funkcje obliczalne są rekurencyjne. Ponadto, wszystkie znane nam metody otrzymywania nowych funkcji obliczalnych z uprzednio danych funkcji obliczalnych prowadzą od funkcji rekurencyjnych do funkcji rekurencyjnych.

*Argument drugi:* próbowano uściślać pojęcie funkcji obliczalnej na wiele bardzo zróżnicowanych sposobów. Definiowane klasy funkcji zawsze okazywały się równe klasie funkcji rekurencyjnych. To dowodzi, że klasa funkcji rekurencyjnych jest bardzo naturalną klasą funkcji. „Trudno sobie wyobrazić, dlaczego miałyby tak być, chyba że jest to rzeczywiście klasa funkcji obliczalnych” (zwieńczyliśmy ten wywód cytatem z książki Schoenfield [1993], s. 27 – L.A.K.).

Argument pierwszy nie zasługuje na dyskusję. Argument drugi jest znacznie ciekawszy, ale dowodzi tylko tego, że klasa funkcji rekurencyjnych jest bardzo naturalną klasą funkcji. *Trudno sobie wyobrazić, dlaczego miałyby z tego wynikać, że jest to rzeczywiście klasa funkcji obliczalnych.*

Istnieją, rzecz jasna, poważniejsze argumenty na rzecz TC. Wspomnimy także o niektórych z nich, najpierw jednak winniśmy skupić się na drugiej z wymienionych powyżej kwestii: co tak naprawdę uznają ci, którzy „przyjmują TC”?

Pytanie jest, być może wbrew pozorom, niebanalne. Normalnie, jeśli ktoś „przyjmuje” jakąś tezę, to uznaje, że jest to teza prawdziwa. Nie jest natomiast *a priori* rzeczą jasną, czy TC może być prawdziwa – w naturalnym tego wyrazu znaczeniu.

Rozważmy bowiem następujące rozumowanie: TC stwierdza równozakresowość pojęcia funkcji rekurencyjnej i pojęcia funkcji obliczalnej. Pojęcie funkcji rekurencyjnej jest precyzyjnym pojęciem matematycznym o ściśle określonym zakresie. Funkcja z liczb naturalnych w liczby naturalne albo jest, albo nie jest rekurencyjna – trzeciej możliwości nie ma. Z drugiej strony, nie jest wcale oczywiste, że pojęcie funkcji obliczalnej ma ściśle wyznaczony zakres. Obliczalność jest pojęciem intuicyjnym, a takie pojęcia (jak wiadomo co najmniej od czasów „paradoksu łysego”) często bywają nieostre. Nie można więc *a priori* wykluczać, że również pojęcie funkcji obliczalnej jest nieostre i ma swoje „przypadki graniczne” (oczywiście ewentualna nieostrość pojęcia obliczalności sprowadzałyby się ostatecznie do – również ewentualnej – nieostrości pojęcia algorytmu; można wyobrazić sobie, że istnieją „przypadki graniczne” pojęcia algorytmu, choć zazwyczaj stosunkowo łatwo jest ocenić, czy jakaś procedura jest algorytmem, czy nie)<sup>9</sup>. Gdyby jednak tak było, to nie mogłoby być ono równozakresowe z pojęciem funkcji rekurencyjnej – czyli TC, rozumiana dosłownie, byłaby fałszywa (ewentualnie, gdyby np. wszystkie funkcje rekurencyjne były na pewno obliczalne, a wszystkie funkcje nierekurencyjne – albo na pewno nieobliczalne, albo „przypadkami granicznymi” obliczalności, można by uznać, że wartość logiczna TC jest „nieokreślona”).

Powyższy wywód winien nam przede wszystkim uświadomić, że należy zastanowić się nad statusem TC, to jest zadać sobie pytanie, czym – jakiego typu zdaniem – TC w istocie jest. Widzimy przynajmniej trzy możliwości odpowiedzi na to pytanie:

(1) Można TC uznawać za definicję syntetyczną. To jest: uznawać, że na mocy TC „obliczalność” znaczy, z definicji i całkiem niezależnie od jakichś wcześniejszych intuicji z tym słowem związanych, to samo co „rekurencyjność” (zwróćmy uwagę, że jest to propozycja bliska jednej z rozważanych przez Ajdukiewicza metod uzasadniania „parafraz”).

(2) Można też uważać, że TC jest tezą faktualną: zakładamy, że obliczalność jest pojęciem ostrym, tj. każda funkcja albo jest obliczalna, albo nie (choć odpowiedź na pytanie, czy jakaś dana funkcja jest obliczalna, nie musi być oczywista; być może nie da się takiej odpowiedzi udzielić, zanim nie zgromadzi się stosownego *quantum* wiedzy na temat tej

---

<sup>9</sup> Shapiro (por. Shapiro [1998], s. 353) podkreśla ponadto, że obliczalność jest, przynajmniej *prima facie*, pojęciem modalnym, a pojęcia modalne mają szczególnie silną tendencję do nieostrości. Oczywiście, pojęcie „praktycznej obliczalności”, uwzględniające ograniczenia czasowe i pamięciowe, którym podlegają faktycznie realizowalne przez nas algorytmy, na pewno ma „przypadki graniczne”. Formułując pojęcie obliczalności (*tout court*), abstrahujemy od tych ograniczeń, nie jest jednak, zdaniem Shapira, całkiem jasne, że uwalniamy się w ten sposób od *wszystkich* „przypadków granicznych”.

funkcji) – i uważamy, że „jest tak”, iż klasy funkcji obliczalnych i funkcji rekurencyjnych są równe (taki pogląd reprezentuje np. Post w pracy Post [1936])

(3) Wreszcie, można TC traktować jako „racjonalną rekonstrukcję”. To znaczy: dopuścić możliwość, że pojęcie obliczalności jest nieostre i zaproponować, by pojęcie rekurencyjności było jego precyzyjnym, naukowym odpowiednikiem. Wymaga się, by wszystkie obiekty podpadające (*resp.* nie podpadające) pod stare pojęcie podpadały (*resp.* nie podpadały) pod jego odpowiednik, nie przesądza się natomiast, jak odpowiednik ma się „zachowywać” w odniesieniu do „przypadków granicznych” starego pojęcia. Stąd, odpowiednik pojęcia nieostrego nie jest wyznaczony jednoznacznie. „Racjonalna rekonstrukcja” może być poprawna (tj. spełniać sformułowane powyżej wymaganie), adekwatna (do takich czy innych celów), ale nie – prawdziwa (chyba, że w poniekąd zdegenerowanym przypadku, kiedy rekonstruowane pojęcie jest ostre).

Ad (1). Oczywiście, jeśli TC jest definicją syntetyczną, to jest też automatycznie prawdziwa – jest to jednak „banalna” prawdziwość. Takie rozumienie TC jest filozoficznie niesatysfakcjonujące, gdyż *de facto* uniemożliwia „redukcję” problemów filozoficznych do matematycznych. Ponadto, należy zwrócić uwagę, że TC miała za zadanie umożliwić wyciąganie z twierdzeń o nierekurencyjności pewnych funkcji (*resp.* relacji) wniosków o nieobliczalności tych funkcji (*resp.* nierozstrzygalności relacji), pojmowanej intuicyjnie. Definicja syntetyczna nie mogłaby spełniać tej roli, stąd też zapewne właściwie nikt nie traktuje TC jako definicji syntetycznej (odnosi się to również do Churcha, który – choć zwał TC „definicją” – starał się przytaczać argumenty świadczące o jej adekwatności, a zatem nie uważał jej za definicję syntetyczną; por. Church [1936], s. 100-102; w gruncie rzeczy punkt widzenia Churcha był zbliżony do (3)).

Ad (2) i (3). Wydaje się, że znakomita większość piszących na temat TC autorów ma na tę tezę pogląd typu (2) i/lub (3). Spójnik „i/lub” jest najwłaściwszy, gdyż zazwyczaj nie wprowadza się precyzyjnego rozróżnienia między poglądem (2) a (3).

Różnica między (2) a (3), choć nie zawsze dostrzegana, może się jednak wydawać filozoficznie istotna. Jeśli bowiem prawdziwa jest TC w wersji (2), to matematyczne problemy dotyczące rekurencyjności można traktować po prostu jako przekłady (na poły filozoficznych) problemów dotyczących obliczalności. Jeżeli natomiast prawdziwe jest stanowisko (3), a pojęcie rekurencyjności jest jedynie wynikiem „racjonalnej rekonstrukcji” pojęcia obliczalności, to wypada np. uznać, że badanie rekurencyjności może nam „zastąpić” badanie obliczalności, ale niekoniecznie jest mu równoważne – albo też przyjąć, że



wnioskowanie z (nie)rekurencyjności o (nie)obliczalności wymaga jeszcze dodatkowych przesłanek (mających np. wykluczyć, że takie czy inne funkcje są „przypadkami granicznymi” obliczalności). Spory między zwolennikami (2) a (3) toczą się niemal od momentu pierwszego sformułowania TC – przykładowo, już w 1937 r., w recenzji dla *Journal of Symbolic Logic* (tekst recenzji znaleźć można w pracy Sieg [1997], s. 177), Church, zwolennik stanowiska (3), ostro skrytykował Post- obrońcę (2), który z kolei już wcześniej (por. Post [1936], s. 291) polemizował z poglądem Churcha, iż TC jest „definicją” (tj., *de facto*, raczej „racjonalną rekonstrukcją” niż definicją syntetyczną – por. uwagę wyżej).

*Opinio communis* autorów uznających różnicę między (2) a (3) zdaje się raczej skłaniać ku stanowisku (3). Stanowisko to bywa jednak krytykowane. Sprzeciwia się mu między innymi Elliott Mendelson w pracy Mendelson [1990]. Swoją sprzeciw opiera Mendelson na przekonaniu, że TC można dowieść. Z kolei według Stewarta Shapira (Shapiro [1981]) do poglądu typu (2) winny prowadzić przynajmniej niektóre ogólne stanowiska w filozofii matematyki – na przykład strukturalizm. Warto odnotować, że zdaniem zarówno Mendelsona, jak i Shapirowskiego strukturalisty, status TC nie jest w żaden sposób wyjątkowy w historii matematyki – jest bowiem taki sam jak status znanych z przeszłości „tez” identyfikujących z pojęciami matematycznymi takie pojęcia prematematyczne jak m. in.: funkcja, prawda, granica, wymiar, koło, trójkąt (Mendelson), ciągłość funkcji, stosunek długości odcinków, definicja (strukturalista z tekstu Shapira) czy pole figury (obaj). Przekonanie, że TC nie jest wyjątkowa, jest zresztą częste (por. np. Sieg [1997], s. 173-174, Kreisel [1987]) i niekoniecznie musi wiązać się z odrzuceniem (3) – można przecież np. utrzymywać, że intuicyjne pojęcie ciągłości funkcji jest nieostre, a znana epsilonowo-deltowa definicja jest wynikiem jego „racjonalnej rekonstrukcji”.

Sytuacja jest więc taka: nie ma pełnej zgody co do statusu TC (nawet jeśli się tę tezę „przyjmuje”), a zatem nie może też być zgody co do statusu rozumowań filozoficznych wykorzystujących TC jako przesłankę. Powinniśmy jednak – nawet zanim spór między (2) a (3) zostanie przekonująco rozstrzygnięty – móc się jakoś odnieść do takich rozumowań i ocenić ich wartość. Mogłoby się wydawać, że ponieważ stanowisko (3) jest ostrożniejsze, tj. słabsze, od (2) (w tym sensie, że tezę faktyczną można traktować jako szczególny, zdegenerowany przypadek „racjonalnej rekonstrukcji”, ale nie na odwrót) – to, przy ocenianiu wartości rozumowań odwołujących się do TC, bezpieczniej byłoby postępować, tak jakby to właśnie stanowisko (3) było właściwe. Osłabiałoby to, rzecz jasna, wymowę owych opartych na TC rozumowań. Uważamy, że jest to w zasadzie słuszny punkt widzenia – z *tym jednakże zastrzeżeniem, że różnica między (2) a (3) straciła istotnie na znaczeniu*. Powód jest

następujący: przy obecnym stanie wiedzy na temat TC, rozsądne argumenty za tym, iż TC jest poprawną „racjonalną rekonstrukcją”, będą również uzasadniały pogląd, że pojęcie obliczalności jest ostre. Postaramy się tę myśl rozwinąć.

\*

Powyżej zastanawialiśmy się nad pytaniem o status TC. Znaleźliśmy dwie *a priori* możliwe odpowiedzi na to pytanie: pogląd, iż TC jest tezą faktualną i pogląd, iż jest ona „racjonalną rekonstrukcją”. Drugi z tych poglądów jest w pewnym sensie słabszy od pierwszego. Twierdzimy jednak, że jeśli nawet jest istotnie słabszy, to nie będziemy tego w stanie pokazać, a dokładniej: jeśli będziemy potrafili pokazać, że TC jest poprawną „racjonalną rekonstrukcją”, to będziemy również potrafili pokazać, że pojęcie obliczalności jest ostre, co – z jedynej poprawnej rekonstrukcji pojęcia ostrego – będzie oznaczało, że możemy TC traktować również jako (prawdziwą) tezę faktualną. Stwierdzenie to można uzasadnić, analizując postać argumentów na rzecz TC.

Zauważmy najpierw, że TC jest równoważnością, czyli koniunkcją dwóch implikacji, i, że stosunkowo łatwo jest uzasadnić jedną z tych implikacji, tę mianowicie, że jeśli funkcja jest rekurencyjna, to jest też obliczalna. Wystarczy w tym celu przyjrzeć się np. podanej w *Dodatku* definicji rekurencyjności. Definicja ta określa zbiór funkcji rekurencyjnych jako najmniejszy zbiór funkcji zawierający pewne dane funkcje podstawowe i zamknięty na pewne metody konstruowania nowych funkcji z uprzednio uzyskanych. Nietrudno znaleźć algorytmy obliczające funkcje podstawowe, równie łatwo przerobić algorytmy obliczające jakieś dane funkcje na algorytm obliczający funkcję uzyskaną z tych funkcji za pomocą jednej z dopuszczalnych metod. Tak więc, każda funkcja rekurencyjna jest obliczalna.

Tu dygresja: można odnieść wrażenie, że z początku nie było wcale dla wszystkich oczywiste, że każda funkcja rekurencyjna jest obliczalna. Pojęcie obliczalności było uważane za nieprecyzyjne pojęcie intuicyjne – i nawet gdy Church po raz pierwszy publicznie proponował utożsamienie obliczalności z rekurencyjnością (było to w kwietniu 1935 r., jeszcze przed powstaniem pracy Church [1936]), jedynym podanym przezeń uzasadnieniem tej propozycji był fakt, że „inne możliwe do przyjęcia (*plausible*) definicje efektywnej obliczalności prowadzą, jak się okazuje, do pojęć albo równoważnych, albo słabszych od rekurencyjności” (cyt. za: Sieg [1997], s. 162). Jak widać, dla Churcha nie było wówczas ważne to, czy wszystkie funkcje rekurencyjne są rzeczywiście obliczalne – tego zagadnienia w swej wypowiedzi nie rozważał – lecz jedynie to, że pojęcie rekurencyjności wystarczało,

jak sądził, do ograniczenia klasy funkcji obliczalnych niejako „z góry” (jest to świadectwo znanego faktu, iż pojęcie rekurencyjności było potrzebne po to, by móc udzielać *negatywnych* odpowiedzi na pytania o obliczalność czy rozstrzygalność). Skądinąd już w Church [1936] stwierdzał, że jest jasne, iż dla każdej funkcji rekurencyjnej istnieje algorytm, który ją oblicza (s. 93).

Jedną z implikacji składających się na TC jest zatem łatwa do uzasadnienia. Trudniej jest uzasadnić implikację odwrotną: jeśli funkcja jest obliczalna, to jest też rekurencyjna. Temu celowi służyć mają tzw. „argumenty za TC”.

Argumentów za TC pojawiło się w ciągu ostatnich kilkudziesięciu lat *multum*. Wygodną klasyfikację tych argumentów zawiera książka Murawski [1990] (s. 64-65). Murawski dzieli argumenty za TC na następujące trzy kategorie: (a) argumenty heurystyczne, (b) argumenty bezpośrednie, (c) argumenty oparte na istnieniu różnych uściśleń pojęcia obliczalności.

Sztandarowym przykładem argumentu typu (a) jest pierwszy z przytoczonych przez nas wyżej „argumentów podejrzanych”. Sztandarowym przykładem argumentu typu (c) jest drugi „argument podejrzany”. Drogą eliminacji można więc dojść do wniosku, że największe nadzieje należy wiązać z argumentami bezpośrednimi. Polegają one zwykle na „dokładnym analizowaniu procesu obliczania i próbach wykazania w ten sposób, że tylko funkcje rekurencyjne mogą być obliczalne” (Murawski [1990], s. 64).

Najsłynniejszy bezpośredni argument za TC zawiera praca Turing [1936]. Turing analizował w niej (s. 135-138) przebieg procesu obliczania – rozbijając ten proces na elementarne „operacje proste” oraz rozważając, jakim ograniczeniom podlega podmiot wykonujący obliczenie. Z analizy wynikało, że każdy proces obliczania może być zasymulowany na maszynie Turinga. Ponieważ obliczalność przez maszyny Turinga jest jedną z równoważnych definicji rekurencyjności (szkic dowodu równoważności tej definicji z definicjami znanymi wcześniej zawiera właśnie praca Turing [1936]), wniosek z rozważań Turinga jest taki, że każda funkcja obliczalna jest rekurencyjna.

Niektórzy autorzy (np. Mendelson) uważają, że rozumowanie Turinga wystarczy, by uzasadnić czy wręcz dowieść TC. Inni jednak sądzą, że rozumowanie to ma ograniczony zakres stosowalności. Robin Gandy na przykład (Gandy [1980]) twierdzi, że analiza Turinga, choć poprawna, odnosi się tylko do obliczeń dokonywanych przez „abstrakcyjnego człowieka postępującego w mechaniczny (*routine*) sposób” (s. 124), nie obejmuje natomiast obliczeń dokonywanych przez maszyny – w szczególności, według Gandy’ego, w wypadku obliczeń maszynowych nie musi istnieć możliwość rozbicia procesu obliczeniowego na elementarne

kroki. Praca Gandy'ego ma właśnie uzupełnić tę lukę w argumentacji Turinga, przynajmniej w odniesieniu do maszyn „dyskretnych i deterministycznych” (w sensie wyjaśnionym w pracy). Gandy konstruuje bardzo ogólną teoriomnogościową metodę opisu maszyn, a następnie formułuje pewne warunki, które powinny spełniać dyskretne maszyny deterministyczne (niemal każde istotne osłabienie tych warunków pozwala znaleźć dla *dowolnej* funkcji maszynę, która spełnia osłabione warunki, a zarazem „oblicza” tę funkcję). Można udowodnić, że każda maszyna spełniająca te warunki oblicza wyłącznie funkcje rekurencyjne. Jest to dalszy argument na rzecz tezy, że wszystkie funkcje obliczalne są rekurencyjne.

Istnieją inne argumenty w duchu Turinga-Gandy'ego. Nie jesteśmy w stanie rozstrzygnąć, czy można uznać, że przesadzają one o prawdziwości TC. Istotne są dwie rzeczy: po pierwsze, są to zdecydowanie najlepsze (by nie rzec: jedyne w miarę dobre) argumenty na rzecz TC, jakimi dysponujemy<sup>10</sup>. Po drugie: argumenty te zmierzają do wykazania, że wszystkie funkcje obliczalne są rekurencyjne.

W tym kontekście można się zastanowić, jak mogłaby przebiegać argumentacja na rzecz stwierdzenia, że TC jest poprawną „racjonalną rekonstrukcją” pojęcia obliczalności. Stwierdzenie takie sprowadza się do koniunkcji dwóch innych: (a) nie ma funkcji rekurencyjnych, które by definitywnie nie były obliczalne i (b) nie ma funkcji definitywnie obliczalnych, które by nie były rekurencyjne. Stwierdzenie (a) można obecnie uznać za uzasadnione, mamy nawet mocniejsze: każda funkcja rekurencyjna jest obliczalna. Jak natomiast można by uzasadnić tezę, iż nie ma funkcji definitywnie obliczalnych, które nie są rekurencyjne? Nie widzimy innej możliwości niż ta, która się narzuca: analizując pojęcie algorytmu tudzież obliczenia, albo też w jakiś całkiem inny sposób, nałożyć pewne minimalne warunki na funkcje, które bylibyśmy gotowi uznać za obliczalne – a następnie pokazać po prostu, że każda funkcja spełniająca te warunki jest rekurencyjna (ten ostatni krok mógłby przebiegać nie wprost, ale to dla nas bez znaczenia). Jednakże ktoś, komu udałoby się pokazać, że każda funkcja obliczalna jest rekurencyjna, tym samym pokazałby, że funkcje obliczalne *tout court* są to dokładnie funkcje rekurencyjne, a zatem, że pojęcie obliczalności jest ostre i że TC, rozumiana jako teza faktualna, a nie tylko „racjonalna rekonstrukcja”, jest prawdziwa.

Skądinąd ewentualne argumenty za tym, że TC nie jest prawdziwa jako teza faktualna, musiałyby zapewne polegać na pokazaniu, że istnieje funkcja obliczalna, która nie jest reku-

---

<sup>10</sup> W pewnym sensie dobry jest również argument „z doświadczenia”, o którym mówi Schoenfield (Schoenfield [1993], s. 27) – „*by a detailed study of recursion theory*”. Nie jest to jednak argument teoretyczny.

recencyjna – co oznaczałoby zarazem, że TC nie można też traktować jako poprawnej „racjonalnej rekonstrukcji”. Przykładowo, najsłynniejszy bodaj argument „przeciw”, sformułowany przez László Kalmára (Kalmár [1959]), polegał na wskazaniu nierekurencyjnej funkcji, którą, zdaniem Kalmára, należało uznać za obliczalną (gdyż uznanie jej za nieobliczalną prowadziło do „dziwnych”, w odczuciu Kalmára, konsekwencji – np. do takiej, że istniałoby nierozstrzygalne zdanie, o którym wiadomo by było, że jest fałszywe<sup>11</sup>). Rzecz jasna, gdyby Kalmár miał rację i gdyby jego funkcja rzeczywiście była obliczalna, obalałoby to TC i w wersji (2), i w wersji (3). Obliczalność rozważanej przez Kalmára funkcji miałyby dla TC takie mniej więcej skutki, jakie – niewątpliwa zresztą – obliczalność funkcji Ackermanna miałyby dla (nigdy nie sformułowanej) tezy identyfikującej klasę funkcji obliczalnych z klasą funkcji pierwotnie rekurencyjnych<sup>12</sup>.

Teoretycznie, mimo wszystko, nadal możliwe jest, że TC jest poprawną „racjonalną rekonstrukcją”, ale nie jest prawdziwą tezą faktualną – mianowicie, możliwe jest, że, choć wszystkie funkcje rekurencyjne są definitywnie obliczalne, i choć nie ma funkcji nierekurencyjnej, która by definitywnie była obliczalna (to drugie jest już, oczywiście, spekulatywnym założeniem), to jednak istnieją funkcje nierekurencyjne, których nie można uznać za po prostu nieobliczalne, i które są „przypadkami granicznymi” obliczalności. Po pierwsze wszelako, jeśli mamy rację, to gdyby taka sytuacja miała miejsce, nie zdołalibyśmy jej wykryć (ciężko wskazać „przypadek graniczny” obliczalności i w dodatku wykazać, że jest on właśnie „przypadkiem granicznym”). A po drugie: gdyby taka sytuacja rzeczywiście zachodziła, to aż narzucaliby się, by uznać – tu faktycznie jest pewien moment decyzji, pewien element „racjonalnej rekonstrukcji” – że wszystkie te „przypadki graniczne” są jednak nieobliczalne. Zwłaszcza, że funkcja będąca „przypadkiem granicznym” obliczalności to jednak konstrukt mocno podejrzany – gdyż podejrzana jest koncepcja algorytmu, który byłby algorytmem „tylko trochę”. Tak więc, różnica między TC jako tezą faktualną i TC jako racjonalną rekonstrukcją nie jest aż tak znaczna, jak mogłoby się wydawać – w szczególności, jak wydawało się Churchowi i Postowi.

---

<sup>11</sup> Zdanie to byłoby nierozstrzygalne w tym sensie, że nie istniałby algorytm, który rozstrzyga, czy jest ono prawdziwe – skądinąd, Kalmár pisze nie o „algorytmach”, lecz o „dowolnych poprawnych metodach”, i w tym być może jest źródło nieporozumienia. Zwięzłą, ale bardzo trafną krytykę wyводу Kalmára podał Kleene w pracy Kleene [1987].

<sup>12</sup> Klasa funkcji pierwotnie rekurencyjnych jest to pewien ważny podzbiór właściwy klasy funkcji rekurencyjnych. Funkcja Ackermanna – to standardowy przykład funkcji, która jest rekurencyjna, ale nie jest pierwotnie rekurencyjna (rośnie szybciej niż wszystkie funkcje pierwotnie rekurencyjne).

#### 4. Zakończenie

Rozważania nad tezą Churcha doprowadziły nas do domniemania, że dwie z pozoru zdecydowanie rozbieżne interpretacje tej tezy nie różnią się zbyt istotnie. Jeśli to domniemanie jest słuszne *i jeśli teza Churcha nie jest po prostu pomyłką*, to traktowanie matematycznych problemów dotyczących rozstrzygalności jako zwyczajnie *równoważnych* filozoficznym problemom dotyczącym obliczalności czy rozstrzygalności jest wprawdzie (być może) trochę nieścisłe, ale nie jest poważnym błędem.

Rozważania nasze odnosiły się bezpośrednio wyłącznie do tezy Churcha i odwoływały się do pewnych szczególnych cech tej tezy, toteż trudno o jakieś naturalne a ciekawe uogólnienia. Wszelkie ogólne wnioski będą tu niemal na pewno banałami i ogólnikami. Niemniej jednak, na dwa banały i ogólniki chcemy sobie pozwolić.

Pierwszy banał jest taki: jest więcej „metod uprawnienia parafraz”, niżli się śniło Ajdukiewiczowi. Widać to dobrze właśnie na przykładzie tezy Churcha. Można powiedzieć, że *wiemy*, że wszystkie funkcje rekurencyjne są obliczalne. Być może kiedyś będziemy *wiedzieli* (a już teraz wielu *sądzi*), że jest też i na odwrót. Nigdy nie doszlibyśmy do tego na drodze samej tylko „analizy znaczeniowej wyrazów mowy potocznej”, nie wspominając o arbitralnych ustaleniach. To, że *wiemy* dziś na temat intuicyjnego pojęcia obliczalności więcej, niż *wiedział*, powiedzmy, Hilbert, zawdzięczamy przede wszystkim temu, że Church czy Turing odważyli się zaproponować ścisłą, matematyczną wersję tego pojęcia – a następnie ta matematyczna wersja „wymusiła” na nas głębsze zrozumienie naszego pierwotnego, intuicyjnego pojęcia, niejako „domyślenie go do końca”.

Obserwacja ta, choć banalna, jest jednak ważna – świadczy bowiem o pewnej przewadze „twardej” analizy filozoficznej (tak zwanej „*horseshoe analysis*”) nad wszelkimi „filozofiami języka potocznego”.

Pora na drugi banał: wyżej wspominaliśmy, że wielu autorów znajdowało w dziejach matematyki analogiczne do tezy Churcha utożsamienia pojęć prematematycznych (czasem ewidentnie filozoficznych) z matematycznymi. Lista, którą podaliśmy, była bardzo przypadkowa i trudno przesądzać, które ze znajdujących się na niej pojęć mają matematyczne równoważniki czy prawie-równoważniki, a które tylko swobodnie rozumiane odpowiedniki. Niemniej jednak, jest to kwestia, którą warto badać, a dzięki wynikom takich badań możemy

odkryć ważne „konsekwencje interpretacyjne” rezultatów matematycznych. Może to być – i niechaj będzie – zachętą do dalszych prób „redukowania filozofii do matematyki”.

## *Dodatek*

### *Kilka definicji*

Poniżej podajemy jedną z możliwych ścisłych definicji klasy funkcji (i relacji) rekurencyjnych. Rozważamy funkcje z liczb naturalnych w liczby naturalne (dopuszczamy funkcje częściowe; zapis „ $f: \omega^n \rightarrow \omega$ ” znaczy jedynie, że dziedzina funkcji  $f$  zawiera się w  $\omega^n$  i że wartości  $f$  są liczbami naturalnymi; funkcję, której dziedzina równa się  $\omega^n$ , nazywamy *całkowitą*) oraz relacje będące podzbiorami potęg kartezjańskich zbioru liczb naturalnych.

*Def.* Zbiór funkcji  $K$  jest *zamknięty na składanie*, jeśli ilekroć funkcja  $g: \omega^n \rightarrow \omega$  i funkcje  $h_i: \omega^k \rightarrow \omega$  (dla  $i = 1, \dots, n$ ) należą do  $K$ , również funkcja  $f: \omega^k \rightarrow \omega$  zdefiniowana wzorem:

$$f(x_1, \dots, x_k) \cong (h_1(x_1, \dots, x_k), \dots, h_n(x_1, \dots, x_k))$$

należy do  $K$  (gdzie symbol „ $\cong$ ” oznacza, że dane funkcje mają tę samą dziedzinę i dla wszystkich argumentów z dziedziny przyjmują te same wartości; innymi słowy, jest to równość w sensie teoriomnogościowym).

*Def.* Zbiór funkcji  $K$  jest *zamknięty na rekursję prostą*, jeśli ilekroć funkcje całkowite  $g: \omega^n \rightarrow \omega$  i  $h: \omega^{n+2} \rightarrow \omega$  należą do  $K$ , również funkcja całkowita  $f: \omega^{n+1} \rightarrow \omega$  zdefiniowana wzorem:

$$f(0, x_1, \dots, x_n) = g(x_1, \dots, x_n),$$

$$f(y+1, x_1, \dots, x_n) = h(f(y, x_1, \dots, x_n), y, x_1, \dots, x_n)$$

należy do  $K$ .

*Def.* Funkcja *charakterystyczna*  $\chi_R$  relacji  $R \subseteq \omega^n$  jest to funkcja całkowita zdefiniowana wzorem:

$$\chi_R(x_1, \dots, x_n) = 1, \text{ jeśli zachodzi } R(y, x_1, \dots, x_n); 0, \text{ w przeciwnym wypadku.}$$

*Def.* Dla relacji  $R \subseteq \omega^{n+1}$  definiujemy  $\mu y R (y, x_1, \dots, x_n)$  jako najmniejsze  $y$  takie, że zachodzi  $R (y, x_1, \dots, x_n)$ . Jeśli nie ma takiego  $y$ , to  $\mu y R (y, x_1, \dots, x_n)$  jest niezdefiniowane.

*Def.* Zbiór funkcji  $K$  jest *zamknięty na operację minimum*, jeśli ilekroć funkcja charakterystyczna  $\chi_R$  relacji  $R \subseteq \omega^{n+1}$  należy do  $K$ , również funkcja  $f: \omega^n \rightarrow \omega$  zdefiniowana wzorem:

$$f(x_1, \dots, x_n) \cong \mu y R (y, x_1, \dots, x_n)$$

należy do  $K$ .

*Def.* Zbiór funkcji  $K$  jest *rekurencyjnie zamknięty*, jeśli spełnia następujące warunki:

1) funkcje całkowite:

i) *Zero* :  $\omega \rightarrow \omega$  (zdefiniowana wzorem  $Zero(x) = 0$  dla każdego  $x$ ),

ii) *Sc* :  $\omega \rightarrow \omega$  (zdefiniowana wzorem  $Sc(x) = x+1$  dla każdego  $x$ ),

iii) dla dowolnych  $i, n$  takich, że  $1 \leq i \leq n$ ,  $I_{i,n} : \omega^n \rightarrow \omega$

(zdefiniowana wzorem  $I_{i,n}(x_1, \dots, x_n) = x_i$  dla wszystkich  $x_1, \dots, x_n$ ),

należą do  $K$ ;

2)  $K$  jest zamknięty na składanie;

3)  $K$  jest zamknięty na rekursję prostą;

4)  $K$  jest zamknięty na operację minimum.

*Uwaga.* Przecięcie dowolnej rodziny zbiorów rekurencyjnie zamkniętych jest zbiorem rekurencyjnie zamkniętym.

*Def.* Zbiór funkcji rekurencyjnych jest to najmniejszy zbiór rekurencyjnie zamknięty. Funkcja jest *rekurencyjna*, jeśli należy do zbioru funkcji rekurencyjnych. Relacja jest *rekurencyjna*, jeśli jej funkcja charakterystyczna jest rekurencyjna.

Powyższa definicja może się wydawać mało intuicyjna. Dlatego podamy również ideę jednej z bardziej intuicyjnych definicji rekurencyjności, a mianowicie- definicji opartej na pojęciu maszyn Turinga. *Maszyna Turinga* jest to dowolna funkcja  $M$ , której dziedzina zawiera się w zbiorze  $\{0, 1, \dots, n\} \times \{0, 1\}$  (dla pewnej liczby naturalnej  $n$ ), a obraz – w



zbiornie  $\{0, 1\} \times \{L, P\} \times \{0, 1, \dots, n\}$ . Nieformalnie, należy sobie maszynę Turinga  $M$  wyobrazić jako maszynę dysponującą głowicą oraz (rozciągającą się po obu stronach w nieskończoność) taśmą. Taśma jest podzielona na kwadraty. W każdym z kwadratów znajduje się jeden z dwóch symboli: 0 lub 1, przy czym w danej chwili tylko skończona liczba kwadratów może zawierać symbol 1. Nad jednym z kwadratów znajduje się głowica – mówimy, że głowica czyta symbol zawarty w tym kwadracie.  $M$  może znajdować się w danym momencie w jednym ze skończonej liczby stanów:  $q_0, q_1, \dots, q_n$ . W każdym kroku swojej pracy maszyna  $M$  wykonuje następujące działania: zapisuje w czytany przez głowicę kwadracie nowy symbol (stary symbol zostaje wymazany, ale nowy symbol może być taki sam jak stary), przesuwa głowicę o jeden kwadrat w lewo lub w prawo, a na koniec przechodzi do nowego stanu (znów: nowy stan może być taki sam jak stary)., to co maszyna robi w danym kroku, jest wyznaczone przez stan, w którym  $M$  znajduje się bezpośrednio przed tym krokiem, oraz symbol czytany przez głowicę. Jeśli w jakiejś chwili maszyna jest w stanie  $q_i$ , a głowica czyta symbol  $s$ , to w następnym kroku maszyna wykona działania określone przez wartość funkcji  $M$  na parze argumentów  $(i, s)$  – mianowicie, jeśli przykładowo  $M(i, s) = (t, L, j)$ , to maszyna zapisze w czytany kwadracie symbol  $t$ , przesunie się w lewo i przejdzie do stanu  $q_j$ . Zauważmy, że funkcja  $M$  może nie być określona dla pary  $(i, s)$  – oznacza to, że, jeśli w którymkolwiek momencie maszyna przeczyta symbol  $s$ , będąc w stanie  $q_i$ , natychmiast zakończy pracę.

Oznaczmy przez  $\lceil x \rceil$  ciąg jedynek o długości  $x+1$ . Mówimy, że maszyna Turinga  $M$  oblicza funkcję  $f: \omega^n \rightarrow \omega$ , jeśli dla dowolnych  $x_1, \dots, x_n$ , ilekroć  $M$  rozpocznie pracę w stanie  $q_0$ , czytając najbardziej wysunięty na lewo symbol ciągu  $\lceil x_1 \rceil 0 \lceil x_2 \rceil 0 \dots 0 \lceil x_n \rceil$  (zakładamy, że cała reszta taśmy jest wypełniona zerami), to: 1) jeżeli  $f(x_1, \dots, x_n)$  jest określone,  $M$  po jakimś czasie skończy pracę, czytając najbardziej wysunięty na lewo symbol ciągu  $\lceil f(x_1, \dots, x_n) \rceil$ ; 2) jeżeli  $f(x_1, \dots, x_n)$  nie jest określone, to  $M$  nigdy nie skończy pracy.

Funkcja  $f$  jest *rekurencyjna* wtedy i tylko, wtedy gdy istnieje maszyna Turinga  $M$ , która oblicza  $f$ . Relacja jest *rekurencyjna*, jeśli rekurencyjna jest jej funkcja charakterystyczna.

Definicja funkcji rekurencyjnych jako funkcji obliczalnych przez maszyny Turinga dobrze oddaje intuicję stojące za pojęciem rekurencyjności. Można jednak myśleć o funkcjach rekurencyjnych w sposób niezbyt ścisły, ale za to być może jeszcze prostszy i jeszcze bardziej intuicyjny: funkcja jest rekurencyjna, jeśli można by było napisać program kompute-

rowy, który ją oblicza, mając do dyspozycji nieograniczoną pamięć. Uwaga: rekurencyjność jest pojęciem ekstensjonalnym, toteż np. funkcja stale równa 1, jeśli hipoteza Goldbacha jest prawdziwa, a stale równa 0, jeśli hipoteza Goldbacha nie jest prawdziwa – jest jak najbardziej rekurencyjna. Program liczący tę funkcję wcale nie musi rozstrzygać hipotezy Goldbacha: odpowiednim programem będzie albo program dający na wyjściu zawsze 1, albo program dający zawsze 0 (to, że nie wiemy, który z tych programów jest właściwy, jest być może przykre, ale nie ma żadnego znaczenia dla rekurencyjności naszej funkcji).

### *BIBLIOGRAFIA:*

AJDUKIEWICZ K.

[1934] *O stosowalności czystej logiki do zagadnień filozoficznych*, przedruk w: Ajdukiewicz [1960], s. 211-214.

[1937] *Problemat transcendentnego idealizmu w sformułowaniu semantycznym*, przedruk w: Ajdukiewicz [1960], s. 264-277.

[1948] *Epistemologia a semiotyka*, przedruk w: Ajdukiewicz [1965], s. 107-116.

[1960] *Język i poznanie*, t. I, PWN, Warszawa.

[1965] *Język i poznanie*, t. II, PWN, Warszawa.

CHURCH A.

[1936] *An Unsolvable Problem of Elementary Number Theory*, przedruk w: Davis (red.) [1965], s. 89-107.

DAVIS M.

(red.) [1965] *The Undecidable. Basic Papers On Undecidable Propositions, Unsolvable Problems And Computable Functions*, Raven Press, New York.

ENDERTON H. B.

[1977] *Elements of Recursion Theory*, w: J. Barwise (red.), *Handbook of Mathematical Logic*, North-Holland Publishing Company, Amsterdam — New York — Oxford, s. 527-566.

GANDY R.

[1980] *Church's Thesis and Principles for Mechanisms*, w: J. Barwise, H. J. Keisler & K. Kunen (red.), *The Kleene Symposium*, North-Holland Publishing Company, Amsterdam, s. 123-148.

HOFSTADTER D. R.

[1979] *Gödel, Escher, Bach. An Eternal Golden Braid*, Harvester Press, New York.  
20<sup>th</sup>-anniversary Edition: Penguin Books 2000.

IMMERMAN N.

[1998] *Descriptive Complexity*, Springer-Verlag, New York.

KALMÁR L.

[1959] *An Argument Against the Plausibility of Church's Thesis*, w: A. Heyting (red.), *Constructivity in Mathematics. Proceedings of the Colloquium held at Amsterdam 1957*, North-Holland Publishing Company, Amsterdam, s. 72-80.

KLEENE S. C.

[1987] *Reflections on Church's Thesis*, "Notre Dame Journal of Formal Logic", 28, 4, s. 490-498.

KREISEL G.

[1987] *Church's Thesis and the Ideal of Informal Rigour*, "Notre Dame Journal of Formal Logic", 28, 4, s. 499-519.

MENDELSON E.

[1990] *Second Thoughts About Church's Thesis and Mathematical Proofs*, "The Journal of Philosophy", LXXXVII, 5, s. 225-233.

MURAWSKI R.

[1990] *Funkcje rekurencyjne i elementy metamatematyki. Problemy zupełności, rozstrzygalności, twierdzenia Gödla*, Wydawnictwo Naukowe UAM, Poznań; II wyd. 1991.

NELSON R. J.

[1987] *Church's Thesis and Cognitive Science*, "Notre Dame Journal of Formal Logic", 28, 4, s. 581-614.

ODIFREDDI P.

[1989] *Classical Recursion Theory*, North-Holland Publishing Company, Amsterdam — New York — Oxford.

POST E.

[1936] *Finite Combinatory Processes. Formulation I*, przedruk w: Davis (red.) [1965], s. 289-291.

REICHENBACH H.

[1951] *The Rise of Scientific Philosophy*, University of California Press, Berkeley and Los Angeles; II wyd. 1954.

SCHOENFIELD J. R.

[1993] *Recursion Theory*, Springer-Verlag, Berlin — Heidelberg.

SHAPIRO S.

[1981] *Understanding Church's Thesis*, "Journal of Philosophical Logic", 10, s. 353-365.

[1998] *Church's Thesis*, w: E. Craig (red.), *Routledge Encyclopedia of Philosophy*, Routledge, London — New York, t. II, s. 351-355.

SIEG W.

[1997] *Step by Recursive Step: Church's Analysis of Effective Calculability*, "The Bulletin of Symbolic Logic", 3, 2, s. 154-180.

TURING A. M.

[1936] *On Computable Numbers, with an Application to the Entscheidungsproblem*, przedruk w: Davis (red.) [1965], s. 116-154.

WOLEŃSKI J.

[1993] *Metamatematyka a epistemologia*, PWN, Warszawa.

**Piotr Labenz**

## **O pozorności sporu realizm vs formalizm**

Punkt wyjścia stanowi referat dr. Krzysztofa Wójtowicza o pozaformalnym uzasadnianiu w matematyce, przedstawiony 25. września 2001 na II Seminarium Lwowsko-Warszawskim Filozofii Nauki. Streszczam niektóre wyrażone w nim poglądy na temat realizmu quine'owskiego i formalizmu. Następnie analizuję je pod względem twierdzeń o istnieniu przedmiotów i prawdziwości zdań matematycznych. W obu wypadkach rozróżniam dwa sensy odpowiedniego wyrażenia – „wewnętrzny” i „zewnątrzny”. Pokazuję, że rozbieżność obu stanowisk bierze się z ekwiwokacji, a zatem jest pozorna. Kończę próbą uzgodnienia poglądów obu stanowisk na pozaformalne uzasadnianie i metafizycznym wnioskiem o pozorności sporów.

### ***1. Zagadnienie uzasadniania***

W referacie i pracy (Wójtowicz [2000]) autor rozważa zagadnienie uzasadniania zdań matematycznych innego niż dowód. Takie uzasadnianie pozaformalne nie zastępuje dowodu, ani nie jest jego nowym rodzajem, ale dotyczy takich zdań, których udowodnić się nie da, mianowicie aksjomatów i zdań niezależnych (tj. takich, że ani danego zdania, ani jego negacji nie można w danej teorii udowodnić). Co do statusu uzasadniania pozaformalnego zachodzą następujące wątpliwości. (i) Czy w ogóle trzeba uzasadniać aksjomaty i zdania niezależne? (ii) Na czym ma takie uzasadnianie się opierać; w szczególności, czy na prawdziwości (rozumianej klasycznie) uzasadnianego zdania?

Odpowiedzi na te pytania zależą od stanowiska filozoficznego. Autor przedstawia, jaki jest status uzasadniania pozaformalnego według skrajnego formalizmu, formalizmu hilbertowskiego, instrumentalizmu, konceptualizmu (czyli ontologicznej wersji intuicjonizmu), realizmu quine'owskiego i realizmu gödłowskiego. Poniżej skupię się na

przedstawieniu statusu uzasadniania pozaformalnego według skrajnego formalizmu oraz realizmu quine'owskiego (za którym skądinąd opowiada się Wójtowicz [1999])<sup>13</sup>.

**1.1. Realizm quine'owski.** Według tego stanowiska, matematyka stanowi część teorii fizycznych. Istnieć może, to co znajduje się w zakresie zmienności zmiennych w danej teorii – a więc, poza przedmiotami fizycznymi - również przedmioty matematyczne. Zatem zdania matematyczne mają ten sam status, co zdania empiryczne<sup>14</sup>, można więc sensownie mówić o ich prawdziwości. To, że zdania prawdziwe to zdania potwierdzalne empirycznie, dotyczy również zdań matematycznych<sup>15</sup>. Uzasadnianie pozaformalne opiera się więc na prawdziwości i jest potrzebne, by uzasadnić aksjomaty i zdania niezależne występujące w matematycznym instrumentarium teorii fizycznych.

Natomiast o prawdziwości tych zdań matematycznych, które nie znajdują zastosowania w teoriach fizycznych, nie można sensownie mówić. Nie odnoszą się one do niczego, czyli przedmioty matematyczne niepotrzebne w teoriach fizycznych nie istnieją. W tym wypadku uzasadnianie pozaformalne jest zbędne – nie ma sensu się nad nim zastanawiać.

**1.2. Skrajny formalizm.** Według tego stanowiska, matematyka zajmuje się przekształcaniem pewnych napisów. Zdania matematyczne są napisami nie odnoszącymi się do niczego. Nie istnieją przedmioty matematyczne. Dlatego też nie ma sensu mówić o zdaniach matematycznych, że są prawdziwe lub fałszywe. Matematyka nie ma semantyki, jest czysto formalnym systemem syntaktycznym<sup>16</sup>. Zatem względy inne niż składniowe są dla matematyki nieistotne, więc nie ma też powodów, by przyjmować takie a nie inne aksjomaty czy zdania niezależne. Pozaformalne uzasadnianie jest właściwie niepotrzebne.

Niemniej można podawać pozaformalne uzasadnienia, kiedy chce się wybrać jakąś aksjomatykę lub dołączyć jakieś zdanie niezależne do teorii, w której chce się pracować. Uzasadnienia takie opierają się na przesłankach heurystycznych – na tym, czy wyniki badawcze uzyskane dzięki przyjęciu uzasadnianych zdań będą interesujące (np. nowe lub ładne twierdzenia) lub pragmatycznych – czy będą użyteczne w zastosowaniach. Nie opierają

---

<sup>13</sup> Przy czym nie chodzi tu o rekonstrukcję historyczną czyichś poglądów, ale o rozważenie pewnych spójnych stanowisk zrekonstruowanych przez autora.

<sup>14</sup> Czyli odrzuca się podział na zdania analityczne i syntetyczne. Holizm stanowi, obok naturalizmu, jeden z dwóch dogmatów quineizmu.

<sup>15</sup> Przy czym oczywiście nie chodzi tu o potwierdzanie bezpośrednio, ale „jako część teorii”, która „odpowiada przed trybunałem doświadczenia”, jak to poetycko napisał Quine, jako całość.

<sup>16</sup> Autor (również w Wójtowicz [1999]) mówi, że według skrajnego formalizmu matematyka to „gra szklanych paciorków”. Jednak określenie to jest chyba nie całkiem *sine ira et studio*.

się one na prawdziwości uzasadnianych zdań, bo ta zdań matematycznych w ogóle nie dotyczy.

## **2. W sprawie istnienia**

Sądzę, że w powyżej przedstawionych stanowiskach używa się wyrażenia „istnieć” w dwóch różnych sensach. Raz chodzi o istnienie w sensie quine’owskim, to jest o bycie wartością zmiennej – czyli istnienie „wewnętrzne”, zrelatywizowane do danej teorii. Kiedy indziej chodzi o istnienie w sensie absolutnym, „bycie realnym” niezależnie od tego, co twierdzą jakiegokolwiek teorie – czyli istnienie „zewnętrzne”<sup>17</sup>. Dla zwięzłości - będę używał określeń odpowiednio „w-istnieć” i „z-istnieć”, definiując pierwsze po quine’owsku, a drugie – następująco: coś z-istnieje, gdy jest elementem rzeczywistości absolutnej w sensie metafizycznym, „obiektywnie istniejącym bytem”. Innymi słowy, z-istnieje, to co występuje w wyróżnionym modelu absolutnym, identycznym z (tak rozumianą) rzeczywistością.

**2.1. Realizm quine’owski** jest przy tych ustaleniach realizmem w sensie wewnętrznym. Przedmioty matematyczne, według realizmu quine’owskiego, w-istnieją. Mówienie o z-istnieniu czegokolwiek jest bez sensu, bo zobowiązania ontologiczne są zaciągane w ramach poszczególnych teorii – i tylko tak.

Jednak wobec tego wydaje się niekonsekwentnym przekonanie, że przedmioty matematyczne niepotrzebne w teoriach fizycznych nie istnieją. Przedmioty te oczywiście w-istnieją, quine’owski realizm zaprzecza zatem albo ich z-istnieniu, albo w-istnieniu niekonsekwentnie zawężonemu w tym wypadku do bycia wartością zmiennej w teorii akurat fizycznej. Wydaje się, że ma miejsce to ostatnie, bo według quine’owskiego realizmu ze względów empirycznych wybiera się to, którą z konkurencyjnych teorii uznać za zinterpretowaną (Wójtowicz [2000], przyp. 26). Z jednej strony mówi się o istnieniu relatywnie do teorii, czyli w-istnieniu, a z drugiej – niektóre teorie są empirycznie lepsze, i to właśnie je uznaje się za zinterpretowane.

To „uznanie za zinterpretowane” ma charakter zewnętrzny – nawet, jeśli nie pociąga postulowania z-istnienia przedmiotów z tak wyróżnionej teorii, to jednak głosi więcej, niż tylko uznanie w-istnienia tych przedmiotów. Gdyby uznanie to było rozstrzygnięciem ontologicznym, kłóciłoby się ono z twierdzeniem, że zobowiązania ontologiczne są zaciągane tylko w ramach poszczególnych teorii. Dlatego sądzę, że uznanie za zinterpretowane nie

---

<sup>17</sup> O „istnieniu relatywnie do danej teorii T” w przeciwieństwie do „istnienia absolutnego” mówi się w Wójtowicz [2000].

pociąga za sobą przyjęcia z-istnienia, a jego negacja - odrzucenia w-istnienia. Chodzi raczej o heurystyczne twierdzenie o tym, która teoria jest właśnie empirycznie lepsza i w której np. warto pracować.

(Oczywiście quine'owski realista istnienie przedmiotów z teorii uznanej za zinterpretowaną uznaje w sensie mocniejszym, niż tylko w-istnienia. Nie jest to jednak z-istnienie, bo model takiej teorii nie jest według quine'owskiego realisty modelem absolutnym. Innymi słowy, uważa on, że teoria uznana za zinterpretowaną mówi prawdę o rzeczywistości (bo „bierze zobowiązania ontologiczne serio”), ale nie, że jest ona izomorficzna z rzeczywistością w ścisłym sensie, tj., że model jest identyczny z rzeczywistością. Zapewne tu quine'owski realista powie, że według niego właśnie na byciu elementem modelu teorii uznanej za zinterpretowaną polega obiektywne istnienie. Ale tu chodzi o obiektywność w sensie mocniejszym, „metafizycznym”, który on – skądinąd słusznie – odrzuca.

**2.2. Skrajny formalizm** z kolei zaprzecza z-istnieniu przedmiotów matematycznych, a nie - ich w-istnieniu. Skrajny formalizm nie zabrania przecież używania kwantyfikatorów, a wystarczy po czymś kwantyfikować, by to mogło w-istnieć.

Formalista może się przed tym wzbraniać, twierdząc, że odrzuca on przekonanie o „egzystencjalnej mocy” kwantyfikatorów. Nie zmieni to faktu, że skoro się kwantyfikuje, to pewne przedmioty matematyczne z definicji w-istnienia - w-istnieją. Podobnie twierdzenie, że się nie „kwantyfikuje”, a tylko stawia znaczki, nie zmieni tego, że postawienie pewnych znaczków wystarczy, by pewne przedmioty w-zaistniały. Odrzucenie owej „mocy” to tyle, co twierdzenie – jak sądzę, słuszne – że w-istnienie nie pociąga za sobą z-istnienia.

**2.3. Podsumowując:** w kwestii istnienia przedmiotów matematycznych można, nie popadając w sprzeczność (i idąc na tylko nieznaczne ustępstwa), być zarazem skrajnym formalistą i realistą quine'owskim<sup>18</sup>. Trzeba tylko co najwyżej konsekwentnie mówić o w-istnieniu, rezygnując z twierdzeń o „istnieniu” w mocniejszym sensie. Przyjmuje się wówczas w-istnienie przedmiotów matematycznych, a nie przyjmuje się ich z-istnienia. Nie rozstrzyga się, czy się ich z-istnieniu zaprzecza, czy też twierdzi, że mówić o nim nie ma sensu. Konsekwentnie, ograniczając się do mówienia o w-istnieniu, nie mówi się również, która teoria „winna zostać uznana za zinterpretowaną” (Wójtowicz [2000], przyp. 26). Nie postuluje się zatem też innego statusu przedmiotów matematycznych niepotrzebnych w teoriach fizycznych.

---

<sup>18</sup> Może to brzmieć nieco dziwnie, ale jest to konsekwencją używania określeń „realizm”, „platonizm” do stanowiska „wewnętrznego”. Tą tyleż efektywną, co mylącą konwencję wprowadził sam Quine, czemu dziwił się już Carnap, gdy ten określił jego logicyzm mianem „platońskiego realizmu” (Carnap [1950], przypis 5).



### 3. W sprawie prawdziwości

Sądzę również, że w powyżej przedstawionych stanowiskach używa się wyrażenia „prawdziwość” w dwóch różnych sensach. Raz chodzi o prawdziwość definiowaną, tak jak w metalogice, czyli za pomocą spełniania w pewnym modelu. Kiedy indziej chodzi o prawdziwość absolutną (w sensie: nierelatywną względem modelu), czyli zgodność z z-istniejącymi faktami. Tę pierwszą będę nazywał wewnętrzną (w-prawdziwością), tę drugą – zewnętrzną (z-prawdziwością). Pierwsza jest relatywna względem - dowolnego modelu, druga wymaga, by z-istniały elementy modelu absolutnego.

**3.1. Realizm quine’owski** mówi o w-prawdziwości, bo zdania matematyczne wygłasza się według niego relatywnie względem danej teorii. Mogą one być prawdziwe w modelu danej teorii, czyli w-prawdziwe, bez roszczeń do z-prawdziwości.

Ponadto jednak twierdzi się, że o prawdziwości zdań matematycznych rozstrzyga potwierdzalność empiryczna teorii fizycznych, w których te zdania występują. Sądzę, że twierdzenie to nie pociąga uznania z-prawdziwości tych zdań, a jego negacja – odrzucenia ich w-prawdziwości. Chodzi tu raczej o „uznanie teorii za zinterpretowaną” w, jak już sugerowałem, sensie tylko heurystycznym.

Natomiast co do zdań o przedmiotach matematycznych niepotrzebnych w teoriach fizycznych sądzę, że nie odrzuca się ich w-prawdziwości<sup>19</sup>. Z definicji tej ostatniej wynika, że wystarczy, by teoria miała model, by jej zdania mogły być w-prawdziwe. Teorie zawierające zdania o przedmiotach matematycznych niepotrzebnych w teoriach fizycznych mają modele (nawet jeśli niezinterpretowane). Zatem zdania te mogą być w-prawdziwe. Co najwyżej zaprzecza się więc ich w-prawdziwości w modelach teorii fizycznych, albo też odrzuca się ich z-prawdziwość. To pierwsze jest banalnie prawdziwym stwierdzeniem, to drugie zgadza się z rezygnacją z roszczeń do z-prawdziwości.

**3.2. Skrajny formalizm** z kolei odrzuca z-prawdziwość zdań matematycznych, bo zaprzecza z-istnieniu przedmiotów matematycznych. Ale co znaczy twierdzenie, że matematyka nie ma semantyki? Jeżeli ma to być twierdzenie o nieistnieniu modeli teorii matematycznych, to jest ono jawnie fałszywe. Jeżeli ma to być twierdzenie głoszące, że modele są „czysto formalne”, „niezinterpretowane”, nie stanowią „referencjalnej semantyki” – to nie neguje ono w-prawdziwości zdań teorii w tych modelach. Bez względu na to, jakimi

---

<sup>19</sup> Mówiąc o „odrzucaniu” z- lub w-prawdziwości jakichś zdań mam na myśli twierdzenie, że nie ma sensu mówić o ich, odpowiednio, z- lub w-prawdziwości i fałszywości (a nie, że są one z- lub w-fałszywe).

epitetami określi się model, spełnianie działa w nim tak samo i wystarczy dla w-prawdziwości.

Natomiast postulowanie rezygnacji z mówienia o modelach i spełnianiu, byłoby ze strony formalisty nieuprawnionym żądaniem: skoro matematyka jest nauką o pewnych operacjach formalnych, dlaczego miałyby się pewnych takich operacji – w szczególności metamatematyki – zabraniać? Sądzę wobec tego, że formalista nie może odrzucać w-prawdziwości, a jedynie - z-prawdziwość. Odpowiednio - twierdzenia o tym, że zdania matematyczne nie odnoszą się do niczego, należy rozumieć w sensie „zewnątrznym”, to znaczy jako odrzucenie modelu absolutnego dla matematyki.

**3.3. Podsumowując:** i realizm quine’owski, i skrajny formalizm odrzucają z-prawdziwość zdań matematycznych. Oba uznają ich w-prawdziwość, której z definicji nie można nie uznawać bez rezygnacji z metamatematyki. Twierdzenia o tym, że nie ma sensu mówić o prawdziwości zdań matematycznych (skrajnego formalisty, że wszystkich, quine’owskiego realisty, że niepotrzebnych w teoriach fizycznych) są niczym więcej, niż odrzuceniem ich z-prawdziwości. Można więc być, nie popadając w sprzeczność, zarazem realistą quine’owskim i skrajnym formalistą w kwestii prawdziwości zdań matematycznych.

#### **4. Wnioski**

Sądzę, że powyższe uwagi pokazują, że spór realizm-formalizm jest w znacznej mierze sporem pozornym, czy też - werbalnym. Przekonania obu stanowisk dają się w dużej części uzgodnić w postaci stanowiska „wewnętrznego”, uznającego w-istnienie przedmiotów i w-prawdziwość zdań matematycznych, a rezygnującego z mówienia o z-istnieniu i z-prawdziwości. Pozornie sprzeczne twierdzenia obu stanowisk o istnieniu i prawdziwości w filozofii matematyki okazują się być sprzeczne właśnie tylko pozornie, z powodu ekwiwokacji.

Oczywiście takie uzgodnienie stanowisk możliwe jest tylko pod warunkiem przyjęcia definicji w-istnienia i w-prawdziwości. Zarówno formalista, jak i realista mógłby twierdzić, że naruszają one jakieś intuicje właściwe ich stanowiskom. Sądzę jednak, że naruszają one jedynie pewne sposoby mówienia o ich stanowiskach. Na przykład – mówienie, że matematyka według formalisty to „gra szklanych paciorków”, „intelektualna łamigłówka”, a realista to „platonik”<sup>20</sup>

---

<sup>20</sup> Por. przypis 6.

**4.1. Uzasadnianie pozaformalne.** Sądzę ponadto, że takie uzgodnienie dotyczy również poglądów zwolenników obu stanowisk na uzasadnianie pozaformalne. Względędy heurystyczne lub pragmatyczne jako podstawę uzasadniania pozaformalnego dopuszcza nawet skrajny formalizm. Sądzę, że również realizm quine'owski dopuszcza względędy heurystyczne nawet w wypadku zdań matematycznych niepotrzebnych w teoriach fizycznych. Natomiast, kiedy quine'owski realista twierdzi, że podstawą pozaformalnego uzasadniania jest prawdziwość, to chodzi mu o prawdziwość w sensie wyboru pewnej teorii, która ma być empirycznie adekwatniejsza – czyli w sensie pragmatycznym, a nie z-prawdziwości. Zatem w obu wypadkach pozaformalne uzasadnianie sprowadza się do względędy pragmatycznych i heurystycznych.

**4.2. Pozorność,** o której mowa, wymaga na zakończenie nieco więcej uwagi. Twierdę, że znaczna część rozbieżności poglądów sprowadza się do różnic w sposobach mówienia, co próbowałem wyżej pokazać. Nie twierdę jednak, że spór realizm vs formalizm jest zupełnym nieporozumieniem. Istnieją w obrębie obu stanowisk istotne różnice, np. taka, że quine'owski realista „bierze serio” zobowiązania ontologiczne teorii, którą uznaje za zinterpretowaną, a skrajnego formalisty interpretowanie teorii w ogóle nie interesuje. Sądę jednak, że są to różnice intuicji, upodobań, i przekonań, nie mających statusu twierdzeń. Mówiąc swobodnie: quine'owski realista ze skrajnym formalistą nie spierają się o to, który z nich wypowiada prawdę o rzeczywistości, ale o to, którego sposób mówienia jest lepszy – wygodniejszy, logicznie elegantszy i ontologicznie oszczędniejszy. (Co do prawdy o rzeczywistości się są zgodni: obaj stoją na wyżej pokazanym stanowisku „wewnętrzny”.)

Zatem pozorność sporu polega na tym, że nie jest on sporem o fakty, ale - o sposób mówienia o nich. (Wydaje się zresztą, że dotyczy to większości sporów filozoficznych<sup>21</sup>.) Słownictwo sporu stwarza złudzenie, że dyskutuje się istnienie pewnych przedmiotów, podczas gdy przecież o istnieniu przedmiotów rozstrzygają nauki szczegółowe. Próbowałem wyjaśnić to powyżej, nie przecząc jednak temu, że spieranie się o sposoby mówienia jest ważne. Przynajmniej dla filozofów, ponieważ jest ono ich głównym zajęciem<sup>22</sup>.

---

<sup>21</sup> Przykładem sporu filozoficznego, ale nie pozornego jest spór o sceptycyzm („mózg w kadzi”, „demon Kartezjusza”), będący sporem o prawdziwość pewnego zdania nieempirycznego. Przykładem innego pozornego w powyższym sensie sporu filozoficznego jest spór, jak to nazwał Kotarbiński, „fakciarze *vs* rzeczowcy”.

<sup>22</sup> Dziękuję serdecznie dr. Krzysztofowi Wójtowiczowi za ciekawe dyskusje i uwagi merytoryczne oraz prof. Jackowi Jadackiemu za cenne uwagi techniczne.

*BIBLIOGRAFIA:*

CARNAP R.

[1950] Empiricism, Semantics and Ontology, Revue Internationale de Philosophie,  
Vol. 11

WÓJTOWICZ K.

[1999] Realizm mnogościowy, WFiS UW, Warszawa

[2000] „O uzasadnianiu w matematyce”

## Czy spór realizm vs formalizm jest sporem pozornym?

W pracy „O pozorności sporu realizm vs formalizm” przedstawiona zostaje pewna interpretacja statusu sporu pomiędzy formalizmem matematycznym a realizmem quine’owskim. Według Autora, różnice między tymi stanowiskami są pozorne, i wynikają *de facto* z różnych sensów, w jakich używane są kluczowe terminy. Spór zatem ma swoje źródło w ekwiwokacji, zaś usunięcie tej ekwiwokacji pozwala również na pogodzenie tych stanowisk.

Niniejszy artykuł stanowi komentarz do pracy (Labenz [2002]). Nie ma on charakteru rozprawy, ale jedynie głosu w dyskusji. Z konieczności więc pewne kwestie będą pominięte, a inne jedynie zasygnalizowane.<sup>23</sup> Z takiego charakteru pracy wynikają również występujące w niej niekiedy powtórzenia i swobodny sposób wyrażania poglądów. W artykule poddaję interpretacji stanowisko Autora. Być może nie zawsze interpretacja ta w pełni zgodna jest z intencjami Autora, mam jednak nadzieję, że pozwoli na lepsze uwypuklenie różnic w naszych stanowiskach (a przede wszystkim - na lepsze zrozumienie natury problemu).

Artykuł podzielony jest na rozdziały odpowiadające w zasadzie poszczególnym, komentowanym tutaj rozdziałom pracy (Labenz [2002]). Będę się odwoływał do numeracji tam przyjętej. Dla wygody Czytelnika, pewne fragmenty komentowanej pracy zostaną tu zacytowane w całości.

### I. Ad „2. W sprawie istnienia”.

Według Autora, dla porównania stanowisk formalisty i quineisty konieczne jest wskazanie dwóch sensów terminu „istnieć”:

- (i) w-istnienie (sens „wewnętrzny”);
- (ii) z-istnienie (sens „zewnętrzny”).

---

<sup>23</sup> Nie dyskutuję tu np. rozbieżności dotyczących problemu uzasadniania pozaformalnego, ponieważ są one pochodne w stosunku do bardziej zasadniczych - a dyskutowanych dalej - różnic.

(i) Pojęcie „w-istnienia” wywodzi się z koncepcji Quine’a zobowiązań ontologicznych. W-istnienie to istnienie zrelatywizowane do danej teorii. Dany obiekt *O* w-istnieje w myśl teorii *T*, jeśli teoria *T* zakłada istnienie tego obiektu. Problem, co dokładnie znaczy zwrot „teoria *T* zakłada istnienie obiektu *O*” Quine wyjaśnia odwołując się do swojego kryterium kwantyfikatorowego. W myśl tego kryterium „wskaźnikiem ontologii” jest kwantyfikacja, zaś ontologia danej teorii może być poprawnie zidentyfikowana po sprowadzeniu jej do „kanonicznej postaci” - postaci teorii elementarnej.

(ii) Z-istnienie to istnienie - w sformułowaniu Autora - jako element rzeczywistości absolutnej w sensie metafizycznym. Z-istnieje, to co „występuje w wyróżnionym modelu absolutnym, identycznym z (tak rozumianą) rzeczywistością.” (Labenz [2002],).

## II. Ad „2.1. Realizm quine’owski”.

Według Autora, realizm Quine’a „jest przy tych ustaleniach realizmem w sensie wewnętrznym. Przedmioty matematyczne, według realizmu quine’owskiego, w-istnieją. Mówić o z-istnieniu czegokolwiek jest bez sensu, bo zobowiązania ontologiczne są zaciągane w ramach poszczególnych teorii – i tylko tak.” [Labenz 2002]. A zatem nie ma sensu mówić o obiektywnym istnieniu przedmiotów matematycznych, gdyż o istnieniu można mówić jedynie w relatywizacji do pewnej teorii, zaś zobowiązania ontologiczne nie mają statusu „metafizycznej absolutności”.

Przy takiej interpretacji, pojęcie „w-istnienia” jest jedynie pewnym czysto technicznym pojęciem metateoretycznym. W myśl tej interpretacji teoria krasnoludków w-zobowiązuje się do istnienia krasnoludków, arytmetyka w-zobowiązuje się do istnienia liczb, teoria mnogości - do w-istnienia zbiorów, fizyka - do w-istnienia półprzewodników i atomów, etc. Jednak fakty te nie mówią nic o rzeczywistości - gdyż fakt „w-istnienia” w myśl pewnej teorii *T* (dalej będę używał skrótu „w-*T*-istnienie”) nie „przenosi się” (swobodnie mówiąc), na z-istnienie.

Warto tu przypomnieć o tym, że Quine jawnie deklaruje brak zainteresowań dla tradycyjnych rozważań metafizycznych, i że jego celem jest podanie kryterium zobowiązań ontologicznych teorii, a nie analizy metafizyczne w tradycyjnym stylu (Quine [1951] s.163). Quine nie twierdzi, że fakt, iż uznane przez nas teorie empiryczne zobowiązują się do istnienia obiektów pewnego typu, implikuje, że obiekty te istnieją „w absolutnym sensie

metafizycznym”. Nie znaczy to jednak, że problem ustalenia ontologii jest według Quine’a problemem jedynie językoznawczym, nie mającym związku z rzeczywistością. Fakt, że teoria zobowiązuje się do istnienia obiektów pewnego typu, nie implikuje ich „absolutnego istnienia”. Nie znaczy to jednak, że problem istnienia jest pozorny. Stanowisko Quine’a jest - w pewnym sensie - stanowiskiem pośrednim, które można schematycznie scharakteryzować w następujący sposób:

1. Spośród istniejących (możliwych) teorii, niektóre teorie uważane są za teorie opisujące rzeczywistość. Np. teoria ciał astralnych nie jest, zaś teoria grawitacji - tak. Kryterium uznania, czy dana teoria empiryczna winna być uznana za właściwą, jest oczywiście zgodność z doświadczeniem.

Ustalona klasa danych empirycznych może być jednak wyjaśniana na wiele sposobów, zaś teorie fizyczne różne pod względem logicznym mogą być empirycznie równoważne.<sup>24</sup> Nasze decyzje dotyczące wyboru teorii, ewentualnej jej modyfikacji, zharmonizowania teorii naukowej z naszymi bodźcami zmysłowymi, podejmowane są więc z uwzględnieniem czynnika pragmatycznego - np. ekonomii myślenia, operatywności teorii czy też - oszczędności ontologicznej.<sup>25</sup>

2. Po dokonaniu wyboru, dana teoria winna być jednak traktowana „poważnie”. Nie ma ona statusu pozbawionego interpretacji instrumentu. Należy ją interpretować realistycznie. Quine zdecydowanie podkreśla, że pomimo, iż mówimy o przedmiotach postulowanych w danej teorii, to nie odbiera im to bynajmniej statusu realności.<sup>26</sup> Jest bowiem jasne, że zawsze o świecie mówimy w kontekście pewnej teorii, że na jakąś teorię musimy się zdecydować - i że zawsze ten wybór opiera się na pewnych kryteriach. Teoria ta nie stanowi jednak „gry pozorów”.

---

<sup>24</sup> „Teoria jest empirycznie niezdecydowana. Nawet gdybyśmy mieli obserwacyjną wyrocznię, umiejacą przypisywać wartość logiczną każdemu stałemu sprawozdaniu z obserwacji, które daje się wyrazić w naszym języku, nadal nie pozwoliłoby to rozstrzygnąć sprawy wyboru jednej spośród wielkiej liczby możliwych teorii fizycznych, z których każda byłaby całkowicie zgodna z ową wyrocznią.” (Quine [1986] s.68).

<sup>25</sup> W eseju „O tym, co istnieje”, Quine dyskutuje problem wyboru między fizykalistycznym a fenomenalistycznym aparatem pojęciowym wskazując na motywy pragmatyczne (Quine [2000] s. 47).

<sup>26</sup> „Wszystko to, czemu przypisujemy istnienie, jest przedmiotem postulowanym z punktu widzenia opisu procesu budowania teorii, a zarazem jest rzeczywiste z punktu widzenia samej tworzonej teorii. Nie powinniśmy również traktować punktu widzenia teorii jako gry pozorów, zawsze musimy bowiem przyjmować perspektywę tej czy innej teorii – najlepszej, na jaką w danej chwili potrafimy się powołać” (Quine [1999] s.37).

3. Skoro więc pewna teoria  $T_r$  zostaje uznana za teorię opisującą rzeczywistość, to istotne jest zidentyfikowanie jej ontologii. Identyfikacja ontologii  $T_r$  jest jedynym sposobem ustalenia tego, jaka jest rzeczywistość. „Wskaźnik” ontologii (kryterium zobowiązań ontologicznych) jest dany za pomocą kwantyfikatorowego kryterium istnienia. Ono pozwala na identyfikację ontologii teorii - w szczególności teorii  $T_r$ .

4. Tym samym, o ile ontologia teorii  $T_r$  zawiera obiekty matematyczne (dzieje się tak np., wtedy gdy teoria empiryczna  $T_r$  odwołuje się do pewnego instrumentarium matematycznego), to konieczne jest uznanie, iż zobowiązujemy się również do uznania istnienia tych obiektów matematycznych.<sup>27</sup>

W myśl koncepcji Quine’a nie ma sensu mówić o istnieniu oderwanym od przyjętej przez nas teorii świata., to co istnieje, jest identyfikowane przez przyjętą przez nas teorię. W tym sensie istnienie absolutne, istnienie „w mocnym sensie metafizycznym” (jako coś wykraczającego poza „istnienie relatywnie do teorii  $T_r$ ”) nie jest pojęciem „kompatybilnym” z koncepcją Quine’a.

W-T-istnienie (istnienie relatywnie do pewnej teorii T) jest pojęciem czysto technicznym, zdefiniowanym w dość prosty sposób dla szerokiej klasy teorii (choć - według Quine’a - nie dla wszystkich, por. uwagi w rozdziale V).<sup>28</sup> Jednak spośród szerokiej klasy możliwych teorii tylko jedna ( $T_r$  - teoria uznana za opisującą rzeczywistość) zostaje uznana za zinterpretowaną. Pojęcie „w- $T_r$ -istnienia” jest niewątpliwie pojęciem wyróżnionym. Quine nie posługuje się pojęciem „istnienia absolutnego”; w jego koncepcji najsilniejszym (i jedynym) sensem terminu „istnieć” jest właśnie w- $T_r$ -istnienie. Jednak, pomimo, iż nie jest

---

<sup>27</sup> Oczywiście, zobowiązani jesteśmy do uznania istnienia obiektów matematycznych, gdy odwołanie się do instrumentarium zakładającego te obiekty jest niezbędne - stąd nazwa „argument z niezbędności” (*indispensability argument*).

<sup>28</sup> Według Quine’a, warunkiem podania dobrego kryterium istnienia jest sformułowanie teorii w postaci teorii pierwszego rzędu. Tylko wtedy dana jest dobrze określona ontologia. „Pytanie o ontologię ... ma określony sens tylko relatywnie do określonych przekładów na notację ontologicznie zdeteminowaną” (Quine [1986] s.135) Za taką notację Quine uważa notację języka predykatów pierwszego rzędu.



ono „metafizycznie absolutne”, odnosi się ono do rzeczywistości i nie ma statusu jedynie pomocniczego pojęcia metateoretycznego.<sup>29</sup>

Autor pisze dalej:

„Jednak wobec tego wydaje się niekonsekwentnym przekonanie, że przedmioty matematyczne niepotrzebne w teoriach fizycznych nie istnieją. Przedmioty te oczywiście w-istnieją, quine’owski realizm zaprzecza zatem albo ich z-istnieniu, albo w-istnieniu niekonsekwentnie zawężonemu w tym wypadku do bycia wartością zmiennej w teorii akurat fizycznej. Wydaje się, że ma miejsce to ostatnie, bo według quine’owskiego realizmu ze względów empirycznych wybiera się, którą z konkurencyjnych teorii uznać za zinterpretowaną. [...] Dlatego sądzę, że uznanie za zinterpretowane nie pociąga przyjęcia z-istnienia, a jego negacja odrzucenia w-istnienia. Chodzi raczej o heurystyczne twierdzenie o tym, która teoria jest właśnie empirycznie lepsza i w której np. warto pracować.

(Oczywiście quine’owski realista istnienie przedmiotów z teorii uznanej za zinterpretowaną uznaje w sensie mocniejszym, niż tylko w-istnienia. Nie jest to jednak z-istnienie, bo model takiej teorii nie jest według quine’owskiego realisty modelem absolutnym. Innymi słowy, uważa on, że teoria uznana za zinterpretowaną mówi prawdę o rzeczywistości (bo „bierze zobowiązania ontologiczne serio”), ale nie, że jest ona izomorficzna z rzeczywistością w ścisłym sensie, tj., że model jest identyczny z rzeczywistością. Zapewne tu quine’owski realista powie, że według niego właśnie na byciu elementem modelu teorii uznanej za zinterpretowaną polega obiektywne istnienie. Ale tu chodzi o obiektywność w sensie mocniejszym, „metafizycznym”, który on – skądinąd słusznie – odrzuca.” (Labenz [2002] )

Takie postawienie sprawy sugeruje, że u Quine’a występują 3 kategorie:

- z-istnienie (istnienie w „mocnym sensie metafizycznym”);
- w- $T_r$ -istnienie (istnienie zgodnie z wybraną teorią rzeczywistości  $T_r$ );
- w-istnienie (istnienie zgodnie z daną teorią).

Według Autora, Quine, mówiąc o w- $T_r$ -istnieniu nie przypisuje mu charakteru absolutnego, ponieważ model teorii  $T_r$  nie jest modelem absolutnym - nie musi być izomorficzny z rzeczywistością. A tylko w wypadku tej ostatniej można mówić o z-istnieniu.

Sądzę jednak, że wprowadzenie rozróżnienia na „z-istnienie” i „w- $T_r$ -istnienie” nie jest zgodne z koncepcją Quine’a. Podobnie nie ma sensu mówienie w ramach tej koncepcji o „modelu absolutnym”. Tym samym zastanawianie się, czy model dla teorii  $T_r$  jest identyczny

---

<sup>29</sup> Sądzę, że nie byłoby nadużyciem terminologicznym stwierdzenie, że Quine mógłby dopuścić pojęcie „z-istnienia” w sensie „w- $T_r$ -istnienia” - gdyż to właśnie pojęcie jest najsilniejszym pojęciem w koncepcji Quine’a. Uwaga ta jednak ma charakter poboczny i w tekście nie będę dokonywał tego utożsamienia.

z modelem absolutnym, nie ma sensu.<sup>30</sup> Sądę, że lepiej jest powiedzieć, iż Quine sam termin „istnienie” charakteryzuje poprzez swoją teorię zobowiązań ontologicznych.

Oczywiście quineista zgodzi się z tym, że być może nasza teoria świata  $T_r$  nie opisuje świata poprawnie, i że „prawdziwa rzeczywistość” jest inna. Powie jednak jednocześnie, że jedyne, o czym można sensownie mówić, to  $w-T_r$ -istnienie. Zarazem jednak stwierdzi, że to pojęcie „ $w-T_r$ -istnienia” nie jest pojęciem wewnątrzjęzykowym, ale że odnosi się do świata (por. przypis 4, w którym mowa jest o tym, że przedmioty postulowane w naszej teorii świata są przedmiotami rzeczywistymi).

### III. Ad „2.2. Skrajny formalizm”

Należy zauważyć, że samo kryterium  $w$ -istnienia jest kryterium czysto technicznym, niezależnym od jakichkolwiek deklaracji ontologicznych. Kryterium to mówi, że identyfikacja zobowiązań ontologicznych odbywa się poprzez stwierdzenie, jakie zdania, rozpoczynające się od znaczka ‘ $\exists$ ’, są wnioskami z teorii  $T$ .<sup>31</sup> Natomiast wiara w egzystencjalną moc kwantyfikatorów opiera się na dwóch założeniach:

(1). Uznania, że teorie mogą być zinterpretowane lub nie. W wypadku teorii matematycznych kryterium tego, czy teoria jest zinterpretowana, jest to, jaką rolę odgrywa jako instrumentarium w teoriach empirycznych.

(2). Uznania, że konieczne jest zidentyfikowanie ontologii tej zinterpretowanej teorii, zaś takim „wskaźnikiem ontologii” jest właśnie kwantyfikator.

Punkt (2) stanowi jedynie techniczne kryterium  $w$ -istnienia. Natomiast uznanie, że to faktycznie kwantyfikacja zobowiązuje ontologicznie (teorię uznaną za zinterpretowaną) wymaga przyjęcia obu założeń.

---

<sup>30</sup> Abstrahuję tutaj od dodatkowych trudności związanych z tym, że teorie pierwszego rzędu nie są kategoryczne, a zatem mają dużo nieizomorficznych modeli (ściślej: jest tak zawsze, gdy mają modele nieskończone).

<sup>31</sup> Można byłoby nawet stwierdzić, że sam zwrot „teoria  $T$  zobowiązuje się do istnienia obiektów typu  $\phi$ ” jest jedynie przeformulowaniem stwierdzenia „zdanie ‘ $\exists x\phi(x)$ ’ jest wnioskiem z teorii  $T$ ”, a nie tezą dotyczącą rzeczywistości.

#### IV. Ad „2.3. Podsumowując.”

Według Autora,

„[W] kwestii istnienia przedmiotów matematycznych można, nie popadając w sprzeczność (i idąc na tylko nieznaczne ustępstwa), być zarazem skrajnym formalistą i realistą quine’owskim. Trzeba tylko konsekwentnie mówić co najwyżej o w-istnieniu, rezygnując z twierdzeń o „istnieniu” w mocniejszym sensie. Przyjmuje się wówczas w-istnienie przedmiotów matematycznych, a nie przyjmuje się ich z-istnienia. Nie rozstrzyga się, czy się ich z-istnieniu zaprzecza, czy też twierdzi, że mówić o nim nie ma sensu. Konsekwentnie ograniczając się do mówienia o w-istnieniu nie mówi się również, która teoria „winna zostać uznana za zinterpretowaną” [...] Nie postuluje się zatem też innego statusu przedmiotów matematycznych niepotrzebnych w teoriach fizycznych.” (Labenz [2002] ).

Argumentację Autora rozumiem w sposób następujący:

- (1) Zarówno formalista, jak i quineista odrzucają z-istnienie przedmiotów matematycznych.
- (2) Obaj akceptują w-istnienie tychże.
- (3) Więc ich stanowiska w kwestii istnienia obiektów matematycznych są zbieżne. Tym samym można być jednocześnie formalistą i quineistą.

Autor utożsamia więc stanowisko, w myśl którego matematyka pozbawiona jest pozajęzykowego odniesienia, ze stanowiskiem, w myśl którego matematyka takie odniesienie posiada. Różnica między tymi stanowiskami byłaby więc pozorna. Moim zdaniem ten pogląd jest błędny.

Niewątpliwie prawdą jest, że zarówno formalista, jak i quineista mogą wprowadzić pojęcie w-istnienia (ściślej: w-T-istnienia) jako pewne czysto techniczne pojęcie (*de facto* skrót notacyjny dla metateoretycznej tezy „zdanie ‘ $\exists x\varphi(x)$ ’ jest wnioskiem z teorii T”). Prawdą jest też, że formalista odrzuca z-istnienie obiektów matematycznych (czyli istnienie „w mocnym sensie metafizycznym” - o ile zgodzi się, że ta kategoria cokolwiek znaczy). Podobne stanowisko można też przypisać quineiście - jeśli bowiem zwrot „obiekt *o* istnieje w mocnym sensie metafizycznym” miałby znaczyć tyle, co „można podać taki argument za istnieniem obiektu *o*, który byłby sformułowany poza jakąkolwiek teorią rzeczywistości”, to quineista faktycznie zgodzi się, że nic nie istnieje w mocnym sensie metafizycznym (w szczególności obiekty matematyczne).

Na tym jednak kończą się podobieństwa. Co więcej, twierdzę, że podobieństwo tych stanowisk w sprawie z-istnienia jest nieistotne. Zgoda na odrzucenie z-istnienia obiektów matematycznych ma inny charakter w wypadku formalisty i quineisty. Rozważmy przykład

obiektów fizycznych, takich jak krzesła czy słonie. Czy status ontologiczny słonia i liczby 2002 jest taki sam? Jak na to pytanie odpowie formalista, a jak - quineista?

Aby udzielić odpowiedzi na to pytanie, należy przypisać formalistcie matematycznemu jakieś stanowisko dotyczące istnienia świata. Jeśli jest np. solipsystą to oczywiście jego stanowisko będzie odbiegać od stanowiska quineisty. Aby więc prowadzone tu rozważania miały sens, należy założyć, że stanowisko formalisty w sprawie istnienia makroskopowych przedmiotów fizycznych jest podobne do stanowiska quineisty (czyli jest to stanowisko zdroworozsądkowe). Dla formalisty istnieją więc krzesła i słonie, natomiast nie istnieje liczba 2002, ani żadna inna. Jeśli więc formalista w ogóle posługuje się pojęciem z-istnienia, to należy przyjąć, że dla niego z-istnieją słonie i krzesła, natomiast nie z-istnieją liczby.

Można tu odpowiedzieć (sądzę, że takie jest stanowisko Autora), że formalista wcale nie musi uznawać z-istnienia krzesła, a tylko - istnienie w sensie  $w-T_r$ -istnienia quineisty. Formalista - w takiej sytuacji - nie jest zobowiązany do wprowadzenia do swojej koncepcji pojęcia „z-istnienia” jako „metafizycznie silnego” sposobu istnienia. Stwierdzi jedynie, że w przyjętym przez niego sensie terminu „istnieć” liczby nie istnieją. To jednak nie usuwa różnic między formalistą a quineistą. Rozważmy dwie możliwości:

(i) Formalista przyjmuje jeden sposób istnienia i twierdzi po prostu, że słonie i krzesła istnieją, ale liczby - nie.

(ii) Formalista przyjmuje różne sposoby istnienia, i twierdzi następnie, że liczby nie istnieją w tym samym sensie, co krzesła i słonie.

ad (ii). W tym wypadku stanowisko formalisty jest niewątpliwie niezgodne ze stanowiskiem quineisty, który przyjmuje tylko jeden sposób istnienia. Nie da się wówczas obronić tezy Autora, iż stanowiska te są możliwe do pogodzenia.

ad (i) W tym wypadku stanowisko formalisty byłoby do pogodzenia ze stanowiskiem quineisty tylko wówczas, gdyby quineista zgodził się na tezę, że  $w-T_r$ -istnieją słonie, ale nie  $w-T_r$ -istnieją liczby.

Tak jednak nie jest. Quine nie odróżnia sposobów istnienia, zaś jedynym kryterium istnienia jest kwantyfikacja, a nie np. rozmiar, zmysłowa postrzegalność czy obecność w czasoprzestrzeni.<sup>32</sup> Jeśli zatem uznaje się istnienie obiektów, do których zobowiązuje się

---

<sup>32</sup> Uwaga: oczywiście nie należy tej uwagi rozumieć w ten sposób, że dla Quine'a świadectwa zmysłowe nie odgrywają roli. Odgrywają - w tym sensie, że każda teoria świata musi być dopasowana do danych zmysłowych. Jednak w momencie, gdy przyjmujemy daną teorię (tj. uznamy że jest ona odpowiednio zharmonizowana z danymi) to kryterium ontologii tej teorii jest właśnie kwantyfikacja. Może się wówczas okazać, że ontologia naszej teorii zawiera obiekty nieobserwowalne, w szczególności np. właśnie abstrakcyjne.

teoria świata  $T_r$ , to będzie tam Księżyc, pole grawitacyjne, liczba 3, etc. Quine nie zgodzi się przy tym z tezą, że Księżyc i liczba 3 istnieją w różnym sensie (choć oczywiście zgodzi się, że Księżyc jest w czasoprzestrzeni, a liczba 3 - poza czasoprzestrzenią). Zgodzi się, że one nie istnieją w „mocnym sensie metafizycznym” (czy raczej: że nie bardzo jest sens się nad tym zastanawiać, bo tego pojęcia u niego nie ma), natomiast wszystkie te obiekty istnieją w tym samym sensie: wszystkie w- $T_r$ -istnieją, a więc istnieją w najmocniejszym (i jedynym) sensie istnienia, jaki jest możliwy do zaakceptowania. Quineista powie więc: jeśli liczba 3 nie istnieje w najmocniejszym sensie tego słowa, to już nic nie istnieje w tym sensie, bo mój najmocniejszy sens istnienia jest relatywny do  $T_r$ .

Oczywiście formalista (będący zarazem zdroworozsądkowym realistą) nie zgodzi się z takim postawieniem sprawy: dla niego Księżyc istnieje w najmocniejszym sensie, ale liczba 3 - nie istnieje. Nie jest przy tym istotne, czy ten najmocniejszy sens istnienia formalisty nazwiemy „istnieniem w mocnym sensie metafizycznym”, czy - „istnieniem relatywnie do teorii, którą uznajemy za empirycznie adekwatną”.

Stanowiska formalisty i quineisty w sprawie statusu ontologicznego liczb różnią się więc w zasadniczy sposób. Autor wskazuje jedynie na podobieństwa dotyczące w-istnienia i z-istnienia obiektów matematycznych, pomija natomiast - kluczowy dla dyskusji - problem w- $T_r$ -istnienia.<sup>33</sup>

## V. Metodologiczny status pojęcia „w-istnienia”.

Warto zwrócić uwagę na różnice stanowisk formalisty i quineisty w odniesieniu do problemu w-istnienia. Nie chodzi tutaj o różnice dotyczące samej definicji tego pojęcia (tu bowiem różnic nie ma), co jego (tj. problemu) poznawczej doniosłości. Dla quineisty pytanie o zobowiązania ontologiczne teorii matematycznych jest istotnym problemem poznawczym

---

<sup>33</sup> Pomijanie tego problemu prowadzić może np. do utożsamienia mojego stanowiska ze stanowiskiem kogoś, kto wierzy w ektooplazmoidalnych kosmitów, którzy sterują naszymi zachowaniami. To utożsamienie wygląda tak:

1. Obaj zgadzamy się co do pojęcia w-istnienia (więc zgadzamy się co do w-istnienia kosmitów relatywnie do teorii  $T_{\text{kosm}}$ ).

2. Obaj odrzucamy „metafizycznie silne” pojęcie z-istnienia (obaj jesteśmy rozsądni, więc nie wygłaszamy zbyt silnych tez metafizycznych dotyczących „prawdziwej natury” świata), a więc obaj odrzucamy z-istnienie kosmitów.

3. Ergo: nasze stanowiska są zgodne.

Utożsamienie to opiera się na wprowadzeniu pewnej nowej, silnej kategorii, którą obaj odrzucamy. Pomija różnice dotyczące w- $T_r$ -istnienia kosmitów.

(ściśle: jest tak wówczas, gdy dana teoria matematyczna stanowi fragment teorii empirycznej uznanej za zinterpretowaną).

Quine, dyskutując problem zobowiązań ontologicznych, opiera się na pewnych założeniach:

(i) „Kanonyczną” formą, w jakiej powinny być formułowane teorie, jest forma teorii pierwszego rzędu. Ma to związek z jego „tezą o logice pierwszego rzędu”, w myśl której to logika pierwszego rzędu jest „prawdziwą logiką”.

Sformułowanie teorii w tej postaci pozwala na zidentyfikowanie ontologii (poprzez wskazanie zdań egzystencjalnych). Według Quine’a nie jest to jednak możliwe dla wielu systemów wykraczających poza logikę pierwszego rzędu. Quine twierdzi, że nie jest jasne, czy dla tych teorii jest w ogóle dobrze określona ontologia, gdyż alternatywne systemy nie muszą być „ontologicznie zdeterminowane”. Według niego:

„[P]rzyjęcie dodatkowych elementów językowych może zniszczyć [...] ontologiczne kryterium. Załóżmy, na przykład, że ktoś przyjmuje explicite operator dla formowania domkniętych iteratów predykatów, zamiast definiować je na gruncie ontologii zbiorów. Czy wolno twierdzić, że ma bardziej oszczędną ontologię? Powiedziałbym raczej, że uchylił on problem ontologiczny przechodząc do języka, który nie ma wyraźnej ontologii. Jego ontologia jest nieokreślona, a co najwyżej relatywna do pewnego uzgodnionego przekładu tej notacji na naszą, standardową notację” (Quine [1986] s. 134).

Wprowadzenie dodatkowej notacji może dawać zubożony obraz ontologii; pewne elementy notacji mogą z kolei działać odwrotnie, sugerując istnienie przedmiotów, których faktycznie nie ma w zakresie zmienności zmiennych.<sup>34</sup>

(ii) A zatem - pojęcie w-istnienia nie jest dobrze określone dla wszelkich możliwych teorii matematycznych, a tylko dla teorii sformułowanych w języku pierwszego rzędu (lub: dopiero po „przełożeniu” na taką kanoniczną notację.) Fakt ten stanowi argument na rzecz metodologicznej tezy dotyczącej rekonstruowania teorii naukowych w określonych ramach formalnych.

Tego typu zagadnienia nie są podejmowane przez formalistę. Formalista nie uważa pytań dotyczących zobowiązań ontologicznych teorii matematycznych za ważne poznawczo. Tym samym problem sprowadzenia teorii do „kanonicznej notacji” w ogóle nie ma znaczenia

---

<sup>34</sup> W [1969], Quine twierdzi, że problem ten nie jest dobrze postawiony dla języków z kwantyfikatorami Henkina, i że zdania zawierające takie kwantyfikatory muszą być najpierw „przetłumaczone” na postać liniową. Dopiero wtedy staje się jasne, jaka jest ich ontologia.

- może być uznany za problem natury estetycznej (jak np. problem sprowadzenia teorii do postaci, w której 55% zdań zawiera symbol  $\alpha$ ), ale nie - merytoryczny. Pojęcie w-istnienia nie jest dla formalisty „światopoglądowo istotne”, więc problem niedookreślenia tego pojęcia nie jest ważny. Dla formalisty pojęcie „w-istnienia” nie ma wyróżnionego sensu, co najwyżej metamatematyczny: zdania o pewnej postaci syntaktycznej (na początku znaczek  $\exists$ ) są formalnie wyprowadzalne z pewnych teorii. Ale to, że na początku jest znaczek  $\exists$  (a nie  $\forall$ ,  $Q_0$ ,  $\diamond$  albo  $@$  czy cokolwiek innego) nie ma tu głębszego znaczenia.<sup>35</sup>

## VI. Ad „3.3. Podsumowując”.

Autor twierdzi, że

„[I] realizm quine’owski, i skrajny formalizm odrzucają z-prawdziwość zdań matematycznych. Oba uznają ich w-prawdziwość, której z definicji nie można nie uznawać bez rezygnacji z metamatematyki. Twierdzenia o tym, że nie ma sensu mówić o prawdziwości zdań matematycznych (skrajnego formalisty, że wszystkich, quine’owskiego realisty, że niepotrzebnych w teoriach fizycznych) są nie więcej, niż odrzuceniem ich z-prawdziwości. Można więc być, nie popadając w sprzeczność, zarazem realistą quine’owskim i skrajnym formalistą w kwestii prawdziwości zdań matematycznych.” (Labenz [2002] )

Nie zgadzam się z tym. Quineista - w przeciwieństwie do formalisty - odrzuca z-prawdziwość, ale przyjmuje w- $T_r$ -prawdziwość zdań matematycznych. Teza Autora opiera się na zatarciu różnic dotyczących w- $T_r$ -istnienia i w- $T_r$ -prawdziwości (por. IV).

## VII. Ad „4. Wnioski.”

Według Autora

„[S]pór realizm-formalizm jest w znacznej mierze sporem pozornym, czy też werbalnym. Przekonania obu stanowisk dają się w dużej części uzgodnić w postaci stanowiska „wewnętrznego”, uznającego w-istnienie przedmiotów i w-prawdziwość zdań matematycznych, a rezygnującego z mówienia o z-istnieniu i z-

---

<sup>35</sup> Uwaga: nie należy tych stwierdzeń oczywiście rozumieć jako stwierdzeń dotyczących praktyki matematycznej. Formalista, będący jednocześnie matematykiem oczywiście stara się np. rozwiązywać otwarte problemy dotyczące istnienia pewnych obiektów matematycznych, problemy, dotyczące tego, jak silne założenia są konieczne, aby udowodnić istnienie pewnego obiektu, etc. Mowa tutaj tylko o filozoficznej interpretacji tego pojęcia.

prawdziwości. Pozornie sprzeczne twierdzenia obu stanowisk o istnieniu i prawdziwości w filozofii matematyki okazują się być sprzeczne właśnie tylko pozornie, z powodu ekwiwokacji.

Oczywiście takie uzgodnienie stanowisk możliwe jest tylko pod warunkiem przyjęcia definicji w-istnienia i w-prawdziwości. Zarówno formalista, jak i realista mogliby twierdzić, że naruszają one jakieś intuicje właściwe ich stanowiskom. Sądzę jednak, że naruszają one jedynie pewne sposoby mówienia o ich stanowiskach.” (Labenz [2002] )

Zgadzam się, że - o ile wprowadzimy do stanowiska formalistycznego i do stanowiska „quineistycznego” pojęcie z-istnienia jako „istnienia w mocnym sensie metafizycznym”, to stanowiska te będą zgodne w kwestii z-istnienia liczb (ale też: z-istnienia krzesel, słoni i w ogóle wszystkiego!). Oba też są zgodne w kwestii w-istnienia, jako skrótu dla pewnego faktu metamatematycznego (choć należy pamiętać o tym, że formalista i quineista nie zgodzą się co do poznawczej i filozoficznej doniosłości tego pojęcia! - por. V)

Jak starałem się wykazać wyżej, stanowiska te jednak istotnie się różnią, przypisując różny status ontologiczny (relatywnie do najsilniejszego akceptowalnego pojęcia istnienia) konkretnym i obiektom matematycznym.

Intuicja formalisty (który wierzy w słonie, krzesła i pola grawitacyjne, ale nie w liczby) jest taka: w moim silnym (potocznym, filozoficznym czy jakimkolwiek innym) sensie istnienia, istnieją stoły i pola fizyczne, a nie istnieją liczby. „Granica istnienia” leży między konkretnymi i abstraktami.

Intuicja quineisty jest natomiast zupełnie inna: teoria jest przyjmowana jako całość, w szczególności, zobowiązania ontologiczne tej teorii winny być całościowo „traktowane poważnie”. Zobowiązania trzeba jakoś zidentyfikować - i sposobem na to jest kwantyfikatorskie kryterium istnienia. To kryterium nie opiera się na podziale przedmiotów na konkretne i abstrakcyjne: wszystkie obiekty teorii istnieją „w takim samym stopniu”. Nie ma różnych sensów istnienia, co najwyżej różne własności obiektów: jedne są obserwowalne, drugie, nie, jedne abstrakcyjne, inne nie - ale tym, co decyduje o istnieniu jest kwantyfikacja. A zatem w najsilniejszym (i zarazem jedynym przyjmowanym w tej koncepcji) sensie istnienia, istnieją zarówno słonie, jak i pola fizyczne i obiekty matematyczne.

Sądzę więc, że utożsamienie stanowisk formalisty i quineisty faktycznie narusza ich intuicje.



## VIII. Ad „4.2. Pozorność”.

Według Autora

„[Z]naczną część rozbieżności poglądów sprowadza się do różnic w sposobach mówienia, co próbowałem wyżej pokazać. Nie twierdzą jednak, że spór realizm vs formalizm jest zupełnym nieporozumieniem. Są istotne różnice, np. taka, że quine’owski realista „bierze serio” zobowiązania ontologiczne teorii, którą uznaje za zinterpretowaną, a skrajnego formalisty interpretowanie teorii w ogóle nie interesuje. Ale sądzę, że są to różnice intuicji, upodobań, przekonań nie mających statusu twierdzeń. Mówiąc nieco luźniej, quine’owski realista z skrajnym formalistą nie spierają się o to, który z nich wypowiada prawdę o rzeczywistości, ale o to, którego sposób mówienia jest lepszy – wygodniejszy, logicznie elegantszy i ontologicznie oszczędniejszy. (Co do prawdy o rzeczywistości się zgadzają: obaj stoją na wyżej pokazanym stanowisku „wewnętrznym”.)

Zatem pozorność sporu polega na tym, że nie jest on sporem o fakty, ale o sposób mówienia o nich. (Wydaje się zresztą, że dotyczy to większości sporów filozoficznych ) Słownictwo sporu zaciemnia to – np. stwarza złudzenie, że dyskutuje się istnienie pewnych przedmiotów, podczas gdy przecież o istnieniu przedmiotów rozstrzygają nauki szczegółowe. Próbowałem to powyżej wyjaśnić, nie przecząc jednak temu, że spieranie się o sposoby mówienia jest ważne; w końcu stanowi ono główną działalność filozofów.” (Labenz [2002] )

Zgadza się co do tego, że fakt, iż quineista „bierze serio” zobowiązania ontologiczne teorii nie ma statusu twierdzenia. Jego stanowiska filozoficznego nie da się udowodnić, można co najwyżej podawać argumenty na rzecz jego zasadności. Nie można się spodziewać, że będą one konkluzywne. W tym sensie różnica między formalistą a quineistą rzeczywiście jest jedynie różnicą upodobań i przekonań. Jest to jednak - o ile poprawnie interpretuję stanowisko Autora - fakt będący szczególnym przypadkiem szerszego zjawiska: wszelkie różnice poglądów filozoficznych są jedynie różnicami przekonań i upodobań, a nie - dowiedzionych twierdzeń. Przypadek formalizmu/quineizmu nie jest więc wyróżniony.

Nie zgodzę się natomiast z tezą Autora, w myśl której formalista i quineista spierają się jedynie o sposób mówienia, a nie o to, jaka jest prawda o rzeczywistości. Autor utożsamia stanowiska formalizmu i quineizmu na podstawie powierzchownych podobieństw dotyczących w-istnienia i z-istnienia - nie bierze natomiast pod uwagę zasadniczych różnic dotyczących w-T<sub>r</sub>-istnienia. Formalista z quineistą spiera się o to, czy istnieją liczby, czy nie! Według formalisty - matematyka jest niezinterpretowanym systemem: ontologia formalisty nie zawiera liczb. Ontologia quineisty obejmuje również byty matematyczne. Różnica jest zasadnicza. Analogiczna sytuacja występuje np. w wypadku sporu między realizmem teoretycznym a instrumentalizmem: to, czy terminy teoretyczne mają desygnaty,

czy nie, nie jest sporem czysto językowym, a sporem o przyjmowaną ontologię - np. o to, czy istnieją elektrony, czy też nie!

Teza ta wydaje mi się być w tak oczywisty sposób prawdziwa, że nie wierzę, aby Autor odrzucał tak rozumianą tezę. Sądzę więc, że różnica naszych stanowisk wynika z głębokich różnic metafizycznych i innego rozumienia pewnych wyrażen (takich jak: „spierać się o to, jaka jest rzeczywistość”). Tym samym w inny sposób rozumiemy termin „być sporem pozornym”.

Autor z pewnością zgodzi się z tym, że formalista zaakceptuje stwierdzenie „liczby nie istnieją”, zaś quineista „liczby istnieją”. Jednak według Autora, różnica ta tak naprawdę nie jest głęboka, i dotyczy jedynie wygody mówienia. Według formalisty, wygodniej jest powiedzieć „moja ontologia nie obejmuje liczb”, zaś według quineisty wygodniej jest powiedzieć „moja ontologia obejmuje liczby”. Jednak tak naprawdę to żadne z tych zdań nie dotyczy rzeczywistości, a jedynie tego, jak należy się o niej wyrażać. Zarówno formalista jak i quineista są więc w gruncie rzeczy indyferentni w sprawie istnienia obiektów matematycznych, zaś ich *credo* można byłoby sformułować jako:

CREDO FORMALISTY: Wygodnie mi powiedzieć „liczby nie istnieją”, ale nie twierdzę, że ten fakt ma jakiś związek ze światem.

CREDO QUINEISTY: Wygodnie mi powiedzieć „liczby istnieją”, ale nie twierdzę, że ten fakt ma jakiś związek ze światem.

Deklaracje te są skrajnie minimalistyczne i - moim zdaniem - nie oddają ani „ducha” stanowiska Quine’a, ani - formalisty.<sup>36</sup>

Teza Autora, iż to jedynie nauki szczegółowe rozstrzygają o istnieniu przedmiotów, zaś spory filozoficzne dotyczą jedynie sposobów mówienia jest - przy minimalistycznej interpretacji - słuszna. Jednak przy tej interpretacji wszystkie spory ontologiczne (a prawdopodobnie w ogóle wszelkie spory filozoficzne), takie jak: spór między instrumentalizmem a realizmem teoretycznym, formalizmem a realizmem w matematyce, czy fenomenalizmem a fizykalizmem, staną się jedynie sporami o sposoby mówienia. Np. spór między fenomenalizmem a fizykalizmem stanie się jedynie sporem o to, czy wygodniej mówić o świecie w kategoriach wiązek wrażeń, czy kategoriach obiektów fizycznych, od których pochodzą dane zmysłowe. Fakt, że wygodniejszy jest język fizykalistyczny, nie

---

<sup>36</sup> Analogiczne deklaracje wygłaszane np. przez instrumentalistę i realistę teoretycznego brzmiałyby odpowiednio: „wolę mówić ‘elektrony nie istnieją’, ale nie sądę, aby miało to jakiś związek ze światem” ; „wolę mówić ‘elektrony istnieją’, ale nie sądę, aby miało to jakiś związek ze światem”. etc. Nie sądę, aby taka interpretacja była poprawna.

będzie uważany za istotny poznawczo, będzie interpretowany jedynie jako pewien obyczaj językowy. O rzeczywistości, tak czy owak, nie da się nic powiedzieć - skoro więc nie da się mówić o świecie, to mówmy tylko o sposobach mówienia o świecie. Tezy filozoficzne przestają mieć wówczas związek z rzeczywistością, stają się jedynie tezami o języku. Przy takim ujęciu wszystkie stanowiska filozoficzne wygłaszające tezy dotyczące świata (istnieją elektrony, duchy, stoły, ciała astralne, liczby niewymierne, wartości moralne, procesy psychiczne, telepatia, entropia, inflacja, pole grawitacyjne, krasnoludki) są zgodne - w tym sensie, że one w ogóle o świecie nie chcą mówić, różnica polega jedynie na wyborze dogodnych form językowych. Obowiązuje Zasada Pozorności Wszelkich Pytań Dotyczących Rzeczywistości - do świata i tak nie ma żadnego dostępu, więc nie ma sensu zastanawiać się nad tym, jaki ten świat jest. Dostęp mamy jedynie do języka i tylko nim warto się zajmować. Autor zdaje się więc odrzucać - jako zbyt ambitny - cel polegający na opisanu rzeczywistości. Fakt, że porozumiewamy się w pewnym języku i że używamy takich a nie innych form językowych, nie upoważnia nas do głoszenia tez na temat świata. Termin „rzeczywistość” oznacza co najwyżej tylko to, o czym mowa jest w naszych teoriach, zaś jedyne kryteria wyboru to kryteria metajęzykowe.

### **IX. Protokół rozbieżności.**

Sądzę, że zasadnicze różnice pomiędzy Autora a moim rozumieniem natury sporu pomiędzy formalizmem a quineizmem można wyrazić w następujących punktach:

1. Autor nie uwzględnia różnicy pomiędzy stanowiskiem formalisty i quineisty odnoszących się do problem w- $T_1$ -istnienia. Według mnie, różnica ta jest zasadnicza, ponieważ według formalisty, w najsilniejszym przyjmowanym przez niego sensie terminu „istnieć”, istnieją słonie, ale nie - liczby. Według quineisty, istnieją zarówno słonie, jak i liczby (choć jedne są konkretne, a drugie, abstrakcyjne).
2. Zarówno formalista, jak i quineista mogą odrzucić z-istnienie liczb. Zgadzam się z tym. Jednak według Autora, fakt ten stanowi istotny argument na rzecz „kompatybilności” obu tych stanowisk. Nie zgadzam się z tym (por. też przypis 11) .
3. Formalista i quineista zgadzają się co do definicji w-istnienia, nie zgodzą się natomiast co do poznawczej doniosłości tego pojęcia.

4. Według Autora, pytania o rzeczywistość w gruncie rzeczy nie mają sensu - sensowne są tylko pytania dotyczące przyjętych form językowych. Tym samym spór realizm *versus* formalizm jest pozorny.

Czy - według Autora - formalista i quineista nie dostrzegają tej pozorności? Sądzę, że jest raczej tak, że przyjmują oni pewne rozumienie terminu „spór pozorny” i inaczej niż Autor rozstrzygają, kiedy spór jest pozorny, a kiedy nie jest pozorny. Oczywiście, żeby spór nie był pozorny, żeby nie był tylko sporem wynikającym z różnic terminologicznych, konieczne jest przyjęcie pewnych założeń - w tym wypadku założeń dotyczących terminów „istnieć”, „rzeczywistość” etc. Nie są to terminy, których znaczenia dają się dobrze i precyzyjnie zdefiniować poprzez odwołanie do pojęć danych uprzednio. Są to terminy pierwotne, które dopiero „nabierają treści”, ich sens ustalamy w trakcie posługiwania się nimi. I właśnie w rozumieniu tych pojęć przez formalistę i quineistę (i mnie) „tkwi” m.in. to, że spór między formalistą i quineistą nie jest sporem pozornym.

Zgodzę się, że można nadać tym terminom pewne „minimalistyczne” znaczenie, przy którym spór ten staje się pozorny. Takie jest - jak sądzą - rozumienie Autora. Twierdzą jednak, że nie jest to rozumienie właściwe dla stanowiska formalisty ani quineisty. Oni (świadomie!) nadają tym terminom takie znaczenia, które powodują, iż stanowiska ich są rozbieżne. W ich (i moim) rozumieniu, ta rozbieżność nie jest pozorna.

## **X. Wniosek końcowy.**

Sądzę, że rozbieżność stanowisk Autora i mojego wyrasta (abstrahując od zgłoszonych wyżej uwag szczegółowych) w gruncie rzeczy z różnicy metafizycznej, a mianowicie - z innego rozumienia zwrotu „spór jest pozorny”. Mogę więc zakończyć konkluzją:

1. W przyjętym przez Autora rozumieniu terminu „spór pozorny”, spór między formalizmem a quineizmem, jest sporem pozornym (ale pozorne są też prawie wszystkie spory filozoficzne).
2. W moim rozumieniu terminu „spór pozorny”, spór między formalizmem a quineizmem, nie jest sporem pozornym.

Jedną z tez tego artykułu jest to, że to drugie rozumienie jest poprawne. Mam nadzieję, że przekonałem do tej tezy przynajmniej część Czytelników.

## *BIBLIOGRAFIA:*

### **Labenz P.**

[2002] „O pozorności sporu realizm vs formalizm”, w niniejszym zbiorze.

### **Quine W.V.O.**

[1951] „O poglądach Carnapa na ontologię” w: Stanosz B. (red.): *Empiryzm współczesny*, Warszawa, Wydawnictwa Uniwersytetu Warszawskiego, 1991, s.163-172.

[1969] „Existence and Quantification”, w: *Ontological Relativity and Other Essays*, New York: Columbia University Press, s. 91-113.

[1986] *Granice wiedzy i inne eseje filozoficzne*, tłum. B. Stanosz, PIW, Warszawa.

[1999] *Słowo i przedmiot*, tłum. C.Cieśliński, Aletheia, Warszawa.

[2000] *Z punktu widzenia logiki*, tłum. B.Stanosz, Aletheia, Warszawa.

## **FILOZOFIA A TEORIA JĘZYKA**

## **Problemy formalnej teorii języka naturalnego<sup>37</sup>**

### ***FORMALNA TEORIA JĘZYKA***

Mówiąc najogólniej, przez formalną teorię języka naturalnego będę rozumiał interdyscyplinarny nurt badań nad językiem potocznym, które prowadzone są przy wykorzystaniu metod matematycznych, głównie logicznych i algebraicznych. Zasadniczym celem tych badań jest skonstruowanie kategorii, które mogłyby posłużyć do ścisłego opisu zależności i praw rządzących językiem naturalnym. Najważniejszym zadaniem jest tutaj podanie adekwatnej gramatyki języka, tzn. takiej, która opisywałaby mechanizm rozumienia i generowania przez ludzi nieskończenie wielu zdań ze skończonego słownika. Badania te inspirowane są gramatyką generatywną Noama Chomsky'ego. I jak dla Chomsky'ego lingwistyka miała stać się częścią psychologii, czyli badania w jej obrębie nie miały być ograniczone tylko i wyłącznie do opisu mechanizmów językowych, lecz miały także wytłumaczyć pewne aspekty funkcjonowania ludzkiego umysłu, tak i dla współczesnych teoretyków języka ich dyscyplina badawcza jest integralną częścią szerszej nauki o mechanizmach poznawczych człowieka (*cognitive science*). W szczególności zgodnie z modularną teorią umysłu zadanie formalnej teorii języka polega na opisaniu najlepiej dostępnego badaniom modułu umysłu, tj. modułu odpowiedzialnego za zdolności językowe człowieka.

Badania te są prowadzone przy wykorzystaniu metod formalnych, ponieważ ostateczne sprawdzenie tego, czy opracowana teoria języka jest adekwatna, czy nie, polega na jej komputerowym zaimplementowaniu, a następnie - na porównaniu „kompetencji językowej” tego komputera z kompetencją językową ludzi. Mówiąc inaczej i troszeczkę na wyrost, celem formalnej teorii języka jest skonstruowanie maszyny, która pozytywnie przesłaby test Turinga, czyli byłaby w stanie konwersować z człowiekiem w sposób nieodróżnialny od rozmowy z innym człowiekiem.

---

<sup>37</sup> Artykuł ten jest skróconą wersją tekstu referatu wygłoszonego przez autora 1 marca 2002 roku na seminarium prof. Jerzego Pelca pt. „Znak- Język- Rzeczywistość”.

Jak w każdej dyscyplinie naukowej, tak i na gruncie teorii języka, dokonuje się pewnych idealizacji oraz ogranicza się zakres przedmiotu badanego. Przede wszystkim, badacze z tego nurtu zajmują się opisem języka wyidealizowanego, czyli języka, którym posługuje się „abstrakcyjny idealny użytkownik”. W kategoriach generatywistycznych możemy powiedzieć, że opisuje się tutaj idealną kompetencję językową przeważnie abstrahując od aspektów wykonania językowego. Generalnie istnieje tendencja do takiego ograniczenia zadania, które pozwoli ominąć problemy chwilowo zbyt trudne, a przy tym nie pozbawi wyników takich badań waloru wyjaśniającego oraz nie podważy ich doniosłości teoretycznej i metodologicznej. W przykłady tego typu postępowania obfituje niniejszy artykuł. Miejmy nadzieję, że przedstawione tutaj problemy, przed którymi staje formalna teoria języka potocznego, doczekają się innego rozwiązania niż po prostu eliminacja.

### **JĘZYK ZINTERPRETOWANY**

Czym jest dla teorii języka jej przedmiot badań? Najlepszą odpowiedzią na te pytanie jest moim zdaniem przytoczenie następującego cytatu z „*Sprache und Sinn*” Kazimierza Ajdukiewicza:

„Język nie charakteryzuje się jednoznacznie tylko swoim zapasem słów i regułami swej składni, lecz także sposobem, w jakim słowom i wyrażeniom przyporządkowane jest ich znaczenie. Gdyby ktoś posługiwał się słowami języka polskiego, lecz łączył z nimi inne niż zazwyczaj znaczenia, z pewnością nie uznalibyśmy jego języka za język polski, lecz za jakiś kryptojęzyk” [Ajdukiewicz 1934: 149].

Takie pojęcie języka ma swój ścisły odpowiednik w teorii modeli, gdzie językowi  $L$  przyporządkowuje się klasę jego modeli  $M^*$  o strukturze  $\langle U, I \rangle$ , gdzie  $U$  jest uniwersum dyskursu, a  $I$  - funkcją przyporządkowującą wyrażeniom prostym języka odpowiednie przedmioty z  $U$ . W efekcie, pojęcia semantyczne, takie jak: prawda, znaczenie, denotowanie, są zrelatywizowane do poszczególnych modeli języka. Jeżeli wydaje się nam to mało intuicyjne, bo uważamy, że pojęcia semantyczne powinny być zrelatywizowane do języka i tylko do języka, to nie wychodząc poza aparat teorii modeli możemy po prostu ustalić sobie, że przez język  $L^*$  rozumiemy parę uporządkowaną  $\langle L, M \rangle$ , gdzie  $M$  jest dowolnie wybranym (rzeczywistym) modelem języka  $L$ . Tak zdefiniowane pojęcie języka jest ścisłym odpowiednikiem intuicyjnego pojęcia języka zinterpretowanego, czyli pokrywa się z tym, co za język uważał zacytowany wyżej filozof. Wydaje mi się, że argument zawarty w



przytoczonym fragmencie pióra Kazimierza Ajdukiewicza w dostatecznym stopniu przekonuje, iż jest to słuszna eksplikacja terminu język.

### **WARUNKI PRAWDZIWOŚCI**

Wydaje się, że na pierwszy plan wśród zadań teorii języka wysuwa się zagadnienie warunków prawdziwości wyrażen języka potocznego. Ponieważ, aby zdanie  $\alpha$  było komunikatywne dla mojego słuchacza, musi on znać warunki prawdziwości  $\alpha$ . Jeżeli nie wiedziałby, jaka sytuacja w świecie czyni  $\alpha$  zdaniem prawdziwym, to nie potrafiłby odgadnąć, jaką informację chcę mu przekazać, stwierdzając  $\alpha$ . Chcąc opisać zjawisko komunikacji musimy ustalić, jaką informację niosą poszczególne zdania, a do tego celu konieczne jest ustalenie ich warunków prawdziwości. Zacytujmy w tej kwestii Adama Nowaczyka:

„Gdy mówię *Ktoś mi zabrał długopis*, to jest to prawdą zawsze i tylko, wtedy gdy **istnieje** człowiek, który mi zabrał długopis. [ . . . ] Jakkolwiek wielką rolę pełniłyby w komunikacji aspekty pragmatyczne wypowiedzi, ich analiza nie zastąpi analizy **warunków prawdziwości** zdania, bowiem to one w pierwszym rzędzie decydują o tym, jaką informację niesie zdanie. Gdybym nie wiedział w jakich okolicznościach zdanie *Ktoś mi zabrał długopis* bywa prawdziwe, słysząc je nie uzyskałbym żadnej informacji, niezależnie od wszelkich okoliczności.” [Nowaczyk 1999: 31].

Jak scharakteryzować pojęcie prawdy dla jakiegoś języka w sposób nie prowadzący do paradoksów oraz zadowalający filozoficzne intuicje, wiemy od czasów publikacji przez Alfreda Tarskiego pracy „Pojęcie prawdy dla języków nauk dedukcyjnych”. Wypracowane przez Tarskiego i innych logików sposoby radzenia sobie z opisem semantyki języków sformalizowanych były inspiracją dla wielu semiotyków logicznych, którzy próbowali wykorzystać aparaturę powstającej teorii modeli do opisu języka potocznego. Język naturalny pod wieloma względami różni się od języków nauk dedukcyjnych, którymi zajmują się matematycy. Centralnym pytaniem jest kwestia, czy różnica ta jest aż tak wielka, że nie może być mowy o logicznej teorii języka naturalnego, która byłaby jakimś rozszerzeniem teorii języków sformalizowanych. Pytanie to, chyba ciągle pozostaje otwarte. Z historycznego punktu widzenia, można powiedzieć tylko tyle, że optymizm co do adekwatnej formalizacji

języka naturalnego wzrastał wraz z rozwojem samej teorii modeli. Spróbujemy uzasadnić ten pogląd cytując reprezentatywne - naszym zdaniem - fragmenty prac wybitnych logików, którzy przyczynili się do rozwoju nie tylko teorii modeli, ale również - formalnej analizy języka potocznego.

W 1933 roku w fundamentalnej dla całej późniejszej logiki pracy A. Tarski tak pisał o semantyce języka naturalnego:

„Ktoś, kto pragnąłby mimo wszelkie trudności uprawiać ścisłymi metodami semantykę języka potocznego, musiałby uprzednio podjąć się niewdzięcznej pracy nad „reformą” tego języka [ . . .]. Wątpić jednak wolno, czy „zracjonalizowany” na tej drodze język potoczny zachowałby swą cechę naturalności [ . . .]” [Tarski 1933: 158].

Niecałe 25 lat później John Kemeny kontynuował ten sceptycyzm w artykule „*Semantics as a Branch of Logic*” kierowanym do szerokiego grona odbiorców niespecjalistów, bo opublikowanym w Encyklopedii Britannica. Kemeny najpierw rozważa pewne antynomie dające się sformułować w języku naturalnym. Na przykład, możemy wypowiedzieć następujące poprawne zdania języka polskiego:

1. Następne zdanie jest prawdziwe.
2. Poprzednie zdanie jest fałszywe.

Wystarczy chwila namysłu, aby spostrzec, że 1 i 2 dają razem antynomię, bo jeżeli 1 jest prawdziwe, to prawdziwe jest również 2, które głosi, że 1 jest fałszywe, natomiast jeżeli 1 jest fałszywe, to znaczy, że 2 nie jest prawdziwe, czyli innymi słowy - 1 jest prawdziwe. Kemeny komentuje tę sprzeczność w następujący sposób:

„Powyższe paradoksy dowodzą, że język angielski umożliwia wyprowadzenie sprzecznych wniosków przy wykorzystaniu praw logiki, co znaczy, że język angielski sam jest sprzeczny. [ . . .] Nie jest też tak, że ową sprzeczność da się usunąć za pomocą jakiegoś prostego zabiegu.” [Kemeny 1957: 313].

Ale już w 1966 roku w słynnym artykule pt. „Do Natural Languages Contain Paradoxes?” wybitny logik Yehoshua Bar-Hillel dobitnie podważa wypowiedź Kemeny’ego:

„Zakrawa na ironię, że Kemeny próbując nas przekonać o tym, że potoczna angielszczyzna nie jest językiem nadającym się do logicznej argumentacji, czyni to za pomocą argumentacji logicznie nienaganej, sformułowanej w potocznym języku angielskim.” [Bar-Hillel 1966: 394].

Po zgoła 40 latach pracy nad rozwojem teorii modeli w roku 1969 w artykule „Prawda i dowód” sam Tarski zmienił stanowisko w kwestii formalizacji języka naturalnego:

„Używając terminu „języki sformalizowane” nie mamy na myśli wyłącznie systemów językowych sformułowanych całkowicie za pomocą symboli i nie mamy też na myśli niczego, co byłoby zasadniczo przeciwstawne językom naturalnym.” [Tarski 1969: 314].

A rok później uczeń Tarskiego, logik, który chyba najwięcej uczynił dla formalnej teorii języka, Richard Montague, w fundamentalnej pracy „*English as a Formal Language*” wprost zadziwia optymizmem:

„Nie zgadzam się z twierdzeniem, że między językiem formalnym a językiem naturalnym istnieje zasadnicza różnica teoretyczna.” [Montague 1970: 188].

W dalszej części pracy staram się uzasadnić pogląd, że optymizm ten jest co najmniej przedwczesny. Podaję przykłady problemów, przed którymi stanęli logicy analizujący język naturalny, i które, bądź nie doczekały się żadnego ostatecznego rozwiązania, bądź, dla których zaproponowane rozstrzygnięcia wydają się niezadowolające z logicznego lub filozoficznego punktu widzenia.

### ***KOMPOZYCYJNOŚĆ***

Co miał na myśli R. Montague, kiedy pisał, że nie zauważa istotnej różnicy między językiem angielskim a językiem logiki elementarnej [Montague 1970: 189] ? Montague był pełen optymizmu, ponieważ uważał, że relacja pomiędzy składnią i semantyką języka naturalnego nie różni się istotnie od relacji, jaka zachodzi pomiędzy składnią i semantyką języków formalnych. Innymi słowy - Montague był przekonany, że język angielski jest językiem kompozycyjnym.

Reguła kompozycyjności została po raz pierwszy sformułowana przez G. Fregego i głosiła co następuje: znaczenie wyrażenia złożonego jest funkcją znaczeń wyrażenń prostych, które go tworzą, oraz reguł syntaktycznych, zgodnie z którymi te wyrażenia proste są

połączone. Pojęcie znaczenia jest chyba jednym z najbardziej niejasnych terminów współczesnej filozofii, dlatego zamiast o znaczeniu może wygodniej będzie myśleć tutaj o dowolnej własności semantycznej. I wówczas powiemy, że język jest kompozycyjny, gdy własności semantyczne przysługujące wyrażeniu złożonemu są funkcją własności semantycznych przysługujących odpowiednim wyrażeniom składowym oraz reguł syntaktycznych, zgodnie z którymi te wyrażenia proste są połączone. Innymi słowy - język kompozycyjny, to język, w którym forma składniowa zdania jednoznacznie determinuje własności semantyczne tego zdania, a w szczególności - jego wartość logiczną.

Wzorcowym przykładem takiego języka jest klasyczny rachunek zdań, którego alfabet jest *explicite* podany: alfabet KRZ = {FA, SL, SP}, gdzie FA = {p, q, r, s, ..., p', q', ...} - przeliczalny zbiór formuł atomowych, SL = {¬, ⇒} - zbiór stałych logicznych, SP = {(, )} - zbiór symboli pomocniczych. A składnia jest zdefiniowana rekurencyjnie przez jednoznaczne określenie tego, jakie wyrażenia, będące ciągami znaków alfabetu, są poprawnymi wyrażeniami rachunku, czyli należą do zbioru wyrażen poprawnie zbudowanych (wff). Wff to najmniejszy zbiór spełniający następujące warunki:

1. Każda formuła atomowa należy do wff;
2. Jeżeli  $\phi$  należy do wff, to również  $\neg\phi$  należy do wff;
3. Jeżeli  $\phi$  i  $\psi$  należą do wff, to również  $\phi \Rightarrow \psi$  należy do wff.

Semantyka jest opisana przez określenie wartości logicznej (jedyna własność semantyczna jaką mogą posiadać wyrażenia tego języka) formuł złożonych jako funkcji wartości wyrażen prostych, które tworzą te wyrażenia złożone, w następujący sposób: niech

$f: FA \rightarrow \{0, 1\}$ , wówczas:

1. Jeżeli  $\phi$  jest FA, to wartość  $\phi$  wynosi  $f(\phi)$ ;
2. Wartość  $\neg\phi$  wynosi 1, gdy wartość  $\phi$  wynosi 0, w przeciwnym przypadku wartość  $\neg\phi$  wynosi 0;
3. Wartość  $\phi \Rightarrow \psi$  wynosi 0, gdy wartość  $\phi$  wynosi 1, a wartość  $\psi$  wynosi 0, w przeciwnym wypadku wartość  $\phi \Rightarrow \psi$  wynosi 1.

Mówiąc inaczej - język klasycznego rachunku zdań spełnia postulat kompozycyjności. Czy postulat ten spełnia również język potoczny, np. język polski? Poddajmy analizie zjawisko wieloznaczności zasięgu zaimków kwantyfikujących w języku polskim, aby zasugerować, że opisanie języka polskiego w formalizmie, który spełniałby postulat kompozycyjności, nie jest zadaniem łatwym.

Rozważmy zdanie:

4. Każdy student kogoś kocha,

któremu przypisałibyśmy najpewniej następującą formę logiczną:

4'.  $\forall x \exists y [S(x) \Rightarrow K(x, y)]$ .

Ale przypuśćmy, że osoba, która wypowiedziała 4 następnie stwierdziła:

5. Ten ktoś to Julia.

Jak wówczas zinterpretujemy zdanie 4? Zapewne pomyślimy, że osoba ta miała na myśli, iż:

6. Jest ktoś taki, kogo kochają wszyscy studenci.

Zdaniu 6 odpowiada następująca forma logiczna:

6'.  $\exists y \forall x [S(x) \Rightarrow K(x, y)]$ .

Innymi słowy, wypowiadając 4 możemy mieć na myśli zarówno 4' jak i 6'; co więcej, zdania te mają inne warunki prawdziwości, niosą więc różną informację. Przykładowo, w modelu  $M = \langle U, I \rangle$ , gdzie  $U = \{a, b\}$  oraz  $I(S) = \{a\}$  i  $I(K) = \{ \langle a, a \rangle, \langle b, a \rangle \}$ , oba zdania są prawdziwe. Natomiast w modelu  $M' = \langle U', I' \rangle$ , gdzie  $U' = \{a, b, c\}$  oraz  $I(S) = \{a, b, c\}$  i  $I(K) = \{ \langle a, b \rangle, \langle b, c \rangle, \langle c, a \rangle \}$ , prawdziwe jest tylko zdanie 4', natomiast 6' jest fałszywe.

Wniosek jest następujący: forma gramatyczna zdania 4 nie wyznacza w jednoznaczny sposób jego semantycznej interpretacji, dopiero analizowanie zdania 4 w szerszym kontekście, w skład którego wchodzi, np. zdanie 5, pozwala nam stwierdzić, co miała na myśli osoba wypowiadająca 4. Innymi słowy, kompozycyjność języka polskiego nie jest tak oczywista jak kompozycyjność języków formalnych przy analizie, które pomijają czynniki pragmatyczne. Dobra formalna teoria języka powinna uwzględnić czynniki pragmatyczne takie jak kontekst wypowiedzi, a następnie - jasno uwypuklić związek syntaktyki z semantyką w taki sposób, żeby składnia determinowała semantykę. Jest to zadanie niesłychanie trudne i

zarazem chyba najważniejsze dla teorii, której głównym celem jest sformułowanie warunków prawdziwości wszystkich wyrażen języka naturalnego.

Niestety problemy z podaniem kompozycyjnej charakterystyki języka naturalnego wpływają jeszcze z wielu innych źródeł. W istocie kwestie opisane w tej pracy można by podciągnąć pod zagadnienie kompozycyjności logiki opisującej język potoczny, lecz ponieważ w historii semiotyki były one stawiane jako odrębne zagadnienia i tutaj rozważono je osobno.

## ***NIESTRUKTURALNY CHARAKTER JĘZYKA NATURALNEGO***

### I. Wyrażenia okazjonalne.

Alfred Tarski twierdził, że aby można było zdefiniować pojęcie prawdy dla jakiegoś języka, musi on spełniać pewne warunki. Przede wszystkim „funkcja i sens danego wyrażenia musi zależeć tylko od jego formy” a co więcej:

„Nie może się zdarzyć, tak aby zdanie jakieś w jednym kontekście było uznane za prawdziwe, podczas gdy w innym kontekście zdanie tego samego kształtu uznane jest za fałszywe. (Wynika stąd w szczególności, że w słowniku tego języka nie powinno być zaimków i przysłówków wskazujących, takich jak „ten” i „tutaj”).” [Tarski 1969: 313].

Tarski żądał więc, aby analizowany język zadośćuczynił zasadzie Fregego, a głównej przeszkody na tej drodze upatrywał w występowaniu wyrażen okazjonalnych. Najogólniej możemy powiedzieć, że *W* jest wyrażeniem okazjonalnym, gdy w pełnieniu przez *W* funkcji semantycznych istotną rolę odgrywa kontekst pragmatyczny. Do wyrażen takich zaliczymy m. in. zaimek osobowy „ja”, zaimek wskazujący „to” i wiele innych. Jednym słowem okazjonalizmy występują powszechnie w języku codziennym, a rola ich jest ogromna, gdyż niesłychanie ułatwiają komunikację. Gdyby nie istniały, zamiast powiedzieć „Ja napisałem tę pracę” musiałbym wygłosić dłuższe zdanie „Jakub Szymanik napisał tę pracę”. Czy możemy w formalnej teorii języka łatwo uporać się z problemem, którego przyczyną są okazjonalizmy?

Y. Bar-Hillel zaproponował, aby wartości logiczne przypisywać nie zdaniom, lecz wypowiedziom (ang. *judgment*), które formalnie są traktowane jako pary uporządkowane o schematycznej budowie  $\langle s, c \rangle$ , gdzie *s* to odpowiedni egzemplarz zdania, *c* - kontekst wypowiedzi [Bar-Hillel 1954: 364]. Wówczas, na przykład zdanie:

7. Jest mi zimno.

mogłoby być częścią następującego wypowiedzenia:

7'. < egzemplarz zdania: "Jest mi zimno", nadawca: E. Hillary, odbiorca: N. Tenzing, miejsce: Mt. Everest, czas: 29 maja 1953 r., godz. 12.16 >

i wówczas byłoby prawdziwe, natomiast wypowiedziane przeze mnie 21 marca o 10.45 w moim mieszkaniu byłoby zdaniem fałszywym, ponieważ, kiedy piszę te słowa jest mi ciepło.

Propozycja Bar-Hillela, która jest fundamentem tak zwanej logiki wyrażeń okazjonalnych, wydaje się mieć wiele wad. Po pierwsze, budzi zastrzeżenia „estetyczne”. Tarski proponował przypisywanie wartości logicznej zdaniom, czyli twórcom precyzyjnie scharakteryzowanym oraz jednolitym pod względem logicznym. Natomiast Bar-Hillelowskie wypowiedzenia są twórami teoriomnogościowymi. W ich skład wchodzi elementy językowe: egzemplarze zdań, jak również najróżniejsze fragmenty kontekstu, czyli przedmioty należące do materialnego otoczenia wypowiedzi. Elementy te są tak różnorodne jak tylko różnorodny jest świat fizyczny. Po drugie, nie umiemy z góry przewidzieć, jakie elementy kontekstu powinny w konkretnych przypadkach zostać uwzględnione, czyli analiza zdania okazjonalnego musi odwoływać się do intuicji językowej badacza. Ostateczną instancją jest tutaj językoznawca lub biegły użytkownik języka, który decyduje, jakie elementy kontekstu są istotne dla warunków prawdziwości danego zdania, a jakie - nie. A to, przynajmniej moim zdaniem, w znaczny sposób obniża wartość tej teorii dla naukowej lingwistyki.

Wobec tego narzucającym się rozwiązaniem jest podanie procedury eliminacji wyrażeń okazjonalnych, czyli wskazanie efektywnej metody przekładu każdego wypowiedzenia okazjonalnego na jego równoważnik nie - okazjonalny. Jak to zrobić? Czy jest to w ogóle możliwe? Pytania te śmiało możemy zaliczyć do zagadnień, które do tej pory nie zostały rozstrzygnięte. Co wcale nie znaczy, że nikt nie próbował dać na te pytania odpowiedzi. W 1954 roku Bar-Hillel postawił następującą hipotezę:

“Nie w każdej sytuacji komunikacyjnej możemy bez straty informacji dowolne zdanie okazjonalne zastąpić nie okazjonalnym, choć z drugiej strony nie ma takiego zdania okazjonalnego,

którego w pewnych szczególnych okolicznościach nie moglibyśmy adekwatnie zastąpić przez nie-okazjonalne.” [ Bar-Hillel 1954: 369].

Niestety Bar-Hillel za swoją tezę nie argumentuje i ogranicza się tylko do podania nic nie rozstrzygającego przykładu. O ile mi wiadomo, w literaturze semiotycznej dyskusja na temat powyższej hipotezy nigdy nie została podjęta. Może przyczyna tkwi w tym, że pełna analiza hipotezy Bar-Hillela wymaga wyjaśnienia, czym są sytuacje komunikacyjne oraz czym jest informacja niesiona przez zdanie. A eksplikacje pojęć sytuacji komunikacyjnej oraz informacji semantycznej nastroczą filozofom i logikom wiele trudności. Można jednak mieć nadzieję, iż w przeciągu najbliższych lat problemy te zostaną rozwiązane, a wówczas dopiero okaże się, czy Bar-Hillel się nie mylił.

## II Zjawisko nieekstensjonalności

Języki formalne są ekstensjonalne, tzn. wszystkie wyrażenia złożone należące do tych języków są ekstensjonalne. Wyrażenie złożone  $W$  jest ekstensjonalne, gdy własności semantyczne przysługujące  $W$  są funkcją wyłącznie własności semantycznych wyrażeń składowych  $W$ . Warunek ekstensjonalności jest więc innym sformułowaniem zasady kompozycyjności. Język potoczny jest nieekstensjonalny, tzn. występują w nim wyrażenia złożone, których własności semantyczne nie są funkcją własności semantycznych wyrażeń prostych je tworzących.

Świadectwem nieekstensjonalności języka polskiego są licznie występujące złożenia przymiotnikowo-rzeczownikowe oraz przysłówkowo-czasownikowe. Złożenia takie występują w zdaniach:

8. Jan szybko jeździ.
9. Paweł jest dobrym mówcą.

To, że Jan szybko jeździ nie znaczy, iż Jan jest szybki i Jan jeździ tak jak z tego, że Paweł jest dobrym mówcą nie wynika, że Paweł jest dobry, choć wynika, że Paweł jest mówcą. Jak powinna wyglądać logika wyrażeń tego typu?

Zadowolające rozwiązanie tego problemu zaproponował Donald Davidson, który postuluje rozszerzenie uniwersum przez wprowadzenie do niego kategorii zdarzeń, dzięki czemu przymiotniki i przysłówki można charakteryzować jako predykaty stosujące się do tych obiektów [Davidson 1984].



Niech  $U = \{\text{zbiór przedmiotów fizycznych, zdarzenia}\}$ ,  $j$  - Jan,  $P(x, y)$  -  $x$  jeździ na  $y$ ,  $Q(x)$  -  $x$  jest szybki; wówczas zdaniu 8. odpowiada:

$$8'. \forall x [P(x, j) \Rightarrow Q(x)]$$

a zdaniu 9 odpowiada:

$$9'. \forall x [M(x, p) \Rightarrow D(x)],$$

gdzie  $p$  - Paweł,  $M(x, y)$  -  $x$  jest mówieniem przez  $y$ ,  $D(x)$  -  $x$  jest dobre. Z formalnego punktu widzenia analiza ta jest wystarczająca: z 8' wynika, że Jan jeździ, lecz nie wynika, że jest szybki; podobnie 9' mówi nam, że Paweł jest mówcą, lecz nic nie stwierdza o jego dobroci. Rozwiązanie Davidsona można więc właściwie przyjąć i dla teoretyków języka jest ono całkowicie zadowalające. Niestety generuje ono pewne problemy filozoficzne, a mianowicie - powstaje pytanie, jakie jest kryterium identyczności zdarzeń. Czym są zdarzenia? Dyskusja tej kwestii daleko wybiega poza ramy tego artykułu.

Innym przykładem nieekstensjonalności języków etnicznych jest występowanie zdań wyrażających, tzw. postawy propozycjonalne. Szczególnym przykładem tego typu zdań są zdania o wiedzy i przekonaniach, jak na przykład:

10. Cezar wiedział, że stolica Republiki leży nad Tybrem.

11. Cezar wiedział, że stolica papieża leży nad Tybrem.

Pomimo że wyrażenie „stolica Republiki” ma identyczną denotację, jak wyrażenie „stolica papieża”, a więc prawdziwość zdania: „Stolica Republiki leży nad Tybrem” pociąga za sobą prawdziwość zdania „Stolica papieża leży nad Tybrem”, to mamy silne intuicje, że 10 jest prawdziwe, podczas gdy 11 jest fałszywe. Jak logik może wyjaśnić ten fenomen?

Kazimierz Ajdukiewicz zaproponował rozwiązanie polegające na eliminowaniu zdań intensjonalnych przez potraktowanie ich jako zdań eliptycznych i dwuznacznych, będących skrótami pewnych rozwiniętych form, które są już ekstensjonalne [Ajdukiewicz 1961]. I tak w naszym przykładzie rozwiniętą formą zdania 10 jest 10' lub 10''.

10'. Cezar wiedział, że stolica Republiki leży nad Tybrem.

- o stolicy Republiki, o relacji leżenia nad i o Tybrze.

10''. Cezar wiedział, że stolica Republiki leży nad Tybrem.

- o stolicy, o Republice, o relacji leżenia nad i o Tybrze.

Warunki prawdziwości tych dwóch interpretacji zasadniczo się różnią. I tak 10' jest prawdziwe, gdy Cezar rozpoznawał jako prawdziwe zdanie, w którym główny operator denotował relację leżenia nad, pierwszy jego argument denotował stolicę Republiki, a argument drugi denotował Tybr. 10'' jest natomiast prawdą, wtedy i tylko, wtedy gdy Cezar rozpoznawał jako prawdziwe zdanie, w którym główny operator denotował relację leżenia nad, pierwszy jego argument składał się z wyrażenia denotującego stolicę i z wyrażenia denotującego Republikę, a argument drugi denotował Tybr.

Jeżeli uznamy, że 10' jest prawidłowym odczytaniem 10, to zdanie 11 musimy uznać za zdanie prawdziwe, bo 10' ma takie same warunki prawdziwości, co 11. Jeśli jednak przyjmiemy, że to raczej 10'' jest dobrą interpretacją tego, co mamy na myśli wypowiadając 10, to 11 nie jest prawdą, lecz też nie otrzymuje się go z 10'' przez wymianę wyrażeń o tej samej ekstensji.

Stanowisko Ajdukiewicza w sprawie zdań propozycjonalnych można sformułować w następujący sposób: w kontekstach zdań wyrażających postawy propozycjonalne wymienialne są tylko wyrażenia ekstensjonalnie izomorficzne, czyli te, które są tak samo zbudowane z wyrażeń prostych o odpowiednio równych ekstensjach.

Podobne do Ajdukiewiczowskiego rozwiązanie problemu wyrażeń propozycjonalnych opisał w „*Meaning and Necessity*” Rudolf Carnap. Jego zdaniem w kontekstach „*belief sentences*” można zastępować wzajemnie tylko wyrażenia izomorficznie intensjonalne (posiadające tę samą strukturę intensjonalną), czyli wyrażenia zbudowane w taki sam sposób z wyrażeń prostych odpowiednio równych sobie pod względem intensji [Carnap 1956: 230-232]. Zgodnie z tą teorią zdania 13 i 14 są sobie równoważne, ponieważ wyrażenia „ $5+4=9$ ” i „ $\text{Sum}(V, IV)=IX$ ” są intensjonalnie izomorficzne, gdyż są zbudowane analogicznie, tzn. z predykatu identyczności, którego pierwszym argumentem jest wyrażenie złożone z dwóch nazw połączonych symbolem funkcyjnym, a drugim argumentem jest nazwa. Natomiast zdanie 12 posiada inną strukturę intensjonalną niż zdania 13 i 14, ponieważ wyrażenie „ $9=9$ ”, choć jest koekstensjonalne z wyrażeniami „ $5+4=9$ ” oraz „ $\text{Sum}(V, IV)=IX$ ”, to nie jest z nimi intensjonalnie izomorficzne, gdyż jest zbudowane w inny sposób - z różnych elementów: nie z wyrażenia złożonego i nazwy, lecz - z dwóch nazw.

12. Jan wie, że  $9=9$ .
13. Jan wie, że  $5+4 = 9$ .
14. Jan wie, że  $\text{Sum}(V, IV)=IX$ .

Stanowisko Carnapa zostało skrytykowane przez Alonzo Churcha, który zauważył, że pewne wyrażenia intensjonalnie izomorficzne nie są wzajemnie wymienne w zdaniach o wierzeniach [Church 1954]. Rozważmy chociażby strukturę semantyczną, w której:  $P(n)$  to tyle, co „ $n$  jest mniejsze od 3” a  $Q(n)$  znaczy, iż „istnieją  $x, y, z$ , spełniające równanie  $x^n+y^n=z^n$ ”. Ponadto wiemy od ośmiu lat, że Ostatnie Twierdzenie Fermata, głoszące, iż dla dowolnej liczby naturalnej  $n$  większej od dwóch nie istnieją liczby naturalne  $x, y, z$ , spełniające równanie:  $x^n+y^n=z^n$ , jest prawdziwe. Wobec tego predykaty  $P$  i  $Q$  są logicznie równoważne, a ponieważ w omawianym systemie semantycznym są wyrażeniami prostymi (symbolami relacyjnymi), to są one w tym systemie intensjonalnie izomorficzne. Można jednak wierzyć, iż:

$$15. \quad \exists n[Q(n) \wedge \neg P(n)],$$

skoro dowód Twierdzenia Fermata jest bardzo długi i niezwykle wyrafinowany, a ponadto większość matematyków go nie rozumie. Innymi słowy, można uważać, że w dowodzie kryje się błąd, nie wierząc równocześnie, iż:

$$16. \quad \exists n[P(n) \wedge \neg P(n)].$$

Nie jest więc tak, że w zdaniach o wierzeniach wyrażenia intensjonalnie izomorficzne są zastępowalne *salva veritate*.

Rozumowanie Churcha uderza jednak raczej w rozwiązanie Ajdukiewicza, a nie jak się dość powszechnie przyjmuje w stanowisko Carnapa. Ponieważ predykaty  $P$  i  $Q$  chociaż są wyrażeniami atomowymi, to jednak mają inną intensję. Natomiast nie da się zaprzeczyć, że  $P$  i  $Q$  są ekstensjonalnie izomorficzne. Przytoczony powyżej przykład Churcha pokazuje więc tylko, że w kontekstach zdań propozycjonalnych nie są wzajemnie wymienne wyrażenia ekstensjonalnie izomorficzne. Natomiast nie mówi nic o wzajemnej wymienności w tych kontekstach wyrażen intensjonalnie izomorficznych.

Problem zdań propozycjonalnych był w literaturze filozoficznej bardzo szeroko dyskutowany; doskonałe omówienie tego sporu zawiera artykuł Barbary Stanosz pt. „Znaczenie i oznaczanie a paradoks intensjonalności” [Stanosz 1965]. Pojawia się pytanie dlaczego, pomimo iż tak wielu wybitnych ludzi pracowało nad formalną teorią „*belief sentences*”, nic takiego nie zostało stworzone. Wydaje się, iż przyczyna tego zjawiska leży w naturze problemu, która jest wybitnie filozoficzna. Moim zdaniem, nie można stworzyć teorii tego typu zdań, dopóki nie ustalimy, jakie znaczenie mają wyrażenia „Jan wierzy, że p” oraz „Jan wie, że p” a następnie nie podamy warunków prawdziwości zdań zbudowanych zgodnie z tym schematem. Problem warunków prawdziwości zdań o wiedzy i przekonaniach to inne sformułowanie starej epistemologicznej zagadki wiedzy i wiary. Problem ten jest przez filozofów dyskutowany od wielu lat i ciągle nie rysuje się na horyzoncie możliwości wypracowania wspólnego stanowiska, które zadowalałoby szerokie grono epistemologów. Innymi słowy, wydaje mi się, że duża część kłopotów związanych ze zdaniem propozycjonalnymi tkwi w ich niejasności. Może dlatego najlepiej w tej sprawie postępować zgodnie z poglądem Quine’a, czyli przyjąć, że zdania wyrażające postawy propozycjonalne należy po prostu eliminować z języka, gdyż nauka zawierająca takie zwroty jest niedojrzała metodologicznie. Ale wówczas pojawiają się inne równie trudne problemy, np. jak formułować twierdzenia dyscyplin humanistycznych oraz nauk społecznych. Dlatego zanim przeprowadzimy taką pochopną redukcję należałoby najpierw zwrócić się w stronę psychologii kognitywnej i zapytać, czy rzeczywiście nie potrafimy podać żadnych własności, które przysługują ludziom, kiedy wypowiadają zdania propozycjonalne, i tylko wtedy. Jeżeli potrafilibyśmy określić takie własności psychiczne, to może udałoby się precyzyjnie sformułować znaczenie oraz warunki prawdziwości zdań o wierzeniach.

## **ZAKOŃCZENIE**

Gdy przyglądamy się logicznym teoriom języka, zauważamy, że chociaż niejednokrotnie pięknie opisują one pewne rozległe fragmenty języka naturalnego, to niestety ciągle nie istnieje uniwersalna teoria logiczna, opisująca wszystkie aspekty tego języka. Nie jest w nauce sytuacją nową to, że różne fragmenty należące do dziedziny badań opisujemy odwołując się do różnych teorii. Za przykład niech posłuży nam fizyka, w której zjawiska mikroświata są opisywane przez teorię kwantową, a zjawiska makroświata objaśniane są w ogólnej teorii grawitacji i chyba ciągle daleko do Wielkiej Unifikacji. Czy możliwa jest unifikacja w logicznej teorii języka? Czy możemy stworzyć rachunek logiczny, w którym w

adekwatny sposób daje się zinterpretować pewien język etniczny? A może nawet rachunek, w którym interpretowalny jest każdy język etniczny? Na te pytania nikt do tej pory nie udzielił odpowiedzi negatywnej, więc nie istnieją racje teoretyczne, aby zaprzestać poszukiwań. Cała trudność sprowadza się chyba do znalezienia takiej formalnej interpretacji składni i semantyki języka naturalnego, która byłaby zarówno kompozycyjna, jak i wyrażona w wystarczająco silnym logicznie języku, aby objąć swoim zakresem większą część kompetencji językowej człowieka.

### *BIBLIOGRAFIA:*

[1957] *Encyclopedia Britannica*, t. 20, Chicago University Press.

AJDUKIEWICZ K.

[1934] *Język i znaczenie*, przedruk w: Ajdukiewicz [1960], ss. 145-175.

[1960] *Język i poznanie*, t. I, PWN, Warszawa.

[1961] *Pewna metoda eliminacji intensjonalnych zdań i formuł zdaniowych*, przedruk w: Ajdukiewicz [1965], ss. 365-371.

[1965] *Język i poznanie*, t. II, PWN, Warszawa.

BAR-HILLEL Y.

[1954] *Indexical Expressions*, "Mind", 63, ss. 359-379.

[1966] *Do natural languages contain paradoxes?*, "Studium Generale", Jahrgang 19, Heft 7.

CARNAP R.

[1956] *Meaning and Necessity*, Chicago University Press.

CHURCH A.

[1954] *Intensional isomorphism and identity of belief*, "Philosophical Studies", vol. V, n. 5.

DAVIDSON D.

[1967] *Logical Form of Action Sentences*, przedruk w: Davidson [1982], ss. 105-149.

[1982] *Essays on Actions and Events*, Clarendon Press, Oxford.

KEMENY J.

[1957] *Semantics as a Branch of Logic*, w: *Encyclopedia Britannica* [1957].

MONTAGUE R.

[1970] *English as a Formal Language*, przedruk w: Thomason [1974].

NOWACZYK A.

[1999a] *Gramatyka i prawda*, BMS 44, Warszawa.

[1999] *Zaimki kwantyfikujące*, w: Nowaczyk [1999a].

PELC J. (red.)

[1994] *Nauka i język*, BMS 32, Warszawa.

STANOSZ B.

[1965] *Znaczenie i oznaczanie a paradoks intensjonalności*, przedruk w: Stanosz [1999], ss. 35-73.

[1999] *Logika języka naturalnego*, BMS 43, Warszawa.

TARSKI A.

[1933] *Pojęcie prawdy w językach nauk dedukcyjnych*, przedruk w: Tarski [1995], ss. 13-173.

[1969] *Prawda i dowód*, przedruk w: Tarski [1995], ss.292-333.

[1995] *Pisma logiczno-filozoficzne*, t. I, PWN, Warszawa.

THOMASON H. T. (red.)

[1974] *Selected Papers of Richard Montague*, Yale University Press.

## **Zdania Moore'a w świetle teorii implikatury<sup>38</sup>**

W pracy zamierzam omówić paradoksalne zdania Moore'a (paradoks Moore'a<sup>39</sup>), zastanowić się, dlaczego wypowiedzi te uważać można za paradoksalne, oraz przy jakich założeniach taka ich ocena jest nieunikniona<sup>40</sup>.

### **1. OKREŚLENIE**

Wszystkie wypowiedzi, których schematem jest „*p*, ale nie wierzę, że *p*”, gdzie „*p*” reprezentuje dowolne, poprawne zdanie języka polskiego, nazywać dalej będziemy zdaniami (lub paradoksalnymi zdaniami) Moore'a; ich równoznacznym odpowiednikiem są zdania typu „*p*, ale nie jestem przekonany, że *p*”. Zdaniami Moore'a, ani ich równoważnymi odpowiednikami nie są wypowiedzi typu: (1) „*p*, ale on nie jest przekonany, że *p*”, ani (2) „*p*, ale nie byłem przekonany, że *p*”, ani (3) „*p*, ale on wie, że nieprawda, że *p*”, ani (4) „kłamie: *p*”, ani też (5) „jestem przekonany, że *p* oraz jestem przekonany, że nieprawda, że *p*” jak też oczywiście wiele innych, których, jako że nie odegrają dalej żadnej roli, nie podajemy.

### **2. PARADOKSALNOŚĆ**

Dlaczego zdania Moore'a uznaje się za paradoksalne, absurdalne lub nieakceptowalne? Wydaje się, że dopatrywać się można czterech możliwych przyczyn; odrzucenie zdania może być motywowane uznaniem go za wewnętrznie sprzeczne lub założenie o jego prawdziwości czy fałszywości prowadzić może do sprzeczności lub też motywowane jest tym, że nikt nie może takiego zdania użyć jako swej szczerzej wypowiedzi lub też ani jako szczerzej, ani jako nieszczerzej (kłamliwej) wypowiedzi.

---

<sup>38</sup> Praca napisana w ramach projektu badawczego Komitetu Badań Naukowych 5 H01A 021 21

<sup>39</sup> Takie sformułowanie w: W. Lenzen [1978], s. 83

<sup>40</sup> Praca przypominać będzie raczej kolóż myśli niżli konsekwentny zmierzający ku jednoznacznej konkluzji wywód; będzie raczej wyrazem przekory wobec stanowiska Hintikki, rozważaniem zakładającym, iż absurdalność zdań Moore'a nie jest oczywista, dana dla kompetentnego użytkownika języka.

Hintikka (Hintikka [1962], ss. 65 – 66) argumentuje w następujący sposób wykazując, że zdania Moore'a nie są wewnętrznie sprzeczne, oceniając je przy tym jako absurdalne: zmiana gramatycznego podmiotu w zdaniach typu " $p$  oraz nieprawda, że  $p$ " nie usuwa ich wewnętrznej sprzeczności, podczas gdy dzięki zmianie osoby w zdaniu Moore'a, otrzymujemy zupełnie naturalne (nieabsurdalne) zdanie, na przykład typu (1). Zauważa też, że poprzedzenie zdania Moore'a spójnikiem "załóżmy" usuwa poczucie paradoksalności związanej z wypowiedzeniem takiego, zmodyfikowanego zdania ("Załóżmy, że (pies pogryzł sąsiada, ale ja w to nie wierzę)"), ale poprzedzenie tym samym spójnikiem zdania wewnętrznie sprzecznego np. "Załóżmy, że pójdę, lecz nie pójdę" jest tak samo logicznie złe ("*wrong logically*") (wyjąwszy dowodzenie nie wprost) jak samo "Pójdę i nie pójdę". Zgadzam się z Hintikką – zdania Moore'a nie są wewnętrznie sprzeczne, jak też sądzimy, że nie są to zdania antynominalne – założenie o ich prawdziwości lub fałszywości nie prowadzi do sprzeczności: zdania te mogą być w zwykłym sensie (o ile taki sens jest) prawdziwe lub fałszywe. Twierdzenie o tym, że zdania Moore'a mogą być prawdziwe pomoże zrozumieć następujący eksperyment myślowy. Załóżmy, że wygłaszam zdanie Moore'a: "Istnieje życie na Marsie, ale w to nie wierzę". Teraz przyjmijmy – faktycznie, nie wierzę, że istnieje życie na Marsie, ale wyobraźmy sobie, że w to nie wierzę, bowiem nie mam dostępu do najnowszych ustaleń – sonda X właśnie dostarczyła niezbitych dowodów na istnienie prymitywnych form życia białkowego na Marsie. Sensacyjne odkrycie staje się niebawem, po wygłoszeniu przeze mnie zdania Moore'a, powszechnie znane – zmieniam zdanie. Odkrywam ponadto, że to dziwne zdanie wygłoszone przeze mnie było w chwili wygłaszania prawdziwe – życie istniało, a ja w to nie wierzyłem.

Na wewnętrzną niesprzeczność zdań Moore'a wskazuje również fakt, że zmiana formy czasownika "wierzyć" z czasu teraźniejszego na przeszły pozbawia zdania absurdalności – wystarczy przyrównać zdania Moore'a do zdań (2).

Zauważmy, że niewątpliwej wewnętrznej sprzeczności zdania "jestem przekonany, że  $p$  oraz jestem przekonany, że nieprawda, że  $p$ " nie wykażemy (podobnie zresztą jak i zdań Moore'a), pozostając na gruncie logiki klasycznej – potrzebujemy do tego innej logiki, na przykład nadbudowanej nad logiką klasyczną logiki przekonania; jednak wewnętrznej sprzeczności zdań Moore'a nie można wykazać na gruncie żadnych ze znanych formalnych systemów przekonania. Zdania (5) *rażą* nas swą <naturalną> (oczywistą) absurdalnością, a zmiana osoby, czasu, czy też poprzedzenie ich frazą "załóżmy, że" tejże absurdalności nie usuwa. Skoro " $p$  ale nie wierzę, że  $p$ " może być bądź prawdziwe, bądź fałszywe, to dla wyjaśnienia



paradoksalności potrzebujemy innych dystynkcji niż prawdziwy – fałszywy. Najwyższy czas, by zaczęło nurtować nas pytanie: dlaczego wypowiedzenie zdania, które może być prawdziwe i może o czymś nas poinformować, miałyby być absurdalne, niedopuszczalne?

### 3. ROZWIĄZANIE HINTTIKI

Hintikka wykazuje (Hintikka [1962], ss. 67 – 71), że następujący typ zdań, ściśle – jak pisze – związany ze zdaniami Moore’a: (6) “Wierzę, że jest tak jak następuje:  $p$  ale nie wierzę, że  $p$ ” jest – na gruncie formalnego systemu wiary/przekonań – wewnątrznie sprzeczny, w przeciwieństwie do zdania “Wierzę, że jest tak jak następuje:  $p$  ale on nie wierzy, że  $p$ ”, którego niesprzeczność na gruncie tego samego systemu wykazuje.

Jaki jest, zdaniem Hintikki, związek między zdaniami Moore’a a zdaniami typu (6)? Otóż związek ten, nazwany dalej przez Hintikkę *implikowaniem doksastycznym* (Hintikka [1962], s. 76), jest następujący: każdy, kto wygłasza swe zdanie, jest przekonany lub co najmniej może być przekonany, że jest, tak jak twierdzi. To dosyć ogólne, jak przyznaje Hintikka, założenie wystarcza, by powiązać zdania Moore’a ze zdaniami o schemacie (6): wypowiadając “ $p$ , ale nie wierzę, że  $p$ ” miałbym być przekonany (zgodnie z ową zasadą), że jest, tak jak następuje:  $p$  ale nie wierzę, że  $p$ . Inaczej sformułowana ta sama “dość ogólna” zasada to: “Absurdem dla każdego jest wypowiedzenie sądu, w który nie można niesprzecznie wierzyć” lub: nie wypowiadaj zdań, co do których nie możesz być przekonany.

“Nie możesz być przekonany” użyte zostało w poprzednim zdaniu w sposób nieco odmienny od tego, z jakim mamy do czynienia w wypowiedziach: “Nie możesz być przekonany, że jest tak a tak – toż to nierozsądne” lub “Nie możesz być przekonany, że jest tak a tak – sam widziałeś, że jest inaczej” lub “Nie możesz być przekonany – sam obiecałeś, że nie dasz temu nigdy wiary”. Gdy mówimy tak, by nie wypowiadać zdań, co do których nie możemy być przekonani, to – w przypadku zdań Moore’a – mamy przede wszystkim na myśli to, że ze zdania stwierdzającego przekonanie co do treści zdania Moore’a wynika (na gruncie ustalonej logiki przekonań) sprzeczność. Nie chodzi w “ogólnej zasadzie” o to, by nie wygłaszać sądów nie mając ku temu np. wystarczających podstaw – aż tak mocno nie musi być ona rozumiana; wystarczy, by przekonanie o tym, co wygłaszane zdanie mówi, nie było wewnątrznie sprzeczne. Jeżeli twoja wypowiedź ma być akceptowalna, to możesz niesprzecznie dać jej wiare.

Można się zatem spodziewać, że każdy, kto podziela owo ogólne założenie wraz z przekonaniem o wewnętrznej sprzeczności (6), oraz uważa, że nie można żywić sprzecznych

przekonań, ten nie wypowie zdania Moore'a – po racjonalnych jednostkach możemy spodziewać się zachowań zgodnych z rozsądnymi wymogami. Oczywiście, mógłby ktoś taki wypowiedzieć jedno z takich zdań będąc do tego jakoś przymuszony, *zartem*, lub – sprzeniewierzając się oczekiwaniom towarzyszącym konwersacji – *kłamiąc*.

Oczywiste jest, że żadne z powyższych trzech założeń nie jest ani powszechnie podzielane, ani, to że formułowane w nich zalecenia są przestrzegane. Inną sprawą jest rozstrzygnięcie, czy racjonalny (*idealny*) podmiot, mówca mógłby racjonalnie odrzucić któreś z tych założeń? Jeśli tak zrobić mógłby, to jakimi przesłankami (jest to podmiot racjonalny, więc nie odrzuca twierdzeń bezpodstawnie) mógłby być motywowany?

Jakkolwiek zdaniem niektórych myślicieli<sup>41</sup>, idealny podmiot może żywić <sprzeczne>, tzn. o postaci “*p i nieprawda, że p*” przekonania, to taką możliwość odrzucamy; interesują nas te możliwości, które zakładają niesprzeczność przekonań.

#### **4. DWIE MOŻLIWOŚCI KWESTIONOWANIA ZASADY**

Pozostały nam w takim razie dwa warianty, które mogą podlegać rewizji lub odrzuceniu. Po pierwsze, czym można by motywować osłabienie “ogólnej zasady”: nie dopuszczaj do tego, że zarazem wypowiedzasz zdanie a nie jesteś przekonany, że jest, tak jak ono głosi? Jakie osłabienie takiego zalecenia można uznać za akceptowalne? Po drugie, jakie założenia pozwalają nam stwierdzić, czy zdania (5) są wewnętrznie sprzeczne? Które z założeń jest najmniej wiarygodne? Jaki jest koszt ewentualnego pozbycia się takiego, najmniej wiarygodnego założenia? Rozpatrywaniem tychże dwóch dróg zajmiemy się niżej.

##### **4.1. MOŻLIWOŚĆ PIERWSZA**

Jeżeli chodzi o sposób pierwszy, to owszem, można argumentować na rzecz jakichś innych *zaleceń* wiążących wypowiedzi mówców z ich przekonaniami, które przedkładałyby ponad “dość ogólne założenie” na przykład formułę następującej treści: (#) jeśli wypowiedzasz: *p*, to nie jesteś przekonany, że *p* (trzeba pokazać, że zdanie Moore'a może być kłamstwem). Trudno byłoby tylko znaleźć dobre i szczerze, racjonalne uzasadnienie takiego polecenia – zbyt przekonani jesteśmy o *dziwaczności* i bezowocności takiego *dictum*. Przeciwno zasadności takiej reguły - przemawia jej *niecodziennosc*; zmiana postępowania na takie, które byłoby najbardziej zbliżone do zaleceń formułowanych w takiej maksymie, wymusiłaby zmianę naszych przyzwyczajęń komunikacyjnych, a efekt nie byłby z punktu widzenia rozpatrywanych tu zagadnień zachęcający. Nawet gdybyśmy jakieś wystarczające dla takiej

---

<sup>41</sup>Omówienie stanowisk w: Poczobut [2000]; tam też liczna bibliografia

zmiany usprawiedliwienie znaleźli, to łatwo sformułować odpowiednik zdania Moore'a dla tych, którzy by taką zasadę przyjęli. Byłaby taką każda wypowiedź typu: "*p* lub jestem przekonany, że *p*". Takie zdanie przysporzyłoby – akceptującym tę odmienną zasadę – tych samych wątpliwości, których nam dostarczają zdania Moore'a – nie mogliby wypowiedzieć go będąc przekonani, że jest fałszywe, tak jak poleca to zasada (#). Reguła taka może mieć jednak ograniczone, lokalne zastosowanie wobec tych interlokutorów, co do których żywimy obawę, że starają się nas zawsze świadomie wprowadzić w błąd, że starają się nas oszukać, okłamać.

Sądzę jednak – co zaprezentuję dalej – że istnieje takie logiczne osłabienie Hintikkowskiego wymogu, które nie musi razić swą kontrowersyjnością, nie musi też zmieniać naszych przyzwyczajęń językowych, a konsekwencją jego przyjęcia (być może dla niektórych wystarczającą do odrzucenia takiej możliwości) jest dopuszczenie zdań Moore'a do rodziny wypowiedzi nieabsurdalnych.

Propozycja nasza przypominać będzie jedną z podmaksym konwersacyjnych Grice'a i brzmieć będzie następująco: nie wygłaszaj zdań, co do których przekonany jesteś, że są fałszem. Jeżeli teraz przyjmiemy, a nie jest to założenie kontrowersyjne, że jeżeli jest się przekonany, że *p*, to nie jest się przekonany, że nieprawda, że *p*, to widać, że – faktycznie – nasza reguła jest logicznym osłabieniem maksymy Hintikkowskiej, z wszystkimi tego konsekwencjami; każda akceptowalna dla Hintikki wypowiedź będzie więc taką dla nas, każda dla nas nieakceptowalna będzie taką dla Hintikki, ale niektóre akceptowalne z naszego punktu widzenia okażą się nieakceptowalne dla Hintikki.

Również naszą propozycję ująć można, jak myślę, przystępniej: jeżeli formułuję (akceptowalną pod każdym innym względem) wypowiedź, to dopuszczam (nie wykluczam), że jest, tak jak mówię. Czy na gruncie takiej zasady zdanie Moore'a jest dopuszczalne, akceptowalne? Wszystko wskazuje na to, że tak – formuła  $\neg B \rightarrow (p \wedge \neg Bp)$  (która jest schematem zdania "nie wykluczam, że jest, tak jak następuje: *p*, a ja w to nie wierzę") nie jest wewnętrznie sprzeczna na gruncie systemu Hintikki (Hintikka [1962]), czy Tokarza (Tokarz [1993]). Skoro tak, to w odpowiednich okolicznościach (wypowiedź musi być przecież chociażby relewantna, etc.) zdanie Moore'a może być akceptowalnie wypowiedziane.

Jak mógłby Hintikka argumentować – nie odwołując się do kwestii potencjalnej nieakceptowalności zdań Moore'a, w której się różnimy – przeciw naszej zasadzie? Musiałby wskazać na jakąś wypowiedź, która byłaby ewidentnie niedopuszczalna, a której wypowiedzenie dopuszczalibyśmy. Być może jedynym argumentem wspierającym zasadę

Hintikki, a przemawiającym na naszą niekorzyść, jest właśnie *absurdalność* zdań Moore'a. Niestety, należy zwrócić uwagę na przykrą możliwość szczerego, zgodnego z przekonaniem wypowiedzenia dwóch wzajemnie sprzecznych sądów; dopuszczamy bowiem, że można nie być przekonanym ani co do prawdziwości (ani co do pewnego zdania), ani - co do fałszywości pewnego sądu (ani co do jego negacji), czego słownym wyrazem – zgodnie z naszą regułą – jest wygłoszenie wzajemnie sprzecznych zdań. Niestety rzadko kiedy uznajemy wypowiedź np. "Pada i nie pada" za wyraz niepewności, co do faktycznego procesu meteorologicznego.

Sprzecznosc nie jest własnością pożądaną, wyjąwszy dowody nie wprost. Sprzecznosc znaleziona w teorii, teorię tę unieważnia. Sprzecznosc czyichś przekonania – wyrażoną w pewien sposób w słowach – w jakiś sposób dezawuuje osobę o takich przekonaniach. Ktoś, kto żywi sprzeczne przekonania wierzy, że jedno zdanie jest (może być) prawdziwe i fałszywe zarazem; teoria sprzeczna to ta, wśród której też znajdziemy zdanie wraz z jego negacją (zaprzeczeniem) – nawet teoria sformułowana przy użyciu jedynie spójnika binegacji może być w tym rozumieniu, dzięki możliwości wzajemnego definiowania spójników zdaniowych, sprzeczna. Wydaje się, że teoretycznie wartościowym, dogodnym i wielokrotnie w dociekaniach logiczno-epistemicznych przyjmowanym założeniem jest ograniczenie badań i analiz do podmiotów, które nie żywią sprzecznych przekonania. Podobnie, warto też jest przyjąć, że są one również dość naiwne – nie dopuszczają myśli, że jeden z nich mógłby mieć sprzeczne sądy. Można domniemywać, że wyrazem takich sprzecznych przeświadczeń są właśnie sprzeczne wypowiedzi.

Zostawmy chwilowo zasadę Hintikki i naszą i zastanówmy się, czy racjonalny podmiot może użyć (albo czy też niekiedy faktycznie używa) jako swej wypowiedzi zdania sprzecznego? W pierwszym odruchu skłonni bylibyśmy udzielić odpowiedzi przeczącej – nie, nie może, bowiem "racjonalny" rozumiemy m.in. jako "unikający sprzeczności". Zastanówmy się jednak dłużej nad tym, czy taka wypowiedź może dostarczyć jakiegś w ogóle informacji? Być może wypowiedź "pada, ale i nie pada" – w pewnych kontekstach – mogła by zostać powiedziana jako żart, zaznaczenie dystansu do tez meteorologicznych oponenta, sarkazm, wytknięcie komuś sprzeczności (wtedy taką wypowiedź można traktować jako cytat (przyczenie) czyjejś wypowiedzi, lub takiej wypowiedzi konsekwencję). Być może wypowiedzi sprzeczną "pada i nie pada" wyrazić ktoś mógłby kłopoty z właściwym rozpoznaniem stanu pogody – pada, czy nie pada? Tłumaczyło to by pewną powszedniość tego typu wypowiedzi bez konieczności przyjmowania tego, że są one wyrazem

patologicznego żywienia niespójnych poglądów. Wypowiedzi takie można traktować jako wyraz przekonań niejasnych, niepewnych – “wziąć parasol, czy nie?”. W ten sposób nie różniłyby się wypowiedzi te od swych semantycznych negatywów, prawd logicznych “pada lub nie pada”, które – w odpowiednich okolicznościach – wyrażają niewiedzę, w tym przypadku, co do stanu pogody. Oczywiście uznanie wypowiedzi sprzecznych jako wyrażających podobne postawy propozycjonalne co tautologie może wydać się pochopte i groźne.

Kolejną nieintuicyjną konsekwencją naszej zasady jest to, że ten, kto jej przestrzega, nie może wypowiedzieć zdania wewnątrznie sprzecznego (znaczyłoby to bowiem, że nie jest przekonany, co do pewnej tautologii (zaprzeczenia kontrtautologii, sprzeczności), a co z kolei jest niezgodne z naszym założeniem), lecz może – bez dawania wyrazu sprzecznym przekonaniom – wypowiedzieć dwa zdania wzajemnie sprzeczne osobno (może być bowiem tak, że  $\neg B\alpha$  i  $\neg B\neg\alpha$  należą do jego przekonań).

Czy przyjmując naszą zasadę można sformułować jakiś odpowiednik zdań Moore’a? Czy potrafimy sformułować taką wypowiedź, która byłaby nieakceptowalna na gruncie naszej maksymy, a której wewnętrznej sprzeczności nie można by wykazać posługując się logiką klasyczną? Oczywiście, każdym takim zdaniem będzie np. **LB**-kontrtautologia (Tokarz [1993], s. 168), np.  $B(p \wedge \neg Bp)$ ; każda wypowiedź **LB**-kontrtautologiczna (np. (6): “Jestem przekonany, że jest, tak jak następuje:  $p$  ale nie wierzę, że  $p$ ”) będzie – przy zaakceptowanej naszej regule – niedopuszczalna. Podobnie niedopuszczalną będzie: “ $p$  ale zarazem nieprawda, że  $p$ ” etc. Pozostaje jeszcze przyznać, że przyjmując zasadę Hintikki zaakceptować można wypowiedź “Dopuszczam, że  $p$  a ja w to nie wierzę”, gdyż można o tym być niesprzecznie przekonanym.

Na gruncie zasady Hintikki, chcąc poprawnie powiedzieć “Jestem przekonany, że  $p$ ”, (to warunek konieczny, niewystarczający) musiałbym być przekonany, że jestem przekonany, że  $p$ . Jednak Hintikka odrzucając logiczną prawdziwość formuły  $BB\alpha \Rightarrow B\alpha$  dopuszcza możliwość, że ktoś jest przekonany, że jest przekonany, że  $p$ , ale nie jest przekonany, że  $p$ . Dzięki temu możemy powiedzieć, że powyższa wypowiedź (“jestem przekonany, że  $p$ ”) może być fałszywa, lecz zarazem – ze względu na spełnienie zasady Hintikki – akceptowalna, nieabsurdalna. Nie wiem, czy taka sytuacja wydałaby się dla Hintikki dziwna – być może trudniej byłoby się pogodzić z tym, że mamy wypowiedzi prawdziwe, lecz nieakceptowalne (ze względu na określone reguły), ale, by w naszych warunkach móc to stwierdzić (*rozdzielić* prawdziwość i akceptowalność) potrzebowalibyśmy (przy regule Hintikki) odrzucenia  $B\alpha \Rightarrow$

$\mathbf{B}\mathbf{B}\alpha$ ; tego jednak Hintikka nie robi. Wydawać się może dziwne jednak to, że nie samo przekonanie, że  $p$  jest wymagane dla dopuszczalności wypowiedzi “Jestem przekonany, że  $p$ ”, lecz niezbędne jest *przekonanie wyższego rzędu*: bycie przekonanym, że jest się przekonanym, że  $p$ , które – zdaniem Hintikki – istotnie różni się od *prostego* bycia przekonanym, że  $p$ . Przy mojej regule powiedzenie: “Jestem przekonany, że  $p$ ” byłoby poprawne, o ile nie byłoby fałszem, że jestem przekonany, że  $p$ . Możemy tak twierdzić, bowiem  $\neg\mathbf{B}\neg\mathbf{B}p \Rightarrow \mathbf{B}p$  (przyjmujemy bowiem  $\neg\mathbf{B}\alpha \Rightarrow \mathbf{B}\neg\mathbf{B}\alpha$ ). Innymi słowy: by akceptowalnie powiedzieć: “jestem przekonany, że  $p$ ”, (to warunek konieczny, niewystarczający) muszę, po prostu, być przekonany, że  $p$ .

Na koniec tej partii rozważań przedstawić chcielibyśmy propozycję zmiany logiki implikatury (Tokarz [1993], s. 168) oraz algorytmu Gazdara (Gazdar [1979], Kański [1991]), uwzględniającą i będącą w zgodzie z zaproponowaną przez nas regułą. Logika implikatury zachowałaby w niezmienionej postaci część określającą funktor  $\mathbf{B}$  (“bycia przekonanym”):

Ax.1.  $\alpha$  jeśli  $\alpha \in \text{Taut}$

Ax.2.  $\mathbf{B}\alpha \leftrightarrow \mathbf{B}\mathbf{B}\alpha$

Ax.3.  $\neg\mathbf{B}\alpha \leftrightarrow \mathbf{B}\neg\mathbf{B}\alpha$

Ax.3.  $\mathbf{B}\neg\alpha \Rightarrow \neg\mathbf{B}\alpha$

Ax.5.  $\mathbf{B}(\alpha \Rightarrow \beta) \Rightarrow (\mathbf{B}\alpha \Rightarrow \mathbf{B}\beta)$

Ax. (O) Jeśli  $\alpha$  i  $\beta$  są formułami logiki  $\mathbf{L}$ , to jeśli  $\text{Var}(\alpha) \supseteq \text{Var}(\beta)$  oraz  $l(\alpha) > l(\beta)$  oraz  $\alpha \vdash_{\mathbf{L}} \beta$ , to jeśli  $\gamma \rightarrow \alpha$  (przez napis “ $\beta \rightarrow \alpha$ ” rozumiemy to, że  $\alpha, \gamma \in \text{Sent}$  oraz  $l(\gamma) \leq l(\alpha)$ ,  $\text{Var}(\gamma) \subseteq \text{Var}(\alpha)$ ) oraz  $\alpha, \gamma \in \text{Sent}$  i  $\gamma \vdash_{\mathbf{LB}} \alpha$  i  $\neg(\alpha \vdash_{\mathbf{LB}} \gamma)$ , to  $\mathbf{U}\alpha \Rightarrow (\neg\mathbf{B}\neg\beta \wedge \neg\mathbf{B}\gamma)$

Schemat aksjomatu (O) jest uogólnionym schematem (U), (HPG) oraz (W)<sup>42</sup>.

W algorytmie Gazdara zmieniamy definicję implikatury jakościowej oraz pre-supozycji pragmatycznej, resztę pozostawiając bez zmian:

$\Phi_{n+1} = (((\Phi_n \cup! I_Q\alpha) \cup! I_C\alpha) \cup! I_S\alpha) \cup! P\alpha$ , gdzie

$I_Q\alpha$  jest implikaturą<sup>43</sup> jakościową wypowiedzi  $\alpha$

<sup>42</sup>Aksjomaty 1 – 5 określają logikę  $\mathbf{LB}$ , zaś powiększając zbiór schematów aksjomatów o kolejne (U) i (HPG) otrzymujemy logikę implikatury Tokarza  $\mathbf{LI}$

(U)  $\mathbf{U}(\alpha \wedge \beta) \Rightarrow (\mathbf{U}\alpha \wedge \mathbf{U}\beta)$

(HPG)  $\mathbf{U}\alpha \Rightarrow (\mathbf{B}\alpha \wedge \neg\mathbf{B}\beta)$  jeśli  $\beta \rightarrow \alpha$  (przez napis “ $\beta \rightarrow \alpha$ ” rozumiemy to, że  $\alpha, \beta \in \text{Sent}$  oraz  $l(\beta) \leq l(\alpha)$ ,  $\text{Var}(\beta) \subseteq \text{Var}(\alpha)$ ) oraz  $\alpha, \beta \in \text{Sent}$  i  $\Box +_{\mathbf{LB}} \Box$  i  $\neg(\Box +_{\mathbf{LB}} \Box)$

(W)  $\mathbf{U}(\mathbf{K}\Box) \Rightarrow \mathbf{U}\Box$

Patrz: Puczyłowski [2000]

$I_C\alpha$  jest im-plikaturą składnikową wypowiedzi  $\alpha$

$I_S\alpha$  jest im-plikaturą skalarną wypowiedzi  $\alpha$

$P\alpha$  jest pragmatyczną pre-supozycją wypowiedzi  $\alpha$

$I_Q\alpha := \{ \neg\mathbf{B}\neg\beta: (i) \beta \text{ jest zdaniem podrzędnym w } \alpha: (ii) \alpha \vdash \beta \}$

$I_C\alpha := \{ \neg\mathbf{B}\neg\beta, \neg\mathbf{B}\beta: (i) \beta \text{ jest zdaniem podrzędnym w } \alpha; (ii) \neg(\alpha \vdash \neg\beta); (iii) \neg(\alpha \vdash \beta); (iv)$

istnieje takie  $\gamma \neq \beta$ , że  $\{ \mathbf{B}\gamma, \mathbf{B}\neg\gamma \} \cap P\alpha[\gamma/\beta] = \emptyset \}$

$I_S\alpha := \{ \mathbf{B}\neg\beta(\xi_j): \text{o ile } \alpha \text{ jest zdaniem złożonym ze zdania } \beta(\xi_{j+1}), \xi_{j+1} \in \Theta, \text{ oraz } \alpha \vdash \beta^{\xi_{j+1}}, \text{ a } \beta(\xi_j) \text{ i } \beta(\xi_{j+1}) \text{ są prostymi zamiennikami wyrażeniowymi ze względu na } \xi_j, \xi_{j+1} \in \Theta \}$

$P\alpha := \{ \neg\mathbf{B}\neg\beta: \beta \text{ jest presupozycją generowaną przez deskrypcje określone i czasowniki faktywne występujące w } \alpha \}$

Zbiór  $\Phi_n$  oraz  $\Phi_{n+1}$  jest zbiorem przekonań odbiorcy, co do przekonań mówcy, odpowiednio, przed i po wypowiedzi  $\alpha$ .<sup>44</sup>

Definicji im-plikatury skalarnej użyliśmy, nie określając wcześniej co rozumiemy przez  $\Theta$ . Niech  $\Theta$  będzie n-ką wyrażen  $\Theta = \langle \xi_0, \xi_1, \dots, \xi_{n-1} \rangle$ . Jeśli  $\Theta$  jest skalą ilościową, to  $\alpha(\xi_i) \vdash \alpha(\xi_{i+1})$ , gdzie  $\alpha(\xi_i)$  i  $\alpha(\xi_{i+1})$  są parą zamienników prostych wyrażeniowych ze względu na  $\xi_i, \xi_{i+1} \in \Theta$ .

Zdania  $z(\alpha)$  i  $z(\beta)$  są zamiennikami wyrażeniowymi ze względu na  $\alpha$  i  $\beta$  wtedy i tylko wtedy, gdy  $z(\alpha)$  i  $z(\beta)$  są identyczne z tym wyjątkiem, że w określonym miejscu, tam, gdzie w  $z(\alpha)$  jest  $\alpha$ , tam w  $z(\beta)$  jest  $\beta$ .

---

<sup>43</sup> Przez *im-plikaturę* danej wypowiedzi rozumiemy każdą *potencjalną implikaturę* tej wypowiedzi, tak tą przez kontekst odwołaną, jak też każdą tą przezeń nie odwołaną. Nie jest więc tak, że każda implikatura jest im-plikaturą; jest oczywiście na odwrót.

<sup>44</sup> Niech  $\Phi$  będzie zbiorem przekonań S-a, do którego należeć będą jego przekonania dotyczące przekonań M-a. Przyjmujemy, że zbiór  $\Phi$  spełniać musi kolejne warunki:

1. Jeżeli  $\alpha \in \text{TAUT}$ , to  $\alpha \in \Phi$
2. Jeżeli  $\alpha \in \text{TAUT}$ , to  $\mathbf{B}_M\alpha \in \Phi$
3.  $\mathbf{B}_M\alpha \in \Phi$  zawsze i tylko, wtedy gdy  $\mathbf{B}_M\mathbf{B}_M\alpha \in \Phi$
4.  $\neg\mathbf{B}_M\alpha \in \Phi$  zawsze i tylko, wtedy gdy  $\mathbf{B}_M\neg\mathbf{B}_M\alpha \in \Phi$
5. Jeżeli  $\mathbf{B}_M\neg\alpha \in \Phi$ , to  $\neg\mathbf{B}_M\alpha \in \Phi$
6. Jeżeli  $\mathbf{B}_M\alpha \in \Phi$  i  $\mathbf{B}_M(\alpha \Rightarrow \beta) \in \Phi$ , to  $\mathbf{B}_M\beta \in \Phi$
7. Jeżeli  $\alpha \in \Phi$ , to  $\mathbf{B}_S\alpha \in \Phi$
8. Jeżeli  $\alpha \in \Phi$  i  $\alpha \Rightarrow \beta \in \Phi$ , to  $\beta \in \Phi$

Zdanie  $z$  jest *proste* ze względu na wystąpienie w nim składowego wyrażenia  $\alpha$  zawsze i tylko, wtedy gdy  $z$  nie zawiera żadnych funktorów logicznych o szerszym zasięgu niż  $\alpha$ .

Definicja operacji  $\cup!$  przedstawia się następująco:  $X \cup! Y = X \cup \{y: y \in Y \wedge (Z \subset X \cup Y)[\text{con}(Z \cup \{y\}) \Leftrightarrow \text{con}Z]\}$ .

## 4.2. KŁAMSTWO

Wróćmy ponownie do problemu *kłamliwości* zdań Moore'a: czy można kłamać wypowiadając takie zdania? Jeżeli tak, to nie są to zdania <zbędne>, tak dalece jak niezbędne są nam kłamstwa. Tak jak <potrzebujemy> zdań fałszywych, czy nawet logicznie fałszywych (np. dla celów ironii, metafory), tak być może ważne jest posiadanie zdań takich jak zdania Moore'a do wypowiadania <ewidentnych> kłamstw. Kłamstwa, gdy rozpoznane, dostarczają jakiejś wiedzy o przekonaniach <kłamcy>, o czymś nas informują, nie są wypowiedziami niezrozumiałymi – nie są bełkotliwe lub nierelevantne. Czy jednak, gdyby kłamstwo było na gruncie logiki przekonań rozpoznawalne, to traktowalibyśmy je jeszcze jako kłamstwo? Zauważmy, jak absurdalne byłoby powiedzenie “jestem i nie jestem przekonany, że  $p$ ”, wypowiedziane, by skłamać – w porównaniu z taką wypowiedzią nawet szczerze powiedziane zdanie Moore'a już tak nie razi.

Nie sądzę, by poprzez logiczno-pragmatyczną analizę zdań Moore'a sprawa skandaliczności tychże wypowiedzi została ostatecznie przez Hintikkę rozwiązana; razić cokolwiek może także tak jednoznaczne, bezwyjątkowe połączenie akceptowalności wypowiedzi z określonymi przekonaniem tego, kto jest określonej wypowiedzi nadawcą: takie powiązanie może wydawać się nadto zdawkowe. Kiedy mówię z przekąsem albo drwiąco – i pod pewnym względem szczerze – “dobry z ciebie przyjaciel”, to oczywiście postępowanie mego słuchacza, adresata tych słów zgodne z Hintikkowską “raczej dość ogólną regułą” doprowadziłoby go do błędnego zrozumienia intencji i przekonań, którym starałem się dać wyraz. Poprzez taki przykład skierowujemy uwagę na wypowiedzi *dosłowne* – kategorię dosłowności traktować musimy, wobec braku adekwatnej definicji, jako pierwotną. Jako jej częściowe określenie przyjąć można proponowaną przez Hintikkę maksymę: ktoś może powiedzieć szczerze:  $\alpha$ , mając przekonanie, że zamierzony słuchacz domyśli się, że mówca jest przekonany, że *nie- $\alpha$*  – wypowiedzi nie można brać dosłownie, tak jakby był on przekonany, że  $\alpha$ . Być może więc lepiej jest patrzeć na zdania Moore'a nie poprzez *prawdę* –



*falsz, czy szczerosc – kłamliwosc*, lecz – dla większej precyzji – konieczne jest <rozszczipienie> wypowiedzi w oparciu o kategorie *dosłownosc – przenosnosc*?

Godzimy się z tym, że oprócz kategorii syntaktycznych i semantycznych o poprawności wypowiedzi decydują również czynniki pragmatyczne – kontekst wypowiedzenia, nadawca, jego przekonania, rozmówca ze swoimi przekonaniem i oczekiwaniami, to co zostało powiedziane, przyczyna i cel rozmowy. Nie są to oczywiście wszystkie aspekty pragmatyczne, może nie są to nawet najważniejsze z tych, które bierzemy pod uwagę oceniając wypowiedź jako akceptowalną. Czy zatem nie postępujemy pochopnie odmawiając zdaniom Moore’a dopuszczalności z powodu logicznej niemożności dosłownego i szczerego ich wypowiedzenia. Może wymagamy zbyt wiele?

Czy można powiedzieć, że zdanie Moore’a równoważne jest wypowiedzi: “Kłamię mówiąc, że *p*”? Czy taką wypowiedź, rozumianą dosłownie, można uznać za szczerą? Myślę, że istotnie w pewnych okolicznościach byłoby to dopuszczalne, szczególnie, jeśli już wcześniej powiedziałem “*p*” lub coś, z czego intencjonalny słuchacz mógł wywnioskować, iż twierdę, że *p*. Inna sprawa, że taką wypowiedzią stawiam mego słuchacza w niedogodnej sytuacji: albo kłamałem wcześniej a teraz sam to przyznaję, albo nie kłamałem - a teraz chciałbym sprawić wrażenie, jakbym wypierał się wcześniejszych wypowiedzi. Nie chodzi o to, że słuchacz może mieć kłopoty z określeniem, która z tych sytuacji ma miejsce, lecz, że w jednym i drugim wypadku sam podważam swą wiarygodność: albo kłamałem kiedyś, albo kłamię teraz, albo mówiłem zgodnie z przekonaniem wcześniej, albo dopiero teraz. Czy można zatem dawać wiarę moim słowom?

Ale jeśli nie powiedziałem, ani nie sugerowałem, że  $\alpha$ ; to co myśleć o mojej wypowiedzi “Kłamię, że  $\alpha$ ”? Czy nie lepiej, by – przy tych przekonaniach, które miałem wypowiadając to zdanie – powiedzieć np.: “ $\alpha$ ” lub “jestem przekonany, że  $\alpha$ ” albo też: “nieprawda, że  $\alpha$ ” lub “nie jestem przekonany, że  $\alpha$ ”? Czy potrafimy powiedzieć, która z tych dwóch możliwości byłaby, z punktu widzenia przekonania lub intencji mówcy, bardziej wskazana? Jeżeli byłem przekonany, że *p*, to oczywiście wypowiedź “*p*” byłaby zgodna z żywym przeze mnie przekonaniem, lecz gdybym chciał skłamać i wprowadzić w błąd mego rozmówcę, to raczej spodziewać się należy, że powiem “nieprawda, że *p*”. Wypowiedź “Kłamię: *p*” wygląda tedy na wyraz schizofrenicznej osobowości – wypowiadam asertywnie jakieś zdanie, uprzedziwszy, że w nie wierzę: przypomina to cytowanie kogoś, z kim się nie zgadzamy. Tak opisane zdania Moore’a wydają się być może zbyt rozsądne, toteż oponent (np. Hintikka) prawdopodobnie zauważyłby, że zdanie Moore’a przypomina raczej wypowiedź, w której

wyrażamy np. i aprobatę, i dezaprobatę zarazem (a nie aprobatę i brak dezaprobaty, lub brak jednego i drugiego), wyrażamy dwie wzajemnie sprzeczne postawy propozycjonalne. Zapewne też zdanie Moore’a wydaje się bardziej absurdalne od “Kłamię:  $p$ ” – ostatnie można bowiem opisać jako wyraz *przekonania*, że nieprawda, że  $p$  oraz *chęci* (która nie może być zrealizowana) dania do zrozumienia słuchaczom, że jest się przekonany, że  $p$ .

Kłamstwo to wypowiedź niezgodna z przekonaniem. W jaki sposób, gdy mówię: “ $p$ ”, mogę mówić zgodnie z własnymi przekonaniem? Nie wydaje się, że znaczy to tyle, że mówiąc: “ $p$ ”, które stwierdza to, że  $q$ , jestem przekonany, że  $q$ . Potrzebujemy szerszego, bardziej uniwersalnego ujęcia. Jeżeli wypowiedź  $\alpha$  (stwierdzająca to, że  $\gamma$ ) wyraża – zgodnie z regułą  $R$  – przekonanie  $\beta$ , to wypowiedzenie  $\alpha$  jest zgodne z przekonaniem mówcy  $\Psi$ , gdy  $\beta \in \Psi$ . Jeżeli więc, mówiąc: “ $p$ ”, wyrażam – zgodnie z regułą Hintikki – przekonanie, że  $p$ , a przekonanie to, jest jednym z faktycznie przeze mnie żywionych, to “ $p$ ” jest wypowiedzią zgodną z moimi przekonaniem. Jeżeli mówiąc: “ $p$ ” wyrażam – zgodnie z zaproponowaną przez nas regułą – nie-wiarę, że nieprawda, że  $p$  (przyzupsczenie, że  $p$ ), a przekonanie to jest jednym z przekonania faktycznie przeze mnie żywionych, to “ $p$ ” jest wypowiedzią zgodną z moimi przekonaniem.

Jak rozpoznać, że wypowiedź “ $p$ ” jest kłamstwem? Załóżmy, że dysponujemy *doksaskopem*, umożliwiającym nam adekwatne opisanie zbioru przekonania  $\Psi$  dowolnego mówcy. Przyjmujemy, że zbiór  $\Psi$  jest zbiorem niesprzecznym, czyli, że jeżeli  $\neg\alpha \in \Psi$ , to  $\neg(\alpha \in \Psi)$  oraz, że  $\Psi$  jest zbiorem niepełnym, tj. istnieje  $\alpha$ , takie, że  $\neg(\neg\alpha \in \Psi)$  i  $\neg(\alpha \in \Psi)$ , domkniętym na **LB**-konsekwencje.

Kłamstwo określimy teraz dwojako: jeżeli wypowiedź – w kontekście przekonania mówcy  $\Psi$  –  $\alpha$  wyraża (przekonanie)  $\beta$ , to:

1.  $\alpha$  jest kłamstwem<sub>1</sub>, gdy  $\neg(\beta \in \Psi)$
2.  $\alpha$  jest kłamstwem<sub>2</sub>, gdy  $\neg\beta \in \Psi$ <sup>45</sup>

Mając dwie reguły – Hintikki oraz naszą – wiążące wypowiedzi z przekonaniem, mamy w konsekwencji – poprzez skrzyżowanie podziałów – cztery rodzaje kłamstw: H-kłamstwo<sub>1</sub>, H-kłamstwo<sub>2</sub>, P-kłamstwo<sub>1</sub>, P-kłamstwo<sub>2</sub>. Uznamy wypowiedź  $\alpha$  za H-kłamstwo<sub>1</sub>, o ile  $\neg(\mathbf{B}\alpha \in$

---

<sup>45</sup> Ogólniejsza definicja:

jeżeli wypowiedź – w kontekście przekonania mówcy  $\Psi$  –  $\alpha$  wyraża (przekonanie)  $\beta_1, \dots, \beta_n$ , to

1.  $\alpha$  jest kłamstwem<sub>1</sub>, gdy dla pewnego  $\beta_i$  ( $1 \leq i \leq n$ )  $\neg(\beta_i \in \Psi)$ .
2.  $\alpha$  jest kłamstwem<sub>2</sub>, gdy dla pewnego  $\beta_i$  ( $1 \leq i \leq n$ )  $\neg\beta_i \in \Psi$

$\Psi$ ), za P-kłamstwo<sub>1</sub>, o ile  $\neg(\neg\mathbf{B}\neg\alpha \in \Psi)$ , za H-kłamstwo<sub>2</sub>, o ile  $\neg\mathbf{B}\alpha \in \Psi$ , za P-kłamstwo<sub>2</sub>, o ile  $\mathbf{B}\neg\alpha \in \Psi$ . Jeżeli jakaś wypowiedź jest P-kłamstwem<sub>1</sub>, to jest też H-kłamstwem<sub>1</sub>. Jeżeli jakaś wypowiedź jest P-kłamstwem<sub>2</sub>, to jest też H-kłamstwem<sub>2</sub>. Oczywiście jeżeli jakaś wypowiedź jest kłamstwem<sub>2</sub> (czy P-, czy H-kłamstwem), to jest też kłamstwem<sub>1</sub> (odpowiednio P-, czy H-kłamstwem).

Dopuszczalna jest możliwość, że jakaś wypowiedź jest H-kłamstwem<sub>2</sub>, lecz nie jest P-kłamstwem<sub>2</sub>. Załóżmy, że  $\Psi = \{\neg\mathbf{B}p\}$  a wypowiedzią jest “ $p$ ”. Słuchacz <kierujący> się zasadą Hintikki z takiej wypowiedzi stwierdzającej, że  $p$ , wywnioskowałby, że mówca jest przekonany, że  $p$  – wypowiedź byłaby tedy H-kłamstwem<sub>2</sub>, bowiem  $\neg\mathbf{B}p \in \Psi$ . Nie byłaby to wypowiedź kłamliwa<sub>2</sub> dla osób <posługujących> się naszą zasadą, bowiem  $\neg(\mathbf{B}\neg p \in \Psi)$ . Taka jednak wypowiedź, przy takim zbiorze przekonań mówcy byłaby dla nas kłamstwem<sub>1</sub>, bowiem  $\neg(\neg\mathbf{B}\neg p \in \Psi)$ . Gdyby zaś  $\Psi = \{\neg\mathbf{B}p, \neg\mathbf{B}\neg p\}$ , a wypowiedź “ $p$ ” wyrażała – zgodnie z maksymą Hintikki –  $\mathbf{B}p$  lub  $\neg\mathbf{B}\neg p$  – zgodnie z naszą – to taka wypowiedź byłaby H-kłamstwem, lecz nie P-kłamstwem.

Jeżeli “ $p$ ” jest przy pewnym  $\Psi$  kłamstwem, to czy - “nieprawda, że  $p$ ” - nie jest, przy tym samym zbiorze przekonań, także kłamstwem? Przyjmując, że zbiór  $\Psi$  jest niesprzeczny i niezupełny, wypowiedź “nieprawda, że  $p$ ” może również być H-kłamstwem<sub>1</sub> (o ile “ $p$ ” jest H-kłamstwem<sub>1</sub>), a jeśli “ $p$ ” jest H-kłamstwem<sub>2</sub> (czyli  $\neg\mathbf{B}p \in \Psi$ ), to “nieprawda, że  $p$ ” również może być H-kłamstwem<sub>2</sub> (bowiem  $\neg\mathbf{B}\neg p$  może należeć do  $\Psi$ ; wtedy  $\Psi = \{\neg\mathbf{B}p, \neg\mathbf{B}\neg p\}$ ). Jeżeli “ $p$ ” jest H-kłamstwem<sub>1</sub> (tzn. jeśli  $\neg(\mathbf{B}p \in \Psi)$ ), to “nieprawda, że  $p$ ” może być H-kłamstwem<sub>2</sub>, bowiem  $\neg\mathbf{B}\neg p$  może należeć do  $\Psi$  oraz *vice versa*: jeżeli “ $p$ ” będzie H-kłamstwem<sub>2</sub>, to “nieprawda, że  $p$ ” może być H-kłamstwem<sub>1</sub>.

Jeżeli “ $p$ ” jest P-kłamstwem<sub>1</sub>, to “nieprawda, że  $p$ ” może być P-kłamstwem<sub>1</sub>, bowiem dopuszczamy sytuację taką:  $\neg(\neg\mathbf{B}\neg p \in \Psi)$  i  $\neg(\neg\mathbf{B}p \in \Psi)$ . Jednak nie możemy przyjąć, że obydwie te wypowiedzi będą P-kłamstwami<sub>2</sub>, bowiem niemożliwe jest by zarazem  $\mathbf{B}\neg p \in \Psi$  i  $\mathbf{B}p \in \Psi$ . Podobnie, nie może być tak, że  $\mathbf{B}\neg p \in \Psi$  i  $\neg(\neg\mathbf{B} \in \Psi)$ , więc jeśli “ $p$ ” jest P-kłamstwem<sub>2</sub>, to “nieprawda, że  $p$ ” nie może być P-kłamstwem<sub>1</sub>. Jeżeli zaś “ $p$ ” jest P-kłamstwem<sub>1</sub>, to “nieprawda, że  $p$ ” nie może być P-kłamstwem<sub>2</sub>, gdyż niemożliwe jest, by przy prawdziwości  $\neg(\neg\mathbf{B}\neg p \in \Psi)$  prawdziwe było  $\mathbf{B}p \in \Psi$ .

Dwa różne podmioty, hołdujące tej samej logice przekonań, lecz rozbieżne w przyjmowanych zasadach łączących przekonania z wypowiedziami, mogłyby zgodzić się, że wypowiedź “ $p$ ” jest H-kłamstwem<sub>1</sub> (lub P-kłamstwem<sub>1</sub>) oraz “nieprawda, że  $p$ ” jest P-kłamstwem<sub>1</sub> (lub, odpowiednio, H-kłamstwem<sub>1</sub>) – innych możliwości *transkłamliwości* tych dwóch wypowiedzi

nie mogą, z pobudek logicznych, dopuścić. Pomimo że rozważania powyższe wydawać się mogą nadto abstrakcyjne i hermetyczne, to pozwalają skonfrontować formalnologiczne dociekania z intuicjami, którym dajemy wyraz np. zastanawiając się nad właściwą odpowiedzią na pytanie: czy jeśli ktoś skłamał mówiąc “Żona Jana jest atrakcyjna”, to byłby szczerzy mówiąc “Żona Jana nie jest atrakcyjna”? Wydaje się, że bliższa intuicjom jest próba scharakteryzowania kłamstwa ujęta w definicjach kłamstwa<sup>46</sup>; pozostaje jednak problem: jak inaczej (jeśli nie jako kłamstwo) określić sytuacje, do których odnosi się definicja kłamstw<sub>1</sub>? Należy też zauważyć, że na gruncie naszego rozwiązania dwa zdania sprzeczne nie mogą być – przy ustalonym zbiorze przekonań mówcy – kłamstwami. Taką sytuację (kłamstwami mogą być równocześnie dwa sądy sprzeczne) dopuszcza Hintikka: uważam to za pewną nieintuicyjną konsekwencję jego propozycji. Powyższe rozważania nie wyczerpują wszystkich subtelnych strategii kłamania. Nadto, niekiedy świadomie wypowiadamy twierdzenia, które uznajemy za fałszywe, lecz nikt nie posądziłby nas o kłamstwo: jest tak, gdy wygłaszamy metaforę, ironizujemy, niekiedy też - gdy przesadnie komplementujemy lub ganimy, czy też wygłaszamy formuły grzecznościowe. Interlokutorzy o tym dobrze wiedzą, o czym wiemy i my, decydując się na określoną wypowiedź. Oczywiście jest też i to, że nie można kłamać, nie znając swych przekonań, jak też nie sądząc, że ktoś z naszej wypowiedzi (kłamliwej lub nie) dojdzie w ustalony sposób do określonych (przez kontekst wypowiedzi i wypowiedź) przekonań, dotyczących przekonań żywionych przez nas.

*Po wprowadzeniu terminów systemu Gazdara można postarać się o trafniejszą charakterystykę pojęcia “kłamstwa”. Nie wykluczamy, że będzie nią poniższa definicja:*

*Wypowiedź  $\alpha$  jest niedopuszczalna (kłamliwa, gdy faktycznie wypowiedziana) w kontekście (zbioru) przekonań nadawcy  $\Psi$  przy przekonaniach odbiorcy  $\Phi$ , gdy istnieje takie (przekonanie)  $x$ , że  $x \in \Phi \cup \mathbf{I}\alpha$ , lecz  $\neg(x \in \Psi)$ .*

*W powyższej propozycji uwzględniliśmy konieczność zawarcia w ocenie wypowiedzi kłamliwej przekonań słuchacza. Definicja ta nie klasyfikuje wypowiedzi  $\alpha$  (przy przekonaniach mówcy  $\Psi$ ) wyrażającej np.  $\beta$  jako kłamliwej, gdy  $\neg(\beta \in \Psi)$ . Odbiorca nie musi wszak pod wpływem wypowiedzi  $\alpha$  (wyrażającej  $\beta$ ) dołączyć  $\beta$  do zbioru swych przekonań: może bowiem wiedzieć, że nadawca  $\alpha$  wcale nie jest przekonany, że  $\beta$ : wiedzieć to*

---

<sup>46</sup> Czy nie jesteśmy zbyt surowi w osądzie? J. Anatas pisze, powołując się na J. Pelca (Pelc [1990]): “[...] Czy nadawca w ogóle musi wierzyć w fałszywość twierdzenia, które wypowiada? Otóż – według niego [J. Pelca – przyp. T. P.] – nadawca może posługiwać się zdaniem, które sam uznaje za prawdziwe, ale równie dobrze może w ogóle nie przypisywać mu żadnej wartości logicznej, normalnie przysługującej sądom” (Anatas [2000] s. 88).

może z wcześniejszych wypowiedzi nadawcy - wypowiedzi  $\alpha$ , a samą wypowiedź  $\alpha$  może potraktować np. jako żart, lub w najgorszym razie, jako niezrozumiałą (w świetle tego, co wcześniej z przekonaniem głosił). Oczywiście w takiej sytuacji – niezależnie czy można traktować  $\alpha$  jako żart – wypowiedź ta niewiele, o ile w ogóle cokolwiek, wnosi w konwersację. Ale może być i tak, że  $\alpha$  wyraża nie tylko  $\beta$ , ale i  $\gamma$ . Odbiorca może zaś wiedzieć (dzięki wcześniejszym wypowiedziom nadawcy), że  $\beta$  nie należy do  $\Psi$  i dlatego nie dołączy  $\beta$  do zbioru swych przekonań. Jeżeli jednak nie będzie żywić podobnych wątpliwości względem  $\gamma$ , to powiększy o to przekonanie zbiór swoich przeświadczeń. Rozsądne jest zatem pytanie, czy wypowiedź  $\alpha$  była w takim kontekście kłamstwem? Wydaje się, że raczej nie, chociaż przecież  $\neg(\beta \in \Psi)$ . Wydaje się więc, że definicja kłamstwa powinna uwzględniać również i skutek, jaki wywarła określona wypowiedź na swym intencjonalnym odbiorcy. Widać też, że definicja nie rozstrzyga tego, czy dwie wypowiedzi sprzeczne mogą być kłamliwe przy określonych zbiorach przekonań. Dopuszcza, że istnieje takie  $x$ , że  $x \in \Phi \cup! \mathbf{I}\alpha$ <sup>47</sup>, lecz  $\neg(x \in \Psi)$  oraz, że istnieje takie  $x$ , że  $x \in \Phi \cup! \mathbf{I}(\neg\alpha)$ , lecz  $\neg(x \in \Psi)$ , co więcej nie można wykluczyć, że dla pewnego  $x$  prawdą jest zarazem:  $x \in \Phi \cup! \mathbf{I}(\neg\alpha)$  i  $x \in \Phi \cup! \mathbf{I}\alpha$ , co (o ile prawdą jest to, że  $\neg(x \in \Psi)$ ) dopuszcza sytuację, że bez względu na to, czy powie się  $\alpha$ , czy też jej zaprzeczenie, to – w pewnym kontekście – wypowiedzi te będą kłamstwami z dokładnie tych samych względów.

Ciekawym przykładem filozoficznego kłamstwa jest tzw. paradoks Gettier (Gettier [1990]). Mówca, którego przekonania składały się na następujący zbiór

$$\Psi = \{\mathbf{B}p, \mathbf{B}(p \vee q), \neg\mathbf{B}q, \neg\mathbf{B}\neg q\},$$

wygłasza zdanie “ $p$  lub  $q$ ”. Zgodnie z przyjętą zasadą Grice’a, zdanie to wyraża nie tylko przekonanie o prawdziwości całego zdania alternatywnego, ale i niewiedzę, co do prawdziwości czy fałszywości składników tego zdania (o ile  $\Phi = \emptyset$ ). Skoro jednak, ani  $\neg\mathbf{B}p \in \Psi$ , ani  $\neg\mathbf{B}\neg p \in \Psi$ , a prawdą jest to, że  $\mathbf{B}p \in \Psi$ , zatem zdanie to traktować możemy jako H-kłamstwo. Opierając się na modyfikacji algorytmu Gazdara dokonanej <w duchu> naszej reguły, wypowiedź “ $p$  lub  $q$ ” ocenić możemy jako P-kłamstwo, dokładnie z tych samych pobudek, którymi kierowaliśmy się oceniając tę wypowiedź jako H-kłamstwo.

Antas pisze:

“<<mówienie prawdy>> to dawanie językowego świadectwa treściom naszych przeświadczeń na jej temat, podczas gdy kłamanie jest dawaniem świadectwa *fałszywym* treściom przeświadczenia na temat

<sup>47</sup> “ $\Phi \cup! \mathbf{I}\alpha$ ” jest skrótem: “((( $\Phi \cup! I_2\alpha$ )  $\cup! I_C\alpha$ )  $\cup! I_S\alpha$ )  $\cup! P\alpha$ ”

tego, w co wierzymy, że jest prawdą. [...] Gdy <<mówimy prawdę>> w rzeczywistości dajemy tylko świadectwo naszym mniemaniom na temat tego, w co wierzymy, że jest prawdą, nie zaś świadectwo faktom. I możemy się pomylić” (Antas [2000], s. 113)

Co w powyższym fragmencie chce powiedzieć Antas, pisząc “I nie możemy się pomylić”? Czy możemy pomylić się co do tego, jakie żyjemy przekonania, czy też, co do tego, w co wierzymy? Czy chodzi tu o to, że mówiąc “Jestem przekonany, że  $p$ ”, mogę nie znajdować się w tym stanie przekonaniowym, czy też, że sąd wyrażany w  $p$  jest fałszywy? Druga możliwość jest banalna: z własnego doświadczenia wiemy, że niekiedy (np. w dzieciństwie) byliśmy faktycznie przekonani, co do pewnych niezbyt rozsądnych, wręcz fałszywych koncepcji (np. co do tego, że istnieje Święty Mikołaj roznoszący w Wigilię dzieciom prezenty). Zasadność tego, co głosi pierwsza możliwość interpretacji stanowiska Antas, wydaje nam się bardziej kontrowersyjna. Mamy bowiem wrażenie, że można utrzymywać, iż pomimo tego, że nie zawsze potrafimy prawdziwie opisać swe przekonania, to będąc przekonani, że  $p$ , jesteśmy zarazem przekonani, że jesteśmy przekonani, że  $p$ . Można też twierdzić, że chociaż nie zawsze udaje nam się trafnie opisać nasze przekonania, to nie oznacza to jeszcze, że kłamiemy. Mogę zupełnie błędnie opisywać swe przekonania (przy tym wcale nie kłamiąc), a to jeszcze nie podważa tezy o tym, że jeżeli ktokolwiek jest przekonany, że  $p$ , to jest też przekonany, że jest przekonany, że  $p$ . Może się jednak mylić, co do tego, czy wierzy, że  $p$  – może głosić, że (wierzy, że)  $p$ , lecz w rzeczywistości nie wierzy, że  $p$ : jego deklaracja nie musi być, powtórzmy, kłamstwem (świadomie złym opisem swych stanów przekonaniowych, mającym wprowadzić intencjonalnego interlokutora). O żywionych przez kogoś przekonaniach dowiadujemy się nie tylko poprzez zadawanie podmiotowi odpowiednich pytań – niekiedy też tego, jakie ktoś przekonania żywi, dowiadujemy się z obserwacji zachowań podmiotu, które należałoby uznać za nieracjonalne (lub “do niego nie podobne”) nie założywszy, że żywi on dane przekonanie.

To, że nie zawsze można polegać na odpowiedziach podmiotu, którego przekonania badamy, pozwoli uzmysłwić następujący przykład. Wyobraźmy sobie kogoś kto mówi: “Ależ ten nasz J-23 ma świetne wyniki, dostarczając nam wielu cennych informacji, a nadto każdy, nawet ten głupi Brunner, myśli, że J-23 jest wzorowym oficerem Abwhery”. Wydaje mi się, że każdy, kto zna serial “Stawka większa niż życie” uznałby tę wypowiedź za prawdziwą, lecz wspomniany Brunner pytany o to, czy wie, że J-23 jest oficerem Abwhery, prawdopodobnie by zaprzeczył.

### 4.3. MOŻLWOŚĆ DRUGA

Zapiszmy zdanie “Jestem przekonany, że jest tak jak następuje:  $p$  ale nie wierzę, że  $p$ ” w dogodnej postaci symbolicznej:  $B(p \wedge \neg Bp)$ . W jaki sposób Hintikka stwierdza, że zdanie to jest wewnętrznie sprzeczne? Po pierwsze, równoważne do  $B(p \wedge \neg Bp)$  jest zdanie koniunkcyjne  $Bp \wedge B\neg Bp$ . Z drugiego z koniunktów wyprowadzić można  $\neg BBp$ , co jest sprzeczne z pierwszym koniunktem, o ile z  $Bp$  wynika  $BBp$ .

Przyjmując inną zasadę Hintikki  $B\neg B\alpha \Rightarrow \neg B\alpha$  można stwierdzić sprzeczność  $B(p \wedge \neg Bp)$ , bez uciekania się do pomocy zasady czy  $B\neg\alpha \Rightarrow \neg B\alpha$ , czy też  $B\alpha \Rightarrow BB\alpha$ . Co ciekawe, Hintikka odrzuca <odwrotną> zasadę  $BB\alpha \Rightarrow B\alpha$  – nie da się jej uzyskać z innych tez jego systemu, nie jest też przyjęta jako aksjomat; Hintikka ocenia ją jako mało wiarygodną. Píše: “Nie jest nienaturalnie powiedzieć o człowieku, że wierzy on, że wierzy, chociaż w rzeczywistości tak nie robi”<sup>48</sup>.

Co zaś z  $B\neg B\alpha \Rightarrow \neg B\alpha$ ? Wydaje się, że jego znaczenie jest bliskie (choć nie jest oczywiście tożsame) znaczeniu  $BB\alpha \Rightarrow B\alpha$ : jeżeli jestem przekonany, że nie jestem przekonany (nie wierzę), że  $\alpha$ , to nie jestem przekonany, że  $\alpha$ , to – mówiąc swobodnie – nie mogę się co do tego mylić, wobec czego mam rację tak sądząc. Dlaczego więc mogę być słusznie (tj. jestem słusznie przekonany, że  $\alpha$ , gdy jest tak jak sędzę, gdy faktycznie  $\alpha$ ) przekonany jedynie co do tego, że nie wierzę, a nie co do wszystkich swych przekonań – tych, że jestem i tych, że nie jestem przekonany? Schemat  $B\neg B\alpha \Rightarrow \neg B\alpha$  jest konsekwencją przyjętych przez Hintikkę tez:  $B\alpha \Rightarrow BB\alpha$  oraz  $B\neg\alpha \Rightarrow \neg B\alpha$  oraz prawa sylogizmu i transpozycji. Mamy bowiem następujący ciąg:  $B\alpha \Rightarrow BB\alpha$ ,  $BB\alpha \Rightarrow \neg B\neg B\alpha$ , (bowiem  $B\neg B\alpha \Rightarrow \neg BB\alpha$ );  $B\alpha \Rightarrow \neg B\neg B\alpha$ ,  $B\neg B\alpha \Rightarrow \neg B\alpha$ . Jeżeli taki typ argumentacji (tj. “nie jest nienaturalnie powiedzieć...”) miałby decydować o przyjmowaniu lub odrzucaniu tez logiki przekonań, to nie widać powodów, by – odrzuciwszy  $BB\alpha \Rightarrow B\alpha$  – przyjmować  $B\alpha \Rightarrow BB\alpha$ . Konsekwencją tego ostatniego jest bowiem  $B\neg B\alpha \Rightarrow \neg B\alpha$ , a przecież (to nasze, nie Hintikki przypuszczenie) “nie jest nienaturalnie powiedzieć o człowieku”, że jest przekonany, że nie wierzy, że  $p$ , podczas gdy w rzeczywistości wierzy, że  $p$ . Myślę, że znaczenie słowa “wierzyć”, przy którym przypuszczenie Hintikki jest prawdziwe, dopuszcza prawdziwość naszego. Wydaje się ponadto, że jeżeli można traktować fakultatywnie  $BB\alpha \Rightarrow B\alpha$ , to – motywując to podobnymi argumentami – podobnie ostrożnie można podchodzić do  $B\alpha \Rightarrow BB\alpha$ ; jednak bez tej zasady rozumowanie Hintikki co do zdań Moore’a upada.

<sup>48</sup> Stanowisko Hintikki w rozdziałach 3.7, 3.8 i 4.3 w: Hintikka [1962]

Oczywiście zdanie Moore'a może być powiedziane jako P-kłamstwo<sub>2</sub>, świadomie niezgodnie z własnymi przekonaniem, tj. zdanie nie jest wewnętrznie sprzeczne na gruncie logiki przekonań Hintikka lub Tokarza:  $B \rightarrow (p \wedge \neg Bp)$ .

Szcątkowa argumentacja motywująca odrzucenie jako tezy  $BB\alpha \Rightarrow B\alpha$ , którą przedstawia Hintikka, może zostać uzupełniona przez wywód D. W. Armstronga dotyczący zagadnienia podważalności świadomości: może analogia nie narzuca się sama, ale jakąś zbieżność dostrzec można. Związek między Armstrongiem a interesującą nas formułą wyjaśnimy omawiając poniższe fragmenty. Najpierw zobaczmy, jakie tezy wydają się Armstrongowi wątpliwe:

“Wielu filozofów [...] uważa, że nie można mylić się co do własnych stanów umysłowych” (Armstrong [1982], s.141)

“[...] Kiedy mówimy, że świadomość introspekcyjna nie podlega korekcie, rozumiemy przez to, że każdy nasz sąd o naszym aktualnym stanie umysłowym jest sądem prawdziwym” (Armstrong [1982], s. 143)

“Gdyby istniała możliwość błędu w ocenie aktualnej treści naszej własnej świadomości, to sensowne byłoby powiedzenie: “Myślę, że czuję teraz ból, ale może się mylę”, albo: “Myślę, że widzę oto coś zielonego, ale może się mylę”. Ponieważ takie zdania są nonsensowne, więc introspekcja jest niepodważalna” (Armstrong [1982], s.141 – 142)

Z taką tezą nie może zgodzić się Armstrong, bowiem teza ta jest niezgodna z głoszonym przez niego stanowiskiem – materializmu stanu centralnego, do którego twierdzeń należy m.in. to, że procesy umysłowe są stanami fizycznymi mózgu.

“Gdyby którakolwiek z tych doktryn okazała się prawdziwa, to między introspekcją i postrzeżeniem powstałaby przepaść. Twierdzenia percepcyjne są wypowiedziami o świecie fizycznym i zawsze jest możliwe, że takie wypowiedzi są błędne. Zdanie “Myślę, że widzę coś zielonego, ale mogę się mylić” jest całkiem sensowne, i nikt, kto postrzega, nie jest w sytuacji logicznie uprzywilejowanej w stosunku do tego, co postrzega. [...] Jeżeli procesy umysłowe są stanami osoby zdolnymi do wywoływania pewnych rodzajów zachowania i jeżeli te stany są rzeczywiście fizycznymi stanami mózgu, to introspekcja, która jest procesem umysłowym, musiałaby być również procesem fizycznym, zachodzącym w mózgu. I od razu staje się jasne, że zawsze jest logicznie możliwe, aby taki samobadający mechanizm podał błędny wynik. Bowiem każdy mechanizm może przestać działać poprawnie.” (Armstrong [1982], s. 145)



Co chce powiedzieć Armstrong? Pozostańmy przy jego określeniu “zdanie percepcyjne” (zdania typu: “widzę, że  $p$ ”) i przyjmijmy też “zdanie wierzeniowe” (“wierzę, że  $p$ ”) – nie należy podziału tego traktować jako wyczerpującego zbiór zdań. Jeżeli przez błędność rozumie tyle i tylko tyle co fałszywość, to zgoda – zdania wierzeniowe mogą być fałszywe (choć, z drugiej strony, skłaniamy się ku idealizacyjnemu założeniu o kompetencji logicznej podmiotów (cel teoretyczny umożliwia tego rodzaju założenie), tzn. przyjmujemy, że każdy taki podmiot jest przekonany co do prawd logicznych): nie podważa to jednak tezy o niekorygowalności tego typu zdań dla podmiotu. Co znaczy to, że zdanie  $p$  jest dla  $A$  niekorygowalne: wyjaśnienia bowiem wymaga, to co mamy na myśli mówiąc, że nie możemy się co do  $p$  mylić. Przytoczmy definicję Armstronga; dalej zastanowimy się nad jej adekwatnością.

“Załóżmy, że  $p$  jest logicznie niepodważalne dla  $A$  wtedy i tylko wtedy, jeżeli:

- (1)  $A$  sądzi, że  $p$
- (2) (sąd  $A$ , że  $p$ ) logicznie implikuje ( $p$ )

[...] Ci, którzy akceptują niepodważalność świadomości introspekcyjnej, że warunki te zostaną spełnione, kiedy  $p$  będzie sądem o rzeczywistym stanie umysłowym.” (Armstrong [1982], s. 143)

Nie jesteśmy pewni, czy właściwie rozumiemy tę definicję: czy możemy się zgodzić na taką jej parafrazę:

$p$  jest logicznie niepodważalne dla  $A$  wtedy i tylko wtedy, jeżeli:

- (1.1)  $B_A p$
- (1.2)  $B_A p \Rightarrow p$ ?

Godząc się jednak na ten przekład, widzimy w końcu związek między tezą o logicznej niepodważalności zdań introspekcyjnych (wierzeniowych) z formułą  $BB\alpha \Rightarrow B\alpha$ . Jeżeli oto zdania o “rzeczywistych procesach umysłowych” są logicznie niepodważalne, to – zgodnie z definicją – prawdziwe jest (na mocy punktu (1.2) definicji) zdanie  $BB\alpha \Rightarrow B\alpha$ .

Czy definicja Armstronga jest poprawna? Nie wiemy, czy dobrze ją odczytaliśmy, zatem nie wiadomo, jak będzie się ona miała do faktycznych przekonań Armstronga. Przekonajmy się rozważając, co znaczy, że  $p$  NIE jest logicznie NIEPODWAŻALNE dla  $A$  (przyjmijmy, że zdanie, które nie jest niepodważalne jest podważalne):

$p$  nie jest logicznie niepodważalne dla  $A$  wtedy i tylko wtedy, jeżeli:

$$(2.1) \neg B_{Ap} \text{ lub}$$

$$(2.2) B_{Ap} \wedge \neg p$$

Zatem, by jakieś zdanie było dla pewnego  $A$  podważalne, wystarczy, by  $A$  nie wierzył (nie był przekonany), że  $p$ . Rzeczywiście, na przykład nie wierzę w to, że mam dziecko (bowiem wierzę, że nie mam dziecka), i – jak sędzę – wystarcza to, by sąd ten (“Tomasz Puczyłowski ma dziecko”) nie był dla mnie, Tomasza Puczyłowskiego, niepodważalny: w dowolnej chwili gotowy jestem się spierać, że nie mam dziecka, czynnie dowodzić swej racji (oczywiście mogę tej racji nie mieć). Teraz jednak załóżmy, że wierzę w to, że mam dziecko, ale w rzeczywistości tak nie jest – nie jestem niczym ojcem (lub, by sytuacja wyglądała bardziej dramatycznie: osoba, którą uważam za swe dziecko, faktycznie nim (o zgrozo!) nie jest). Czy wiedząc o tym, że Tomasz Puczyłowski jest przekonany co do swego ojcostwa, lecz nie wie jak się fakty mają, zgodzilibyśmy się – zgodnie z definicją – na to, że sąd “Tomasz Puczyłowski jest ojcem” nie jest dla Tomasza logicznie niepodważalny? Cóż, dla nas jest to sąd logicznie podważalny, ale dlaczego być by musiał takim dla Tomasza? Dlatego, że wierzy on w coś, co prawdą nie jest? A gdyby sądził, że nie może być inaczej niż myśli? I dlaczego w końcu, przygodny fakt, o którym wie ze szkolnych lektur np. - wygrana Sobieskiego pod Wiedniem (co do czego Tomasz jest przekonany) - miałby być dla niego logicznie niepodważalny?

Adekwatniejszą propozycją definicji niepodważalności logicznej wydaje się poniższe zdanie:  $p$  jest *logicznie niepodważalne dla  $A$  zawsze* i tylko wtedy, jeżeli:

$$(3.1) B_{Ap} \text{ oraz}$$

$$(3.2) B_A(B_{Ap} \Rightarrow p)$$

Przyjmując powyższą definicję niepodważalności, uzyskujemy następujące, bliższe potocznemu rozumieniu określenie tego, że sąd nie jest niepodważalny:

$p$  nie jest *logicznie niepodważalne dla  $A$  zawsze* i tylko wtedy, jeżeli:

$$(1) \neg B_{Ap} \text{ lub}$$

$$(2) \neg B_A(B_{Ap} \Rightarrow p)$$

Warunek (2) odczytać można następująco: nie jestem przekonany, co do tego, co następuje: (jakieś zdanie jest prawdziwe dlatego tylko, iż sędzę, że tak jest (że zdanie to jest prawdziwe)). Albo: dopuszczam, że jestem przekonany do jakiegoś zdania, które nie jest prawdziwe; jeżeli jakiś sąd nie jest niepodważalny, to dopuszczam, że mogę być co do niego przekonany, jak też dopuszczam to, że nie mogę co do niego nie być przekonany (“nie ma zniewalających racji by nakłonić mą wolę, by uwierzyła”). To, kiedy sąd nie może być dla

mnie podważalny rozpoznaję stwierdzając, albo to, że jestem co do tego sądu przekonany, albo, że przekonany jestem, iż nie wierzę, że jest, tak jak on głosi. Dostrzec to łatwo na gruncie logiki Tokarza **LB**, gdzie mamy następującą tezę:  $\neg B_A(B_{Ap} \Rightarrow p) \Rightarrow (\neg B_A \neg B_{Ap} \wedge \neg B_{Ap})$ .<sup>49</sup>

Jak rozumiem to, że coś jest dla mnie logicznie niepodważalne? Sądzę, że mianem takich przekonań określiłbym najszybciej te, co do których prawdziwości sądę, że nie mogę się mylić, te, o których sądę, że muszą być prawdziwe. Podałbym więc, jako przykład, znane mi prawa klasycznego rachunku zdań, znane mi fakty z historii, jak np. ten, że jednym z prezydentów Finlandii był Carl Mannerheim. Są to dla mnie sądy niepodważalne. Nie wiem co prawda, czy mógłbym się zgodzić z tym, że są to sądy LOGICZNIE niepodważalne. Nie wiem – Mannerheim nie musiał być przecież prezydentem Finlandii, historia mogła – potoczyć się inaczej, mógł np. zginąć w trakcie Rewolucji Październikowej. Czy tezy klasycznego rachunku zdań są podważalne logicznie? Chociaż wydaje się to trudne do zrozumienia, lecz Armstrong pisze, że tak, bowiem ktoś nie musi być co do tego rachunku zdań przekonany; pojawić się więc musi natychmiast trudne pytanie: czym zatem wtedy niepodważalność logiczna różni się by miała od niepodważalności psychologicznej? Rozbudowując dalej definicję o pojęcia ‘filozoficznie podejrzane’ można zaryzykować następujące określenie:

*p* jest logicznie niepodważalne dla *A* zawsze i tylko wtedy, jeżeli:

(1)  $B_A \Box p$  oraz

(2)  $B_A(B_{Ap} \Rightarrow p)$ ,

gdzie punkt pierwszy odczytujemy następująco: *A* jest przekonany, że *p* jest koniecznie prawdziwe. Wprowadzenie kategorii konieczności wprowadza groźbę zwiększenia niejasności w definicji. Można sobie wyobrazić, że ktoś twierdzi, iż coś jest dla niego logicznie niepodważalne, chociaż zarazem twierdzi, że nie uważa tego za koniecznie prawdziwe lub też twierdzi szczerze zarazem, że nie ma zdań koniecznie prawdziwych. Być może lepszą modyfikację definicji uzyskamy powiększając definicje niepodważalności o kolejne pojęcia epistemiczne:

---

<sup>49</sup> Pojawia się pytanie, czy  $B_A \alpha \Rightarrow B_A \alpha \vee B_A \beta$  jest tezą systemu **LB**? Jeżeli tak, to skoro tezą jest  $\neg B_A(B_{Ap} \Rightarrow p) \Rightarrow (\neg B_A \neg B_{Ap} \wedge \neg B_{Ap})$ , więc też i  $(B_A \neg B_{Ap} \vee B_{Ap}) \Rightarrow (B_A(B_{Ap} \Rightarrow p))$  jest tezą. A skoro tak, to na mocy prawa sylogizmu hipotetycznego mamy  $B_A \alpha \Rightarrow (B_A(B_{Ap} \Rightarrow p))$ . Zatem, należy uznać – na mocy **LB** oraz definicji – że każde moje przekonanie jest dla mnie niepodważalne. Przyznaję, że trudno jest przyzwyczaić się do tego odkrycia.

$p$  jest logicznie niepodważalne dla  $A$  wtedy i tylko wtedy, jeżeli:  $B_A K_A p$  ( $A$  jest przekonany, że wie, że  $p$ )

Przyjmując powyższą definicję oraz klasyczną definicję wiedzy, aby,  $p$  było niepodważalne dla jakiegoś podmiotu, muszą spełnione być trzy warunki: (i) podmiot musi być przekonany, że  $p$ , (ii) musi być świadomy tego, że jest świadomy, że  $p$ , (iii) podmiot musi być przekonany, że ma wystarczające uzasadnienie przekonania, że  $p$ . Dlaczego przy tej definicji mamy przyjmować, że zdania wierzeniowe nie są niepodważalne logicznie? Rozpatrujmy, jako zdanie wierzeniowe, w dalszym ciągu zdania typu: “jest przekonany, że  $p$ ”. Zdanie “ $Bp$ ” będzie podważalne dla ustalonego podmiotu, gdy  $\neg BBp$  lub  $\neg BBBp$  lub  $\neg BUzBp$ . Widzimy, że przyjmując taką definicję niepodważalności logicznej, łatwo możemy się zgodzić na to, że zdania wierzeniowe nie muszą być niepodważalnymi; definicja ta sama z siebie nie wyróżnia żadnego typu zdań jako niepodważalnych (tak jak to miało miejsce przy przyjęciu definicji, gdzie w wyniku podstawienia za “ $p$ ” wyrażenia “ $Bp$ ” otrzymywaliśmy zdanie prawdziwe  $BBp \Rightarrow Bp$ ).

Nie przeczę, że zdania takie jak “Sądzę, że  $p$ ” czy też “Wierzę, że jestem przekonany, że  $p$ ” mogą być fałszywe, podtrzymujemy jednak, że jeżeli prawdziwe jest zdanie “Wierzę, że jestem przekonany, że  $p$ ”, to prawdziwe jest zdanie “Jestem przekonany, że  $p$ ”. Sądzę, że jest to kwestia rozstrzygnięć pojęciowych, tak daleko jak przyjmujemy, np., że jeżeli ktoś wie, że  $p$ , to jest też przekonany, że  $p$ . Nawet jeżeli w przeważającej mierze zdania dotyczące stanów umysłu są podważalne, to nie możemy sobie wyobrazić sytuacji, w której jesteśmy przekonani, że wierzymy, że  $p$ , ale w rzeczywistości tak się nam tylko wydaje – nie wierzymy, że  $p$  (lub, co gorsza, wierzymy, że nie- $p$ ). Oczywiście może nie być prawdą, że zawsze, kiedy jestem przekonany, że  $p$ , to jestem również przekonany, że jestem przekonany, że  $p$ .

Na koniec proponujemy zmienić pierwotne określenie zdania Moore’a na bardziej ogólne:

Niech  $R$  będzie funkcją przyporządkowującą, w pewnym kontekście  $\Phi_n$ , dowolnej niesprzecznej wypowiedzi (mówcy  $M$ )  $\alpha$  pewien zbiór przekonań ( $M$ -a)  $\Phi_{n+1}$  (symbolicznie  $R(\alpha, \Phi_n) = \Phi_{n+1}$ ):  $\beta$  jest zdaniem Moore’a w kontekście  $X$  zawsze i tylko wtedy, kiedy  $R(\beta, X) = J$ .

## **BIBLIOGRAFIA**

Ajdukiewicz K.

[1985a] *Okres warunkowy a implikacja materialna*, w: Ajdukiewicz [1985]

Ajdukiewicz K.

[1985b] *Język i poznanie*, tom 2, PWN, Warszawa

Antas J.

[2000] *O kłamstwie i kłamaniu*, Universitas, Kraków

Armstrong D. M.

[1982] *Materialistyczna teoria umysłu*, PWN, Warszawa

Gazdar G.

[1979] *Pragmatics: Implicature, Presupposition, and Logical Form*, Academic Press, London

Gettier E.

[1990] Is Justified True Belief Knowledge? (Czy prawdziwe i uzasadnione przekonanie jest wiedzą?, tłum J. Hartman, J. Rabus), *Principia*, tom 1

Hintikka J.

[1962] *Knowledge and Belief*, Cornell University Press

Kański Z.

[1991] Znaczenie jako kontekst w pragmatyce Gazdara, w: Pelc [1991]

Lenzen W.

[1978] *Recent Work in Epistemic Logic*, Acta Philosophica Fennica, vol. XXX

Pelc J.

[1990] O pojęciu kłamstwa – z punktu widzenia semiotyki, *Studia Semiotyczne XVI–XVII*, ss. 289–297

Pelc J.

(red) [1991] *Prace z pragmatyki, semantyki i metodologii semiotyki*, Ossolineum, Wrocław

Poczobut R.

[2000] *Spór o zasadę niesprzeczności*, Towarzystwo Naukowe KUL, Lublin

Puczyłowski T.A.

[2000] Problem Gettier'a a logika przekonań, *Edukacja Filozoficzna* 31

Tokarz M.

[1993] *Elementy pragmatyki logicznej*, PWN, Warszawa

## Teza o niezdeterminowaniu przekładu jako przykład Quine'owskiej koncepcji lingwistyki

Quine'owska teza o niezdeterminowaniu przekładu jako koncepcja bardzo szeroko dyskutowana w różnych kręgach może być rozpatrywana na wielorakich płaszczyznach: od filozofii lingwistyki i metod badania języka aż po problemy *stricte* filozoficzne, związane z epistemologiczną tezą o niedookreśleniu teorii naukowych przez dane empiryczne oraz z relatywizmem ontologicznym. Wszystkie te koncepcje będę się starał pokrótce przedstawić, jednakże największy nacisk chciałbym położyć na Quine'owską koncepcję języka i metodologię badań nad językiem a zilustrować ją właśnie tezą o niezdeterminowaniu przekładu.

Dlatego też na wstępie zaprezentuję pokrótce poglądy Quine'a na historyczny rozwój badań nad językiem. Za punkt zwrotny dla tego rozwoju uznawał Quine tzw. eksternalizację empiryzmu, czyli uniezależnienia się koncepcji empirystycznych od teorii idei i metody introspekcji, ku którym skłaniali się wcześnie empiryści, tacy jak np. Locke. Za pierwszą wyraźną oznakę eksternalizacji empiryzmu uznawał Quine publikację angielskiego filologa Johna Horne'a Tooke'a z 1786 roku (za: Quine [1977]), w której pisze on „[...] wszystko, co dotyczy [...] abstrakcji, złożoności, generalizacji, relacji [...] idei, w istocie odnosi się do języka.” Od tej pory, zdaniem Quine'a, idee zostały zdegradowane do podrzędnej roli dodatków do słów, przetrwały jako znaczenia, intensywne, konotacje słów. Na ów „przewrót kopernikański” w badaniach naukowych (jak zwykł eksternalizację nazywać Quine) złożyły się dwa czynniki: filozoficzne przeniesienie uwagi z idei na słowo oraz zaobserwowanie lingwistycznego faktu, że najwyraźniej i najłatwiej rozumiane są słowa odnoszące się do przedmiotów zewnętrznych., to co wewnętrzne, wymaga od tej pory zewnętrznych oznak istnienia. Nowe spojrzenie na koncepcję języka nazywa Quine behawioryzmem. Należy zauważyć, że behawioryzm lingwistyczny nie wiąże się bezpośrednio z tym, co zwykliśmy nazywać behawioryzmem lingwistycznym w psychologii i co od czasów krytyki Chomsky'ego uznawane jest za pogląd nietrafny; Quine twierdził, że behawioryzm lingwistyczny nie implikuje behawioryzmu psychologicznego, czego jego poglądy są

najlepszym przykładem. Mawiał on, że w psychologii można być behawiorystą lub nim nie być, ale w kwestii badań nad językiem nie ma innego wyjścia - lingwista jest niejako behawiorystą z urzędu.

Głównym przeciwnikiem nowej koncepcji behawiorystycznej miał być wszelkiego rodzaju mentalizm, reprezentowany przez tzw. mit muzeum – koncepcję, zgodnie z którą desygnatami słów są eksponaty a ich znaczeniem - etykiety pod eksponatami. Na mocy tej tezy przekład z jednego języka na inny miał być jedynie zmianą etykietek.

Niewłaściwe jest kojarzenie opozycji lingwistycznego behawioryzmu i mentalizmu z rozróżnieniem na to co wyuczone i to co wrodzone. Quine otwarcie przyznaje, że „jego” behawioryzm uznaje za warunek konieczny wszelkiego rodzaju uczenia się istnienie pewnych wrodzonych dyspozycji (bez względu na ich charakter) a w przypadku uczenia się języka będzie to zespół dyspozycji niezwykle bogaty.

Współczesna lingwistyka, pomimo dokonania w wyniku eksternalizacji niejakiego postępu, ciągle, rzec by można, „siedzi okrakiem na murze” pomiędzy słowem a ideą, a o przechyleniu na jedną lub drugą stronę świadczy stosunek poszczególnych badaczy m.in. do pojęcia analityczności - im bardziej nieufny, tym większe przechylenia badacza na stronę werbalną.

Jak widać Quine w specyficzny dla siebie sposób wpisuje się we współczesne koncepcje lingwistyczne (jak choćby tezy Chomsky’ego), ale interpretuje je po swojemu i snuje radykalne plany „nowej lingwistyki”.

Tyle pokrótce na temat ogólnych poglądów Quine na badania lingwistyczne; przejdźmy teraz do fundamentalnego przykładu Quine’owskiego behawioryzmu - niezdeterminowania przekładu.

Teza Quine’a o niezdeterminowaniu przekładu przedstawiona pierwotnie w jego książce *Słowo i przedmiot* to, obok poglądów tego autora na analityczność i synonimiczność (wyłożonych z kolei w słynnym artykule *Dwa dogmaty empiryzmu*), jedna z najgłośniejszych i wzbudzających najwięcej kontrowersji opinii w obrębie filozofii języka. Podobnie jak *Dwa dogmaty...* także niezdeterminowanie przekładu wzbudziło niezliczoną liczbę polemik. Jednak m.in. w związku z nieporozumieniami, jakie teza o niezdeterminowaniu wywołała, Quine chętniej wypowiadał się w tej kwestii niż w sprawie polemik dotyczących jego poglądów na analityczność. Zresztą z pewnych opinii zawartych w *Słowie i przedmiocie* autor wycofał się, przedefiniował pewne pojęcia (m.in. definicje kluczowego pojęcia znaczenia bodźcowego), obstając jednak przy zasadniczym rdzeniu tezy o niezdeterminowaniu przekładu.

Na użytek tej tezy Quine ukuł pojęcie *radykałnego przekładu*. Zachodzi on wówczas, gdy tłumacz nie znając języka tubylców (żyjących na dodatek w kulturowej izolacji - tak by wykluczyć ewentualne wpływy oraz istnienie osób dwujęzycznych znających język owego tubylczego plemienia) musi dokonać tłumaczenia z języka źródłowego (tubylczego) na język przekładu (najczęściej jest nim ojczysty język tłumacza) jedynie w oparciu o bazę empiryczną. Zdaniem Quine'a dwaj niekontaktujący się ze sobą tłumacze mogą stworzyć dwa różniące się od siebie (w sposób zasadniczy) słowniki języka tubylców, które jednocześnie będą zgodne ze wszelkimi dostępnymi danymi empirycznymi. Krótko mówiąc, przekład nie jest zdeterminowany przez dane empiryczne i to, z grubsza rzecz biorąc, głosi teza Quine'a.

Język jako uniwersalne narzędzie, którym ludzie posługują się niemalże we wszystkich dziedzinach swojego życia, jest w dużym stopniu niedostępny dla pozostawionego samemu sobie wśród obcych tłumacza. Jednak wśród całej masy wypowiedzi pojawiających się na co dzień, istnieje pewna kategoria zdań, które stanowią, według Quine'a, furtkę do języka. Są to *zdania obserwacyjne* - elementarne sprawozdania ze świadectw dotyczących świata. Tłumacz po prostu kojarzy próbnie pewne zdanie, które uzna za obserwacyjne, ze zdaniem, które sam w podobnej sytuacji wypowiedziałby w swoim własnym języku, a następnie, gdy sytuacja się powtórzy wypowie owo zdanie w języku tubylców oczekując potwierdzenia lub zaprzeczenia. Jeśli tubylcy potwierdzą wypowiedziane przez tłumacza zdanie, wówczas wstępnie może on przyjąć tożsamość owych dwóch zdań. Warunkiem powodzenia tego eksperymentu jest oczywiście trafne rozpoznanie przez lingwistę oznak aprobaty i negacji u tubylczego plemienia, co jednak nie jest sprawą zbyt skomplikowaną, to co łączy zdanie w języku tubylców i zdanie w języku tłumacza nazywa Quine *znaczeniem bodźcowym*, gdyż wiąże się z aprobatą lub zanegowaniem owych zdań w wyniku takich samych pobudzeń zmysłowych. Na tej samej podstawie powstaje pojęcie *synonimiczności bodźcowej*: zdania obserwacyjne są dla danego użytkownika języka bodźcowo synonimiczne, jeśli ich znaczenia bodźcowe są dla niego takie same. Z kolei dwa zdania są bodźcowo synonimiczne dla danej wspólnoty, wtedy gdy są bodźcowo synonimiczne dla każdego jej członka. Są to punkty zaczepienia, na których oprzeć się może radykalny przekład. Jednak są to również punkty, w których pojawiają się pierwsze kłopoty związane z tezą Quine'a. Jeszcze w *Słowie i przedmiocie* autor nie zauważa problemu intersubiektywnej tożsamości pobudzenia, jaką postuluje jego definicja. Nie wiadomo po prostu, czym taka tożsamość miałaby być i na czym należałoby opierać tezę, że w ogóle coś takiego istnieje. Dlatego Quine



przedefiniował swoje pojęcie zdania obserwacyjnego dla pojedynczego użytkownika języka i zdania obserwacyjnego dla całej wspólnoty. W książce *Na tropach prawdy* pisze: „Jeśli zapytanie o rozważane zdanie wywołuje potwierdzenie ze strony danego użytkownika przy jednej okazji, to wywołuje ono potwierdzenie z jego strony przy każdej innej okazji, gdy pobudzony zostanie ten sam zbiór jego receptorów; i analogicznie w przypadku zaprzeczenia” (Quine [1997] s. 70). Następnie definiuje zdanie obserwacyjne dla całej wspólnoty językowej jako takie, które jest obserwacyjne dla każdego członka tej wspólnoty. Niestety nie rozwiązuje to problemu intersubiektywnej tożsamości pobudzeń, bo jak zasugerowano autorowi, zdania typu „Poczułem, że ryba bierze” jest obserwacyjne dla każdego z członków wspólnoty z osobna, ale nie - dla całej wspólnoty. Pewnym rozwiązaniem byłoby tu proponowane przez Donalda Davidsona (por. Quine [1997]) przeniesienie bodźca z powierzchni ciała i umieszczenie go na zewnątrz jako najbliższą wspólną przyczynę określonych zachowań członków wspólnoty. Quine jednak, ze względu na swoje nastawienie behawiorystyczne, odrzuca takie rozwiązanie i w charakterystyczny dla siebie sposób proponuje, by obywać się jakoś bez pojęcia intersubiektywnego podobieństwa pobudzeń. Twierdzi, że pokrewieństwa zdań należy szukać w zewnętrznych aspektach komunikacji oraz zdać się po prostu na empatię. „My wszyscy posiadamy tajemniczą zdolność wczuwania się w cudzą sytuację percepcyjną, nie znając wcale mechanizmu fizjologicznego czy optycznego tej percepcji. Jest to dar niemal porównywalny z naszą zdolnością rozpoznawania twarzy, której wcale nie towarzyszy umiejętność zobrazowania tych twarzy na rysunku czy w opisie słownym.” (Quine [1997] s.73) Na wizerunku Quine’a-behawiorysty powstaje w tym miejscu pewna rysa: dla podparcia swej behawiorystycznej tezy z zakresu filozofii języka sięga on do arsenału argumentów wziętych z filozofii umysłu w bynajmniej nie-behawiorystycznym ujęciu. Być może nie jest to powód, by odrzucić tezę o niezdeteminowaniu, jednak po raz pierwszy widać tu Quine’a przypartego do muru.

Wróćmy jednak do radykalnego przekładu: kolejnym krokiem jest tłumaczenie spójników logicznych - lingwista zauważa, że pewne zdania obserwacyjne złożone są z innych zdań obserwacyjnych połączonych spójnikami. Znaną już metodą przyglądania się reakcjom tubylców może on skonstruować matryce dla tych spójników i tłumaczyć je na spójniki swojego własnego języka posiadające takie same matryce. Jednak ta operacja nie jest całkowicie wolna od niezdeteminowania.

Obok pojedynczych zdań obserwacyjnych w pełni zdeterminowane, zdaniem Quine’a, są językowe odpowiedniki negacji logicznej: bodziec powinien wywoływać uznanie negacji zdania zawsze i tylko ,wtedy gdy wywołuje on odrzucenie samego zdania i odwrotnie.

Oczywiście, słowa lub zwrotu odpowiadającego logicznej negacji nie należałoby bezwzględnie tłumaczyć jako „nie”, bowiem „nie” ze względu na pewną nieregularność (choćby powszechne w języku polskim zjawisko podwójnego przeczenia) nie spełnia ściśle kryterium negacji. Nie przeszkadza to jednak uznać operacji negacji za w pełni zdeterminowaną.

Inaczej jest już jednak w kwestii spójników koniunkcji i alternatywy. Częściowe kryterium semantyczne dla tych spójników otrzymuje się w sposób analogiczny jak kryterium dla negacji (dla koniunkcji: uznanie całego zdania przy akceptacji obu jego zdań składowych; dla alternatywy: odrzucenie całego zdania przy odrzuceniu obu jego zdań składowych). Jednak jest to tylko zdeterminowanie częściowe, gdyż odrzucenie lub akceptacja poszczególnych alternatyw czy koniunkcji może mieć miejsce również , wtedy gdy , wtedy gdy dana osoba nie ma podstaw ani do odrzucenia ani do zaakceptowania zdań składowych, np. zdanie „Jan urodził się w Warszawie i Jan urodził się w Nowym Jorku” jest ewidentnie fałszywe i zostanie odrzucone, nawet jeśli nie mamy pojęcia, gdzie urodził się Jan. Z kolei przy założeniu, że nasz rozmówca nie potrafi orzec o prawdziwości lub fałszywości zdań składowych zdania „Jan urodził się w Warszawie lub Jan urodził się w Nowym Jorku” może zarówno taką alternatywę odrzucić jak i zaakceptować lub pozostawić nierozstrzygniętą. Quine twierdzi, że „zależy to zarówno od przekładu, jak i od pewnych innych okoliczności.” (Quine [1977] s. 510)

Częściowo niezdzeterminowany jest również pewien sposób kwantyfikacji (uznany m.in. przez Stanisława Leśniewskiego) nazywany przez Quine’a kwantyfikacją substytucyjną. Polega on na tym, że kwantyfikacja egzystencjalna jest prawdziwa wówczas, gdy opuszczając kwantyfikatory i podstawiając za zmienną pewne wyrażenie można otrzymać zdanie prawdziwe, natomiast kwantyfikacja ogólna jest fałszywa, gdy po opuszczeniu kwantyfikatora i podstawieniu za zmienną pewne wyrażenie można otrzymać zdanie fałszywe. Niezdeterminowanie pojawia się tutaj wówczas, gdy ma się podstawy do uznania kwantyfikacji egzystencjalnej, nawet, jeśli nie uznaje się żadnego określonego podstawienia za zmienną. (Quine [1977] ss. 511-513). Quine nie podaje przykładu, ale łatwo go sobie można wyobrazić: zdanie „Syn sąsiadów kradnie nam jabłko z sadu” jest zdaniem, które z pewnością uznamy za prawdziwe, widząc owego chłopca na naszej jabłoni nawet, jeśli nasz sąsiad ma dwóch synów-bliźniaków, których nie potrafimy od siebie odróżnić. Zdanie jest bodźcowo prawdziwe, chociaż nie potrafimy wybrać właściwego podstawienia dla wyrażenia „syn sąsiada”.

Mówienie o spójnikach logicznych w języku naturalnym często budzi sprzeciw, wszak czysto ekstensjonalne spójniki klasycznego rachunku zdań, choć wzorowane na języku naturalnym, nie oddają wszystkich aspektów spójników języków etnicznych, najczęstszymi przykładami są tu problemy przemienności koniunkcji oraz paradoksu implikacji. Jednak w Quine'owskiej koncepcji wyszukiwania spójników logicznych nie chodzi wcale o bezwzględnie ekstensjonalne traktowanie spójników języka naturalnego a jedynie - o wykrycie takich struktur tam, gdzie jest to możliwe (a nie ulega przecież wątpliwości, że często używamy spójników językowych, tak jak używa się spójników logicznych), w celu wykluczenia pewnych oczywistych sprzeczności w języku źródłowym, widocznych już na poziomie syntaktycznym. Quine był bowiem przeciwnikiem postulowania tzw. kultur prelogicznych - twierdził, że jeżeli nie możemy założyć, że system przekonań tubylców jest niesprzeczny (lub stawiając słabszą tezę: że ów system nie jest jawnie sprzeczny), to przekład języka owych tubylców będzie zależał wówczas już tylko od fantazji tłumacza. Innymi słowy - założenie niesprzeczności jest nam potrzebne, by móc w ogóle rozpocząć radykalny przekład.

Jednak większość wypowiedzi, w przeciwieństwie do zdań obserwacyjnych, nie da się skorelować z czynnikami zewnętrznymi. W takiej sytuacji tłumacz może owe wypowiedzi próbować dzielić na segmenty, rozpoznawać wśród nich fragmenty zinterpretowanych już zdań obserwacyjnych, uznawać je za pojedyncze wyrazy i próbować je łączyć tak, by odpowiadały pewnym wypowiedziom jego ojczystego języka. Połączenia takie Quine nazywa *hipotezami analitycznymi*, ich tworzenie polega do pewnego stopnia na zgadywaniu, lecz na jego podstawie tłumacz będzie sukcesywnie rozbudowywał swój słownik, aż stwierdzi, że przy jego pomocy jest zdolny przełożyć dowolne zdanie języka tubylców. W wybieraniu hipotez analitycznych pomagają tłumaczowi pewne kanony lingwistyczne, jednak w tym miejscu zaczyna się już, zdaniem Quine'a, niezdeterminowanie przekładu. Oczywiście dotyczy to tylko przekładu radykalnego, ale też i tylko taki przekład „ukazuje ubóstwo ostatecznych danych, które są podstawą identyfikacji znaczeń.” (Quine [1997] s. 78).

Na czym jednak polega ów kanon lingwistyczny, na którym tłumacz może się z dużym prawdopodobieństwem oprzeć? Jest to pewien zespół założeń, które umiarkowanie wykorzystywał już wcześniej przy interpretacji zdań obserwacyjnych: (1) ciągłość konwersacji - rozsądnie jest przyjąć, że kolejne wypowiedzi tubylców mają ze sobą jakiś związek, (2) hipotezy psychologiczne, czyli zespół przekonań, które mogą żywić tubylcy - niewkładanie im w usta wypowiedzi w oczywisty sposób fałszywych i założenie, że mówią

oni na co dzień o równie banalnych rzeczach co my; (3) sporządzany przez siebie słownik, który tłumacz nieustannie rewiduje, o którego poprawności świadczą gładkie i skuteczne konwersacje.

Spójrzmy teraz na możliwości, jakie nam daje opisana powyżej metoda przekładowa: (1) możemy przekładać zdania obserwacyjne; (2) możemy przekładać funkcje prawdziwościowe; (3) możemy rozpoznawać zdania bodźcowo analityczne i bodźcowo sprzeczne.

Założmy teraz, że dodajemy jeszcze jeden, czwarty, warunek poprawności przekładu. Może być to jakaś hipoteza psychologiczna lub jeden z kanonów lingwistycznych. Założmy też, że dodanie owego warunku powoduje, że dla dowolnego zdania języka źródłowego otrzymujemy dokładnie jeden przekład tego zdania w języku docelowym. Quine powie wówczas, że nie redukuje to niezdeterminowania przekładu, gdyż za koniecznością warunków (i)-(iii) przemawia doświadczenie, natomiast nie można tego powiedzieć o warunku (iv), bowiem jego akceptacja lub odrzucenie „nie jest sprawą faktów.”

Wszystko jednak, o czym była mowa do tej pory, jest tylko eksperymentem czysto teoretycznym nie mającym wiele wspólnego z rzeczywistością i sam Quine oczywiście zdaje sobie z tego sprawę:

„Nie jest prawdopodobne, by niezdeterminowanie przekładu ujawniło się w praktyce, nawet w przekładzie radykalnym. [...] Tym, na co ma zwrócić uwagę teza o niezdeterminowaniu przekładu, jest fakt, że radykalny translator musi założyć równie wiele, jak odkrywa.” (Quine [1993a] s. 126).

Jednak nie jest to jedyny wniosek Quine’a z tego eksperymentu myślowego. Głównym jego celem było wykazanie, że teorie znaczenia wykorzystujące pojęcie sądu i oparte na różnych koncepcjach mentalistycznych w rzeczywistości nie mają oparcia w empirii, a podejście behawiorystyczne w teorii języka jest, zdaniem Quine’a, jedynym dopuszczalnym i uprawnionym. Jednym słowem - znaczenie bodźcowe nie jest behawioralną namiastką jakiegoś *prawdziwego* znaczenia, lecz jedynym znaczeniem, jakie w rzeczywistości występuje w języku.

Na podstawie tego, co do tej pory było powiedziane, można by wnioskować, że niezdeterminowanie przekładu jest tylko pewnym szczególnym przypadkiem tego, co znamy

pod nazwą niedookreślenia teorii naukowej przez dane empiryczne. A jest to przecież koncepcja, która nie wzbudza specjalnych emocji i nie prowokuje polemik. Głosi jedynie, że mogą istnieć dwie sprzeczne ze sobą teorie naukowe, które jednocześnie są zgodne ze wszystkimi danymi empirycznymi. Quine jednak twierdzi, że niezdeteminowanie przekładu jest rzeczą poważniejszą ze względu na niezdeteminowanie odniesienia i stworzoną przez niego na użytek tego twierdzenia tezę o relatywizmie ontologicznym.

Niezdeteminowanie odniesienia, to poniekąd językowy odpowiednik niedookreślenia teorii naukowych; posłużę się tu klasycznym, stworzonym przez Quine'a przykładem: na widok przebiegającego królika tubylcy krzyczą "Gavagai!", co pozwala sądzić, że jest to zdanie obserwacyjne, które można przetłumaczyć jako "(Patrz) królik!" Jest to skrajny przypadek, gdyż zdanie składa się tu z jednego wyrazu. Nie jest to jednak powód, zdaniem Quine'a, by termin "gavagai" tłumaczyć jako "królik", gdyż równie dobrze w obrębie tego zdania (jak i wszystkich dostępnych zdań o królikach) "gavagai" może oznaczać "stadium królika", "część królika" lub nawet "cały kosmos poza tym królikiem". I, o ile w obrębie jednego zdania możemy mieć do czynienia z takim różnym rozumieniem poszczególnych słów, które w sumie dają zdanie o takim samym znaczeniu, to w obrębie całego języka prowadzi to do tego, że pewne inne terminy zmuszeni będziemy tłumaczyć w inny sposób. Quine mówi, że odniesienie jest tutaj zrelatywizowane do podręcznika przekładu, w którym termin "gavagai" jest tłumaczony jako „królik”. W istocie nie jestem w stanie przeprowadzić ostrej granicy pomiędzy niezdeteminowaniem odniesienia a relatywizmem ontologicznym. Usprawiedliwia mnie jednak fakt, że sam Quine przyznaje się do podobnego problemu. Mowa w każdym razie o tym, że decydując się na jakiś sposób przekładu dokonujemy arbitralnego wyboru jednej z ontologii, które są empirycznie równoważne, gdyż nie ma empirycznych podstaw, by wybrać raczej jedną niż drugą, a w ontologii empirycznie istotny jest tylko jej "wkład neutralnych węzłów" (tzn. przedmiotów) do "struktury teorii" (utkanej ze zdań obserwacyjnych i teoretycznych). (Quine 1997: 56-60) Tak oto Quine szuka usprawiedliwienia dla swej tezy o niezdeteminowaniu przekładu w metafizyce.

Częstym zarzutem przeciwko Quine'owskiej filozofii języka jest jej skrajnie empirystyczna orientacja - znaczenie zredukowane do „znaczenia bodźcowego” ze względu na swe ubóstwo nazywane jest „wegetariańską semantyką” i pod jej adresem kierowane są najczęściej zarzuty o metodologiczną niepoprawność. Stąd też dziwić może fakt, że pewne ciekawe uwagi metodologiczne sformułowano z pozycji empirystycznych. Barbara Stanosz ([1991]: 86-7) zauważyła, że Quine poczynił pewne bezzasadne, z empirycznego punktu

widzenia, założenia: sytuacja radykalnego przekładu wyklucza nie tylko dwujęzyczność badacza i pomoc dwujęzycznych tłumaczy, ale i również - Quine'owskie techniki badania dyspozycji użytkowników języka do zachowań werdyktowych. Nie mają one bowiem czysto eksperymentalnego charakteru, zakładają pewien rodzaj kooperacji pomiędzy tłumaczem a użytkownikami badanego języka. Mimo że ludzie mają naturalną dyspozycję do uznawania lub odrzucania pewnych zdań pod wpływem określonych pobudzeń zmysłowych, to nie mają jednak naturalnej dyspozycji do potwierdzania lub negowania tychże zdań. Innymi słowy: zakłada się tu prawdomówność i życzliwość wobec badacza, co oczywiście jest założeniem niezgodnym z wytycznymi metodologicznymi naukowego empiryzmu. (Stanosz 1991) Skądinąd wiemy jednak, że Quine'owi na owej naukowości badań lingwistycznych zależy, gdyż lingwistyka jest dla niego częścią „nauk behawioralnych”. (Quine [1977] s. 180)

Uwagi Barbary Stanosz mają poważniejsze konsekwencje niż z pozoru może się wydawać. Można bowiem na ich podstawie zapytać, czy człowiek w ogóle ma jakiegokolwiek dyspozycje do takich a nie innych zachowań werbalnych, wszak prawie nigdy nie jesteśmy w stanie przewidzieć, co ktoś powie w danej sytuacji. Stąd wniosek, że semantyka Quine'owska jest nie „wegetariańska”, lecz wręcz przeciwnie – „grzeszy ona przeciwko zasadom empiryzmu”. (Stanosz [1991] s. 87)

W związku z tym rozwiązania tej sytuacji autorka szuka na gruncie dyspozycji zachowaniowych odbiorców, nie zaś - nadawców. Jako że odbieranie sygnałów dźwiękowych jest, obok doświadczenia, jednym z głównych źródeł ludzkiej wiedzy i przekonań, te zaś warunkują zachowanie każdego człowieka, to przy założeniu racjonalności zachowań ludzi (akceptowanym przez wszystkie dyscypliny społeczne i humanistyczne) można uzyskać zadowalający ekwiwalent Quine'wskich dyspozycji do zachowań werbalnych.

„Załóżmy, że osoba  $o$  odebrała sygnał dźwiękowy (zdanie)  $s$  języka  $J$  w sytuacji, w której nie działał żaden inny czynnik przekonaniotwórczy. Niech  $B_1$  będzie zbiorem zdań reprezentujących przekonania osoby  $o$  przed odebraniem  $s$ ,  $B_2$  zaś zbiorem zdań reprezentujących jej przekonania po odebraniu  $s$ . Zakładając, że  $B_1$  i  $B_2$  są systemami dedukcyjnymi (każdy z nich zawiera wszelkie konsekwencje logiczne), mamy:

$$\text{Cn}(B_1 \cup \{s'\}) = B_2$$

gdzie  $C_n$  jest funkcją konsekwencji logicznej, a  $s'$  wyraża przekonanie, o jakie wzbogaciła się osoba  $o$  przez odebranie  $s$ . Jeśli  $s'$  oraz elementy zbiorów  $B_1$  i  $B_2$  są zdaniami języka  $J'$ , to  $s'$  jest przekładem zdania  $s$  języka  $J$  na język  $J'$ . Innymi słowy,  $s'$  w  $J'$  jest przekładem  $s$  w  $J$ , wtedy gdy  $s'$ , dołączone do zbioru zdań  $J'$  wyrażających przekonanie użytkownika  $J$  przed odebraniem  $s$ , aksjomatyzuje zbiór zdań  $J'$  wyrażających jego przekonania po odebraniu  $s$ ." (por. Stanosz [1991])

Zatem jeśli tłumacz potrafi opisać w swoim języku stany przekonaniowe członków badanej społeczności językowej, to potrafi również, obserwując zmiany ich dyspozycji zachowaniowych, przełożyć niektóre zdania języka tej społeczności, a konkretnie te, które zmieniają stany przekonaniowe odbiorców. Metodologiczne usztywnienie Quine'owskiej koncepcji nie podważa oczywiście samej tezy o niezdeterminowaniu. Wręcz przeciwnie - z empirystycznego punktu widzenia ugruntowuje ją jeszcze bardziej. Modyfikacja następuje właściwie jedynie w pierwszym kroku - zdania obserwacyjne nie są już rekcją werbalną na bodźce zewnętrzne, lecz bezpośrednią i widoczną oznaką zmiany stanów przekonaniowych odbiorców. Będą to wszelkiego rodzaju instrukcje wydawane podwładnym przez zwierzchników w rozmaitych przedsięwzięciach zbiorowych, ostrzeżenia o niebezpieczeństwie i informacje o faktach czy zjawiskach, z których można odnieść natychmiastową korzyść. Dalszy ciąg eksperymentu myślowego przebiegałby podobnie jak u Quine'a.

Na poparcie Quine'owskiej tezy o niezdeterminowaniu przekładu Gilbert Harman ([1969] s. 49) wymyślił argument pokazujący to zjawisko na gruncie przekładu zdań w obrębie języków formalnych. Jako że sam Quine mówił od początku o niemożności zobrazowania jakimkolwiek przykładem niezdeterminowania w językach naturalnych (przykład z królikiem częstokroć mylnie brany za takowy jest w rzeczywistości przykładem niezdeterminowania odniesienia), pomysł Harmana okazał się użyteczny dla obrońców tezy Quine'a, stał się jednak również obiektem ataków ze strony jej krytyków.

Pomysł Harmana jest prosty: posłużył się on różnymi wersjami teoriomnogościowej definicji liczby porządkowej. Definicje te dołączone do jakiejś aksjomatyki teorii mnogości tworzą razem z nią różne systemy, w których zdanie „ $3 \in 5$ ” przyjmuje różne wartości logiczne. W standardowej definicji Zermela zero jest utożsamiane ze zbiorem pustym, natomiast każda inna liczba naturalna  $n$  - z pojedynczym zbiorem, którego jedynym elementem jest  $n-1$ . Z kolei w definicji von Neumanna, każda liczba naturalna  $n$  utożsamiana

jest ze zbiorem liczb mniejszych od  $n$ . Tym samym zdanie „ $3 \in 5$ ” na gruncie teorii Zermela jest fałszywe, natomiast na gruncie teorii von Neumanna - prawdziwe. Fakt ten ma służyć jako przesłanka uzasadniająca tezę o niezdeterminowaniu przekładu: oto pewne zdanie przełożone na język teorii mnogości w zależności od przyjętych wcześniej przekładów pojęć podstawowych ma różną wartość logiczną. Jednak przede wszystkim ma pokazywać, że gdy mamy do czynienia z tłumaczeniem, to “nie jest sprawą faktów”, jaki wybierzemy podręcznik przekładu. Oczywiście przykład nie jest pozbawiony mankamentów - zdanie „ $3 \in 5$ ” poza teorią mnogości jest bezsensowne, nie można o nim powiedzieć, czy jest prawdziwe, czy fałszywe. Jednak w przypadku ilustrowania niezdeterminowania wyboru słownika przez fakty empiryczne nie odgrywa to większego znaczenia.

Zarzuty przeciwko argumentowi Harmana sformułował Richard Rorty ([1993] ss. 149-153). Twierdzi on, że przykład niezdeterminowania na gruncie języków formalnych eliminuje pewne ważne aspekty przekładu języków naturalnych. Rorty mówi, że zwykle tłumacz ma pewne intuicje związane z przekładanymi zdaniami i wyrażeniami, natomiast trudno mówić o intuicjach w przypadku przekładania liczb na język teorii mnogości - żadna intuicja nie powie nam czy liczba 3 jest singletonem liczby 2 czy też 3 jest zbiorem liczb 0, 1 i 2 - tu rzeczywiście nie ma “sprawy faktów”, co nie znaczy, że w języku naturalnym jest podobnie. Poza tym - przykład Harmana pozbywa się również tego, co Rorty nazywa “ukrytymi kanonami”, czyli - pewnych ogólnych reguł, którymi posługuje się tłumacz. Do takich reguł należy np. zasada tłumaczenia wyrażenia jednowyrazowego w języku źródłowym na jednowyrazowe w języku przekładu. O tym także wspominał sam Quine, jednak uznał te reguły za jedno ze źródeł niezdeterminowania. Rorty chciałby po prostu pewne intuicje i regularność podnieść do rangi praw naukowych lingwistyki. Od takich praw jest jednak, jak wiadomo, niezliczona liczba wyjątków, a przecież prawa, od których są wyjątki, nie są żadnymi prawami. Jeśli więc chcemy traktować lingwistykę jako naukę empiryczną, a na takim stanowisko stoi Quine, wówczas nie możemy się przychylić do uwag Rorty’ego.

Wśród antybehawiorystycznie nastawionych polemistów Quine’a panuje opinia o zbyt wąskim pojęciu świadectw empirycznych. Zarzutom tym na ogół jednak nie towarzyszą konkretne kontrproponycje. Wyjątkiem jest tutaj Katz - jest on skłonny przyznać, że w sytuacji radykalnego przekładu badacz nie może jednoznacznie ustalić znaczeń wyrażenia języka, który chce przełożyć. Twierdzi jednak, że scenariusz zarysowany przez Quine’a różni się w sposób istotny od wszelkich rzeczywistych sytuacji przekładowych, przez co nie może być ich przypadkiem granicznym. Katz [1978] twierdzi, że we wszystkich rzeczywistych



sytuacjach na którymś z etapów dostępna jest pomoc osób dwujęzycznych. Quine'owski radykalny przekład jest nieuprawnioną idealizacją sytuacji przekładowych.

Gdy zaś mamy do dyspozycji osoby dwujęzyczne, wówczas można, zdaniem Katza ([1978] s. 254), testować hipotezy translatorskie przez zadawanie owym osobom pytań o podobieństwo odpowiednio dobranych próbek materiału językowego. Na przykład na właściwe znaczenie słowa „gavagai” - jako równoznacznego ze słowem „królik” a nie: „faza czasowa królika” czy „część królika” - wpaść można dzięki odpowiedziom na pytania, czy „gavagai” jest bardziej podobne do słów „dzieciństwo” i „młodość”, czy do słów „dziecko” i „młodzieniec” a także na pytanie, czy „gavagai” jest bliższe słowom „gałąź” i „ręka”, czy też słowom „drzewo” i „ciało”.

Pytanie o podobieństwo czegoś do czegokolwiek innego wymaga relatywizacji - wskazania względu, pod którym podobieństwo ma być oceniane. Do tego potrzebne byłoby jednak użycie pojęć semantycznych, czego Katz właśnie chciał uniknąć. “Póki zaś wzgląd nie jest wskazany, pytanie pozostaje bezsensowne, a odpowiedzi na bezsensowne pytania mogą być diagnostyczne co najwyżej dla psychiatry czy psychologa; dla badacza języka są nie tylko niediagnostyczne, lecz wręcz zwodnicze.” (Stanosz [1991] s. 85)

Teza Quine'a ma pewne słabe punkty, w które uderzają jego krytycy. Wcześniej już zasugerowałem, że jednym z takich punktów jest postulowany przez niego relatywizm ontologiczny, bez którego niezdeteminowanie przekładu przestaje być w gruncie rzeczy kontrowersyjną i jednocześnie tak istotną tezą. Innym takim punktem jest Quine'owskie pojęcie znaczenia bodźcowego atakowane jako nie spełniające warunków, jakie nakładamy na pojęcie znaczenia, a o którym skądinąd wiemy, że istnieje (Searle [1993] s. 158) lub, że jest wręcz „samobójcze” - wykorzystywane bywa na gruncie teorii, która prowadzi do konkluzji, że owo pojęcie znaczenia jest pozbawione sensu (Zabłudowski [1993] s. 182). W istocie znaczenie bodźcowe nie przypomina niczego, co bylibyśmy skłonni uznać za *znaczenie* - odwołuje się do korelacji między bodźcami zewnętrznymi i dyspozycjami do zachowań językowych, co prowadzi do tego, że pewne wyrażenia jak „królik” i „stadium królika”, o których wiemy, że *znaczą* co innego, są bodźcowo synonimiczne – „zdania mające to samo znaczenie bodźcowe - jak mówi Searle - nie mają tego samego znaczenia w żadnym przyjętym sensie słowa «znaczenie»” (Searle [1993] s. 159). Relatywizm ontologiczny nie rozwiązuje tego problemu - mówienie, że odnosimy się do królika lub do stadium królika jedynie na mocy arbitralnie przyjętej ontologii a w rzeczywistości nie jesteśmy w stanie rozróżnić królika od jego stadium, gdyż te bodźce są synonimiczne, jest zdaniem Searle'a,

zwykłym absurdem, gdyż wówczas w ogóle nie bylibyśmy w stanie zrozumieć powyższego wywodu. Przyznać trzeba, że krytyka dokonana przez Searle'a jest posunięciem dość śmiałym, wręcz brawurowym - autor nie odwołuje się tu do formalnego wnioskowania, lecz sprowadza rozumowanie do prostych przekonań człowieka dotyczących znaczenia. Stąd jest to argument śmiały, ale i niezwykle sugestywny.

Jednakże przy bliższym przyjrzeniu się argumentacji Searle'a zauważyć można błędy w rozumowaniu wynikające z niezrozumienia idei Quine'a. Różne tłumaczenia wyrażenia „gavagai” są rezultatem użycia różnych, niezgodnych ze sobą podręczników. Gdy będziemy łączyć wyrażenia dwóch różnych języków zgodnie z jakimś słownikiem przekładu może się okazać, że poszczególne podręczniki korelują z danym wyrażeniem języka obcego różne wyrażenia naszego języka, np. z wyrażeniem „gavagai” odpowiednio wyrażenia „królik”, „nieodłączna część królika”, „faza czasowa królika”, itd. Owe słowniki będą się różnić (zgodnie z ideą holizmu) również w wielu innych miejscach. Należy jednak pamiętać, że decyzję o wyborze podręcznika przekładu podejmuje się raz i nieodwołalnie - nie można korzystać jednocześnie z wielu podręczników. Na tym polega błąd brawurowej argumentacji Searle'a. (Føllesdal [1990] s. 105)

Dla prowadzenia badań nad językiem nie jest w gruncie rzeczy istotne to, czy uznajemy samą tezę o niezdeterminowaniu przekładu czy też nie. Ważne jest jednak to jak do tych badań podchodzimy - być może najbardziej istotnym osiągnięciem dyskusji wokół niezdeterminowania było ponowne wywołanie sporu o metodę badań nad językiem. A na gruncie owej dyskusji można wyodrębnić dwie zasadnicze grupy zwolenników konkretnej wizji filozofii lingwistyki i filozofii języka. Zwolennicy Quine'a chcieliby widzieć w lingwistyce prawdziwą naukę empiryczną - *hard science*. Z kolei jego przeciwnicy dopuszczają pewne nienaukowe metody zbliżając tym samym lingwistykę do dyscyplin społeczno-humanistycznych. Podobny podział widać również w dyskusji na temat analityczności czy holizmu semantycznego. Stąd wniosek, że nie są to tylko dyskusje dotyczące wąskich tematów filozoficznych znanych jedynie specjalistom - problemy analityczności czy niezdeterminowania przekładu dotyczą również fundamentalnych zagadnień filozofii języka i lingwistyki a konsekwencje mają również w wielu innych dziedzinach, takich jak np. filozofia umysłu czy psychologia kognitywna. Dlatego też spory dotyczące tych tematów toczą się od tak wielu lat z taką zaciętością.

Quine'owska koncepcja lingwistyki nie zyskała takiej akceptacji, jakiej pewnie życzyłby sobie jej autor. Trudno się zresztą dziwić - jego projekt był niezwykle radykalny, a

surowe wskazówki metodologiczne zmusiłyby wielu badaczy do zdecydowanego zwrotu w ich pracy. Jednakże sam pomysł uczynienia z lingwistyki nauki ścisłej jest wciąż żywy - wielu filozofów, jak np. Keith Devlin, uznaje wręcz obecną lingwistykę formalną za naukę matematyczną (Devlin [1999] ss. 136-7) Z kolei psycholingwistyka kognitywna, w której Quine pokładał wielkie nadzieje, od lat już rozwija z sukcesami swoją metodę badawczą, choć może nie według surowych Quine'owskich reguł. Jedynym ogniwem koncepcji „nowej” nauki o języku, która czeka na swój przełom, jest bodaj neurolingwistyka, choć i tu owego przełomu można się w najbliższej przyszłości spodziewać.

Wszystko to pozwala twierdzić, że coraz większej liczbie badaczy bliska jest myśl o lingwistyce jako prawdziwej, ścisłej nauce empirycznej i że niewiarygodna z pozoru koncepcja Quine'a znajdzie swoją realizację, choć może na innej drodze, niż przewidywał to Quine.

#### *BIBLIOGRAFIA:*

BARRETT R. B., GIBSON R. F.

(red) [1970] *Perspectives on Quine*, Basil Blackwell, Cambridge, Massachusetts.

DAVIDSON D., HINTIKKA J.

(red) [1969] *Words and Objections. Essays on the Work of W. V. Quine*, Dordrecht, D. Reidel Publishing Company.

DEVLIN K.

[1999] *Żegnaj, Kartezjuszu*, Prószyński i S-ka, Warszawa.

FØLLESDAL D.

[1990] *Indeterminacy and Mental States*, w: Barrett, Gibson [1990].

GUENTHER F., GUENTHER-REUTTER M.

(red) [1978] *Meaning and Translation*, New York University Press, New York.

HARMAN G.

[1969] *Introduction to Chapter Two of 'Word and Object'*, w: Davidson, Hintikka

[1969]

KATZ J.

[1978] *Effability and Translation*, w: Guenther F., Guenther-Reutter M. [1978].

- [1990] *The Refutation of Indeterminacy*, w: Barrett, Gibson [1990].
- PUTNAM H.
- [1975] *Mind, Language and Reality*, Cambridge University Press.
- QUINE W.V.O.
- [1977] *Filozoficzny postęp w teorii języka*, w: Stanosz [1977].
- [1993] *Raz jeszcze o niezdeterminowaniu przekładu*, w: STANOSZ [1993].
- [1993a] *Trzy niezdeterminowania*, w: STANOSZ [1993].
- [1997] *Na tropach prawdy*, Aletheia, Warszawa.
- [1999] *Słowo i przedmiot*, Aletheia, Warszawa.
- RORTY R.
- [1993] *Niezdeterminowanie przekładu i prawdy*, w: STANOSZ [1993].
- SEARLE J.R.
- [1993] "Niezdeterminowanie, empiryzm i pierwsza osoba", w: Stanosz 1993;
- STANOSZ B.
- (red) [1977] *Lingwistyka a filozofia*, PWN, Warszawa.
- [1991] *10 wykładów z filozofii języka*, Biblioteka Myśli Semiotycznej 19., Warszawa.
- (red) [1993] *Filozofia języka*, Aletheia - Spacja; Warszawa.
- ZABŁUDOWSKI A.
- [1993] *O Quine'a teorii niezdeterminowania*, w: STANOSZ [1993].

# **FILOZOFIA A BIOLOGIA**

## **Próby przewycięzenia filozoficznego paradoksu Empedoklesa w protobiologii**

### *1. Cele artykułu.*

Protobiologia, nazywana także nauką o biogenezie, jest interdyscyplinarną dziedziną badań prowadzonych w zakresie nauk szczegółowych. W jej ramach próbuje się uzyskać odpowiedź na zasadnicze pytania: jaka jest geneza życia na Ziemi; jaką formę miały pierwsze organizmy żywe? Prostota tych pytań jest tylko pozorna - wystarczy wspomnieć, że od 1957 roku, kiedy zwołano pierwszą międzynarodową konferencję poświęconą studiom nad pochodzeniem życia, powstało około stu dwudziestu ugruntowanych eksperymentalnie i teoretycznie koncepcji biogenezy.

Celem niniejszego artykułu będzie pokazanie, że przyrodnicy, próbujący dać odpowiedź na fundamentalne pytania protobiologii, stawali wobec problemu o charakterze czysto filozoficznym, który będę nazywać paradoksem powstawania celowości lub skrótowo – paradoksem Empedoklesa<sup>50</sup>. Na wybranych przykładach spróbuję pokazać, że naukowcy osiągają różne efekty w przewycięzaniu tego filozoficznego paradoksu między innymi dlatego, że ich badaniom towarzyszą odmienne przesłanki filozoficzne, często entymematycznie zawarte w ich teoriach. Postaram się, aby zostały one wyrażone *explicite*, korzystając przy tym z wyników pracy W. Ługowskiego<sup>51</sup>, za którym zresztą definiuję filozoficzne przesłanki teorii jako jej przesłanki metodologiczne i ontologiczne.<sup>52</sup> Rozwijając wątki zawarte u Ługowskiego, wskażę na to, że pewne przesłanki filozoficzne w teoriach

---

<sup>50</sup> Właśnie Empedokles był pierwszym filozofem, który próbując rozwikłać aporie eleackie – z jednej strony przekonanie o niezmienności bytu, z drugiej dane doświadczenia zdające się potwierdzać jego zmienność – rozważał problem powstawania i ginięcia. Ostatecznie Empedokles przyznał pierwszeństwo elementom, tłumacząc powstawanie wszelkich całości jako łączenie się czterech podstawowych i jakościowo niezmiennych elementów. Widać zatem, że Empedokles jako pierwszy sformułował problem, choć nie uświadamiał sobie jeszcze jego paradoksalnego charakteru. Por. Reale [1994] s. 171-181.

<sup>51</sup> Por. Ługowski [1987] oraz Ługowski [1995].

<sup>52</sup> Por. Ługowski [1995] s. 130.

biogenezy pojawiają się jako konsekwencja konkretnego typu definicji życia, przyjmowanej na gruncie tych teorii. Korzystając z poczynionych ustaleń, poruszę problem wartości eksplanacyjnej teorii protobiologicznych.

Dla porządku należy wspomnieć, że przy omawianiu technicznych aspektów koncepcji protobiologicznych A.I.Oparina i S.W.Foxa nie sięgałem do źródeł, bazując na pracach W.Ługowskiego. Sytuacja ta nie dotyczy koncepcji biogenezy J.M.Smitha.

## ***2. Sformułowanie paradoksu Empedoklesa na gruncie biogenezy.***

Paradoks ten w swoim filozoficznym sformułowaniu przyjmuje postać następującego pytania: co istniało wcześniej, część czy całość? Jeśli spróbujemy odpowiedzieć, że część, pojawi się niechybnie wątpliwość, że skoro część ma być częścią, to musi być ona częścią jakiejś całości, a zatem część nie może w porządku genetycznym wyprzedzać całości. Jeśli natomiast odpowiemy, że wcześniej istniała całość, pojawi się podobna wątpliwość, że skoro całość ma być całością, to musi być ona złożona z jakichś części, a zatem całość nie może w porządku genetycznym poprzedzać całości. Po zastosowaniu takiego rozumowania wyjściowe pytanie zdaje się nie mieć dobrej odpowiedzi.

Na gruncie protobiologii podobny paradoks powstaje wówczas, gdy układom żywym przypiszemy pewną konstytutywną cechę „celowości” organizacji elementów. Celowość ta ma się przejawiać na dwa sposoby: po pierwsze – jako przystosowanie elementów układu do pełnionych przez nie funkcji w ramach tego układu, po drugie – jako przystosowanie całego układu do życia w danym środowisku. Łatwo zauważyć, że, jeśli zaczniemy rozważać genezę pierwszego w historii układu żywego, z konieczności uwikłamy się w paradoks Empedoklesa. Jeżeli pierwszy żywy układ miałby celowo funkcjonować jako całość, musiałby powstać z celowo przystosowanych elementów. Gdybyśmy się jednak na to zgodzili, należałoby również przyznać, że źródłem celowości elementów układu, tzn. ich funkcjonalnego przystosowania do pełnienia określonej roli w ramach układu, jest rozpad istniejącego wcześniej całego organizmu (a to nie wchodzi w rachubę, gdyż rozważamy obiekt genetycznie pierwszy) albo istniejący uprzednio plan. Jak inaczej moglibyśmy bowiem wytłumaczyć to, że nagle dochodzi do spontanicznego samomontażu elementów, z których każdy już „na wejściu” jest doskonale przystosowany do roli, jaką ma pełnić w ramach całości? Założenie rozpadu całego organizmu wikła nas jednak w regres, a założenie istnienia

planu nie jest zadowalające z punktu widzenia kogoś, kto chce sprostać kryteriom wyjaśnienia naukowego. Rozważając w następnych paragrafach wybrane teorie protobiologiczne wskażę na to, że próby przewycięzania paradoksu Empedoklesa w protobiologii mogą występować co najmniej na trzy sposoby:

- (1) poprzez odrzucenie założenia o tym, że cechą konstytutywną układów żywych jest funkcjonalna celowość ich elementów (rozwiązanie J.M.Smitha)
- (2) poprzez przyznanie prymatu całości pozbawionej funkcjonalnie określonych części oraz wskazanie na mechanizm, który mógłby być odpowiedzialny za ostateczne funkcjonalne ukształtowanie części (rozwiązanie A.I.Oparina)
- (3) poprzez modyfikację kryteriów wyjaśnienia naukowego, tak aby dopuścić możliwość istnienia planu (rozwiązanie S.W.Foxa)

### *3. Metodologiczny status definicji życia.*

Po zapoznaniu się różnymi teoriami biologicznymi, a w szczególności - z teoriami protobiologicznymi, można zauważyć, że stosowane w nich definicje życia (występujące zazwyczaj jako nominalne definicje nazwy „organizm żywy”) ze względu na ich zadanie można zaliczyć do regulujących definicji projektujących. Z definicją projektującą mamy do czynienia, wtedy gdy ustalamy znaczenie jakiegoś słowa w ramach naszego przyszłego projektu. Jej regulująca odmiana występuje wówczas, gdy podczas projekcji znaczenia uwzględniamy już zastane – choć zdecydowanie nieostre – znaczenie danego wyrazu<sup>53</sup>. „Układ żywy” jest niewątpliwie nazwą niewyraźną, a wiązane z nią potocznie znaczenia nie są wystarczające dla celów teoretycznych. Dlatego przyrodnicy precyzują to znaczenie na gruncie swoich teorii. Efekt jest taki, że poszczególni naukowcy mogą posługiwać się takimi samymi terminami, przypisując im różne znaczenia. W niektórych przypadkach ma to swoiste konsekwencje dla ustalenia cezur życia – ten układ, który według jednego badacza spełnia definicyjnie wyznaczone kryteria układu żywego, przez innego badacza nie może być uznany za żywy. Pokażę, że tego typu sytuacja ma miejsce w teoriach Smitha i Oparina.

W przypadku Smitha można przyjąć, że z definicji uznaje on za żywy taki układ, który charakteryzuje się własnościami powielalności, dziedziczności oraz zmienności. Dla porządku zapiszmy definicję Smitha w stylizacji semantycznej i oznaczmy indeksem:

---

<sup>53</sup> Por. Ziemiński [2001] s. 46-47.



*Df(S): Zwrot „organizm żywy” oznacza wszelkie byty posiadające cechy powielalności, dziedziczności oraz zmienności.*

W przypadku Oparina definiowanie życia przebiega nieco inaczej, a lista cech typowych dla układów żywych jest różna od zaproponowanej przez Smitha<sup>54</sup>. W stylizacji semantycznej ma ona następującą formę:

*Df(O): Zwrot „organizm żywy” oznacza wszelkie byty posiadające cechy celowego przystosowania budowy swoich elementów do pełnionych przez nie funkcji, dziedziczności oraz zdolności do przeciwstawiania się wzrostowi entropii.*

W przypadku Foxa nie dysponuję materiałem źródłowym, który umożliwiłby zacytowanie jego definicji życia, lecz w ramach teorii umieszcza on cezurę życia w ten sposób, że rekonstruując na jej podstawie definicję życia, otrzymalibyśmy wersję zbliżoną do Df(O).

Rozważając treść tych definicji łatwo zauważyć, że Df(S) jest szersza od Df(O), tzn. że zakres jej definiendum obejmuje więcej przedmiotów, niż zakres definiendum Df(O). Df(O) nakłada bardziej restrykcyjne warunki na obiekty, które chcielibyśmy nazwać układami czy organizmami żywymi, wymagając od nich oprócz dziedziczności (która zakłada jakąś formę powielania) także - celowego przystosowania elementów oraz trwałości w czasie. W kolejnym paragrafie pokażę, w jaki sposób definicje życia pociągają za sobą konieczność przyjęcia konkretnych reguł metodologicznych oraz wrócę do sposobów przewyżczania paradoksu Empedoklesa, wymienionych w paragrafie 2.

#### ***4. Merytoryczna zawartość teorii i ich***

##### ***założenia filozoficzne***

Koncepcje protobiologiczne powstają w ramach naukowego paradygmatu, którego wspólne przyrodnicze przesłanki właściwie ograniczają się do następującej listy:

- założenie, że u podstaw powstawania życia miała miejsce długotrwała ewolucja chemiczna, rozumiana jako proces powstawania prostych substancji organicznych na drodze abiotycznej
- założenie o radykalnej odmienności pierwotnego środowiska (tzw. pierwotny bulion) od obecnego

---

<sup>54</sup> Por. Ługowski [1995] s. 108.

- hipoteza o heterotroficznym charakterze pierwszych organizmów
- dane biochemiczne o wspólnym budulcu organizmów (20 aminokwasów, 4 nukleotydy)
- dane kosmochemiczne, że pierwiastki fundamentalne dla życia są powszechnie dostępne w kosmosie<sup>55</sup>.

Zbiór tych faktów i hipotez jest przyjmowany – z nielicznymi wyjątkami – właściwie przez wszystkich badaczy. Inaczej ma się sytuacja z przesłankami filozoficznymi teorii. W tej kwestii różnicowanie wśród naukowców jest większe, co postaram się wykazać, rozważając trzy wspomniane już przypadki.

Przyjmowana na gruncie teorii Smitha Df(S) pozwala uznać, że pierwszymi organizmami żywymi były podlegające replikacji dwuniciowe polinukleotydy. Wszak spełniały one trzy zasadnicze warunki: powielalności (poprzez replikację), zmienności (za sprawą mutacji) oraz dziedziczenia (za sprawą komplementarności zasad)<sup>56</sup>. Jednocześnie te właśnie molekuly - dokładnie ze względu na te same wymienione powyżej cechy - stawały się bytami, na których zaczynał działać Darwinowski mechanizm doboru naturalnego. To właśnie elegancki w swojej prostocie mechanizm doboru naturalnego jest w teorii Smitha jedynym czynnikiem, odpowiedzialnym za zdumiewające różnicowanie wszystkich późniejszych form życia. Jest on także na gruncie tej teorii jedyną zasadą wyjaśniającą to różnicowanie. Filozoficzna perspektywa w myśleniu Smitha, ujawnia się poprzez jego ewentystyczne nastawienie do problemu powstania minimalnego organizmu oraz procesualne nastawienie do problemu następstw tego zdarzenia. Inaczej mówiąc, Smith początków życia szuka w korzystnym przypadku (jeśli nawet nie jednym, to kilku), którego skutkiem było powstanie molekuly o wspomnianych własnościach układu żywego. Molekuła ta jest na tyle nieskomplikowana, że uwzględniając wszystkie przyrodnicze założenia teorii, jej powstanie - choć przypadkowe – jest jednak w jakimś stopniu prawdopodobne. Od momentu swojego powstania staje się ona jednostką doboru naturalnego i jej dalsze losy są podporządkowane regułom tego procesu. Koncepcja Smitha jest z gruntu neodarwinowska, a jej newralgicznym punktem jest szeroka Df(S). To właśnie jej przyjęcie umożliwia Smithowi przezwycięzenie paradoksu Empedoklesa, ponieważ tak zdefiniowany organizm żywy jest na tyle prosty, że nie posiada żadnych organelli, których celowe przystosowanie do pełnionych funkcji należałoby rozważać. Przy pewnym ujęciu zabieg ten może zostać uznany nie za

---

<sup>55</sup> Por. Ługowski [1995] s.13.

<sup>56</sup> Por. Smith [1993] s. 113.

przewyciężenie problemu, lecz za jego ominięcie, jako że istota paradoksu jest sformułowana w nieco innych kategoriach.

Teoria Oparina jest w swych filozoficznych fundamentach zdecydowanie procesualna, cały wysiłek jej autora skierowany został na uniknięcie rozważania genezy życia w kategoriach szczęśliwych zdarzeń, czy serii przypadków o kluczowym znaczeniu. Wąska Df(O) wymaga, aby organizm żywy miał już pewien stopień strukturalnej złożoności i funkcjonalnego zróżnicowania części. Według Oparina pojedynczy fragment dwuniciowego polinukleotydu (przeciwnie niż u Smitha) nie zasługuje na to, by nazwać go organizmem żywym. Protoorganizmy miały charakter najprymitywniejszych komórek prokariotycznych, posiadających w swej strukturze organelle i podstawowe białka funkcjonalnie przystosowane do pełnienia zadań katalitycznych. Aby wyjaśnić powstanie tego protoorganizmu, unikając odwołania do korzystnego przypadku i jednocześnie unikając paradoksu Empedoklesa, Oparin przyjmuje bardzo istotne założenie o charakterze czysto filozoficznym. Jego podstawową regułą metodologiczną staje się ekstrapolacja mechanizmu doboru naturalnego na materię nieożywioną. Wyjaśnienie Oparina przyjmuje następującą postać: owszem, protoorganizmy miały już złożoną postać, lecz stało się to możliwe dzięki temu, że dobór naturalny działał najpierw na prostych systemach wyodrębnionych ze środowiska, a składających się z niespecyficznych polinukleotydów i polipeptydów (białek). Na początku była zatem całość, która w toku ewolucji (ekstrapolowanej poza sferę organizmów żywych) doprowadziła do celowego przystosowania swoich części. Pierwotna całość (wyodrębniony ze środowiska system białek i kwasów nukleinowych) miała niespecyficzne, funkcjonalnie nieokreślone części (te właśnie polinukleotydy i polipeptydy), które wskutek doboru naturalnego ewoluowały w kierunku systemów o coraz lepiej przystosowanych elementach, zapewniających większą stabilność w tym samym środowisku. Dopiero takie byty możemy nazywać pierwszymi organizmami i od momentu ich pojawienia się przyjmujemy, że nadal podlegają one doborowi naturalnemu w jego podstawowym, Darwinowskim ujęciu (jako mechanizmu kształtującego rozwój *organizmów żywych*).

Teoria Foxa, którą tu zreferuję, jest pochodzącym z 1980 roku rozwinięciem jego wcześniejszych badań szeroko akceptowanych w kręgach akademickich. Pewne jej osobliwości czysto filozoficznej natury przyczyniły się do tego, że w kręgach naukowych spotkała się z jawną krytyką i została uznana za przejaw regresu w porównaniu z wersjami poprzednimi. Kandydatami na pierwsze żywe organizmy są w koncepcji Foxa protokomórki, rozumiane jako otoczone membraną mikrosfery, zdolne do reprodukcji i prowadzenia procesów przemiany materii. To właśnie wyjaśnienie faktu nabycia przez protokomórki

zdolności katalitycznych jest dla Foxa głównym problemem biogenezy. Protokomórki mają znaczne kompetencje biochemiczne, będące wynikiem funkcjonalnego zgrania ich części. Pytając o źródła tego funkcjonalnego uporządkowania protokomórek, Fox staje się kolejnym badaczem stającym w obliczu paradoksu Empedoklesa. Jednak sposób przewycięzenia paradoksu okazuje się dość zaskakujący. Kluczem do swojej teorii czyni Fox pojęcie porządku (*order*), który okazuje się siłą organizującą kolejne poziomy biologicznych kompetencji różnych bytów. Protokomórki przejęły swoje zdolności do metabolizmu bezpośrednio od prostszych makrocząsteczek, wykazujących już pewne zdolności katalityczne (tzw. proteinoidów). Te natomiast przejęły wspomniane zdolności katalityczne od jeszcze prostszych aminokwasów, które organizowały się w proteinoidy zgodnie z konstytutywną dla nich własnością samoinstrukcji (*self-instructing power*), przez którą rozumie Fox zakodowaną zdolność do rozwoju zgodnego z planem. I w tym właśnie miejscu koncepcji ujawnia się jej szczególny charakter:

„... cała argumentacja autora sprowadza się do wykazania, że aktywność biochemiczna protokomórek (zwłaszcza – zdolność do metabolizmu) istniała już w gotowej postaci na poziomie prostych cząsteczek aminokwasów. One właśnie – posiadając zdolność do „samoinstrukcji” – mogły nawzajem dla siebie pełnić funkcje kodu genetycznego; im przysługiwało uporządkowanie (*order*), które przekazały następnie bardziej złożonym układom; one także zapewniły istnienie metabolizmu *ab initio*” (Ługowski [1995] s. 45)

Widać już jasno, że próbując przewyciężyć paradoks Empedoklesa, powołuje się Fox na kategorie czysto metafizyczne: metafizyczny porządek realizuje się w kontinuum bytów poprzez metafizyczną zdolność do samoinstrukcji. Dobór naturalny jest w tej koncepcji niemal nieobecny i schodzi do rangi nieistotnego czynnika pobocznego. Powstanie i ewolucja żywych organizmów zostają tutaj potraktowane jako konieczny etap realizowania się metafizycznego planu, a cała teoria niebezpiecznie przypomina Leibnizjańską koncepcję harmonii przedustawnej. Widać teraz jasno, co miałem na myśli, pisząc w paragrafie 2., że sposobem Foxa na przewycięzenie paradoksu Empedoklesa jest taka modyfikacja standardów wyjaśnienia naukowego, aby dało się w nich umieścić założenie o istnieniu metafizycznego planu. Problem tylko w tym, że tak rozumiana nauka sadowi się niebezpiecznie blisko religii.

## *5. Rola przypadku w protobiologii a wartość eksplanacyjna teorii.*

W poprzednim paragrafie zrealizowałem właściwie podstawowy cel tego artykułu, którym było wykazanie, na czym polega i jak próbuje się przewycięzać pewien teoretyczny problem, który nazwaliśmy paradoksem Empedoklesa, wskazując jednocześnie na filozoficzne uwarunkowania teorii protobiologicznych, w których się tego dokonuje. Przy okazji tych rozważań zasygnalizowany został inny problem, do którego wrócę w tym miejscu, co pozwoli mi sformułować pewne uwagi końcowe. W poprzednim paragrafie wskazywałem na to, że w teorii Smitha zakłada się przypadkowe powstanie protoorganizmu, rozumianego jako dwuniciowy polinukleotyd. W przeciwieństwie do takiego postawienia sprawy, Oparin starał się za wszelką cenę nie powoływać w wyjaśnieniu genezy życia na przypadek, postulując, że protoorganizm – rozumiany jako funkcjonalnie zróżnicowana protokomórka – powstał wskutek działania mechanizmu doboru naturalnego. Od razu pojawiają się pytania o to, czy Oparinowi istotnie udało się uniknąć „znamiona przypadkowości” w jego teorii oraz o to, czy w ogóle powoływanie się na kategorię przypadku jest jakimś zagrożeniem dla naukowego statusu teorii protobiologicznych? Co do pierwszego z tych pytań, widać od razu, że Oparin tylko przesunął problem. Również na gruncie jego teorii zasadne jest rozważenie genezy pierwszych dwuniciowych polinukleotydów (choć bytów tych Oparin nie uznaje za żywe). I w tej kwestii Oparin, podobnie jak Smith, przyznałby zapewne, że ich powstanie było dziełem przypadku. Zanim postaram się odpowiedzieć na drugie z pytań, przyjrzę się najpierw pewnym intuicjom związanym z pojęciem przypadkowości. Kiedy mówimy o czymś, że zaszło przez przypadek, myślimy najczęściej o tym, że nie sposób wskazać przyczyn, których obecność koniecznie pociąga zajście tego zdarzenia. Jeśli tylko wprowadzimy parametr prawdopodobieństwa, zyskamy sposób porównywania skądinąd przypadkowych zdarzeń. Wylosowanie w rzucie sześcienną kostką ścianki oznaczonej numerem jeden jest oczywiście bardziej prawdopodobne, niż wylosowanie ścianki z numerem jeden na kostce dwunastościennej, a oba zdarzenia są przypadkowe. Jeśli zatem Smith przyznaje, że powstanie dwuniciowego polinukleotydu miało charakter przypadkowy, a ktoś inny postulowałby, że przypadkowo powstała cała protokomórka, widać od razu, która z tych koncepcji jest bardziej prawdopodobna.

Wrócę teraz do kwestii tego, czy tak rozumiana przypadkowość stanowi jakiś metodologiczny problem dla eksplanacyjnej wartości teorii protobiologicznych. Najbardziej

rozpowszechnionym modelem wyjaśnienia naukowego jest model wyjaśnienia przez prawa, którego opracowanie zawdzięczamy C.G.Hemplowi, P.Oppenheimowi i K.R.Popperowi. W ramach tego modelu wyróżnia się, obok wyjaśnień nomologiczno-dedukcyjnych i statystyczno-dedukcyjnych, także statystyczno-indukcyjne, stosowane do wyjaśniania indywidualnych zdarzeń. Stefan Amsterdamski tak streszcza ich charakter:

„...eksplanans wyjaśnienia indukcyjno-statystycznego nie implikuje logicznie eksplanandum, lecz – jak powiada się – uzasadnia je indukcyjnie. Zawiera on mianowicie prawa statystyczne, stwierdzające wysokie prawdopodobieństwo zajścia określonego rodzaju zdarzenia w jakichś warunkach, oraz odpowiednie zdania, stwierdzające, że warunki te miały właśnie miejsce. Tego typu wyjaśnienia wskazują w gruncie rzeczy na to, że w danych warunkach wyjaśniane zdarzenie powinno było czy też mogło zajść, a nie, że zajść musiało” (Amsterdamski [1983] s.100)

Taki właśnie model wyjaśniania towarzyszy naukowcom, którzy, wykorzystując statystyczne prawa chemii i termodynamiki oraz empiryczne dane o warunkach początkowych (tzn. geologicznym i atmosferycznym stanie Ziemi sprzed kilku miliardów lat), starają się wykazać, że pewne zdarzenia w świecie abiotycznych, a później już - biotycznych, częściak zachodziły z niemałym prawdopodobieństwem. Teorie protobiologiczne zawierają prawa podobne formalnie, o ogólnej postaci typu „w takich to a takich okolicznościach zachodzą zazwyczaj takie to a takie zdarzenia, a te okoliczności miały w tym przypadku miejsce”, a różnią się materialną treścią tych praw oraz danymi empirycznymi, które przywołuje się, aby potwierdzić historyczną obecność zawartych w prawie warunków początkowych. Ambicją naukowców jest to, aby ich model charakteryzował się jak najwyższym prawdopodobieństwem zajścia kluczowego zjawiska, a jednocześnie uwzględniał dostępny materiał empiryczny. Jednocześnie wszyscy oni (bądź prawie wszyscy) zdają sobie sprawę, że wyjaśniane zdarzenia mają znamię przypadkowości, pojawiającej się wszędzie tam, gdzie rozważamy zdarzenia słabo zdeterminowane. Widać zatem, że obecność kategorii przypadku nie jest zagrożeniem dla naukowego statusu teorii protobiologicznych, jeśli tylko ktoś uzna wyjaśnienie statystyczno-indukcyjne za satysfakcjonujące z naukowego punktu widzenia.

## SŁOWNICZEK PODSTAWOWYCH TERMINÓW:

**aminokwasy:** cząsteczki wchodzące w skład polipeptydu (białka); rozmaite układy 20 różnych aminokwasów ułożonych w różnej kolejności; tworzą białka wszystkich żywych organizmów

**DNA (kwas deoksyrybonukleinowy):** spolimeryzowana cząsteczka złożona z czterech różnych jednostek zwanych nukleotydami (skrót: A, T, C, G) i niosąca informację genetyczną

**makrocząsteczki:** bardzo duże spolimeryzowane cząsteczki, zbudowane z wielu podobnych elementów budulcowych, są nimi np. DNA (złożony z nukleotydów) oraz białka (złożone z aminokwasów)

**mutacja:** zmiana w sekwencji DNA, może spowodować zmianę produktu genu lub inną regulację ekspresji genu, co w efekcie może prowadzić do zmiany cechy

**nukleotydy:** pojedyncze składniki budulcowe DNA i RNA (występują w czterech różnych typach, zakończone różnymi zasadami: adeniną (A), cytozyną (C), guaniną (G), tyminą (T))

**peptyd:** cząsteczka składająca się z dwóch lub więcej aminokwasów powiązanych silnym wiązaniem chemicznym

**polipeptyd:** długi peptyd, zwykle ponad 20 aminokwasów

**prokarionty:** organizmy jednokomórkowe nie zawierające jądra

*BIBLIOGRAFIA:*

AMSTERDAMSKI S.

[1983] *Nauka a porządek świata*, Warszawa.

BERG P., SINGER S.

[1997] *Język genów. Poznawanie zasad dziedziczenia*, Warszawa.

ŁUGOWSKI W.

[1987] *Paradoks powstawania życia*, Warszawa.

[1995] *Filozoficzne podstawy protobiologii*, Warszawa.

REALE G.

[1994] *Historia filozofii greckiej*, tom I, Lublin.

SMITH M.J.

[1993] *The Theory of Evolution*, Cambridge.

ZIEMBIŃSKI Z.

[2001] *Logika praktyczna*, Warszawa.



## **FILOZOFIA A TEORIA UMYŚLU**

## **Kilka uwag o poznawaniu cudzych stanów psychicznych**

Celem niniejszego artykułu jest próba krytycznego spojrzenia na klasyczne już teorie, starające się uchwycić i wyjaśnić zjawisko poznawania cudzych stanów psychicznych. Wskażemy tu na pewne problemy, w które uwikłane są wszystkie przedstawiane stanowiska. Następnie zaś zaprezentujemy alternatywną możliwość rozwiązania tytułowego zagadnienia.

Nasze codzienne, przednaukowe nastawienie, zdaje się przemawiać za tym, że potrafimy jakoś poznawać, przynajmniej niektóre, stany psychiczne innych ludzi. Ktoś mógłby na wstępie zapytać, po co w ogóle podejmować całą tę kwestię. Jeżeli nie wiadomo co przeżywa inna osoba, to zawsze przecież można ją o to zapytać i w większości przypadków udzieli nam ona prawdziwej odpowiedzi. Jednak nie taki sposób poznawania cudzych stanów psychicznych będziemy tutaj rozważać. Aby pokazać przypadki, o które nam chodzi, rozważmy przykład. Rozmawiam z kimś. W pewnej chwili mówię mu coś przyjemnego i nagle widzę, że jest bardzo uradowany. Nie informuje mnie o tym werbalnie, ale gdy patrzę na niego, widzę po prostu, że się cieszy.

Zaskakujące jest to, że w czasie rozmowy z drugim człowiekiem nie musimy co chwilę jej przerywać, aby dowiedzieć się, co właśnie przeżywa. Po prostu wiemy, że teraz się smuci, wstydzi czy też raduje. Naszym celem jest próba zbadania tych specyficznych sytuacji, w których bez cudzego poinformowania i bez specjalnego wysiłku intelektualnego jakoś wiemy co przeżywają inni. Nie chcemy na razie zaklasyfikować omawianego tu rodzaju poznania jako pośredniego lub bezpośredniego. Kwestię tę poruszymy nieco później. Na razie niechaj nam wystarczy wymieniony przykład do zarysowania przedmiotu naszych badań.

Istnieje wiele teorii, które próbują wyjaśnić opisane tu zjawisko. Teorie te (nazwijmy je teoriami wyjaśniającymi - TW) starają się pokazać jakieś ukryte mechanizmy, czy też władze - pozwalające nam uzyskiwać wiedzę na temat psychiki innych ludzi. Poniżej postaramy się krótko przedstawić kilka z owych teorii. Należy tu jednak zaznaczyć, iż zestaw ten nie obejmuje wszystkich możliwych rozwiązań, które pojawiły się w filozofii. Ograniczymy się tu do przytoczenia tylko tych najbardziej rozpowszechnionych.

- Teoria wnioskowania przez analogię (m.in. J. S. Mill)<sup>57</sup>.

Głosi ona, że postrzegamy cudze zachowanie lub wyraz twarzy. Z własnego doświadczenia zaś wiemy, że, gdy nasze ciało znajduje się w podobnym do spostrzeganego stanie, to mamy wtedy określone przeżycia psychiczne. Z tego wnioskujemy przez analogię, że i inni przeżywają to samo co my, gdy zachowują się w dany sposób lub mają określony wyraz twarzy.

Zwolennicy tej teorii twierdzą, że jest to jedyny sposób poznawania cudzych stanów psychicznych, a w przypadkach takich jak ten przytoczony przez nas na początku, wnioskowanie odbywa się tak szybko, że jest wręcz automatyczne, odruchowe.

- Teoria naśladowania (C. Groos)<sup>58</sup>.

Wedle tej teorii, gdy postrzegamy cudze zachowania, gesty wyrazy twarzy, to mimowolnie je naśladujemy. Często zdarza się to “naprawdę” tzn. wykonujemy takie gesty jak obserwowany przez nas człowiek, a czasami (w większości przypadków) naśladujemy go jedynie “w myśli”, bez fizycznych ruchów. Podczas owego naśladowania doświadczamy pewnego przeżycia psychicznego, które przypisujemy drugiemu człowiekowi. Złudnie jawi nam się ono jako cudze.

- Teoria kojarzenia<sup>59</sup>.

Gdy spostrzegamy cudzy gest, zachowanie czy wyraz twarzy, to kojarzymy to z wyobrażeniem własnego, przeszłego gestu, zachowania i wyrazu twarzy. To zaś powoduje, że przeżywamy jakiś stan np. uczucie, związane ze skojarzonymi własnościami ciała. Jednakże przeżywamy je jako cudze, ponieważ nie jest ono umotywowane jakimś naszym wcześniejszym doświadczeniem, ani też nie przejawia się w żadnym naszym ruchu cielesnym. Dlatego też to powstałe uczucie wydaje nam się być cudzym.

- Teoria wczucia wg Teodora Lippsa<sup>60</sup>.

W tej teorii również wychodzi się od faktu spostrzeżenia cudzego ciała. Uczucia, które budzą się w nas na ten widok “wczuwamy” - czyli jak gdyby projektujemy, “wkładamy” w cudze ciało. Mamy tu do czynienia z własnymi odczuciami, które są reakcjami na ruchy drugiego człowieka. Poprzez “wczucie” czyli owo “włożenie” tychże odczuć w inną osobę, wydaje

---

<sup>57</sup> Por. też: Ingarden [1971].

<sup>58</sup> Por. Ingarden, tamże.

<sup>59</sup> Por. Ingarden, tamże.

<sup>60</sup> Por. Ingarden, tamże.

nam się, iż mamy do czynienia właśnie z jej, a nie naszymi, przeżyciami.

- Teoria wczucia wg Edmunda Husserla<sup>61</sup>.

Teoria wczucia Husserla różni się znacznie od wyżej przytoczonego stanowiska Lippsa. U twórcy fenomenologii wczucie jest właściwym doświadczeniem drugiego człowieka (*Fremderfahrung*). Nie polega ono na projektowaniu własnych przeżyć w cudze ciało, lecz - na uchwytowaniu jakoś psychiki innego człowieka. Wczucie jednakże jest niesamodzielnym doświadczeniem. Aby mogło się uobecnąć, potrzebne jest jeszcze spostrzeżenie, z którym się splata.

Husserl nie wyjaśnia do końca, na czym ma polegać wczucie. W *Ideach II* opisuje ten proces tak:

*“Przenoszę się w inny podmiot: przez wczucie uchwytuję, co i jak mocno, z jaką siłą go motywuje [...] Uzyskuję te motywacje, gdy się przenoszę w jego sytuację, na jego poziom wykształcenia, w jego rozwój w młodości etc. i w tym przeniesieniu w głąb muszę się z tymi motywacjami solidaryzować; czuję się nie tylko, jak bym wchodził w jego myślenie, czucie, działanie, lecz muszę w tym iść za nim, jego motywy stają się quasi-motywami [...] Jestem razem z nim w jego pokusach, w jego mylnych wnioskach, a w tym <<razem>> tkwi wewnętrzne współprzeżywanie czynników motywujących [...]”*  
(Husserl [1974])

Można by chyba zinterpretować Husserlowską teorię wczucia jako pewną czynność “przenoszenia się” (Husserl nie wyjaśnia, na czym ma to polegać) w sytuację drugiego człowieka. Dzięki temu mielibyśmy dotrzeć do jego czynników motywujących i solidaryzować się z nimi. Ta solidaryzacja, jak zauważa Adam Węgrzecki<sup>62</sup>, nie ma polegać na uznaniu cudzych motywów za własne, lecz mają się one stać quasi-motywami dla poznającego. Może on podążać za nimi, aż dojdzie do “współprzeżywania” czegoś, co należy do cudzego życia osobowego.

- Teoria wczucia wg Edyty Stein.

Jak Stein sama zaznacza, w pierwszym rozdziale swej rozprawy doktorskiej *Zum Problem der Einfühlung*, chce używać terminu wczucie “abstrahując od wszelkich historycznych

---

<sup>61</sup> Zob. Husserl [1974].

<sup>62</sup> Węgrzecki [1992].

tradycji, które obciążają to słowo”<sup>63</sup>. Jej koncepcja jest w gruncie rzeczy podobna do poglądu Husserla. Mimo to, należy jej stanowisko omówić jako odrębne.

Stein przybliży wczucie, czyli doświadczenie, w którym dane są cudze stany psychiczne, poprzez pewną analogię. Otóż porównuje ona wczucie do procesu przypominania sobie czegoś. W obu przypadkach przedmiot dany jest nieźródłowo, w przeciwieństwie do własnych przeżyć terażniejszych podmiotu, czyli np. uczuć przeżywanych aktualnie albo spostrzegania.

Cudze stany psychiczne, w szczególności np. uczucia, przejawiają się, zdaniem Stein, w tzw. fenomenach wyrazowych. Można w nich wyróżnić wyraz (np. zaczerwienienie na twarzy) i, to co wyrażane (wstyd). Obie strony tego fenomenu stanowią naturalną całość. Zaczerwienienie jest zewnętrzną stroną wstydu. Mimo że istnieje tak silne powiązanie między wyrazem, a tym co wyrażane, to samo spostrzeżenie cudzego ciała nie wystarcza do uchwycenia jego stanu psychicznego. Otóż po spostrzeżeniu cudzej miny czy gestu, następuje wczucie - czyli przeniesienie się w cudze ciało i spełnienie przeżycia, które było już pusto współdane z ową miną tudzież gestem.

Można podać szereg trudności, jakie powodują wyżej przytoczone teorie. Gruntownej krytyki pierwszych czterech z nich dokonuje Roman Ingarden<sup>64</sup>. Nie chcemy jednakże polemizować tutaj z każdą z nich i wykazywać fałszywości jakiś określonych założeń czy też przytaczanych faktów. Postaramy się raczej wskazać na inną, naszym zdaniem zasadniczą ich słabość.

Nie bez powodu nazwaliśmy powyższe teorie wyjaśniającymi. Mają one, bowiem na celu wyjaśnienie właśnie, opisanego zjawiska. W tym celu, opierając się na obserwacjach, częstokroć bardzo trafnych, postulują istnienie pewnych specyficznych procesów czy też władz poznawczych, które to mają być odpowiedzialne za postrzeganie cudzych stanów psychicznych (albo mają nas łudzić, że takie stany poznajemy). Trudność polega na tym, iż nie jesteśmy w stanie w żaden sposób określić prawdziwości zdań opisujących działanie postulowanych właściwości naszej psychiki. Już sama próba wyobrażenia sobie takich procesów, które rzekomo mają zachodzić w codziennym życiu, wydaje się dość osobliwa. Najlepiej chyba zobrazują to uwagi Ludwiga Wittgensteina zawarte w *Kartkach* i *Last Writings on the Philosophy of Psychology*:

---

<sup>63</sup> Stein [1988].

<sup>64</sup> Ingarden [1971].

*“Interpretuję słowa; dobrze - ale czy interpretuję również miny? Czy interpretuję pewien wyraz twarzy jako groźny bądź jako przyjemny? [...]”*

Gdybym powiedział: <<Nie wystarczy, że postrzegam groźny wyraz twarzy, muszę go najpierw zinterpretować>>. - Ktoś dobywa na mnie noża, a ja mówię: <<Pojmuję to jako groźbę>>.” (Wittgenstein [1999] nr 218)

“Czy patrzysz w siebie, aby rozpoznać złość w jego twarzy? Jest ona tam równie wyraźna jak w twojej własnej piersi.” (Wittgenstein [1999] nr 220)

“Czy oprócz tak zwanego smutku, ujawnianego w rysach twarzy, spostrzegam również smutek jego umysłu? Czy dedukuję go z twarzy? Czy mówię: <<Wyglądał na zasmuconego, zachowywał się jak zasmucony, więc prawdopodobnie sam też był smutny?>>” (Hacker [1997])

Skoro sama próba przełożenia takiego np. procesu wnioskowania, naśladowania czy wczucia, na konkretne przykłady z życia codziennego, wydaje się absurdalna, to znaczy, że coś jest chyba nie w porządku z TW, które postulują zachodzenie tychże zjawisk. Jeżeli zachodziłyby one rzeczywiście, to byłyby oczywiste i wystarczyłoby zwyczajne odwołanie się do codziennego doświadczenia, a tam bez problemu moglibyśmy się przekonać o ich istnieniu. Tymczasem nic takiego nie ma miejsca. I gdy próbujemy się przyjrzeć sytuacjom, w których po prostu postrzegam czyjś smutek, gniew, wstyd, to trudno jest się tam dopatrzeć jakiegoś wnioskowania, kojarzenia, czy wczuwania. Nie doświadczamy też żadnego “przeniesienia się” w cudze ciało, ani nie naśladujemy cudzego gestu (cóż to by miało znaczyć, że naśladujemy go “w myśli”?).

Skąd się biorą takie nieadekwatne TW? Przyczyny można by chyba dopatrywać się w tym, iż wszystkie one bazują niewątpliwie na pewnych rzeczywistych zjawiskach. Często zdarza się przecież tak, iż np. obserwując jakiegoś człowieka wnioskujemy coś na temat jego właściwości psychicznych. Ingarden przytacza następujący<sup>65</sup> przykład: z obserwacji kogoś, kto czyta gazetę, idąc przez ruchliwy plac, wnioskujemy, iż jest on lekkomyślny, albo że zajęło go coś interesującego w lekturze. Tak samo zdarza nam się czasami przypisywać nasze uczucia innym, albo naśladować mimowolnie (lecz realnie) cudzą minę (gdy np. ktoś wykrzywia twarz z bólu, patrząc na niego, możemy robić to samo). Owszem, takie przypadki mają miejsce. Jednakże błąd TW polega na tym, iż te sytuacje generalizują całość przejawów

---

<sup>65</sup> Zob. Ingarden [1971].

tę, co nazywamy poznawaniem cudzych stanów psychicznych.

TW popełniają jeszcze jeden błąd. Do danych faktycznych, dorabiają pewne hipotezy czysto teoretyczne - w różny sposób przystające do tych danych. Henry Le Roy Finch w ślad za Wittgensteinem nazywa to chorobą, polegającą na wyjaśnianiu, która nie pozwala "po prostu patrzeć"<sup>66</sup>. Sam autor *Dociekań filozoficznych* przedstawia to następująco:

"I co teraz powiem? Że twarz innego człowieka pobudza mnie do naśladowania, że zatem odczuwam na własnej twarzy drobne ruchy i napięcia mięśni i mam na myśli ich sumę? Bzdura. Bzdura - czynisz bowiem założenia, zamiast po prostu opisywać. Komu po głowie grasują wyjaśnienia, ten zaniedbuje myślenie o najważniejszych faktach." (Wittgenstein [1999] nr 220)

Nieweryfikowalność przedstawionych tutaj TW sprawia, że stają się one praktycznie bezużyteczne. Możemy skonstruować sobie szereg innych rozwiązań równie ciekawych jak wyżej wymienione. Jesteśmy na gruncie filozofii w o tyle gorszej sytuacji, niż na przykład psychologia, w której również postuluje się często pewne byty czysto hipotetyczne, że nie możemy wysunąć argumentu z użycia. Postulowane bowiem przez psychologię takie konstrukty teoretyczne jak inteligencja czy osobowość, pozwalają na rozwiązanie szeregu praktycznych spraw (a więc są użyteczne). Poza tym, posiadamy określone kryteria ich wyznaczania i specjalne techniki ich badania (jak np. kwestionariusze osobowości czy testy inteligencji). W naszym przypadku niestety TW pozostają jedynie filozoficznymi domysłami, możliwymi tylko wyjaśnieniami omawianego tu zagadnienia.

Cóż więc ma począć filozof w takiej sytuacji? Naszym zdaniem należy pójść w badaniach nad poznawaniem cudzych stanów psychicznych, takim tropem, jaki podjęli m.in. Wittgenstein i Ingarden. Ich dociekania skupiają się przede wszystkim na faktach i opisach konkretnych przypadków. Autor *Traktatu* właściwie na tym poprzestaje i w jego pismach z późnego okresu znajdujemy szereg obserwacji różnych konkretnych przypadków z życia codziennego. Ingarden, z kolei, idzie w swoich rozważaniach dalej. Dla niego sam opis nie wystarcza i oprócz tego stara się także wyodrębnić i nazwać pewne aspekty procesu poznawania cudzych stanów psychicznych. Spróbujemy pójść w ślad za Wittgensteinem, skupiając się głównie na analizie i opisie (przy czym będziemy też czerpali w tej kwestii z dorobku Ingardena, mimo że polski fenomenolog nie zgodziłby się na poprzestanie na analizie i opisie). Nie może jednak dziwić fakt, że poniżej odwołujemy się do filozofów "wyjaśniających", ponieważ - jak już zaznaczyliśmy - w wielu przypadkach warstwa empiryczna ich teorii jest trafna.

---

<sup>66</sup> Finch [1996].

Rozważmy taki przykład. Idę ulicą. Nagle spotykam starego znajomego. Spotkanie jest niespodziewane. W tym momencie narzuca mi się ogromna radość, którą dostrzegam w spotkanym przeze mnie człowieku. Nie wywnioskowałem tego z jego miny czy też z jego zachowania. Po prostu widzę, że się cieszy i jestem tego pewien. Nie muszę sobie mówić: “Przypuszczam, że jest wesoły”, ja to po prostu wiem. Wiem też bez wątplenia, że jest to jego uczucie, a nie moja projekcja. Łatwo to stwierdzić, gdy dookreślimy nasz rozważany przykład. Otóż może zdarzyć się tak, iż ja jestem pogrążony w smutku, z powodu, powiedzmy, czyjejś śmierci. Smutek ów jest tak wielki, że spotkanie mojego znajomego nie jest w stanie wywołać we mnie żadnych pozytywnych uczuć. Jednocześnie postrzegam jego radość. Czy w tej sytuacji dokonałem jakiejś projekcji lub skojarzyłem jego twarz z moim uczuciem? Wydaje się to niedorzeczne. Jego radosny nastrój należy do niego. Ja nie mam z nim w tym momencie nic wspólnego. Jego radość może mnie nawet irytować. To zaś znaczy, iż narzuca ona mi się jakoś z zewnątrz. Możemy więc sformułować tu tezę następującą: są pewne przypadki w których spostrzegamy jakoś, nie wymyślony przez nas, cudzy stan psychiczny.

Za tą tezą przemawiają jeszcze inne fakty. Czasami zdarza się tak, iż zdajemy sobie sprawę, że pewien wyraz twarzy drugiej osoby nie odpowiada jej stanowi psychicznemu. Oczywiście często stwierdzam taki stan rzeczy na podstawie pewnego wniosku. Zastanówmy się jednak nad takim przykładem. Rozmawiam z kimś. Ten ktoś stara się być dla mnie miły tzn. spełnia wszystkie warunki, że tak powiemy “cielesne”, abym uznał, że żywi wobec mnie pozytywne uczucia. Ale widzę jasno sztuczność jego zachowania, a przez to narzuca mi się jakoś raczej jego niechęć w stosunku do mnie.

Max Scheler przytacza jeszcze jedną dość ciekawą obserwację<sup>67</sup>, naszym zdaniem sporną. Otóż twierdzi on, iż możemy rozumieć i postrzegać stany psychiczne innych ludzi, stany - których my nigdy nie przeżywaliśmy. Scheler podaje tu następujący przykład: jesteśmy w stanie rozumieć czyjś lęk przed śmiercią, mimo iż go nigdy nie przeżywaliśmy. Jest to pogląd nieco dyskusyjny. Wydaje nam się, że w takich przypadkach jak opisany przez Schelera rzeczywiście narzuca się nam jakiś cudzy stan psychiczny, ale wtedy raczej stajemy wobec czegoś niezrozumiałego. Zachowanie danej osoby przejawia jakiś fenomen psychiczny, ale nie wiemy, co to jest. Możemy jedynie domyślać się tego z sytuacji w jakiej się ona znajduje.

---

<sup>67</sup> Zob. Scheler [1986].



Jest jeszcze coś osobliwego w tym jak ujawnia się psychika drugiego człowieka. Zwraca na to uwagę Ingarden<sup>68</sup>. Otóż wbrew powszechnym poglądom wcale nie jest tak, iż cudze ciało mamy dane wcześniej od jego psychiki (chodzi tu o pewne tylko specyficzne przypadki). Zastanówmy się nad czymś takim. Spoglądam na kogoś bardzo rozbawionego. Dostrzegam u niego radość. Patrzyłem na niego tylko przez bardzo krótką chwilę, a mimo to trafnie osądziłem jego stan psychiczny. Gdy mam teraz, zrelacjonować co widziałem na jego twarzy, opisać ją dokładnie, to nie jestem w stanie tego zrobić. Widziałem tylko ogólnie jej zarys i np. uśmiech na ustach, ale gdybym miał opisać szczegóły, to nie potrafię. Ingarden zauważa, że najpierw uderzył mnie jakiś fenomen psychiczny drugiego człowieka, oczywiście na podłożu jego miny, gestów i zachowania, ale na szczegóły tego cielesnego podłoża mogę zwrócić uwagę dopiero później, kiedy mam na to czas i ochotę.

Ingarden zdaje się mieć tu rację. Niech przemówi za tym jeszcze jeden przykład. Przebywam przez pewien czas w towarzystwie jakiejś osoby. Dostrzegam dużo radości w jej zachowaniu. Jednakże w czasie obcowania nie zastanawiam się nad jej stanem psychicznym. Po prostu przeżywam dany okres, oddając się chwili. Kiedy ktoś później zapyta jak się zachowywała tamta osoba przeżywając radość tzn. jakie wykonywała gesty, ruchy, jakie robiła miny, to niewiele mogę na to pytanie odpowiedzieć.

To co się narzuca jako stan psychiczny drugiego człowieka (nazwijmy to fenomenem psychicznym) Ingarden chce nazywać wyrazem<sup>69</sup>, natomiast podłoże cielesne, w którym on się przejawia - czynnikami wyrażającymi. Pomiędzy tymi dwoma aspektami zachodzi ścisły związek. Jeden nie może obyć się bez drugiego. Czynniki wyrażające są jednakże tylko podłożem do przejawienia się wyrazu. Świadczy o tym fakt, że spostrzegamy je tylko mimochodem, bez koncentrowania uwagi na nich. Fenomen psychiczny drugiego człowieka, który spostrzegamy w jego twarzy i zachowaniu nie da się sprowadzić do żadnego szczegółu jego cielesnego wyglądu. Nie mogę wziąć czyjejś twarzy i powiedzieć: "To jest smutek". Bo czy wtedy np. odpowiednie napięcie mięśni na jego twarzy byłoby smutkiem?

Konkluzja naszych analiz jest więc taka: istnieją niewątpliwie takie przypadki, w których jakoś są nam dane cudze stany psychiczne, nie będące naszą projekcją. Te stany, które nam się narzucają w czasie przebywania w towarzystwie drugiego człowieka, są nieredukowalne do sfery jego cielesności.

---

<sup>68</sup> Ingarden [1971].

<sup>69</sup> Ingarden [1971].

Powstaje teraz problem, który sygnalizowaliśmy już wcześniej. Otóż, czy te cudze fenomeny psychiczne postrzegamy pośrednio czy też bezpośrednio? Faktem jest, na co się wszyscy zgadzają, że nie przeżywamy ich tak samo jak własnych stanów. Czy więc są nam one dane pośrednio? Różne były odpowiedzi na to pytanie. Wśród filozofów nie ma zgody na ten temat. Naszym zdaniem jednak problem tkwi gdzie indziej. Chodzi tu o kwestię zdefiniowania poznawania bezpośredniego. Możemy starać się odrzucić klasyczną definicję Locke'a, iż jest to uchwycenie przedmiotu przez któryś z pięciu zmysłów lub "refleksję", czyli introspekcję. Tak czyni np. Ingarden<sup>70</sup>, który próbuje zdefiniować poznawanie bezpośrednie ogólniej, jako pewien rodzaj spostrzeżenia naocznego, źródłowego i dokonującego się z pełnym przeświadczeniem o istnieniu tego, co spostrzegamy. Oczywiście, tak ogólne ujęcie bezpośredniości pozwala tu zakwalifikować także i ten rodzaj doświadczenia, który nazywamy poznawaniem cudzych stanów psychicznych.

Naszym zdaniem próba zakwalifikowania tegoż poznania jako pośredniego, bądź też bezpośredniego wymaga dokładniejszych badań z tego zakresu<sup>71</sup>. Jednakże nasze intuicje, uwzględniające powyższe rozważania, skłaniają się ku uznaniu tegoż poznania za bezpośrednie.

Oprócz tych kwestii, które poruszyliśmy powyżej, istnieje jeszcze bardzo dużo zagadnień dotyczących poznawania cudzych stanów psychicznych. Nie będziemy ich tu jednak szczegółowo przedstawiali, lecz tylko zasygnalizujemy możliwe problemy do analizy.

W ostatnich latach rozwija się dość mocno tzw. filozofia spotkania ( m.in.: Józef Tischner, Adam Węgrzecki). Jak zauważył Roman Ingarden w *Wykładach i dyskusjach z estetyki*, ludzie wzajemnie na siebie oddziałują, w obcowaniu ze sobą nawzajem siebie jakoś "podniecają". Filozofowie spotkania starają się właśnie uwzględnić sytuację, w jakiej dochodzi do poznawania drugiego człowieka. Zdarzają się przecież przypadki, w których druga osoba może się zamknąć w sobie i starać się nie ujawniać swych stanów psychicznych. Dochodzi tu do tzw. niedopuszczenia, które wyznacza granicę poznania. W trakcie poznawania drugiego człowieka mogą zajść trzy możliwe zachowania ze strony poznawanego: odmowa uczestnictwa w procesie poznawania, akceptacja, pozorna akceptacja. Filozofia spotkania bada wszystkie te okoliczności i analizuje sytuację poznawczą

---

<sup>70</sup> Ingarden, tamże.

<sup>71</sup> Kwestię bezpośredniego i pośredniego poznania omawia obszernie, w swojej monografii, Józef Dębowski: zob. Dębowski [2000].

obu stron tj. zarówno poznającego jak i poznawanego (warto zauważyć, że często w codziennym życiu jesteśmy jednocześnie i jednym i drugim).

Ten, jak i szereg jeszcze innych problemów, wymaga dokładnego zbadania, abyśmy mogli stworzyć dokładny obraz kwestii poznawania cudzych stanów psychicznych. Jak już to było powiedziane, nie może nas jednak w dociekaniach nad postawionym zagadnieniem ponosić wyobraźnia. Często wychodzą z tego różne teorie, które w ogóle nie przystają do rzeczywistości. Tymczasem powinniśmy, jak mówi Ingarden w *Dążeniach fenomenologów*<sup>72</sup>, “słuchać, co mówi nam doświadczenie”. Rola filozofii byłaby tu następująca. Należy przyjrzeć się konkretnym przypadkom zachodzącym w doświadczeniu i starać się je zanalizować. Takie działanie filozofów może się stać o wiele bardziej pożyteczne i inspirujące dla nauk szczegółowych zajmujących się człowiekiem, zwłaszcza dla psychologii, która niejednokrotnie czerpała z obserwacji filozoficznych.

#### *BIBLIOGRAFIA:*

DĘBOWSKI Józef

[2000] *Bezpośredniość poznania*, Wydawnictwo UMCS, Lublin.

FINCH Henry Le Roy

[1996] *Fenomenalizm fizjologiczny*, w: *Metafizyka jako cień gramatyki*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław.

HACKER Peter

[1997] *Wittgenstein*, Wydawnictwo Amber, Warszawa.

HUSSERL Edmund

[1967] *Idee I*, PWN, Warszawa.

[1974] *Idee II*, PWN, Warszawa.

---

<sup>72</sup> Ingarden [1963].

INGARDEN Roman

[1963] *Dążenia fenomenologów*, w: *Z badań nad filozofią współczesną*, PWN, Warszawa.

[1971] *U podstaw teorii poznania*, PWN, Warszawa.

SCHELER Max

[1986] *Istota i formy sympatii*, PWN, Warszawa.

STEIN Edyta

[1988] *O zagadnieniu wczucia*, Znak, Kraków.

TISCHNER Józef

[1978] *Fenomenologia spotkania*, w: *Analecta Cracoviensia*, t. X.

WĘGRZECKI Adam

[1992] *O poznawaniu drugiego człowieka*, Kraków.

WITTGENSTEIN Ludwig

[1999] *Kartki*, Wydawnictwo Spacja, Warszawa.

[2000] *Dociekania filozoficzne*, Wydawnictwo Spacja, Warszawa.

## **Searle i jego gwoździe**

Standardową strategią nowoczesnych podejść do problemu miejsca jaźni w świecie przedmiotów nieożywionych jest proponowanie jakiejś wersji materializmu. John Searle przeciwstawia się temu rozwiązaniu. Jego zdaniem pomija ono fakt, że świadome procesy psychiczne stanowią rzeczywisty i centralny fenomen świata. Szczególnie interesujące i szeroko dyskutowane są jego poglądy na komputerowe modele umysłu. W tej pracy postaram się przeanalizować argumenty, przedstawione w pracy Searle [1999]. W przedmowie książki tej autor stwierdza: “Chciałbym również wbić ostatni gwoździe do trumny teorii głoszącej, że umysł jest programem komputerowym”. Postaram się wykazać, że przy stwierdzaniu zgonu Searle postępował wielce nonszalancko – wobec czego potrzebna jest zdecydowana interwencja. Trumna zawierać może żywe i zdrowe organizmy!

W rzeczonyj publikacji odnajdziemy trzy argumenty przeciw utożsamianiu myślenia z operacjami obliczeniowymi. Na początek dokonam ich rekonstrukcji, jak również postaram się wskazać, w jakie stanowiska są one wymierzone. Następnie - postaram się poddać badaniu prawomocność krytyki, jaką przeprowadza Searle.

### **I. Argument “Chińskiego Pokoju”.**

Argument ten wymierzony jest w silny program AI – twierdzenie, że myślenie jest procesem polegającym na przetwarzaniu danych o charakterze syntaktycznym, do którego zdolne są komputery. Wywód Searle’a można zrekonstruować w postaci następujących kroków:

1. Komputer przetwarza dane w oparciu o ich struktury syntaktyczne;
2. Każdy układ, który przetwarza dane poprzez transformacje syntaktyczne, jest w istocie komputerem cyfrowym;
3. Wyobraźmy sobie człowieka, zamkniętego w pokoju, do którego dostarczane są kartki ze słowami czy zdaniami w języku chińskim. Człowiek nie rozumie ich, ale bierze słownik i znajduje odpowiedniki polskie. Przepisuje je i oddaje na innej kartce;

4. Człowiek z owego “chińskiego pokoju” dokonuje operacji polegającej na transformacjach syntaktycznych;
5. Jednak człowiek ten nie rozumie chińskiego: dokonuje on jedynie automatycznego przekładu;
6. Jako że jest on *de facto* komputerem cyfrowym, to nie można powiedzieć, żeby komputer wykonujący odpowiednie działania rozumiał chiński;
7. (Wniosek końcowy) *Rozumienie* jest czymś, co nie redukuje się do procesów syntaktycznych. Tak więc, tego rodzaju stany mentalne ludzi nie polegają tylko na wykonywaniu programów<sup>73</sup>.

## II. Argument modalny.

Ta grupa argumentów wymierzona jest w funkcjonalistyczne odmiany materializmu, oraz w coś, co Searle nazywa “projektami naturalizowania treści”. Ogólnie rzecz biorąc, funkcjonalizm jest złagodzoną wersją behawioryzmu, jako że poszukuje determinantów treści semantycznej stanu mentalnego w jego konsekwencjach. Nie ogranicza się jednak jedynie do konsekwencji w postaci zewnętrznych zachowań, lecz uwzględnia również interakcje: dwa stany są tego samego typu, jeśli pozostają w tych samych relacjach do stanów wyjściowych i wejściowych organizmu *oraz* powodują te same stany wewnętrzne. W projektach naturalizowania treści – których semantyka funkcjonalistyczna stanowi w gruncie rzeczy odmianę – chodzi z grubsza o to, by wyjaśnić zjawiska w rodzaju znaczenia czy intencjonalności w kategoriach porządku naturalnego: wyjaśnić posiadanie znaczenia w kategoriach relacji naturalnych, na przykład sprowadzić znacznie myśli ‘Mruczek jest biały’ do związków przyczynowych pomiędzy jej podmiotem a przypadkami białości i Mruczkowości<sup>74</sup>. Dzięki tego rodzaju redukcji uzyskalibyśmy możliwość traktowania zjawisk mentalnych jako elementu pola badań nauk przyrodniczych.

Zdaniem Searle’a program naturalistyczny jest jednak niewykonalny z pewnych zasadniczych przyczyn. Istnienie relacji funkcjonalnych, przyczynowych, czy -

---

<sup>73</sup> Patrz Searle [1999], str. 71: “Najgłośniejszym zarzutem wobec mocnej wersji SI był .... argument <<chińskiego pokoju>>, który *wykarzuje* (kursywa moja – MG), że ... system mógłby realizować ... program, nawet jeśli ... w ogóle nie rozumiałby języka chińskiego”.

<sup>74</sup> Taka kauzalna strategia naturalizacji obrona została przez takich autorów, jak Fodor [patrz np. Fodor [2001]] czy Dretske (patrz Dretske [1981]). Inną odmianą tego projektu będą podejścia poszukujące wyjaśnienia teleologicznego, jak Milikian (patrz Mlikian [1984]) czy Papineau (patrz Papineau [1993]).

teleologicznych, pomiędzy elementami systemu, nie gwarantuje własności bycia procesem umysłowym, gdyż nie determinuje pojawienia się świadomości:

“Przypuśćmy, [...] że dysponujemy wyczerpującym, eksternalistycznym, kauzalnym opisem przekonania, że woda jest mokra. Nasuwa się tu jednak pewna oczywista trudność: system ten mógłby pozostawać we wszystkich tych relacjach, a mimo to wcale nie mieć przekonania, że woda jest mokra”<sup>75</sup>.

Mamy tu więc do czynienia ze swoistą odmianą modalnego argumentu na rzecz braku identyczności pomiędzy zjawiskami mentalnymi a fizycznymi. Jest to podejście, które opiera się na zastosowaniu prawa Leibniza: skoro przedmioty mają różne własności, to są różnymi przedmiotami. Pomysł polega na odwołaniu się nie do zwykłych cech przedmiotu, lecz - do różnic zachodzących w innych możliwych światach. Pozwala to dokonać rozróżnienia pomiędzy *współwystępowaniem* a *identycznością* własności (dwie cechy mogą ściśle współwystępować w naszym świecie, a jednak różnić się zakresami w sytuacjach kontrfaktycznych). Najbardziej znane wersje tego rozumowania przedstawili Kripke (Kripke 1988) – wskazuje on, że istnieją różnice pomiędzy modalnymi własnościami bólu i odpowiedniego stanu neuronalnego - oraz Nagel (Nagel 1997), który dowodził, że stany mentalne i odpowiednie stany fizyczne różnią się pewnymi swoistymi własnościami epistemicznymi.

### III. Argument z względności własności syntaktyczno – obliczeniowych.

Ta warstwa dyskusji wymierzona jest w coś, co Searle nazywa *kognitywizmem* – zbioru założeń, stojących u podstaw dominujących programów badawczych kognitywistyki (*Cognitive Science*). Na str. 259/260 pracy Searle [1999] autor podaje wygodną charakterystykę tego rodzaju podejść. Można ją streścić następująco:

Procesy poznawcze są zlokalizowane w mózgu; wywierają też wpływ na świadomość. Te cechy nie są jednak dla nich konstytutywne. Te same zjawiska mogłyby być realizowane poprzez zupełnie odmienne układy fizyczne. Zasadnicze aspekty procesów przetwarzania informacji (operacje na regułach gramatyki uniwersalnej Chomsky’ego, języku myśli Fodora, czy regułach widzenia Marra) są niedostępne dla świadomości. Mózg jest komputerem, a procesy mentalne mają charakter obliczeniowy. Nie chodzi tu o *utożsamienie* procesów mentalnych z realizacją *software’u* (wówczas można by tu zastosować argument “chińskiego pokoju”), lecz wskazanie, że “stany mentalne są, co najmniej, procesami obliczeniowymi, a

---

<sup>75</sup> Searle [1999] str. 79. Warto może zauważyć, że jeśli jest to wystarczający argument, to stosuje się on do wszystkich prób utożsamienia zjawisk mentalnych z niementalnymi.

procesy mentalne są procesami obliczeniowymi przeprowadzanymi na formalnej strukturze owych stanów mentalnych”, mimo że “programy, same przez się, nie konstytuują zjawisk mentalnych”<sup>76</sup>. Fakt ten uzasadnia się wskazując, iż realizowanie procesów o charakterze syntaktyczno – obliczeniowym jest jedynym dostępnym modelem procesu przekształcania obiektów wyposażonych we własności semantyczne, który tłumaczy relewantność owych transformacji dla wartości logicznej przekształcanego obiektu. Postulowanie obliczeniowego poziomu opisu zjawisk umysłowych pozwala nam więc wytłumaczyć, dlaczego proces zastanawiania się prowadzi systematycznie do poprawnych wniosków<sup>77</sup>.

Konsekwencją takiego podejścia do sprawy jest przyjęcie tezy, że bardzo odmienne w sensie fizycznym zjawiska są w stanie realizować ten sam program (algorytm). Możemy skonstruować układ myślący zarówno w oparciu o białko, krzemowe tranzystory, jak i być może złożony układ puszek i drutów. Decyduje o tym nie tyle użyty materiał, co sposób jego zorganizowania. Istnieje tu, rzecz jasna, daleko idąca zbieżność z własnościami *napisów*: w tym wypadku również zdają się nie istnieć ograniczenia co do własności fizycznych materiału piśmienniczego.

Searle wyciąga stąd wniosek, że każdy odpowiednio złożony przedmiot posiada opis jako realizacja dowolnego programu:

“Gdy wychodzimy od definicji obliczania, trudno dostrzec, w jaki sposób można by uniknąć następujących wniosków: 1. Dla każdego przedmiotu istnieje pewien opis tego przedmiotu, na gruncie którego jest on komputerem cyfrowym. 2. Dla każdego programu i każdego dostatecznie złożonego przedmiotu istnieje jakiś opis tego przedmiotu, na gruncie którego realizuje on ów program. Na przykład ściana, o którą się w tej chwili opieram, realizuje program WordStar, ponieważ istnieje pewna konfiguracja ruchów cząstek, która jest izomorficzna z formalną strukturą tego programu”<sup>78</sup>

Autor zdaje się twierdzić, że na mocy założeń kognitywistycznych zmuszeni jesteśmy przyjąć, iż przypisywanie stanów mentalnych obiektom jest zależne wyłącznie od naszego kaprysu. Gdyby procesy mentalne były tożsame z obliczeniowymi, to moglibyśmy przypisywać stany mentalne dowolnym przedmiotom, gdyż dowolny przedmiot można dowolnie interpretować obliczeniowo. Bycie procesem obliczeniowym nie wyróżnia procesów mentalnych spośród innych zdarzeń świata, gdyż nie wyróżnia niczego.

---

<sup>76</sup> Searle [1999], str. 264.

<sup>77</sup> Klasycznym przedstawicielem tak rozumianego kognitywizmu jest Fodor. Patrz Fodor 2001rozdz. 1. gdzie znajdziemy omówienie tego rodzaju programu badawczego i diagnozę stojących przed nim problemów.

<sup>78</sup> Searle [1999], str. 274.



\*

Pomiędzy tymi rozumowaniami zachodzą różnorakie związki. Jeśli argument modalny jest skuteczny w przypadku redukcjonizmu naturalistycznego, to niewątpliwie spełni swą rolę również przy refutacji silnej AI. Argument chińskiego pokoju jest lematem w argumentacji modalnej – stwierdza, że możliwe jest dokonywanie obliczeń bez świadomości, stanowiąc swoisty odpowiednik takich narzędzi heurystycznych, jak analizowane przez m.in. przez Zombie Chalmersa czy Bagnolud Davidsona<sup>79</sup>. Postaram się najpierw poddać krytyce założenia stojące za postulowanymi przez Searle’a intuicjami modalnymi. Następnie zaś poddam analizie krytykę kognitywizmu.

Argumentacje modalne stanowią dość kontrowersyjne zagadnienie. Ich status jest niejasny, choćby dlatego że kluczowe pojęcia - jak na przykład możliwy świat, czy konieczność *de re* - nie posiadają jednoznacznej i powszechnie akceptowanej analizy. W tej pracy pozostawię na boku ogólne zagadnienie prawomocności argumentacji modalnej, skupiając się nad użyciem, jaki czyni z niej Searle.

Proponowane przez niego rozumowanie stoi przed pewnym zasadniczym dylematem. Wskazuje przecież tylko, że przedstawione teorie *nie wyjaśniają zjawiska świadomości*. Jednak redukcjonista czy zwolennik programu badawczego AI nie musi mieć takich aspiracji. Może on traktować swoje przedsięwzięcie w sposób skromny: próbować badać, jakie cechy musimy przypisać umysłowi, by wyjaśnić pewne znane nam skądinąd jego własności. Byłaby to swoista odmiana argumentacji transcendentalnej. Tezy wielu autorów da się ujmować w tych kategoriach. Quine i Davidson głoszą, że własności mentalne muszą przysługiwać bytom materialnym, inaczej bowiem nie byłibyśmy w stanie wyjaśnić przyczynowego oddziaływania stanów psychicznych. Fodor i inni zwolennicy obliczeniowo – kognitywistycznych modeli myślenia dodadzą, opierając się na wynikach badań AI, że własności intencjonalne przysługują bytom fizycznym za pośrednictwem poziomu obliczeniowego – dlatego, że jest to *jedyne dostępne wyjaśnienie zachowywania wartości semantycznych przez procesy psychiczne*.

Searle zdaje się stać na stanowisku, że taki zabieg jest nieuprawniony. Teorie redukcjonistów są *błędne*, dlatego, że nie wyjaśniają świadomości. Ma ona być zasadniczą cechą życia psychicznego, jak to wskazywać ma rozdział VII pracy Searle [1999]. Wszelka próba dania opisu umysłu bez uwzględnienia tego faktu ma być *zasadniczo niepełna*.

---

<sup>79</sup> W sprawie zombie patrz Chalmers [1996]; co do Bagnoluda zaś – Davidson [2001]

Łatwo byłoby przyjąć taki zarzut, jeśli miałby on być jedynym. Nie ulega wątpliwości, że świadomość jest zjawiskiem ciekawym, które dobrze byłoby wyjaśnić. Nikogo zaskakiwać nie powinny jednak próby naświetlenia innych aspektów życia psychicznego. Przecież “niepełność” nie musi przerażać: mocne stanowisko Searle’a zdaje się zaś wymagać, by teorie, które eksplikują intencjonalność czy semantyczność bez odwołań do świadomości/fenomenalności, były nie tyle *niekompletne*, co *falszywe*.

Searle przedstawia rozumowanie, które ma wskazać, że pojęcie nieświadomego stanu mentalnego będzie zrozumiałe jedynie jako pojęcie stanu, który potencjalnie mógłby być świadomy. Ma nas ono przekonać, że analiza pomijająca fenomenalny aspekt zjawisk mentalnych będzie niezrozumiała. Zrekonstruować je można w następujący sposób<sup>80</sup>:

1. Przekonania, pragnienia, i inne stany, które miałyby być wyjaśniane poprzez odwołanie do stanów nieświadomych, mają charakter “aspektowy”. Łączy się z tym *intensjonalność*: fakt, że ich warunki spełniania obejmują sposób przedstawienia przedmiotu. Na przykład pragnienie wody nie zawsze musi być pragnieniem H<sub>2</sub>O, mimo że woda = H<sub>2</sub>O.
2. Aspektowość nie przysługuje mózgowi na mocy jego publicznie dostępnych właściwości, jak budowa neurologiczna czy zachowanie organizmu.
3. Nieświadome stany mentalne to zjawiska o charakterze wyłącznie neurologicznym.
4. Z punktów 1 oraz 2 wynika, że aspektowość przysługuje stanom mózgu na mocy ich własności fenomenalnych. Skoro aspektowość przysługuje stanom świadomym, a nie przysługuje analogicznym stanom dostępnym publicznie, to jest ona własnością (drugiego rzędu) fenomenalnych cech mózgu.
5. Przypisywanie jednostkom przekonań, i innych stanów potocznie uznawanych za nieświadome staje się zrozumiałe, jeśli przyjmiemy, że możemy przypisywać mózgowi własności dyspozycyjne względem świadomości. Przypisywanie osobie śpiącej bez snów przekonania, że Paryż jest stolicą Francji sensowne jest, dlatego że mózg tej osoby jest w stanie, *który w odpowiednich warunkach będzie odpowiadał temu przekonaniu*.

Wobec tego rozumowania można jednak sformułować następujące zarzuty:

Po pierwsze, kroki 1-5 mają uderzać w modele obliczeniowo/ funkcjonalistyczne. Tymczasem przesłanka 3 – twierdzenie, że stany nieświadome nie mają żadnych własności psychologicznych – zaprzecza temu stanowisku już w punkcie wyjścia. Program funkcjonalistyczny polega na twierdzeniu, że istnieje poziom pośredni pomiędzy psychologią

---

<sup>80</sup> Searle [1999], str. 209 – 217.

a neurologią. Argumentacja opiera się więc na błędnym kole, zakładając wynik rozważań już w punkcie wyjścia<sup>81</sup>. Uderza ona jedynie w zwolenników teorii identyczności typów lub eliminatywistów – tylko oni byliby skłonni przyjąć przesłankę 3. Ci pierwsi nie zgodzili by się jednak na 2, a drudzy - uznając intencjonalny poziom wyjaśniania za zbędny - nie wchodziliby w ogóle w tę dyskusję. Centralnym punktem spornym jest stanowisko funkcjonalistyczne (w jego licznych wariantach i odmianach), a przedstawiona linia rozumowania wymaga założenia jego fałszywości.

Po drugie, na poważne kłopoty narażona jest argumentacja za twierdzeniem 2. Głosi ono, że aspektowość nie przysługuje zjawiskom na mocy ich własności neurologicznych. Searle twierdzi, że nawet gdybyśmy potrafili odczytać treść ze stanów mózgu, to potrzebne byłoby do tego wnioskowanie:

“Ktoś, kto dysponowałby idealną wiedzą o związkach przyczynowych, mógłby wyprowadzać wnioski dotyczące tego, co intencjonalne, na podstawie tego, co neurofizjologiczne (...) ale (...) w takich wypadkach pojawia się *wnioskowanie*, a opis faktów neurofizjologicznych nie jest jeszcze opisem intencjonalności”<sup>82</sup>.

Można interpretować ten zarzut na dwa sposoby. Odwołanie do związku przyczynowego odczytywać można jako zaprzeczenie istnieniu relacji pojęciowych czy identyczności: miałyby by to wskazywać na fakt, że zjawiska, między którymi zachodzi ta relacja, są w jakiś zasadniczy sposób odrębne. Takie podejście czyniłoby argument Searle’a kolistym: tezę o fenomenalnym charakterze odrębności intencjonalności i fizykalności mielibyśmy jako jedną z *przesłanek*. Nie sądzę, żeby to właśnie miał Searle na myśli. Bardziej naturalne zdaje się odczytanie tego fragmentu jako inferencji własności metafizycznych (prawdziwości/konieczności) zdań z jego atrybutów epistemicznych (tego, że potrzebne jest wnioskowanie, by na podstawie własności neurofizjologicznych stwierdzić własności intencjonalne). Tego typu rozumowanie jest jednak bardzo wątpliwe. Analogicznie moglibyśmy powiedzieć, że nie można przypisywać własności bycia H<sub>2</sub>O wodzie – przejście

---

<sup>81</sup> Takie ujęcie sprawy nie jest do końca ścisłe. Kontrowersyjna przesłanka nie stoi, dokładnie rzecz biorąc, w sprzeczności z programem funkcjonalisty. Trudno jednak wyobrazić sobie, dlaczego miałby on przyjmować takie założenie: przecież poszukuje on determinantów stanu mentalnego w aspektach przyczynowo – obliczeniowych, a nie w powiązanych z nim stanach świadomości. Niemożliwe jest, wedle tego podejścia, zachodzenie odpowiednich własności semantycznych bez odpowiadających im relacji funkcjonalno – obliczeniowych; żaden z elementów naturalistycznego programu badawczego nie skłania do przyjęcia postulowanej przez Searle’a tezy. W tym ujęciu błąd traci charakter błędnego koła, polegając zaczyna natomiast na skrajnym wypaczeniu poglądów oponenta.

<sup>82</sup> Searle 1999, str. 213.

od przekonań o wodzie do przekonań  $H_2O$  wymaga bowiem wnioskowania. Wymienialność terminu w kontekstach epistemicznych nie może stanowić kryterium tożsamości. Jak to wyłożył Achillesowi żółw, przejście od przekonania do przekonania zawsze “wymaga wnioskowania”. Nie przejdziemy jednak automatycznie od tego stwierdzenia do tezy, że wnioski są zawsze niepoprawne.

Po trzecie – wniosek 5 nie przekreśla programu funkcjonalisty. Może on przecież uznać, że stany, o których mowa, mają pewien element fenomenalny, i w tym sensie są stanami fenomenalnymi. Nie musi to jednak znaczyć, że treść intencjonalna stanu zależy od jego własności fenomenalnych. Mogą one zjawiskom intencjonalnym przysługiwać, nie mając wpływu na ich treść – dokładnie, tak jak każdemu zjawisku fizycznemu przysługuje umiejscowienie geograficzne, zupełnie dla jego natury nieistotne. To stanowisko nie klóciłoby się ze stwierdzeniem, że każdy stan intencjonalny jest (potencjalnie lub faktycznie) świadomy. Stany świadome byłyby tożsame z intencjonalnymi, lecz *własności* intencjonalne byłyby tożsame nie z fenomenalnymi, lecz - z syntaktycznymi. Do wyjaśnienia treści potrzebna byłaby więc taka właśnie teoria, jaką postulują krytykowie przez Searle’a autorzy. Jeśli – w zgodzie z 4 – każdy stan intencjonalny musi mieć charakter fenomenalny, to wartość modelu obliczeniowego możemy zachować, postulując przygodność związku pomiędzy fenomenalnym a intencjonalnym aspektem danego zjawiska czy stanu. Istota tezy funkcjonalisty tkwi twierdzeniu, że treść przysługuje stanom mentalnym na mocy własności (na przykład) obliczeniowo – przyczynowych; uznanie, że przysługują im również z konieczności różnorakie inne cechy (fenomenalne, przestrzenne) nie stanowi żadnego problemu. Rozwiązanie to – które jest w istocie przetransponowaniem “monizmu anomalnego” Davidsona (patrz Davidson 1991) z poziomu fizyczno/intencjonalnego na fenomenalno/intencjonalny – znakomicie uzupełnia zauważany przez Searle’a fakt, że własności intencjonalne mają mieć charakter *dyspozycyjny* wobec świadomości: tym, co wspólne świadomemu stanowi myślenia, że *p*, z dyspozycją do znalezienia się w nim są własności syntaktyczno - obliczeniowe. Możemy więc wyjaśnić, dlaczego dany stan neurologiczny ma takie a nie inne własności dyspozycyjne względem świadomości. Warto również zwrócić uwagę na fakt, że nieświadome przekonania przypisujemy nie tylko w przypadku stanów braku świadomości (przypadek snu). Znacznie bardziej istotne jest przypisywanie przekonań, że *p*, osobom w pełni świadomym, które jednak nie rozważają aktualnie kwestii, czy *p*. Związek pomiędzy stanami świadomości a przekonaniami zdaje się być w tych wypadkach właśnie *przygodny*: dana osoba ma dane przekonanie niezależnie od

tego, czy akurat je sobie uświadamia, czy nie. Dokładnie w ten sposób ujmuje to powyżej zaproponowane rozwiązanie.

Jak sądzę, przedstawiona analiza wskazuje, że teza o nierozdzielalnym związku pomiędzy intencjonalnością a świadomością jest nieuzasadniona. Nie powinno to specjalnie zaskakiwać: praktyka tak nauki, jak i poznawania świata w ogóle, ma charakter zasadniczo pluralistyczny. Opiera się na szeregu podejść, i abstrahowaniu od jednych własności obiektów przy badaniu innych. Nie ma powodu, żeby w przypadku psychologii miało być inaczej; w szczególności analizy przedstawione w 'Umyśle na nowo odkrytym' nie skłaniają do jakiegóż szczególnego jej traktowania.

\*

Kolejnym krokiem w kampanii Searle'a przeciw obliczeniowym modelom umysłu jest jego krytyka kognitywizmu. Argumentacja, którą zaliczyłem we wstępie do grupy III, ma pokazać, że nieporozumieniem jest uznanie zachodzenia procesów obliczeniowych za *warunek konieczny* występowania zjawisk mentalnych (a w szczególności - intencjonalnych). O dowolnym obiekcie powiedzieć można, że realizuje dowolny proces obliczeniowy: własność ta nie wyróżnia więc własności psychicznych spośród reszty świata.

Dlaczego możemy przypisywać własności obliczeniowe dowolnym przedmiotom? Żeby odpowiedzieć na to pytanie, trzeba najpierw choćby wstępnie określić, co rozumiemy przez proces obliczeniowy. Można powiedzieć, że jest to ciąg zdarzeń, który polega na zmianie własności syntaktycznych. Przesunięcie kartki z napisem 'król jest żywy' nie stanowi procesu obliczeniowego. Będzie nim natomiast dopisanie do tego zdania wyrażenia 'i nie jest żywy'. Co istotne, definicja procesu obliczeniowego opiera się pojęciu *własności syntaktycznej*. Procesy obliczeniowe są ciekawe przez to, że jesteśmy w stanie skonstruować maszynę, która będzie je realizować. Umiemy też opisać dokonywanie operacji logicznych i rachunków arytmetycznych za pośrednictwem transformacji syntaktycznych. Maszynami, które wykonują tego rodzaju działania, są komputery.

Searle dąży do zakwestionowania tego ostatniego twierdzenia. Sądzi on, że dokonywanie obliczeń nie jest własnością obiektów fizycznych, jak komputery, lecz raczej układu obejmującego urządzenie i interpretującego je w pewien sposób obserwatora. A to, dlatego, że dowolny obiekt możemy uznać za realizację dowolnej struktury syntaktycznej. Składnia nie jest, jak pisze Searle, kwestią własności fizycznych. Wydzielenie własności relewantnych syntaktycznie jest działaniem, które musi podjąć obserwator.

Zazwyczaj uważamy, że konstytutywne dla składni są własności graficzno – geometryczne napisu. Ale nie musi przecież tak być. Jeśli uznamy, że położenie przestrzenne

jest relewantne dla składni, a przesunięcie napisu stanowi zamianę zdania w koniunkcję z jego negacją, to – wbrew naszym intuicjom – okaże się, że ów ruch był procesem o charakterze obliczeniowym. Można sobie wyobrazić jakiegoś rodzaju grę, której elementem byłoby umieszczanie kartki z napisem na odpowiednim polu, w zależności od czego zmieniałoby się jej znaczenie. Będziemy mieli pole koniunkcji, pole negacji, itd. Oczywiście, napisy określające reguły transformacji nie musiały by się znajdować na planszy – wystarczy, by grający *znali je na pamięć*. To na mocy naszej umowy nadajemy elementom świata interpretacje syntaktyczną – tak więc, *to stany mentalne, idee i wyobrażenia stojące za tymi umowami konstytuują własności syntaktyczne obiektów, a nie - odwrotnie*. Idąc dalej tym tropem zobaczymy, że do tego, by móc stwierdzić, że dany proces ma określony charakter obliczeniowy, potrzebna jest jedynie odpowiednia złożoność obiektu czy zestawu obiektów. Nadanie (bądź nie) interpretacji syntaktycznej tym obiektom jest wyłącznie kwestią decyzji interpretatora czy umów zawartych w danym społeczeństwie. Ma to prowadzić do wniosku, że dowolny obiekt, przy odpowiednim sposobie opisu, jest częścią procesu obliczeniowego. Tak więc, nie ma sensu postulowanie szczególnej roli faktu realizacji transformacji syntaktycznej dla bycia podmiotem stanów mentalnych. Odwrócenie rzeczywistego porządku zjawisk ma zmuszać kognitywistę do przyjęcia swoistego panpsychizmu.

Argumentacja ta jest sprzeczna z różnymi intuicjami. Nawet w epoce internetu myśl, że świat mógłby być do tego stopnia nasycony komputerami, wydaje się dziwaczna. Wedle tego wizerunku, konstruktorzy komputerów i programiści zajmują się nie tyle wymyślaniem czy konstruowaniem nowych maszyn liczących tudzież - tworzeniem obsługujących je aplikacji, co odnajdywaniem ich w przyrodzie, czy też może nauczaniem użytkowników takiego patrzenia na rzeczywistość, by udzielała ona stosownych informacji, przechowywała je, itd. Wydaje się to absurdalne. Gdzie tkwi błąd?

Powróćmy do przykładu z dokonywaniem transformacji syntaktycznej poprzez przesuwanie kartki z napisem. Dlaczego w procesie tym dochodzi do przetwarzania informacji syntaktycznej? Otóż syntaksa układu: kartka – położenie, określona jest w następujący sposób:

- a) Składnią wstępną zdania jest standardowo rozumiany napis na kartce.
- b) Składnią pełną jest składnia, jeśli kartka znajduje się na polu A, a jeśli zaś na polu B – koniunkcja składni wstępnej i jej negacji.

Nasz “komputer” przetwarza dość dziwne informacje: połączenia standardowo rozumianego napisu i jego położenia. I tym różni od zwykłego komputera, że ten jest w stanie

operować na napisach rozumianych standardowo. Tu tkwi klucz do zagadki: postulując obliczeniowy poziom funkcjonowania mózgu, nie chodzi nam o to, że można go opisać jako maszynę przetwarzającą *jakieś* informacje, lecz o to, że *znając skądinąd budowę organizmu ludzkiego i charakter informacji, jakimi dysponuje jednostka, można będzie wyjaśniać rzeczywiście dokonywane działania poprzez odwołanie się do syntaktycznej transformacji tych danych*. Pytanie, czy umysł ma strukturę obliczeniową, jest pytaniem, czy można opisać zależności i związki przyczynowe, jakie zachodzą w naszym organizmie, tudzież np. zależności epistemologiczne pomiędzy obserwacją, namysłem, i przeprowadzaniem wnioskowań, *jako realizację procesów obliczeniowych*. Nie chodzi o to, że możemy dokonać opisu życia jednostki myślącej jako procesu przetwarzania informacji syntaktycznej poprzez wydzielenie *a posteriori* pewnych jej elementów jako składników przetwarzania danych. Searle nie zauważa, że nakładamy bardzo konkretne warunki na wyjaśnienie obliczeniowe. Interpretację podstawowych pojęć - wejść, wyjść, natury danych - otrzymujemy z psychologii, biologii, itd. Teza ‘kognitywizmu’ ma charakter naturalistyczny – zakłada całość informacji o świecie i *ma stanowić jego część*.

\*

Dokonując krytyki obliczeniowych podejść do umysłu, Searle przypisuje swym oponentom zbyt podniosłe aspiracje i cele. Funkcjonalista czy fizykalista nie dąży do wyjaśnienia wszystkiego, co kiedykolwiek miało związek z myśleniem – nie musi mu więc przeszkadzać brak rozwiązania zagadki świadomości. Searle’owski ‘kognitywista’ nie aspiruje do tego, żeby teza o obliczeniowym charakterze zjawisk mentalnych była wszystkim, co potrzebne do opisanego Natury Umysłu – wręcz przeciwnie, stwarza ramy badania związków umysłu ze światem zewnętrznym i przebiegu procesów wewnętrznych. Argumenty z książki ‘Umysł na nowo odkryty’ stosowałyby się do badaczy, którzy traktowałiby analizowane stanowiska jako odpowiedzi na pytanie o ostateczną naturę Myśli i Uczucia. Oczywiście – stwierdzenie ‘Umysł to maszyna Turinga’ byłoby tu *niepełne lub fałszywe*: niepełne z powodów, jakie wskazuje Searle, a fałszywe, dlatego, że odpowiedź na nie jest inna.

*BIBLIOGRAFIA:*

Chalmers, David

[1996] *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press Inc, New York.

Davidson, Donald

[2001] *Knowing One's Mind*, w: Subjective, Intersubjective, Objective, Oxford University Press Inc, New York.

Dretske, Fred

[1982] *Knowledge and the Flow of Information*, MIT Press, Cambridge, MA.

Fodor, Jerry

2001] *Eksperci od więzów*, tłumaczył Marcin Gokieli, Wydawnictwo Aletheia, Warszawa.

Millikan, Ruth

[1984] *Language, Thought, and Other Biological Categories*, MIT Press, Cambridge, MA.

Nagel, Thomas [1997]

*Jak to jest być nietoperzem*, w: *Pytania ostateczne*, tłumaczył Michał Szczubiałka, Wydawnictwo Aletheia, Warszawa.

Papineau, David

[1993] *Philosophical Naturalism*, Blackwell Publishers, Oxford.

Searle, John

[1999] *Umysł na nowo odkryty*, tłum. Tadeusz Baszniak, Państwowy Instytut Wydawniczy, Warszawa.



## **Czy istnieje granica między kognitywistyką a filozofią umysłu?**

*Filozofia jest ściślej związana ze sztuczną inteligencją niż z pozostałymi naukami. Obie te dziedziny wymagają formalizacji wiedzy zdroworozsądkowej i usuwają jej słabości.*

John McCarthy [1977]<sup>83</sup>

Zdawałoby się, że gdyby pod uwagę wziąć jedynie naturalistyczne nurty w filozofii umysłu, to odpowiedź na postawione w tytule pytanie musiałaby być jednoznacznie przecząca, jako że współczesny naturalizm odrzuca ścisły rozdział filozofii od nauki. Zamierzam jednak odejść od takich ogólnikowych rozważań. Twierdzę, że granica dzieląca filozofię i kognitywistykę istnieje, choć czasem trudno ją elegancko wytyczyć. Na konkretnym przykładzie pewnego problemu chcę pokazać zawile relacje między współczesną filozofią umysłu a naukami o poznaniu (kognitywistyką). Co więcej, pod uwagę będę brał przede wszystkim nurty naturalistyczne w filozofii.

### **1. Kognitywistyka**

Przez kognitywistykę rozumiał będę naukę zajmującą się modelowaniem procesów poznawczych. Modelowanie to może polegać na budowaniu modeli ludzkiego umysłu poznającego — w takim wypadku kognitywistyka jest kontynuacją nurtów psychologii kognitywnej (poznawczej). Może jednak polegać również na próbie konstrukcji umysłu sztucznego. Te nurty w kognitywistyce znane są szeroko pod nazwą badań nad sztuczną inteligencją, a wywodzą się z informatyki. Niewątpliwie w naukach o poznaniu mamy do czynienia z interdyscyplinarnością, a więc - z różnymi metodologiami badań.

---

<sup>83</sup> Większość przywoływanych tu tekstów McCarthy'ego można znaleźć w jego witrynie internetowej pod adresem <http://www-formal.stanford.edu/jmc>.

Przedmiot badań kognitywistów jest dla filozofa swojski; idzie wszak o to, czym interesowali się teoretycy poznania od samego zarania filozofii. Kognitywistyka nie jest jednak ani spełnieniem najgorszych przeczuć przeciwników psychologizmu w epistemologii, ani wypełnieniem prorocत्व neopozytywistów. Ze względu na interdyscyplinarność nie daje się nawet zaklasyfikować ściśle jako nauka empiryczna. Przecież nawet informatyki nie sposób zaliczać do nauk empirycznych<sup>84</sup>. Zbyt wiele badań w informatyce prowadzi się metodami ściśle matematycznymi. Zbyt wiele zastosowań ma wymiar czysto inżynierski: opiera się na technice generowania i testowania kolejnych rozwiązań. Wymykająca się nieco podziałowi na nauki dedukcyjne i empiryczne kognitywistyka jest, być może, przykładem dziedziny właśnie inżynierskiej.

Może więc kognitywistyką powinna zajmować się nie filozofia nauki, lecz - filozofia inżynierii? Filozofia inżynierii, dziedzina niemalże nieznaną, uprawiana przez kilku zaledwie filozofów analitycznych, przede wszystkim przez Daniela Dennetta, interesuje się naukami, które odpowiadają na pytania: „jak to działa?” lub „jak to zrobić”<sup>85</sup>. Sztuczna inteligencja pyta, o działanie inteligencji ludzkiej i próbuje zbudować na tej podstawie działający, sztuczny model.

## **2. Rozterki sztucznego inteligenta**

W 1969 roku w badaniach nad sztuczną inteligencją odkryto pewien problem. Niektórzy twierdzą, że to trudność nie mająca rozwiązania, a wręcz głęboki problem filozoficzny. Inżynierowie, przeciwnie, narzekają na filozofów, twierdząc, iż to trudność zaledwie techniczna.

Problem zrodził się w logistycznym nurcie badań nad sztuczną inteligencją<sup>86</sup>. W tym nurcie tworzy się systemy dedukcyjne, które mają dostarczać narzędzi reprezentacji świata, zaś dzięki adekwatnej reprezentacji maszyna ma podejmować inteligentne decyzje. Reprezentacja wiedzy ma charakter zdaniowy i opiera się najczęściej na jakiejś wersji logiki pierwszego rzędu. Fakty w świecie wyrażane są przez tezy budowane przy użyciu

---

<sup>84</sup> Klasycy sztucznej inteligencji, Newell i Simon są przeciwnego zdania. Jednak ich stanowisko należy rozpatrywać w innym świetle, a mianowicie jako polemikę ze ściśle deduktywistycznym ujęciem informatyki. Por. Newell i Simon [1997].

<sup>85</sup> Z punktu widzenia Dennetta biologia ewolucjonistyczna w paradygmacie adaptacjonistycznym jest właśnie inżynierską nauką. Por. Dennett [1995].

<sup>86</sup> Konkurencyjny wobec logistycznego jest nurt heurystyczny, w którym problemy reprezentacji są pomijane.

zdefiniowanych aksjomatycznie predykatów. Na podstawie zbioru aksjomatów, opisujących podstawowe regularności w świecie (np. prawa mechaniki) i typy działań podmiotu inteligentnego (np. przesuwanie przedmiotów), wprowadzonych do bazy danych faktów i reguł działania „sztuczny inteligent” ma planować racjonalne i inteligentne decyzje. Innymi słowy, racjonalna decyzja to rezultat przeprowadzenia formalnego dowodu pewnych tez dotyczących działań podmiotu.

J        John McCarthy opracował własny system zaliczający się do logiki pierwszego rzędu, mający reprezentować wiedzę robota, tzw. rachunek sytuacji. Rachunek sytuacji miał uwzględniać zmiany zachodzące w świecie. Ogólnie rzecz biorąc, zmiany w czasie opisywano przez użycie predykatu *Zachodzi(zdanie, sytuacja\_określona\_czasowo)*, np.:

*Zachodzi(pada\_deszcz(x,s), w\_czasie(p, s ,x))* gdzie *s* jest sytuacją, a *p*- osobą w miejscu *x*.

Niemniej, pojawił się pewien problem. Polega on na modelowaniu zmiany w świecie. W oryginalnym sformułowaniu problem ten został określony mianem „**problemu ramy odniesienia**” (ang. *frame problem*), bowiem McCarthy usiłował go rozwiązać przy użyciu aksjomatów ramowych (ang. *frame axioms*) (Hayes i McCarthy [1969])<sup>87</sup>. Zbiór pojęć podzielono z użyciem różnych „ram odniesienia” na kategorie, które umożliwiały określenie, iż np. żadne działanie polegające na przesuwanie przedmiotu nie powoduje zmiany własności koloru tegoż przedmiotu (Hayes [1987], s.130). Niestety, aksjomatów musi być prawie tyle, co egzemplarzy działań, a więc szybko zauważono, że formalizm jest niezadowolający:

- (1) zapisanie wszystkich aksjomatów jest niezwykle czasochłonne;
- (2) system zawierający miliony oczywistych aksjomatów jest nieelegancki;
- (3) ogromna liczba aksjomatów spowalnia działanie systemu;
- (4) aksjomaty ramowe są często intuicyjnie fałszywe, zwłaszcza w systemach dopuszczających działania równoczesne (Morgenstern [1996])<sup>88</sup>.

Powstało zatem pytanie: czym zastąpić te aksjomaty? Na poszukiwaniu odpowiedzi polega właśnie problem ramy odniesienia w wersji pierwotnej.

Problem jest niezwykle kłopotliwy dla entuzjastów kognitywistyki, bo inteligentnym ludziom się nie przytrafia. Nawet małe dziecko wie, że - kiedy przesuwa spodek z filiżanką - to wówczas przesuwa się też i filiżanka. Tymczasem w rachunku sytuacji należałoby

---

<sup>87</sup> Proszę zwrócić uwagę na tytuł tekstu, w którym po raz pierwszy opisano problem ramy odniesienia. Artykuł jest zresztą znakomity.

<sup>88</sup> Fałszywość wypływa z prostego faktu, że rachunek sytuacji nie uwzględnia działań równoczesnych. Przykłady innych fałszywych — w świetle zdrowego rozsądku, a nie w rachunku, rzecz jasna — aksjomatów ramowych podaje Morgenstern [1996]. Por. także niżej problem strzelaniny z Yale.

wprowadzić dodatkowy aksjomat, a raczej bardzo liczną ich grupę aksjomatów. Okazuje się więc, że w logice stosowanej reprezentacja zmieniającego się świata jest bardzo trudna, i że pojawiają się trudności wymykające się intuicjom.

Owe niezgodności z intuicjami popchnęły, jak sędzę, (niektórych) filozofów w innym kierunku niż badaczy. Podczas gdy kognywiści precyzyjnie definiowali problemy (problem ramy odniesienia wskazuje na pewne kolejne problemy reprezentacji wiedzy zdroworozsądkowej), to filozofowie próbowali sprowadzić ten problem np. do problemu indukcji. Trudność wiążąca się z indukcją jest bowiem zrozumiała i przyswojona intuicyjnie.

### 3. Czy *Ćwirek lata*?

Szybko pojawiły się różne metody usuwania problemu<sup>89</sup>. Po pierwsze, McCarthy w ramach rachunku sytuacji opracował pewien formalizm wnioskowań niemonotonicznych, będący modelem rozumowań zdroworozsądkowych (McCarthy [1980]). Z jego badań wyrósł cały, niezwykle ciekawy nurt w zakresie badań nad rozumowaniami niemonotonicznymi, o którym napiszę szerzej. Do tego nurtu należy logika rozumowań zaocznych (*default logic*) Raymonda Reitera, gdzie rozwiązanie polegało na przyjęciu założenia o tym, że domyślnie zmienia się minimalnie mało aspektów świata. Współcześnie, najbardziej rozbudowane rozwiązania niemonotoniczne tworzą Vladimir Lifschitz (Lifschitz [1995]) i Leora Morgenstern (Morgenstern [1991], [1996]).

Po drugie, istnieją próby ucieczki od paradygmatu logistycznego. I tak Drew McDermott, dawniej zwolennik rozwiązań McCarthy'ego<sup>90</sup>, porzucił badania nad reprezentacją wiedzy. Obecnie zajmuje się sztuczną inteligencją w wymiarze proceduralnym, czyli koncentruje się na heurystycznych aspektach inteligencji. Po trzecie, istnieją rozwiązania oparte na programowaniu w logice (języku Prolog) i rachunku zdarzeń (Shanahan [1997]). Po czwarte, rozważane są też metody statystyczne i probabilistyczne.

Zanim przejdę do krótkiego omówienia rozumowań niemonotonicznych, wspomnę o pokrewnych - wobec problemu ramy odniesienia - trudnościach. **Problem konsekwencji**

---

<sup>89</sup> Pomijam rozwiązania podobne do aksjomatów ramowych i wczesne próby z lat siedemdziesiątych (np. system STRIPS Fikesa i Nilsona, por. Fikes i Nilson [1971]). Jako całkowicie chybione pomijam też pewne czysto inżynierskie propozycje, np. Lenting i Braspenning [1993], gdzie autorzy po prostu nie zamierzają budować autonomicznych systemów reprezentujących świat, lecz roboty przyjmujące nakazy człowieka.

<sup>90</sup> Por. McDermott [1987], s. 113, gdzie twierdzi, że problem ramy odniesienia nie istnieje, a więc nie ma sensu go rozwiązywać. Później, w 1986, McDermott pod wpływem własnego kontrprzykładu wobec znanych rozwiązań, przeszedł na stronę heurystycznych rozwiązań. Por. np. Hanks, McDermott [1986].

(*ramification problem*) polega na próbie znalezienia sposobu niejawnego reprezentowania konsekwencji działania. Otóż, każde działanie ma nieskończenie wiele konsekwencji (to oczywiste przy reprezentacji zdaniowej: każde zdanie pociąga nieskończenie wiele innych zdań), a żadna elegancka i obliczeniowo użyteczna reprezentacja nie może ich uwzględniać.

**Problem kwalifikacji** (*qualification problem*) polega na trudności określenia warunków koniecznych i wystarczających każdego działania. Żeby małpa mogła zjeść banana, między nią a bananem nie może stać szklana szyba, odnóża małpy nie mogą być przywiązane do detonatora bomby atomowej, małpa nie może się zamienić nagle w babcię Jerry'ego Fodora itd. Choć istota tych trzech problemów: ramy odniesienia, konsekwencji i kwalifikacji, jest nieco inna, to jednak rozwiązania zwykle stosują się do wszystkich trzech. Wszystkie bowiem związane są z reprezentacją i rachunkiem logicznym.

Tutaj zajmę się przede wszystkim rozumowaniami niemonotonicznymi, bo najbardziej zbliżają się one do wzorca zarysowanego w motcie tego artykułu: chodzi w nich o formalizację zdrowego rozsądku (ogólne filozoficzne wprowadzenie do problematyki logiki niemonotonicznej znajduje się w Fuhrmann [1998]; w języku polskim dobre wprowadzenie znajdzie Czytelnik w Malinowski [1997]). Rozumowania niemonotoniczne można formalizować w klasycznej logice pierwszego rzędu (cyrkumskrypcja McCarthy'ego, patrz niżej) lub w logice niemonotonicznej „pełną gębą”.

Czym jest logika niemonotoniczna? Zwykła logika jest monotoniczna: zbiór wniosków jest monotoniczny względem zbioru założeń. Obowiązuje w niej zasada monotoniczności (M):

(M) Dodanie przesłanek nigdy nie powoduje podważenia wyciągniętych już wniosków.

Na przykład relacja konsekwencji w klasycznej logice jest w tym sensie monotoniczna. Logika niemonotoniczna to taka, w której zasada (M) nie obowiązuje<sup>91</sup>. Przyjęcie dodatkowego założenia może spowodować usunięcie pewnych elementów ze zbioru wniosków. Istnieją bowiem wnioskowania domyślne, zwykle czy też normalne. Jeśli Ćwirek jest ptakiem, to wnioskujemy zwykle, że Ćwirek lata. Jeśli dodamy przesłankę, że Ćwirek jest pingwinem, to już nie wnioskujemy, że lata.

Z punktu widzenia klasycznej logiki, w rozumowaniach zdroworozsądkowych występują liczne entymematy i niedopowiedzenia. Należy je usuwać, aby uzyskać właściwy tok rozumowania, badany przez logika. Tymczasem w logice niemonotonicznej nie uzupełnia

---

<sup>91</sup> Warto może dodać, że z tego względu logika niemonotoniczna *nie* podaje reguł poprawności rozumowań; stanowi jedynie formalizację zasad wnioskowań zdroworozsądkowych. Uważa się, że to zbliża ją do kognitywistyki i oddala od matematyki (Malinowski [1997]).

się tych brakujących informacji, gdyż często są one całkowicie niedostępne. Można więc rozumować tylko na podstawie niepełnego i niedoskonałego zbioru przesłanek. Innymi słowy, w logice niemonotonicznej występują założenia domyślne, które uznaje się za prawdziwe przy braku informacji, że są fałszywe (Lifschitz [1995b]). Wydaje się zresztą, że tak właśnie - zgodnie ze zdrowym rozsądkiem - rozumiemy.

Jak więc wygląda rozwiązanie problemu ramy odniesienia? W logice niemonotonicznej definiuje się własności domyślne sytuacji, nie zmieniające się pod wpływem określonych działań. Przeto *explicite* trzeba definiować tylko zmiany istotne. Weźmy dla przykładu (i ułatwienia) cyrkumskrypcję (*circumscription*) McCarthy'ego, czyli jeden z formalizmów rozumowań niemonotonicznych — cyrkumskrypcja nie jest logiką niemonotoniczną, lecz umożliwia uwzględnienie niemonotonicznych rozumowań w klasycznej logice pierwszego rzędu<sup>92</sup>. Intuicje, które kryją się za cyrkumskrypcją, są następujące: ludzie często uogólniają pewne spostrzeżenia na temat pewnych obiektów na obiekty należące do tej samej klasy. Nie jest to dedukcja, lecz raczej - zawodny sposób wnioskowania; często jednak jedyny, jakim dysponują. Warto więc sformalizować dochodzenie do takich – często intuicyjnie zasadnych – przypuszczeń.

Cyrkumskrypcja jest **regułą tworzenia przypuszczeń** stosowaną wraz z innymi regułami logiki pierwszego rzędu; przy czym jest ona przekształceniem syntaktycznym określonych zdań (ściślej: rodziną takich przekształceń). Rezultatem zastosowania cyrkumskrypcji do zbioru  $A$  faktów w bazie danych (aksjomatów) jest schemat zdaniowy, który stwierdza, że predykat  $P(x, \dots, z)$  jest spełniany tylko przez te  $n$ -tki, które muszą go spełniać ze względu na zdania należące do  $A$ . Dodanie kolejnych zdań do zbioru  $A$  mogłoby spowodować spełnienie  $P$  przez więcej  $n$ -tek, dlatego też cyrkumskrypcja nie jest monotoniczna. Wnioskami wyprowadzonymi przy użyciu cyrkumskrypcji są przypuszczenia, że  $A$  zawiera wszystkie istotne fakty, i że obiekty, których istnienie wynika z  $A$ , są wszystkimi istotnymi obiektami (McCarthy [1980]). Istotne jest tutaj to, że pod uwagę bierze się nie dowolne modele danego zbioru aksjomatów, lecz tylko - modele spełniające określone warunki minimalności.

Podstawową zasadą stosowania cyrkumskrypcji do rozwiązania problemu ramy odniesienia jest odpowiednie ograniczenie ekstensji predykatu (lub zbioru predykatów) przez podanie pewnych deklaracji na temat stosowania reguł cyrkumskrypcyjnych do

---

<sup>92</sup> Tu muszę abstrahować od ściśle formalnej strony zagadnienia, jako że jest ona dosyć złożona. Zainteresowanych odsyłam do najlepszego i zwięzłego opisu cyrkumskrypcji w podręcznikowym omówieniu — Lifschitz [1993]. Por. też McCarthy [1980], McCarthy [1986].

aksjomatów<sup>93</sup>. Wskazuje się mianowicie względy, pod którymi dany predykat jest nienormalny, tj. kiedy dany predykat nie może być prawdziwie orzekany. Na przykład ptaki latają, o ile nie są nienormalne pod jakimś względem:

$$\forall x(((ptak(x) \wedge \neg nienormalny(x)) \Rightarrow ptak(x))$$

Tego rodzaju formalizację można zastosować w rozumowaniach temporalnych:

$$\forall f \forall z \forall s (Zachodzi(f,s) \wedge \neg nienormalny(f, z, s)) \Rightarrow Zachodzi(f, Rezultat(z, s))$$

gdzie  $f$  jest funkcją zmieniającą wartość zgodnie z indeksem czasowym (tzw. fluentem),  $z$  - zdarzeniem, a  $s$  - sytuacją. Jeśli więc fluent  $f$  jest prawdziwy<sup>94</sup> w określonej sytuacji  $s$  i zachodzi zdarzenie  $z$ , które nie jest nienormalne względem tej sytuacji, to fluent  $f$  będzie nadal prawdziwy w sytuacji wynikłej z zajścia zdarzenia  $z$ . Cyrkumskrypcje mogą zatem posłużyć do rozwiązania problemu ramy odniesienia i pokrewnych trudności.

Warto jednak wskazać, że z oryginalnym rozwiązaniem McCarthy'ego wiąże się pewien kłopot. Otóż, nawet w minimalnych modelach formuły rachunku sytuacji z cyrkumskrypcją mogą być spełnione w sposób sprzeczny z intuicjami. W pewnych modelach aksjomaty cyrkumskrypcyjne (tj. aksjomaty, do których stosuje się reguły cyrkumskrypcji) trzeba będzie uznać za fałszywe w świetle zdrowego rozsądku. Klasycznym przykładem tego rodzaju trudności jest tzw. problem strzelaniny z Yale, opisany przez McDermotta<sup>95</sup>. Występuje ona zresztą, dlatego że rozumowanie niemonotoniczne dopuszcza niezgodne informacje jako przesłanki. Posługując się zdrowym rozsądkiem, rozumujemy po otrzymaniu niezgodnych informacji; czasem dochodzimy do określonych wniosków, czasem - nie. Niestety, ta własność rozumowania niemonotonicznego działa na jego zgubę, bo w pewnym momencie odchodzimy od zdroworozsądkowych intuicji.

Oto **problem strzelaniny z Yale** (*Yale shooting problem*). Wiadomo, że broń jest naładowana w chwili 1, i że w chwili 5 wypala w stronę Freda. Załadowanie broni powoduje, że jest załadowana, a wypalenie załadowanej broni w stronę danej osoby powoduje jej śmierć. Fluenty *żywy* i *załadowany* są prawdziwe, dopóki nie wydarzy się coś nienormalnego. Co można wywnioskować o stanie Freda w chwili 6? W rachunku sytuacji pytanie brzmi, czy następujący fluent jest prawdziwy ( $S0$  jest sytuacją wyjściową):

<sup>93</sup> Cyrkumskrypcja całej dziedziny, zwana także inferencją minimalną, podpada pod cyrkumskrypcję predykatów.

<sup>94</sup> W tym artykule będę mówił wyłącznie o fluentach dwuwartościowych, przyjmujących wartości prawda i fałsz.

<sup>95</sup> Właśnie z powodu odkrycia tego kontrprzykładu McDermott zmienił front i został proceduralistą. W opisie korzystam z jasnego streszczenia Morgenstern [1996].

*Zachodzi(żywy(Fred), Rezultat (Wypalenie, Rezultat (Czekanie, Rezultat (Czekanie, Rezultat (Załadowanie, S0))))))?*

Zdrowy rozsądek podpowiada, że ten fluent jest fałszywy, bo Fred zszedł. Ale modele są dwa. W modelu oczekiwanym najdłużej, prawdziwy jest fluent *naładowany*. Broń jest więc załadowana do chwili, gdy zostanie wypalona w stronę Freda i biedny Fred umiera. Wypalenie z broni w chwili 5 jest nienormalne względem własności bycia żywym, która do tej pory przysługiwała Fredowi. W modelu nieoczekiwanym najdłużej prawdziwy jest fluent *żywy*: Fred żyje po wypaleniu broni. Fluent *naładowany* nie zachowuje więc prawdziwości; oznacza to, że broń musiała zostać w jakiś sposób rozładowana. W pewnej sytuacji od chwili 2 do 5 fluent *Czekanie* jest nienormalny względem naładowania broni. Który z modeli wybrać? Przecież Fred albo żyje, albo nie.

Przy oryginalnym ujęciu cyrkumskrypcji nie można na to pytanie odpowiedzieć. Mamy niezgodne wnioski i nie wiadomo, który z nich wybrać. Proponowano różne rozwiązania tego dylematu; najczęściej opierały się one na chronologii wydarzeń lub na przyczynowości. Rozwiązanie chronologiczne: model oczekiwany jest zgodny z rozumowaniem od pierwszej chwili do ostatniej (nie dobiera się więc dodatkowych przesłanek, np. o rozładowaniu broni); rozumowanie w odwrotnym kierunku generuje model nieoczekiwany. Przyczynowe podejście daje lepsze rezultaty, ale nie radzi sobie z działaniami równoczesnymi i nieznanymi<sup>96</sup>. Nie rozwiązuje więc problemu ramy odniesienia. Morgenstern zaproponowała rozwiązanie oparte na zdroworozsądkowej zasadzie, że nieoczekiwane — wedle wiedzy działającego podmiotu — zdarzenia zachodzą najrzadziej (np. ufoludki nie startują w wyborach na Sejm; por. Morgenstern [1991]). Jej teoria działań oczekiwanych (*Motivated Action Theory, MAT*) jest w stanie uporać się z pewnymi trudnościami; niestety, polega na wprowadzeniu aksjomatów ramowych (lecz w niewielkiej liczbie). Warto jednak dodać, że to rozwiązanie można połączyć z teoriami rozumowania przyczynowego Lifschitza i jego zagnieżdżonymi teoriami anormalności (czyli teoriami dotyczącymi predykatu *nienormalny* w sensie podanym wyżej), które w sposób dosyć zadowalający rozwiązują problem ramy odniesienia. Lifschitz bowiem zdołał znaleźć metodę tworzenia eleganckich bloków aksjomatów ramowych, dzięki czemu można je automatycznie generować<sup>97</sup>. Niestety, nie ma tu miejsca, aby ją choćby w przybliżeniu przedstawić.

---

<sup>96</sup> Dokładniejsze omówienie trudności znaleźć można u Morgenstern [1996].

<sup>97</sup> Szczegóły techniczne można znaleźć w wielu pracach Lifschitza; m.in. w Lifschitz i Kartha [1995].



Wydaje się, że – obecnie - stan badań jest znacznie bardziej zaawansowany niż w na początku lat osiemdziesiątych, gdy powstały najśmielsze interpretacje filozoficzne problemu ramy odniesienia.

#### ***4. Tego nie można zostawić hakerom!***

Filozoficzne stanowiska w sporze o wagę problemu ramy odniesienia można podzielić zasadniczo na trzy rodzaje. Po pierwsze, stanowisko, zgodnie z którym istnienie problemu ramy odniesienia świadczy o tym, że reprezentacje mentalne nie mają charakteru zdaniowego. Po drugie, ujęcie problemu ramy odniesienia jako problemu redukowalnego do problemu indukcji. Po trzecie zaś, uznanie występowania problemu za charakterystyczny rys wszelkich systemów racjonalnych reprezentujących wiedzę. Dodatkową opcją jest też opinia, że istnienie problemu ramy odniesienia świadczy o holizmie przekonań podmiotu racjonalnego lub o niemożności formalizacji holistycznych przekonań w postaci zdaniowej (dotyczy to najczęściej stanowiska pierwszego i trzeciego).

Za przedstawicieli pierwszego stanowiska można uznać zarówno zwolenników koneksjonizmu (Churchland [1997], s. 254), jak i zagorzałych przeciwników klasycznej sztucznej inteligencji, do której należy zaliczyć logistyczne badania McCarthy'ego i logiki niemonotoniczne (oczywiście, można być i zwolennikiem sieci neuronowych, i przeciwnikiem logistycznej SI — takie postawy bynajmniej się nie wykluczają). Kto przy zdrowych zmysłach przypuszczałby, że Hubert Dreyfus, zagorzały przeciwnik „sztucznych inteligentów”, działających na podstawie ściśle sformułowanych reguł, będzie załamywał ręce z powodu problemu ramy odniesienia<sup>98</sup>?

Dreyfus, jak większość przeciwników logiki stosowanej, podkreśla, że inteligentne działanie nie polega na wykonywaniu reguł, lecz na umiejętnościach wykształczanych z czasem. Twierdzi, że umiejętności nie wypływają automatycznie z wiedzy. Nic to nowego — już Kant podkreślał wagę władzy sądenia; już Ryle opisywał wiedzę-jak w odróżnieniu od wiedzy-że. Istotnie, ludzie nie przechowują swoich umiejętności w pamięci w postaci zdaniowej<sup>99</sup>, lecz uczą się działać i wiedzą jak to robić, mając trudności z werbalizacją — jeździć na rowerze uczymy się jeżdżąc, a nie, zapamiętując zdania z podręcznika.

---

<sup>98</sup> A jeśli kto wątpi, niech przeczyta tekst Dreyfus i Dreyfus [1987].

<sup>99</sup> Ciężar dowodu, że jest inaczej, przerzucam na mojego oponenta: wystarczy wskazać choćby jedną osobę, która przechowuje np. swoją czynną kompetencję językową w postaci zdaniowej. Twierdząc, że takich osób nie ma.

Zarzuty Dreyfusa wobec logistycznego nurtu AI nie są, moim zdaniem, rozstrzygające. A po prawdzie, nawet tak lubiane przez Dreyfusa nieklasyczne rozwiązania koneksjonistyczne, oparte na sieciach neuronowych, realizowane są w językach programowania, w których formułuje się całkowicie jasne reguły działania. Możliwe więc, że wiedza-jak może być reprezentowana w wiedzy-że o postaci zdaniowej (choć niekoniecznie są to zdania oznajmujące). Nadto w rachunku sytuacji McCarthy i Hayes uwzględnili wiedzę-jak, analizując ją jako umiejętność (wyrażaną predykatem *potrafi*). Nawet jeśli przyznamy rację koneksjonistom i „dreyfusistom”, twierdząc, że umiejętności nie są zapamiętywane (przez ludzi) w postaci zdań oznajmujących w logice pierwszego rzędu, to nie podważa to wcale programu klasycznej sztucznej inteligencji. Idzie w nim wszak o reprezentację wiedzy-że, a nie, umiejętności; koneksjoniści muszą dopiero udowodnić, że wiedzy-że nie sposób reprezentować w postaci zdaniowej (dowód taki próbuje przeprowadzić Churchland[1997]). To jednak sprawa zupełnie odrębna wobec problemu ramy odniesienia; można więc podsumować, że dla przedstawicieli pierwszego stanowiska w sporze kłopoty McCarthy’ego stanowią często jedynie przykład ilustrujący tezę o miałości klasycznej sztucznej inteligencji.

Ze względu na niechęć do klasycznej sztucznej inteligencji także John Haugeland, którego do dogmatycznych przeciwników „sztucznych inteligentów” zaliczyć nie sposób, należy do reprezentantów stanowiska pierwszego w odmianie holistycznej: reprezentacja wiedzy w postaci zdaniowej jest według niego nieadekwatna, bo nie radzi sobie z holizmem; w reprezentacji zdaniowej istnieje podział na treści jawne i implikowane. Haugeland twierdzi, że sztuczna inteligencja dostarczyła epistemologii nową trudność (Haugeland [1987]). Przyznaje, że najlepszą dotychczas metodą jej rozwiązywania była strategia „śpiącego psa”: wszystko jest jak jest, dopóki nie ma powodu, żeby sądzić inaczej (Haugeland [1987], s. 86). Jako rozwiązanie alternatywne Haugeland proponuje reprezentację graficzną wiedzy. W modelach graficznych („obrazy myślowe”) wszystkie relacje są reprezentowane niejako wprost („komplicytnie” w jego terminologii), więc problem być może nie powstaje.

Z punktu widzenia kognitywisty logistycznego Haugeland nieco się pospieszył z oddaleniem reprezentacji zdaniowej jako nieużytecznej. Cały problem w tym, że reprezentacje graficzne (modele w różnej skali) nie poddają się tak łatwo formalizacji, więc wcale nie wiadomo, czy i one nie niosą ze sobą istotnych trudności formalnych. Przy tym Haugeland sprowadza, jak się zdaje, problem ramy odniesienia do problemu holizmu w rozumowaniu: gdzież w opisach technicznych znajdziemy wzmianki o informacjach

zawartych *implicite* i *explicite*<sup>100</sup>? Jak się zdaje, sądzi on, że problem ramy odniesienia wiąże się z trudnością przekształcenia zawartych w systemie implicytnych informacji o zmianie w informacje eksplicytne, czyli wyodrębnienia ze zbioru wszystkich zdań w systemie (uczynienia eksplicytnymi) tych wniosków (zdań implikowanych, czyli implicytnych informacji), które są istotne. Zapewne tak można opisywać to, z czym borykają się badacze sztucznej inteligencji. Wydaje mi się jednak, że taki opis nie jest ani w pełni adekwatny — nie mówi się, pod jakimi względami pewne informacje mają być istotne — ani precyzyjny (równie dobrze mógłby to być problem wyszukiwania istotnych, lecz nie zmieniających się w czasie informacji w bazie danych).

Uważam, że zarysowane wyżej pierwsze stanowisko jest raczej nie do utrzymania: opiera się na mętym opisie problemu ramy odniesienia; co więcej, argumenty za zmianą systemu reprezentacji zdaniowej na graficzną czy też koneksjonistyczną są zbyt wątle. Nie ma bowiem danych po temu, aby sądzić, że określone modele reprezentacji — graficzne, koneksjonistyczne czy inne niezdaniowe struktury danych — są doskonałe jako formalne modele reprezentacji wiedzy-że. Przypuszczalnie mogą być one wolne od pewnych trudności reprezentacji zdaniowej, kto jednak może dać głowę, że będą wolne od wszelkich wad? Optowanie za koneksjonizmem czy też - za „obrazami myślowymi” wydaje się być po prostu swoistym hura- optymizmem.

Przejdźmy do drugiego stanowiska: do redukowania problemu ramy odniesienia do problemu indukcji. Jego przedstawicielem jest m.in. Daniel Dennett<sup>101</sup>. Przy okazji Dennett rozważa niezwykle ciekawą kwestię stosunku fenomenologii umysłu do sztucznej inteligencji. Jego zdaniem w ramach badań nad sztuczną inteligencją założeniem jest modelowanie działania ludzkiego umysłu. Tymczasem formalizacje takie jak rachunek sytuacji czy cyrkumskrypcja nie znajdują odpowiednika w fenomenologii procesów poznawczych. Mamy do czynienia zatem z rozdźwiękiem między zaawansowanymi teoriami matematycznymi a tym, co inteligentni ludzie mówią o swoich działaniach<sup>102</sup>. Dennett

---

<sup>100</sup> Taki zarzut stawia Haugelandowi Eric Lormand, który, mimo że sam filozof, podważa filozoficzną istotność problemu ramy odniesienia. Por. Lormand [1996] Wcześniejsza wersja tego samego tekstu jest łatwiej dostępna (Lormand [1990]).

<sup>101</sup> Artykuł Dennetta pochodzi z 1984 roku. Obecnie, o czym świadczy jego komentarz przy przedruku w zbiorze *Brainchildren*, nie podtrzymuje w pełni zawartych w nim opinii na temat istoty problemu ramy odniesienia. Por. Dennett [1998].

<sup>102</sup> Fenomenologia w wydaniu Dennetta jest intersubiektywna (tzw. heterofenomenologia) i polega na badaniu sprawozdań podmiotów, nie zaś treści świadomości samego badacza. Por. Dennett [1991].

przyrównuje „wynałazki” formalne w kognitywistyce do wynalazku koła: koło jako takie nie występuje w przyrodzie, lecz dzięki niemu można poruszać się szybciej niż na nogach. Problem polega na tym, że jeśli będzie za dużo takich „kognitywnych kółek”, to sztuczna inteligencja przestanie pełnić rolę modelu umysłu ludzkiego.

Z drugiej jednak strony, w formalizacjach można odkryć, zdaniem Dennetta, istotne problemy filozoficzne. Takim problemem jest dla niego problem ramy odniesienia, który rozumie jako problem wyznaczenia istotnych konsekwencji ze zbioru wszystkich konsekwencji przyjętego zbioru przesłanek. Stara się zresztą sprowadzić ten problem do problemu predykcji, a następnie do problemu indukcji, komentując, iż może nie jest to duży postęp, ale przynajmniej wiadomo, o jaki problem chodzi. Oczywiście, przedstawiciele logistycznych badań nad sztuczną inteligencją zaprotestowali (Hayes [1987]; Lormand [1996]). Problem ramy odniesienia jest węższy i dotyczy, ich zdaniem, reprezentacji szybko zmieniającego się świata. Dennett obecnie nie upiera się jednak przy swoim.

Wskazane przez Dennetta niejako przy okazji zagadnienie „kognitywnych kółek” jest godne uwagi. Przecież motywacją nurtu logistycznego było zbudowanie adekwatnego modelu zdrowego rozsądku, opartego na fenomenologii (w sensie Dennettowskim: prowadzone badania kognitywne polegały często na zadawaniu pytań badanym, którzy mieli opowiedzieć, jak sami rozwiązują problemy, por. Newell i Simon [1998]); współczesne nurty koneksjonistyczne natomiast oparte są na modelu neurofizjologicznym. Clark Glymour, komentując tekst Dennetta (Glymour [1987]), przypomina, że wcale nie wiemy, czy ludzkie sposoby myślenia są najlepsze z możliwych. W końcu koło jest lepsze nawet od stu tysięcy par odnóży. Dlaczego więc sztuczna inteligencja ma koncentrować się na modelowaniu ludzkiego umysłu, skoro nawet nie wiadomo, jak on działa? Rzecz jasna, Glymourowi można odpowiedzieć: właśnie o to chodzi, że nie wiadomo. Modele psychologiczne obywające się bez formalizacji są znacznie mniej precyzyjne od technicznych opisów działania umysłu, a takie właśnie podają badacze sztucznej inteligencji. Należałoby więc liczyć na to, że zbudują precyzyjny matematycznie model umysłu ludzkiego. Nic jednak, moim zdaniem, nie wyklucza możliwości, że to właśnie odejście od modelowania pozwoli zbudować inteligencję sprawniejszą od ludzkiej.

Dennetowska redukcja wydaje się nieudana, choć towarzyszące jej refleksje są trafne. Problem indukcji, czyli problem znajdowania istotnych wniosków przez indukcję, prawdopodobnie też nie ma apriorycznych rozwiązań. Czy to jednak nie przesada, domagać się apriorycznych granic dla rozumowania, które z zasady nie ma wiele wspólnego z

apriorycznymi granicami? Wielu elementów dedukcyjnych nie znajdziemy w indukcji, a to właśnie dlatego, że nie jest ona dedukcją.

Czas na deser, czyli na opowieść o Fodorze, śpiących psach, muzyce sfer i lodówce (Fodor [1987]). Stanowisko Fodora, które zaliczam do trzeciej grupy filozoficznych eksplikacji problemu ramy odniesienia, wywołuje wściekłe ataki informatyków. Hayes (Hayes [1987]) wskazuje, że sam Fodor się gubi i przedstawia trzy interpretacje problemu ramy odniesienia w dziesięciostronicowym artykule. Po pierwsze, problem ramy odniesienia rozumie Fodor jako problem Hamleta z punktu widzenia inżyniera: kiedy przestać myśleć? Czyli, mówiąc słowami Fodora, jak zbudować urządzenie, „które jest racjonalne w tym sensie, że jego mechanizmy ustalania przekonań nie są zahermetyzowane” (Fodor [1987], s. 140) w odrębnych modułach? Fodor odpowiada: nie wiem, gdybym wiedział, byłbym sławny i bogaty. Twierdzi, że problem ramy odniesienia dotyczy samej istoty racjonalności, a pojawienie się tego problemu jest objawem racjonalnego przetwarzania informacji; jeśli występuje on w danym systemie, znaczy to, że system ten jest racjonalny przynajmniej pod tym względem, iż nie jest zahermetyzowany w oddzielnych modułach. Inteligentne procesy, zdaniem Fodora, nie są modularne, a to prowadzi do problemu ramy odniesienia.

Problem ramy odniesienia jest tak głęboko powiązany z istotą racjonalności, że jest jak muzyka sfer: nie słycać jej, bo jest wszędzie. Nie można twierdzić, że strategia „śpiącego psa” działała: przecież dotychczasowe „programy inteligentne” są niewiarygodnie głupie, dodaje Fodor. Uważa, że przyjęcie zasady inercji w myśleniu przesunęło problem ramy odniesienia z dziedziny ustalania przekonań w dziedzinę ontologii, a tu już czyhają kolejne niebezpieczeństwa. Należy zbudować bardzo bogatą reprezentację kanoniczną faktów, która musi pozostawać niezmienna przy wystąpieniu większości zdarzeń. Tymczasem kanoniczną notację<sup>103</sup> bardzo trudno stworzyć. Oto przykład Fodora. Filozof włącza i wyłącza lodówkę, a to powoduje nieskończenie wiele zmian w jego przekonaniach: jeśli włączy lodówkę, uważa, że warzywa się nie zepsują, piwo się schłodzi, rachunek za prąd będzie wyższy, w nocy coś będzie buczeć itd. Jeśli zdefiniować relacyjną własność cząstek fizycznych po włączeniu lodówki Fodora jako „lodówkowłacz”, to po włączeniu lodówki zmienia się stan każdej cząstki fizycznej we wszechświecie, bo wszystkie mają własność bycia lodówkowłaczem. A można stworzyć dowolnie wiele takich predykatów, jak „lodówkowłacz”; ich uwzględnienie w kanonicznej notacji prowadzi do powstania ogromnej liczby faktów, które się zmieniają przy każdym działaniu. Innymi słowy, bez dobrej reprezentacji kanonicznej niemożliwe jest

---

<sup>103</sup> Zdaje się, że Fodor zamiennie używa wyrażen „reprezentacja” i „notacja”.

zastosowanie strategii „śpiącego psa”. A takiej reprezentacji jeszcze nie ma, bo nie stworzono modelu zdroworozsądkowej ontologii świata; dlatego też nie ma rozwiązania problemu ramy odniesienia.

To jednak nie koniec wywodów Fodora. Twierdzi on przy tym, że problem istotności indukcji jest pod istotnym względem identyczny z problemem ramy odniesienia. Oddam mu głos:

Problem ramy odniesienia [...] jest problemem wnioskowania niededukcyjnego (*nondemonstrative inference*); zaś problem wnioskowania niededukcyjnego [...] jest problemem, jak działa umysł poznający. Przykro mi, że McDermott stracił cierpliwość dla filozofów; ale szczerze mówiąc, problem ramy odniesienia jest zbyt ważny, aby go zostawić hakerom (Fodor [1987], s. 148).

Na zarzut Hayesa, że zbyt szeroko potraktował problem ramy odniesienia, Fodor odpowiada, iż oczywiście można nazwać ten problem inaczej, ale jeśli nie nazwiemy raka rakiem, to i tak na niego umrzemy.

Rzadko kto w logistycznym nurcie AI negował, że problem ramy odniesienia wiąże się z reprezentacją świata. Rzecz w tym, że to właśnie bardzo swoista reprezentacja okazała się ułomna i szukano sposobów jej „załatania”. Z tego punktu widzenia Fodor ma rację mówiąc, że problem ramy odniesienia jest problemem reprezentacji; nie ma jednak racji, sądząc, że nie da się go rozwiązać nie rozwiązując przy okazji problemu adekwatności tej reprezentacji w ogóle. Nie od razu Rzym zbudowano. Może rzeczywiście informatycy powinni użyć innego modelu świata zamiast mechanistycznej metafizyki rodem z XVII w., aby lepiej reprezentować zmiany i dynamikę świata (por. Janlert [1987], s. 32). To jednak nie znaczy, że ontologię należy przedstawić od razu w ośmiu tomach w twardej oprawie i z indeksem rzeczowym. Fodor używa, moim zdaniem, nazwy problemu ramy odniesienia jak metonimii — po prostu ma na myśli wszelkie problemy związane z reprezentacją wiedzy przez system racjonalny w postaci zdaniowej. To jednak bardzo ogólne, żeby nie powiedzieć - przesadnie ogólne, ujęcie.

Fodorowski problem „kiedy przestać myśleć” ma inną naturę niż problem ramy odniesienia. Po pierwsze, nie sędzę, aby można było *a priori* odpowiedzieć na pytanie „kiedy przestać myśleć?”. Może jedyną odpowiedzią jest zasada ekonomii myślenia; moim zdaniem, jeśli już chcemy zadawać takie ogólne pytania, możemy zapytać, kiedy należy zacząć myśleć. W każdym razie pytanie to nie ma wiele wspólnego z tworzeniem modelu zmieniającego się świata. Fodor drapie tam, gdzie nie swędzi. Przy tym drapie - nie we własną nogę.

Czy nie można jednak jakoś bronić stanowiska drugiego i trzeciego? I czy za opcją holistyczną nie stoją jakieś dodatkowe argumenty?

Zenon Pylyshyn - we wstępie do wielokrotnie tu przywoływanego zbioru o dylemacie robota - pisał, że problem ramy odniesienia jest to bardziej techniczne ujęcie problemu holizmu w rozumowaniu (Pylyshyn [1987b], s. VIII-IX). Co więcej, podkreślał, że - być może - istotą problemu ramy odniesienia jest problem istotności wnioskowań. Jednym słowem, zgadzał się z większością filozofów.

Problemem holizmu w rozumowaniu nie będę się tutaj zajmował, gdyż nie dostrzegam żadnego związku, wbrew temu, co twierdzi Pylyshyn, między problemem ramy odniesienia a holizmem. Nie jestem także przekonany, czy rozumowania jakiegokolwiek podmiotu inteligentnego muszą być rzeczywiście holistyczne (dobre argumenty przeciwko holizmowi przedstawia m.in. Millikan [1984]).

Natomiast filozoficzne poszukiwanie niejako apriorycznego kryterium istotnych wnioskowań, kamienia probierczego, który pozwoliłby raz na zawsze odróżnić stany świata zmieniające się pod wpływem działań robota, jest być może skazane na porażkę. Dlatego też stanowiska drugie i trzecie, nie mówiąc o całkowicie nieprecyzyjnym stanowisku pierwszym, są moim zdaniem nie tylko błędne, lecz także - często nie na temat (błędne rozumienie istoty problemu).

Zdroworozsądkowa zasada „śpiącego psa” w rozumowaniach niemonotonicznych jest do tej pory najlepszym przybliżeniem poprawnego rozwiązania. Jest to jednak rozwiązanie typowo robocze; po prostu nie ma nic lepszego. Dla filozofa powstaje zadanie: zastanowić się, czy to nie kontekst każdorazowo decyduje o istotności wniosku rozumowania indukcyjnego<sup>104</sup>. I tak na przykład: zbiór istotnych w danej sytuacji warunków wstępnych danego działania jest zawsze zbiorem pewnych określonych cech, lecz istotność nie jest odrębną cechą. Mówiąc językiem współczesnej ontologii, istotność superwenuje na tych cechach, jest własnością pochodną wobec własności w danej sytuacji. Pytać, jak wyróżnić istotne warunki danego działania to, innymi słowy, nic innego jak zadawać pytanie typu: „Co jest kryterium występowania zdarzenia w czasoprzestrzeni?”, „Co jest kryterium prawdy?”, „Co jest kryterium dobra?”. Jak wiadomo, nie istnieją tego typu uniwersalne kryteria<sup>105</sup>. Takie

---

<sup>104</sup> To oczywiście niezbyt jasne sformułowanie; poprawna odpowiedź musi zawierać także pewne uściślenia. „Kontekst” rozumiem tak samo jak McCarthy [1989]; nie podaję definicji, jako że nawet McCarthy ogranicza się tylko do przykładów.

<sup>105</sup> Pokazali to odpowiednio Kant (krytyka dowodu ontologicznego), Tarski (definicja prawdy) i Moore (definicja dobra; może nie jest ona adekwatna, lecz jest jasne, że nie istnieje proste naturalistyczne kryterium dobra).

rozważania nie należą do kognitywistyki, lecz mogą stanowić wkład filozofa w jej rozwój. Twierdzą, że nie warto szukać apriorycznych, uniwersalnych kryteriów istotności pewnych zdań opisujących zmieniający się świat. Istotność tych zdań nie może być bowiem wyznaczona w sposób niezależny od znajomości całej sytuacji, której opis nigdy nie jest dany a priori.

### **5. *Ontologowie do komputerów, hakerzy do fenomenologii?***

Dotychczasowe filozoficzne ujęcia problemu ramy odniesienia u Haugelanda, Dennetta i Fodora nie wydają się szczególnie płodne. Filozofowie albo sprowadzali problem do innych, znanych trudności (np. problem istotnych wniosków indukcyjnych), albo proponowali całkowitą rewolucję: zmianę systemu reprezentacji lub stworzenie reprezentacji kanonicznej od nowa. Niekiedy rozkładali ręce, żałując, że to nie oni się wzbogacą i nie oni przejdą do historii jako twórcy rozwiązania problemu ramy odniesienia.

Czy mogło być inaczej? Czy rozsądne było oczekiwanie na to, że filozofowie przysiądą fałdów i zaproponują pewne konkretne rozwiązania? Na pierwszy rzut oka wydawać się może, że to właśnie rola inżyniera. Filozof jedynie może spekulować, stawiać nowe hipotezy i próbować wyciągać wnioski z dotychczasowych badań; filozoficzny doradca kognitywisty to po prostu sofista, ale nigdy - inżynier. Czy na pewno?

Oto pewna historia. Wczoraj dostałem pocztą elektroniczną wiadomość od znajomej, która informowała mnie, że razem z kolegami przystosowała własną ontologię formalną do wymagań zespołu badaczy sztucznej inteligencji. Częściowo automatycznie wygenerowana ontologia formalna posłużyła im do zbudowania pełnego systemu reprezentacji, który jest wprawdzie uproszczony, ale przynajmniej - pełniejszy od systemów znanych w programach symulujących inteligencję. Perspektywy są obiecujące, bo stworzono dwa równoległe systemy, ontologię zdroworozsądkową opartą na „fizyce naiwnej” Pata Hayesa i wersję ontologii ewentyzmu punktowego.

Czy w tej opowieści jest coś, co przeczyłoby roli filozofa? Kto wie, może ta fikcja kiedyś okaże się czymś programem badawczym. Tak czy siak, narzędzia sztucznej inteligencji są niezwykle interesujące dla ludzi uprawiających filozofię formalną. Z jednej strony, mariaż z kognitywistyką może usprawiedliwić owo dziwactwo ciągłej formalizacji, tak niemiłe wielu innym filozofom, w tym również - analitycznym. Z drugiej strony, powstające formalizacje można oceniać pod kątem przydatności w badaniach kognitywistycznych. Cykl generowania rozwiązań i testowania można, innymi słowy, wykorzystać w ontologii.



Kognitywiści nie mają wielkiej wprawy w budowaniu tego rodzaju formalizmów metafizyczno-epistemologiczno-semantycznych, jakie budują filozofowie analityczni. Fodor miał rację, pewnych kwestii nie można po prostu pozostawić hakerom. Całe szczęście, Janlert ukuł już odpowiedni termin dla tego rodzaju filozofii: ontologia eksperymentalna (Janlert [1987], s. 35). Tyleż eksperymentalna, co inżynieryjna, dodam od siebie. Oczywiście, o epistemologii eksperymentalnej, rozumianej jako formalizacja wiedzy dla potrzeb sztucznej inteligencji, też nie wolno zapominać.

Warto przy okazji wspomnieć o zarzucie Dennetta. Badania nad sztuczną inteligencją należy jednak ściślej połączyć z badaniami fenomenologicznymi w swoistym sensie: samorozumienie inteligentnych ludzi powinno jakoś współgrać z tworzonymi systemami. To oczywiście odrębna dziedzina: sztuczna inteligencja fenomenologiczna, jeśli wolno zaproponować nową nazwę.

## **6. Jak wyznaczyć granicę?**

Być może próba ogólnego wyznaczenia granicy między kognitywistyką a filozofią jest skazana na niepowodzenie. Może jest to tak samo trudny problem, jak znalezienie istotnych wniosków z rozumowania indukcyjnego. Zwolennik Quine'a zapewne rzuci uwagę, że trudność taka wypływa z ogólnej niemożności rozgraniczenia pojęć empirycznych i teoretycznych i z zamazania granic między filozofią a naukami przyrodniczymi.

Sytuacja w kognitywistyce jest jednak nieco bardziej złożona. Po pierwsze, badania nad sztuczną inteligencją należą nie tylko do nauk teoretycznych, ale także - do inżynierii. Co więcej, podpowiedzi filozofów, jeśli mają być przydatne w badaniach, też będą należeć do inżynierii. Po drugie zaś, niemałe filozoficzne znaczenie mają wyniki uzyskiwane w naukach kognitywnych. Modele rozumowania zdroworozsądkowego, sztuczny umysł, konieczne i wystarczające warunki poznania to typowe tematy filozofii. Więcej nawet, gdyby pewne przedsięwzięcia informatyków się powiodły, dowodziłoby to pewnych tez metafizycznych.

Zrealizowanie umysłu na różnych komputerach mogłoby stanowić ogromne uprawdopodobnienie funkcjonalistycznej tezy o wielorakiej realizacji stanów mentalnych na przedmiotach fizycznych. Tymczasem samą tezę uznajemy za filozoficzną. Na jakiej podstawie? Czy w naturalistycznej filozofii jedynym kryterium podziału jest inne przyporządkowanie instytucjonalne<sup>106</sup>? Wydaje się, że nie.

---

<sup>106</sup> To nie jest dobre kryterium, bo np. Dennett kieruje ośrodkiem badań kognitywistycznych na uniwersytecie w Tufts.

Ważną, jak sędzę, cechą współczesnego naturalizmu w filozofii umysłu jest odrzucenie w nim dogmatyzmu metafizycznego. Przyjmuje się w nim, co prawda, założenia naturalizmu jako takiego, ale są one poniekąd i tak roboczymi założeniami naukowca-przyrodnika. Warto przypomnieć, czym różni się metafizyka dogmatyczna od krytycznej. Chodzi mi oczywiście o rozróżnienie Kantowskie. Po prostu krytyczna metafizyka opiera się na usprawiedliwieniu zwanym przez Kanta „dedukcją transcendentálną”. W największym uproszczeniu: argument transcendentálny jest swego rodzaju wnioskowaniem redukcyjnym, czyli dobieraniem pewnych racji do znanego wniosku. Ów wniosek musi być powszechnie uznany i stosowany, np. we współczesnej nauce empirycznej (tak jest np. w *Krytyce czystego rozumu*, gdzie wnioskiem tym jest przede wszystkim teza „wykonujemy doświadczenie empiryczne”) i jest przyjmowany jak coś oczywistego. Warto podkreślić, że odrzuca się w ten sposób wszelki uprzywilejowany dostęp epistemiczny, przysługujący rzekomo filozofowi. Punkt wyjścia filozofowania, ów tak ważny w systemach metafizycznych punkt archimedesowy, nie jest tutaj niczym niezwykłym. To po prostu powszechnie znany fakt. Podejście takie oczywiście nie czyni z Kanta naturalisty, ale z naturalistą ma Kant wiele wspólnego. Zmiana poglądów na temat niezmiennych cech doświadczenia, np. odkrycie innych systemów geometrycznych (jedna geometria przestała być oczywistością dostępną powszechnie i intersubiektywnie), powoduje, że system metafizyczny traci swe usprawiedliwienie. Nauka może zmienić stopień uprawdopodobnienia metafizyki, mimo że ta uzasadniana jest niezależnie od doświadczeń naukowych.

We współczesnym naturalizmie relacja między filozofią a nauką empiryczną wygląda podobnie. Jeśli się zmieniają założenia robocze, stosowane w praktyce naukowej (np. gdyby najlepsze predykcje uzyskiwano na całym świecie dzięki radom jasnowidza), to uwzględnić je będzie musiała cała teoria poznania. Jednak sprawa gmatwa się w związku z inżynieryjnym aspektem kognitywistyki i filozofii. Naturalny dobór lepszych rozwiązań przez ich generowanie i testowanie (rzecz znana także z ewolucji biologicznej) stanowi swoistą cechę wspólną obu tych dziedzin.

Granica przebiega zatem tak. Kognitywiści tworzą modele umysłu poznającego; budują teorie empiryczne lub prototypy funkcjonalne. Filozofowie szukają tych niewielu elementów, które w takich badaniach uznaje się za oczywiste i inwariantne. Niemniej, przy braku zgody wśród naukowców i niewielkiej ilości danych, skazani są na stawianie własnych hipotez. Mogą zatem brać udział w samych badaniach, sugerując nowe nurty badań. Zadają więcej pytań, niż znają odpowiedzi. Jednocześnie — aby jednak udzielić jakichś odpowiedzi — korzystają oni z wyników nauk o poznaniu. Logiczna analiza wiedzy zdroworoządkowej -

przy zderzeniu z fenomenologią - może stać się punktem wyjścia dla wielu rozważań filozoficznych, zwłaszcza w nurcie formalnym. Do roboty!

## BIBLIOGRAFIA

CHURCHLAND, P. M.

[1997] *On the Nature of Theories: A Neurocomputational Perspective*, w: Haugeland [1997].

DENNETT D.C.

[1991] *Consciousness Explained*, Little, Brown, Boston 1991.

[1995] *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Simon and Schuster, New York.

[1998] *Cognitive Wheels: The Frame Problem of AI* w: Dennett D.C., *Brainchildren*, MIT Press, s. 181-205.

DREYFUS H. L., DREYFUS S. E.

[1987] *How to Stop Worrying about the Frame Problem Even though It's Computationally Insoluble*, w: Pylyshyn [1987].

FIKES R., NILSON N.,

[1971], *STRIPS: A new approach to the application of theorem proving to problem solving*, *Artificial Intelligence*, 3, s. 189-208.

FODOR J.

[1987] *Modules, Frames, Fridgeons, Sleeping Dogs, and the Music of the Spheres*, w: Pylyshyn [1987], s. 139-149.

FUHRMANN A.

[1998] *Nonmonotonic Logic*, *Routledge Encyclopedia of Philosophy*, Vol. 7, s. 30-35, Routledge, London.

GABBAY D.M., HOGGER C.J., ROBINSON J.A.

(red.) [1993] *The Handbook of Logic in AI and Logic Programming*, Oxford University Press.

GLYMOUR C.

[1987] *Android Epistemology and the Frame Problem: Comments on Dennett's "Cognitive Wheels"*, w: Pylyshyn [1987], s. 65-75.

HANKS S., MCDERMOTT D.

[1986] *Default reasoning, nonmonotonic logic, and the frame problem* w: *Proceedings of the 5th National Conference on Artificial Intelligence*, MIT Press, Cambridge, Mass., ss. 328-333.

HAUGELAND J.

[1987] *Overview of the Frame Problem*, w: Pylyshyn [1987].  
[1997] (red.), *Mind Design II*, MIT Press, Cambridge, Mass.

HAYES P.

[1987], *What the frame problem is and isn't*, w: Pylyshyn [1987].

HAYES P., MCCARTHY J.

[1969] *Some philosophical problems from the standpoint of artificial intelligence*, w: Meltzer i Michie [1969].

HENRICH D.

[2002] *Struktura dowodowa dedukcji transcendentnej Kanta*, Przegląd Filozoficzno-Literacki 3/2002 (w przygotowaniu)

JANLERT L.-E.

[1987] *Modeling Change — The Frame Problem*, w: Pylyshyn [1987].

LENTING, J. H. J., BRASPENNING P. J.

[1993] *The Frame Problem from an Engineering Perspective*, *International Journal of Software Engineering and Knowledge Engineering*, 3, nr 2 (1993), s. 257-285.

LIFSCHITZ V.

[1993] *Circumscription* w: Gabbay, Hogger, Robinson [1993], tom 3, s. 298-352.

[1995a], *Nested abnormality theories*, *Artificial Intelligence*, 74, s. 351-365.

[1995b] *Logic of Common Sense*, *ACM Computing Surveys*, 27, 3.

LIFSCHITZ V., KARTHA G. N.

[1995] *A Simple Formalization of Actions Using Circumscription* w: *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, Chris Mellish(red.), Morgan Kaufmann, San Francisco, s. 1970-1975.

LORMAND E.

[1990] *The Holorobophe's Dilemma*, *Synthèse*, 82, s. 353-374.

[1996] *The Holorobophe's Dilemma*, w Pylyshyn, Ford [1996].

MALINOWSKI J.

[1997] *Logiki niemonotoniczne*, *Przegląd Filozoficzny – Nowa Seria*, 1997, R. VI, Nr 1 (21), s. 31-53.

MCCARTHY J.

[1977], *Epistemological Problems of Artificial Intelligence*, *Proceedings of the Fifth Joint Conference on Artificial Intelligence*, MIT Press, Cambridge, Mass.

[1980] *Circumscription — a form of non-monotonic reasoning*, *Artificial Intelligence*, 13, s. 27-39.

[1986] *Applications of circumscription to formalizing common sense knowledge*, *Artificial Intelligence*, 26, 3, s. 89-116.

[1989] *Artificial Intelligence, Logic, and Formalizing Common Sense* w: *Philosophical Logic and Artificial Intelligence*, Richmond Thomason (red.), Kluwer Academic, Dordrecht.

MCDERMOTT D.

[1987] *We've Been Framed: Or, Why AI Is Innocent of the Frame Problem*, w: Pylyshyn [1987].

MELTZER B., MICHIE D.

(red.) [1969], *Machine Intelligence 4*, Edinburgh University Press, Edinburgh.

MILLIKAN R.G.

[1984] *Language, Thought, and Other Biological Categories*. New Foundations for Realism, MIT Press, Cambridge, Mass.

MORGENSTERN L.

[1991] *Knowledge and the Frame Problem*, International Journal of Expert Systems, 3(4).

[1996] *The Problem with Solutions to the Frame Problem* w: Pylyshyn, Ford [1996].

NEWELL A., SIMON H.,

[1997] *Computer Science as Empirical Enquiry: Symbols and Search* w: Haugeland [1997].

PYLYSHYN Z.

(red.) [1987] *The Robot's Dilemma. Frame Problem in Artificial Intelligence*, Ablex, Norwood.

[1987b] *Preface* w: Pylyshyn [1987].

PYLYSHYN Z., FORD K.,

(red.) [1996] *The Robot's Dilemma Revisited: The Frame Problem in AI*, Ablex, Norwood.

SHANAHAN M.

[1997] *Solving the Frame Problem*, MIT Press, Cambridge, Mass.

REITER R.

[1980] *A logic for default reasoning*, Artificial Intelligence, 13, s. 81-132.

# **FILOZOFIA A NAUKI SPOŁECZNE**

## **Czy współcześnie można być etnologicznym ewolucjonistą?**

Ewolucjonizm sprawia wiele kłopotu etnologom. To właśnie ta teoria zapoczątkowała naukową etnologię, a w odniesieniu do przedstawicieli tego nurtu używa się często określenia „ojcowie założyciele”. Określenie to podszyte jest jednak pewną dozą ironii, ponieważ powszechnie wypowiada się opinię, że: „dziś ewolucjonizm jest już historią etnografii i nikt nie mógłby serio uprawiać naszej nauki w oparciu o hipotezy, jakimi się posługiwali dziewiętnastowieczni ewolucjoniści” (Waligórski [1973] s. 195). Nie byłoby nic dziwnego w tym, że fundatorzy dyscypliny naukowej odchodzą z czasem do muzeum, niekiedy nawet do lamusa, lecz - choć odesłano tam pewnych badaczy - mniej jasno rysuje się stosunek wobec ich teorii. Większą część terminów i pojęć używanych przez etnologów wypracowali właśnie ewolucjoniści, co nietrudno sprawdzić przeglądając etnologiczne podręczniki, słowniki, czy encyklopedie. Choć tę teorię odsądzono od czci i wiary, terminy, którym nadaje ona sens, i które zawdzięczają jej moc eksplanacyjną, zachowano w niezmienionej postaci. Nie będę rozwijał tu problemu eklektyzmu pojęciowego i metodologicznego, w który popada etnologia, co z pewnym zdziwieniem konstatują nawet sami etnologowie. Chcę natomiast wskazać, że zarzucenie koncepcji ewolucyjnego rozwoju kultury, jako rzekomo nienaukowego, może okazać się zbyt pochopne. Przemawia za tym na przykład fakt, że nadal używane pojęcia są po prostu efektywne i – mimo odrzucenia teorii, która je uprawomocniała i nadawała im pełny sens – nadal trafnie opisują przynajmniej część rzeczywistości. Sądzę, że można podać argumenty wskazujące, że w pewnej interpretacji ewolucjonizm może być wciąż atrakcyjną propozycją myślenia o kulturze i kulturach.

Moim zamiarem w niniejszym artykule jest próba odpowiedzi na pytanie: czy współcześnie można być ewolucjonistą? Najpierw przedstawię w skrócie „podręcznikową” wersję ewolucjonizmu abstrahując od poglądów poszczególnych badaczy. Następnie zreferuję najważniejsze zarzuty, jakie w ogólności postawiono temu programowi badawczemu oraz zastanowię się nad ich trafnością i możliwością ich uwzględnienia przez ewolucjonizm. Jeżeli zarzuty stawiane teorii ewolucjonistycznej mogą zostać przez nią uwzględnione bez naruszenia jej spójności, to może ona stanowić teorię konkurencyjną wobec innych -



powszechnie akceptowanych, takich jak funkcjonalizm czy strukturalizm. Aktualne wówczas staje się pytanie o kryteria wyboru między konkurencyjnymi teoriami. Rozpatrzenie tego zagadnienia będzie stanowić zakończenie niniejszych rozważań.

### **1. Co to jest ewolucjonizm w etnologii?**

Mianem ewolucjonizmu w etnologii określa się najogólniej koncepcję uznającą za „podstawowy paradygmat teorii kultury koncepcję ewolucji kulturowej” (Posern-Zieliński [1987b] s. 116). Przedmiotem zainteresowania tak pojętego ewolucjonizmu może być kultura jako całość, kultury poszczególnych grup ludzkich, czy instytucje wchodzące w ich skład ujmowane pod kątem zmienności, co do której postuluje się istnienie pewnych regularności. W takim ujęciu ewolucjonizm jest nastawieniem badawczym, uwrażliwieniem badacza na pewne zagadnienia i nie musi wiązać się z określonymi założeniami teoretycznymi i metodą badawczą. Przedmiotem mojego zainteresowania będzie natomiast teoria ewolucyjnego rozwoju kultury, która wywarła decydujący wpływ na ukonstytuowanie się etnologii jako nauki. Okres, na który przypadła dominacja ewolucjonizmu w etnologii, to lata 1860-1890 (Waligórski [1973] s. 141). Dodajmy, że jest to zarazem czas kształtowania się instytucjonalnego zaplecza „nauki o ludach”.

Ewolucjonizm w tym znaczeniu przyjmuje określone założenia teoretyczne. Syntetyczne opracowania dotyczące tego kierunku z reguły wymieniają podobny zestaw:

- uznanie istnienia obiektywnych praw rządzących rzeczywistością kulturową,
- przekonanie o jedności psychicznej gatunku ludzkiego,
- założenie racjonalności działań ludzkich,
- uznanie ewolucji kulturowej za podstawową determinantę rzeczywistości kulturowej,
- uznanie postępu za główny czynnik wyznaczający kierunek ewolucji,
- uznanie za cel etnologii rekonstrukcję przebiegu dziejów kultury przez określenie następstw jej form (por. Posern-Zieliński 1987b, s. 117).

Pierwsze z wymienionych założeń często było poddawane w wątpliwość, choć przyjęcie istnienia obiektywnych praw rządzących rzeczywistością kulturową wydaje się warunkiem sine qua non uprawiania jakiegokolwiek nauki o kulturze. Jednak w okresie tzw. przełomu antypozytywistycznego pojawiły się propozycje metodologiczne, zgodnie z którymi etnologia, podobnie jak inne nauki o społeczeństwie, należeć miała do Geisteswissenschaften. Ewolucjonizm natomiast traktowano jako naukę przyrodniczą i lokowano w dziale nauk formułujących prawa. Etnologowie chętnie odwołują się do wspomnianych propozycji. Nietrudno jednak zauważyć, że rozróżnienie nauk idiograficznych i nomotetycznych jest typologią pozwalającą uporządkować nauki według roli, jaką odgrywają w nich: opis i

formułowanie praw. Trudno zresztą wyobrazić sobie naukę – o ile w ogóle byłaby to nauka – posługującą się tylko opisem. Spośród nauk o tak zwanych kulturach pierwotnych w największym stopniu opisem posługuje się etnografia, którą nie uważa się za odrębną naukę, lecz za działalność przygotowawczą do refleksji naukowej. „Odpowiada ona pierwszym fazom pracy badawczej: obserwacji i opisowi, pracy w terenie (field-work)”. Natomiast „etnologia jest pierwszym krokiem ku syntezie” i „obejmuje etnografię, jako swój pierwszy krok i stanowi jej przedłużenie” (Lévi-Strauss [1970] s. 454).

Należy wskazać nadto na nieporozumienie rozpowszechnione w obiegowych opiniach dotyczących ewolucjonizmu, spotykane także wśród etnologów. Omawiany kierunek refleksji nad kulturą bywa uważany za spadkobiercę i kontynuatora dziewiętnastowiecznych nauk biologicznych. Przeczą temu wypowiedzi samych ewolucjonistów (por. Tylor [1896] t. 1, s. 13). Jan Lutyński w swojej pracy poświęconej ewolucjonizmowi wymienia trzy najbardziej rozpowszechnione błędy dotyczące poglądów ewolucjonistów:

„1 – uznanie ewolucjonizmu w etnologii za rezultat przeniesienia do tej gałęzi wiedzy założeń darwinowskiej teorii rozwoju lub nawet innych biologicznych czy w ogóle przyrodniczych teorii i koncepcji, 2 – wyprowadzenie stanowiska Tylora i Morgana z oddziaływania tych teorii i w związku z tym stawianie ewolucjonizmu w etnologii na równi z takimi naturalistycznymi kierunkami jak antropologiczny czy geograficzny, 3 – identyfikacja ewolucjonizmu w etnologii z ewolucjonizmem Spencera” (Lutyński [1956] s. 147).

Według ewolucjonistów społeczeństwo nie jest wytworem sił przyrody, a kultura - po prostu kolejnym stopniem ewolucji biologicznej. Człowiek jest w przeciwieństwie do pozostałych organizmów biologicznych twórcą kultury, czyli „środowiska”, w którym żyje (Tylor [1923] s. 121-123, 184). Edward Burnett Tylor ujął to dobitnie: „istnieje niezmierna różnica między najbardziej pierwotnym człowiekiem i najwyższym ze zwierząt” (cyt. za Lutyński [1956] s. 155). Kulturę będącą podstawowym przedmiotem badania etnologii, ewolucjoniści definiowali jako „zdolności i przyzwyczajenia zdobyte przez człowieka jako członka społeczeństwa” (Tylor [1896] t. 1, s. 15). „Naturalizm” ewolucjonistów odwołuje się do pojęcia natury ludzkiej, nawiązując do antropologicznych wątków myśli filozoficznej XVIII wieku, a nie współczesnych mu teorii biologicznych. Był to zresztą powód licznych krytyk pod ich adresem w okresie powszechnego panowania myśli pozytywistycznej.

Jedność natury ludzkiej upatrywali ewolucjoniści w identyczności właściwości psychicznych ludzi. Wszyscy ludzie posiadają identyczne dyspozycje psychiczne, takie jak zdolność obserwowania zjawisk, kojarzenia wrażeń lub wyobrażeń, gromadzenia

doświadczeń. Należy to do ich natury, a poza tym jest warunkiem biologicznego przetrwania i kontynuacji gatunku. Struktura ludzkich władz psychicznych posiada uwarunkowanie biologiczne – Henry Lewis Morgan uznawał, że procesy psychiczne są „istotnym przymiotem mózgu samego” (Morgan [1887] s. 64), to one sprawiają, iż obserwujemy „powszechne podobieństwo natury ludzkiej” (Tylor [1896] t. 1, s. 19). Kulturotwórcze działanie bezpośrednio zależy od własności psychicznych biorących w nim udział ludzi. Jest to niezmienny czynnik warunkujący kulturę twórczość człowieka, powodujący, że w jednakowych warunkach ludzie reagują lub działają jednakowo. Czynnikiem zmiennym są warunki, czyli środowisko naturalne. Przy czym chodzi o środowisko nie w stanie „czystym”, lecz takie, w którym ludzie żyją i, które swoją działalnością przekształcają czy współkształtują.

Natura ludzka nie daje się zdaniem ewolucjonistów sprowadzić jedynie do wrodzonych zdolności lub dyspozycji psychicznych. Wrodzone zdolności lub dyspozycje psychiczne, uwarunkowane biologicznie, stanowią by mogły *differentia specifica* gatunku ludzkiego w świecie przyrody. Ewolucjoniści jednak uważali ludzi za istoty zeń wyemancypowane. Naturalizm w stylu oświeceniowym domagał się jednocześnie przyjęcia założenia pewnej swoistej i nieredukowalnej własności stanowiącej istotę człowieczeństwa. Dlatego zakładano, że istnieją „nieliczne zarodki pierwotne myśli, powstałe w skutek potrzeb pierwotnych ludzkich” (Morgan [1887] s. 278). Skoro wszyscy ludzie posiadają w zasadzie takie same zdolności intelektualne, dzięki którym mogą sprawnie działać w świecie, to zważywszy na różnorodność warunków, w jakich żyją ludzie, przynajmniej część z nich musi być niezależna od indywidualnego i społecznego doświadczenia i stanowić ich podstawę i warunek możliwości. Morgan i Tylor zaliczali do nich zdolność rozumowania według podobnych zasad, określaną jako „logika umysłu ludzkiego” czy „zasada inteligencji”, na przykład ujmowanie zdarzeń w kategoriach przyczyny i skutku (por. Morgan [1887] s. 17, 64, 606; Tylor [1896] t. 1, s. 26; [1923] s. 342-347).

W przytoczonych wyżej poglądach odnaleźć można w uproszczeniu następującą wizję rozwoju kultury: dzięki – zawsze i wszędzie takim samym, bo biologicznie uwarunkowanym – władzom psychicznym ludzie wytwarzają elementy kultury. Efekt twórczości kulturowej ludzi zależy przy tym od warunków, w jakich działają: podobieństwo warunków daje efekt podobieństw kulturowych, ich rozbieżność – różnic. Ponieważ mówi się o ludziach, zakłada się, że ich działania niezależnie od czasu i miejsca, w którym żyją oraz zgromadzonych indywidualnie lub społecznie doświadczeń, mają charakter racjonalny.

Podkreślić należy, że nie chodzi tu o jakąś szczególnie wyrafinowaną umiejętność myślenia. Racjonalność rozumiana jest w sensie potocznym, jako zdolność rozpoznawania napotkanych problemów i radzenia sobie z nimi. Ludzie nieustannie gromadzą zdobytą wiedzę o świecie, gdyż należy to do ich natury, a poza tym jest po prostu warunkiem przetrwania. Warunki, w jakich żyją i działają ludzie, zmieniają się przez wpływ czynników naturalnych i na skutek ludzkiej działalności. Zmiany wywołane przez działalność człowieka zależą od jego wiedzy, zatem można przyjąć, że mają również charakter kumulatywny (przynajmniej do pewnego stopnia). Ciągłość i kumulatywność zmian zachodzących w kulturze decydują o tym, że uznaje się je za ewolucyjne.

Współczesne słowniki etnologiczne określają ewolucję jako „serię procesów dynamicznych zmian kulturowych o uporządkowanej naturze, zachodzących w długich przedziałach czasowych” (Posern-Zieliński [1987a] s. 113). Czynniki, które wywołują wspomniane zmiany, dzieli się na endogenne, czyli działające w zmieniającej się kulturze (na przykład: innowacja, modyfikacja, adaptacja) oraz egzogenne, czyli działające na kulturę z zewnątrz (na przykład: dyfuzja, kontakt kulturowy). Przyjmuje się, że istotą ewolucji kulturowej jest „ciągły, w dużej mierze kumulatywny i ukierunkowany rozwój, cechujący się zmianami jakościowymi, w wyniku których zjawiska kulturowe ulegają transformacjom zwiększając stopniowo swoją złożoność, stopień wewnętrznej integracji oraz udoskonalając funkcjonalną efektywność, co w konsekwencji prowadzi do coraz wyższego poziomu zaawansowania cywilizacyjnego” (Posern-Zieliński [1987a] s. 113).

Rozwój kultur ma charakter postępowy: „dążnością człowieczeństwa podczas jego długiego trwania było przechodzenie od stanu dzikości do cywilizacji” (Tylor [1896] t. 1, s. 39). Dokonuje się to przez przechodzenie kolejnych stadiów rozwoju kulturowego. Wyższość lub niższość danego stadium, jego postępowość lub zacofanie, można ocenić przy pomocy kilku kryteriów. Według kryterium ilości i zróżnicowania elementów kultury, o zaawansowaniu etapu rozwoju świadczy liczba elementów, takich jak: narzędzia, techniki, umiejętności i instytucje społeczne funkcjonujące w danej kulturze. Według kryterium doskonałości elementów, o wyższości danego stadium świadczy to, że elementy kultury są „doskonalej uporządkowane, czyli lepiej uorganizowane do swoich celów”, lepiej odpowiadają „potrzebom społeczeństwa” (Morgan [1887] s. II; Tylor [1896] t. 1, s. 68; [1923] s. 15). O doskonałości szczebla rozwoju stanowi również „rozległość wiadomości”, na przykład „znajomość praw fizycznych rządzących światem i możliwość stosowania sił przyrody do celów praktycznych” (Tylor [1896] t. 1, s. 35). Ocenie pod tym kątem, podlegać miały również zwyczaje, wzory moralne i wierzenia, a o ich wyższości rozstrzygać miało

podobieństwo do analogicznych elementów kultury, której reprezentantami byli ewolucjoniści (por. Tylor [1896] t. 1, s. 35, 38; Morgan [1887] s. 37).

Według ewolucjonistów poszczególne szczeble w rozwoju kultury można oceniać według względnej doskonałości. Wyższe szczeble są przy tym logicznie zależne od niższych – aby opisać stan kultury na wyższym szczeblu rozwoju, trzeba założyć szczebel niższy: „zanim człowiek osiągnął stanu cywilizacji, musiał koniecznie posiadać wszystkie jej pierwiastki” (Morgan [1887] s. 29). Zarazem poszczególne stadia kultury zależne są od siebie w realnym następstwie, co uwarunkowane jest między innymi kumulatywnym rozwojem wiedzy – „żaden szczebel cywilizacji nie powstał samodzielnie, lecz wyrósł, czyli rozwinął się, z poprzednich” (Tylor [1923] s. 20). Ewolucjoniści wprost stwierdzali, że jedno z zadań etnologii upatrują w formułowaniu praw historycznych (Tylor [1896] t. 1, s. 16). Kultury ludów „dzikich”, stanowiących główny przedmiot zainteresowania etnologii, zawierają głównie elementy właściwe niższemu, a tym samym - dawniejszemu fazom rozwoju kultury. „Był terażniejszy dzikich – pisał John Lubbock – wiernie przedstawia różne fazy, które w swym postępowym rozwoju przechodził cały ród ludzki” (Lubbock [1873] t. 3, s. 162).

Dzięki przyjęciu wymienionych założeń etnologia mogła przeprowadzić „rekonstrukcję przebiegu dziejów kultury przez określenie następstwa [jej – uzupełnienie SS] form” (Posern-Zieliński [1987b] s. 117). Dokonywano tego przez zestawianie schematów obrazujących kolejne stadia rozwoju kultury<sup>107</sup>. W każdym stadium dziejów kultury pojawiają się właściwe dla niego elementy kultury, takie jak: typy narzędzi, gospodarowania, osadnictwa, organizacji społecznej, form pokrewieństwa i wierzeń. Możliwe jest także pojawienie się obcych danemu stadium elementów kultury w postaci tzw. przeżytków, czyli takich elementów kultury, które „niegdyś miały cel praktyczny lub przynajmniej obrzędowy w miejscu i czasie swojego powstania, a stały się obecnie niedorzeczne wskutek tego, że przeniesione zostały do nowego społeczeństwa, które zatraciło ich sens właściwy” (Tylor [1896] t. 1, s. 290; por. Posern-Zieliński [1987d]). Schematy rozwoju mają wymiar uniwersalny i stosują się do wszystkich bez wyjątku kultur.

## **2. Co zarzucano ewolucjonizmowi?**

„Jeśli zapytamy, dlaczego ewolucjonizm został przez tak przeważającą większość uczonych zajmujących się człowiekiem odrzucony, odpowiedź nasza musi być sformułowana nie tylko w terminach metodologii. Musi ona wziąć w pełni pod uwagę zmianę klimatu opinii w czasie, kiedy rozwinęła się opozycja. Gdy dla Europy epoka

---

<sup>107</sup> Przykłady rekonstrukcji niektórych schematów rozwojowych L. H. Morgana i E. B. Tylora podaje w swojej pracy Jan Lutyński (Lutyński [1956] s. 176-188).

wiktoriańska w coraz większym stopniu stawała się przeszłością i gdy w Ameryce granica sięgnęła Pacyfiku, pojawiły się wątpliwości w słuszność ogólnej koncepcji postępu, która dominowała w myśli społecznej w okresie panowania ewolucjonizmu” (Herskovits [1956] s. 473).

Takie sformułowanie postawy wobec teorii ewolucjonistycznych bardzo często pojawia się w wypowiedziach dwudziestowiecznych etnologów. Trudno jednak uznać, że formułuje się w nim ważki zarzut przeciwko omawianej teorii. Po pierwsze: jest to konstatacja należąca raczej do socjologii wiedzy, niż etnologii. Po drugie: nacisk kładziony na społeczne okoliczności tworzenia wiedzy etnologicznej, jej zależność od zmieniających się zewnętrznych czynników politycznych, społecznych i gospodarczych nie otwiera obiecującej perspektywy badawczej. Takie postawienie zagadnienia nieuchronnie musi doprowadzić do analizowania okoliczności, w których etnolog zdobywa wiedzę, przy zaniedbaniu rozwijania wiedzy o badanych kulturach.

Jednak opozycja antyewolucjonistyczna nie ograniczyła się jedynie do wyrażenia sprzeciwu o charakterze światopoglądowym – przeciwko teorii ewolucjonistycznej wysunięto poważne zarzuty. Najważniejsze z nich to:

- krytyka naturalistycznego podejścia ewolucjonistów,
- zakwestionowanie wartości schematów rozwojowych kultury,
- poddanie w wątpliwość jednakowości postępu kultury,
- poddanie w wątpliwość koncepcji wyższości i niższości stadiów rozwoju kultury.

Krytyka naturalizmu ewolucjonistów przybierała różne formy. W najbardziej spośród nich rozpowszechnionej poddaje się w wątpliwość istnienie uniwersalnej, ahistorycznej natury ludzkiej, czyli nieredukowalnej własności decydującej o byciu człowiekiem, od poznania której zależy możliwość formułowania praw dotyczących cech, własności lub działań ludzkich. Takie podejście rozpowszechniło się zwłaszcza we współczesnej etnologii. Jak oświadczają reprezentanci takiego podejścia: „postawa taka odrzuca przede wszystkim racjonalistyczną retorykę oświeceniową mówiącą o »wspólnej naturze ludzkiej«, w której jakoby osadzone są »prawa«, to metafizyczne podłoże miało być zawsze pierwotne względem »kulturowych superstruktur«” (Burszta, Piątkowski [1994] s. 95). Tego typu krytyka prowadzi – jak się wydaje – donikąd. Przyjęcie istnienia wspólnej natury ludzkiej jest niezbędne, by projekt naukowego poznania innych kultur, nie tyle przeprowadzić, lecz - by go w ogóle sformułować. Albo etnologia bada swoisty przedmiot, czyli kulturę lub kultury zasadzające się na nieredukowalnych własnościach ludzi, albo nie jest nauką we właściwym tego słowa sensie. W tym kontekście jest sprawą drugorzędną, w czym będziemy upatrywać

wspólnej wszystkim ludziom właściwości, czy będą to potrzeby i utrwalone sposoby ich zaspokajania, instynkt językowy, głębokie struktury myślenia, czy cokolwiek innego.

Ważniejsze wydają się wątpliwości związane z ahistorycznością podejścia ewolucjonistów, wynikającego z koncepcji wspólnej natury ludzkiej. Najistotniejszy zarzut – z którym w pełni należy się zgodzić – padający z tej strony, to wskazanie dowolności, z jaką ewolucjoniści posługiwali się faktami zaobserwowanymi w obcych kulturach (Lutyński [1956] s. 212-213). Istotnie ewolucjoniści, przekonani o zasadniczej jedności ludzkiej natury, bez zbytniego skrupowania zestawiali ze sobą fakty z najodleglejszych epok i regionów świata, które stanowić miały wytwór zawsze jednakowych władz psychicznych człowieka. Zauważmy jednak, że zarzut ten trafia przede wszystkim w praktykę badawczą ówczesnych etnologów, a nie - w teorię ewolucjonistyczną. Można wskazać skąd bierze się nietrafność wymienionego zarzutu. W mniemaniu oponentów, ewolucjoniści uznawali swoją dziedzinę za naukę indukcyjną, a schematy ewolucji kultury i jej elementów budowali w oparciu o empiryczne badanie i indukcyjne uogólnienie. To przekonanie wydaje się wątpliwe. Aby wyjaśnić istniejący stan badanej kultury, ewolucjoniści odwoływali się do poprzedzającego go – faktycznie lub hipotetycznie – stanu. Analizując obserwowane elementy kultury, zastanawiali się, jakie inne elementy były koniecznymi warunkami ich pojawienia się (Morgan [1887] s. 424-427, 440-457, 546, 547; Tylor [1923] s. 16). Poza tym, warto wskazać w tym miejscu na fakt, że Tylor w konstruowaniu schematów rozwoju kultury posługiwał się – niedoskonałymi jeszcze – metodami statystycznymi, przy pomocy których w oparciu o badania empiryczne ustalał prawidłowości we współwystępowaniu elementów kultury (Tylor [1889] s. 621-630). Teorie ewolucjonistów wydają się być raczej konstrukcjami, a ich zadaniem było ułatwienie zrozumienia przemian, którym podlegają kultury. W takiej interpretacji schematy rozwoju kultury są modelami rozwoju kultury, a nie po prostu - opisami jej przeobrażeń (por. Marciszewski [1988] s. 128). Zauważmy, że takie ujęcie nie zamyka możliwości badania kultur w ich historycznej zmienności, choć wiąże się z koniecznością odrzucenia naiwnego realizmu historycznego, czyli przekonania, że historyczna zmienność społeczeństw jest dana bezpośrednio, a do jej ujęcia wystarcza zwykły opis. Jeśli zgodzić się na takie podejście, schematy rozwoju kultury są konstrukcjami, budowanymi w oparciu o zdobytą wiedzę empiryczną, ale wiedzę empiryczną zdobywa się sformuławszy wcześniej pewne hipotezy. Opis faktów dokonany przez etnologa nie jest punktem wyjścia dla formułowania hipotez o zmienności badanej kultury, a wręcz przeciwnie. Podlega on interpretacji przez pryzmat uprzednio sformułowanych hipotez. Hipotezy te przez konfrontację z opisami faktów podlegają sprawdzaniu i mogą zostać przeformułowane.

Przy takiej interpretacji upada także zarzut, że ewolucjoniści traktują cały świat pozaeuropejski jako ogromne muzeum, w którym można odnaleźć eksponaty dające obraz kultury europejskiej sprzed wieków. „Dzicy” nie są, jak to złośliwie ujmowali krytycy ewolucjonizmu, naszymi żyjącymi przodkami. Natomiast pod pewnymi – prawdopodobnie pod wszystkimi – względami mogą być tacy jak oni. Nie jest prawdą, że zatrzymali się niejako w pewnym momencie rozwoju. Ale ich kultura daje się zrozumieć dopiero przez odniesienie występujących w niej zjawisk do sformułowanej już uprzednio teorii, która umożliwia także zrozumienie zjawisk zachodzących w naszej kulturze.

Z zagadnieniem historycznej zmienności społeczeństw pierwotnych wiąże się jeszcze inny ważny problem. W najprostszym ujęciu polega on na tym, że etnologowie często dysponują bardzo niezadowolającymi źródłami informującymi o przeszłości badanej kultury. Wpływają na to rozmaite czynniki. Kultury badane przez etnologię z reguły nie znają pisma, a ich tradycja opiera się na pamięci i ustnym przekazie. Dominujący w tych kulturach tradycjonalistyczny lub mitologiczny typ światopoglądu również nie sprzyja zachowaniu wiadomości o zdarzeniach historycznych. Tryb życia niektórych ludów zbieracko-koczowniczych skutecznie uniemożliwia zachowanie świadectw ich przeszłości. Niekiedy przyczyny tego tkwią w języku – słynny jest przykład mieszkańców wysp Nikobarów leżących na Oceanie Indyjskim, którzy po śmierci członka społeczności poddawali gruntownej reformie język, usuwając z niego określenia związane ze zmarłym, co „radycznie uniemożliwia »tworzenie historii«, a w każdym razie - przekazywanie historycznych opowieści” (Frazer 1996, s. 213).

Etnolog z reguły znajduje się w sytuacji nie do pozazdroszczenia. Sami badani nie często mogą opowiedzieć mu coś interesującego na temat historii (tak jak rozumie ją etnolog) swojej kultury. Poza tym - z reguły nie dysponują oni pojęciami, które by pozwalały uczynić to w sposób zadowolający z punktu widzenia etnologa. Nie może on przyjąć, że badana kultura się nie zmienia, gdyż przeczy to najbardziej oczywistym intuicjom. Jeżeli nie dysponuje jakimiś świadectwami pośrednimi, na przykład dostarczonymi przez badania archeologiczne, to dzieje i przemiany badanej kultury są przed nim po prostu zamknięte. Jeżeli natomiast w jego dyspozycji pozostają świadectwa pośrednie, to – ze względów wymienionych wyżej – włączenie ich w rekonstruowaną przeszłość badanej kultury napotyka na ogromne przeszkody. Wydaje się, że podejście ewolucjonistyczne może usunąć przynajmniej pewne trudności z tym związane. Warto przy tym zauważyć, że współczesna etnologia jest w sytuacji bez precedensu – to chyba jedyna nauka, której przedmiot badania po prostu znika. Kultury określane mianem „prymitywnych” lub „tradycyjnych” powoli



wymierają, a etnologia, przez próby przededefiniowania przedmiotu swoich badań, dość rozpaczliwie poszukuje wyjścia z sytuacji, w której się znalazła (por. Krzemień-Ojak [1997]). Jeżeli jednak przynajmniej część z etnologów pozostanie przy tradycyjnym przedmiocie badań, teoria ewolucjonistyczna może okazać się cenną propozycją teoretyczną pozwalającą porządkować i łączyć ze sobą fakty, które bez niej traktować by należało, jako zupełnie ze sobą nie powiązane.

Z przedstawionymi zarzutami łączyło się kwestionowanie przekonania o powszechności postępu. Utrzymywano, że skoro ewolucjoniści w tak dowolny sposób dobierali fakty, by konstruować swoje schematy, to ich roszczenie do powszechnego obowiązywania wyznawanych teorii nie jest uprawnione. Jeżeli uwzględnić poczynione wyżej przeformułowanie statusu schematów, taki zarzut można pominąć. Innym rozwiązaniem omawianego zagadnienia jest sformułowane w połowie XX wieku rozróżnienie ewolucji uniwersalnej i konkretnej, w myśl którego:

„dzieje poszczególnych kultur różnią się w sposób istotny i nie muszą przechodzić przez analogiczne studia rozwojowe, ale kultura ludzka jako całość przeobraża się z pewną ogólną regularnością [...]. Poszczególne społeczeństwa zachowując swoją specyfikę rozwojową wnoszą pewne istotne elementy do ogólnoludzkiego postępu, którego wznosząca się linia stanowi jakby wypadkową osiągnięć i doświadczeń poszczególnych kultur” (Posern-Zieliński [1987a] s. 114-115; por. [1987c]).

Podobne poglądy wypowiadali niekiedy także klasycy ewolucjonizmu, na przykład Tylor-pisząc o „powszechnej linii ewolucji” lub głosząc, że postęp „odbywa się w kierunku jednej linii ogólnej” (Tylor [1896] t. 1, s. 43, 68; podkreślenie SS). Te zagadnienia należałoby uwzględnić przy budowaniu modelu rozwoju kultury.

Często podejmowana była kwestia nieuprawnionego oceniania stadiów rozwoju kultury jako niższych i wyższych. O takim wartościowaniu mówiono w dwóch aspektach: ewolucji doskonałości instytucji kulturowych oraz ewolucji postaw i wartości etycznych. Pod pierwszym względem ocena kultury przez przyporządkowanie do niższego lub wyższego stadium jest moim zdaniem jak najbardziej uprawniona. W przytoczonej wyżej definicji ewolucji wyróżniono trzy kryteria zaawansowania kultury:

- złożoność,
- stopień wewnętrznej integracji
- funkcjonalną efektywność.

Wydaje się, że można mówić o złożoności zarówno kultury, jak i jej elementów. Przez złożoność kultury rozumieć należy liczbę elementów, na przykład instytucji kulturowych,

wchodzących w jej skład - oraz rodzaj łączących je relacji. Natomiast złożoność elementów kultury różnić się będzie w zależności od ich typu. W przypadku - na przykład - narzędzi będzie to stopień skomplikowania ich konstrukcji i obsługi, w przypadku form gospodarowania - liczba osób zaangażowanych w produkcji i zależności między wykonywanymi przez nich czynnościami. Sam w sobie wzrost zróżnicowania stanowić może wręcz słabość pewnego elementu kultury, dla przykładu można wymienić współczesną biurokrację czy system feudalny w XV wieku. A jako zaawansowany traktowany jest taki element kultury, w którym zróżnicowanie łączy się ze wzrostem funkcjonalnej efektywności, która wyraża się w stosunku nakładów - potrzebnych do wykonania zadania stojącego przed daną instytucją - do osiągniętych efektów. Narzędzi oceny dostarczyć może między innymi prakseologia. W przypadku kultury dotyczyć będzie to stosunków zachodzących między jej elementami. Wzrost funkcjonalnej efektywności wydaje się natomiast nie do pomyślenia bez wzrostu integracji wewnętrznej, czyli eliminacji sprzeczności pragmatycznych między działaniami wykonywanymi w ramach badanego elementu lub sprzeczności między uznawanymi wartościami, w wypadku kultury. Zauważmy, że wspomniane kategorie, których wprowadzenie uznaje się za dokonanie szkoły funkcjonalistycznej, zostały sformułowane - co prawda o wiele mniej precyzyjnie - już przez ewolucjonistów, czyli znacznie wcześniej niż się powszechnie przyjmuje.

Pozostaje jeszcze - najbardziej kontrowersyjne - zagadnienie ewolucji postaw i wartości etycznych. Jest oczywiste, że nie do utrzymania jest pogląd głoszony przez klasyków ewolucjonizmu, iż najdoskonalszym, pod względem uznawanych i realizowanych wartości etycznych, stadium rozwoju kultury jest kultura anglosaska drugiej połowy XIX wieku. Jednak warto zwrócić uwagę, że rosnąca złożoność i doskonałość instytucji kultury pociąga za sobą wzrost liczby ról społecznych funkcjonujących w ramach danej kultury. Nie pozostaje to bez związku z omawianym zagadnieniem. Jeśli przyjąć, że jednym z warunków osiągnięcia doskonałości etycznej jest możliwość rozwijania wrodzonych talentów i umiejętności, to im więcej ról społecznych, tym więcej możliwości ich doskonalenia. Uważam, że między innymi z tego względu, stadia rozwoju można oceniać w kategoriach postępu i zacofania.

Jest jeszcze inny aspekt tego zagadnienia. Należy, jak sądzę, założyć, że pewne wartości, muszą być realizowane w każdej kulturze, gdyż bez nich żadna ludzka społeczność nie mogłaby po prostu funkcjonować. Wymienić tutaj można na przykład - odwagę. Zarazem można wskazać pewne wartości, których realizowanie uzależnione jest ściśle od doskonałości instytucji kulturowych zaspokajających ludzkie potrzeby, czego przykładem może być bezinteresowność. Oczywiście standardy bezinteresowności są zmienne, jednak trudno w

ogóle realizować tę wartość w społecznościach, w których większość działań nakierowanych jest na zapewnienie biologicznego przetrwania, bezpośrednio od tych działań zależnego. Możliwość realizacji przynajmniej pewnych wartości etycznych jest zatem bezpośrednio związana ze stopniem zaawansowania kultury, na przykład pod względem rozwoju gospodarczego. Sądzę, że także pod tym kątem można oceniać kultury jako reprezentujące niższe lub wyższe stadium rozwoju.

Powyższe rozważania prowadzą do wniosku, że część zarzutów stawianych teorii ewolucyjnej jest po prostu nietrafna. Inne wskazują na istotne słabości referowanej koncepcji. Jeżeli przyjąć, że schematy ewolucji kultury lub jej elementów są modelami ich zmian, można w większości te zarzuty odsunąć.

### **3. Po co ewolucjonizm?**

Na koniec wypada postawić pytanie: po co prezentować – nawet w takiej postaci – ewolucjonizm jako aktualną propozycję teoretyczną dla etnologii? Sądzę, że tak ujęty ewolucjonizm wykazuje przynajmniej pod pewnymi względami przewagę nad konkurencyjnymi propozycjami teoretycznymi w etnologii, takimi jak funkcjonalizm i strukturalizm.

Jeżeli chodzi o kwestię zmienności – także historycznej – kultur, wymienione wyżej kierunki nie oferują obiecującej perspektywy badawczej. Jak przyznają zwolennicy kierunku funkcjonalnego: „w problematyce zmiany kulturowej najwyraźniej ukazują się słabe strony teorii kultury właściwej funkcjonalizmowi”. Pod tym względem teoria ta „okazała się nieadekwatna wobec rzeczywistości kulturowej, także rzeczywistości kultury pierwotnej” (Paluch [1981] s. 116). Jakkolwiek dostarczyła bardzo cennych narzędzi do opisu i analizy kultur pierwotnych i kultury w ogóle.

Podobnie jest w wypadku strukturalizmu, który deklaruje co prawda, że:

„tym, co interesuje etnologa, nie jest uniwersalność funkcji, która bynajmniej nie jest pewna i która nie może być stwierdzona bez uważnego zbadania wszystkich obyczajów z danego zakresu i ich rozwoju historycznego; interesuje go właśnie to, że te zwyczaje są bardzo zmienne” (Lévi-Strauss [1970] s. 68; podkreślenie CLS).

Jednak, pod tym względem, poza sferę ogólnych deklaracji nie wykracza. Ewolucjonizm natomiast, uznając zmienność za immanentną cechę kultury, daje nadzieję na zbudowanie interesującej koncepcji zmian w kulturze.

Ponadto, teoria ewolucjonistyczna, dzięki założeniu ahistorycznej natury, unika słabości skrajnego historyzmu, rozumianego jako pogląd głoszący, że każde zdarzenie zrozumiałe jest

jedynie w swoim kontekście, a zmienność kulturowa uwarunkowana jest jedynie swoistymi prawami historycznymi. Przez to założenie włącza się również w długą tradycję filozoficzną, czego niewątpliwie brakuje współczesnej etnologii.

Innym pożytkiem płynącym z proponowanego ujęcia ewolucjonizmu jest otwarcie możliwości skonstruowania uniwersalnych kryteriów oceny kultur. Przy panującej obecnie w etnologii modzie na skrajny relatywizm, ze wszystkimi destrukcyjnymi tego konsekwencjami, i może to stanowić zaletę tej propozycji.

Inną mocną stroną interpretacji, na którą wskazują niniejsze rozważania jest to, że można zasadnie przypuszczać, iż przedstawiony sposób rozumienia ewolucjonizmu prowadzi do skonstruowania znacznie pełniejszej, to znaczy obejmującej więcej aspektów, teorii kultury, niż konkurencyjne kierunki badawcze. Ewolucjonizm porusza na przykład zagadnienia postaw i wartości etycznych oraz ich miejsca w kulturze, którego ani funkcjonalizm, ani strukturalizm właściwie nie podejmuje.

Nie bez znaczenia jest również, o czym wspominałem wyżej, to, że znika tradycyjny przedmiot badań etnologii. Społeczeństwa dotychczas silnie tradycjonalistyczne zmieniają się w ogromnym tempie, asymilując treści i mechanizmy zasadniczo im obce, dotychczas charakteryzujące tak zwaną kulturę euro-amerykańską. Etnologia- badając kultury małych społeczności, w których stosunki społeczne opierały się najczęściej na pokrewieństwie i więzi lokalnej, a tradycję przekazywano ustnie, wypracowała narzędzia, które w opisanej sytuacji mogą okazać się nieprzydatne i być może - będzie musiała zwrócić się w stronę badań historycznych. Ponadto pewne grupy etniczne (na przykład australijscy Aborygeni), w których zachodzą wspomniane procesy, dążą do świadomego rekonstruowania tradycyjnej kultury. Mogą być w tym pomocne badania ewolucji kultur prowadzone przez etnologów. Biorąc to pod uwagę, warto rozważyć przedstawioną tu propozycję.

Z pewnością nie poruszyłem w tym szkicu wszystkich kwestii związanych z proponowanym przeformułowaniem koncepcji ewolucyjnego rozwoju kultury i nie wskazałem wielu konsekwencji z niego płynących. Sądzę jednak, że te zaprezentowane uzasadniają wzięcie pod uwagę wymienionej propozycji w dyskusji nad kształtem współczesnej etnologii.

## *BIBLIOGRAFIA:*

BURSZTA WOJCIECH J., PIĄTKOWSKI KRZYSZTOF

[1994] *O czym opowiada antropologiczna opowieść*, Warszawa, Instytut Kultury.

FRAZER JAMES GEORGE

[1996] *Złota gałąź*. Przełożył Henryk Krzeczkowski. Przedmowę napisał Jan Lutyński, Warszawa, PIW.

HERSKOVITS MELVILLE J.

[1956] *Man and His Work. The Science of Cultural Anthropology*, New York, Alfred A. Knopf.

KRZEMIENŃ-OJAK SŁAW

[1997] *Żegnanie postmodernizmu? Rozważania o przyszłości antropologii*, „Sztuka i Filozofia”, nr 13, s. 74-83.

LÉVI-STRAUSS CLAUDE

[1970] *Antropologia strukturalna*. Wstępem poprzedził Bogdan Suchodolski. Przełożył oraz *Słownik pojęć antropologii strukturalnej* opracował Krzysztof Pomian, Warszawa, PIW.

LUBBOCK JOHN

[1873] *Początki cywilizacji. Stan pierwotny człowieka i obyczaje dzikich współczesnych*. Tłumaczenie z francuskiego przekładu drugiego wydania Edmunda Barbier przez B. i Z. [Daniel Zgliński], t. 1-3, Warszawa, Redakcja „Przeglądu Tygodniowego”.

LUTYŃSKI JAN

[1956] *Ewolucjonizm w etnologii anglosaskiej a etnografia radziecka*, Łódź, Zakład Imienia Ossolińskich we Wrocławiu.

MARCISZEWSKI WITOLD

(red.) [1988] *Mała encyklopedia logiki*. Redaktor naukowy: Witold Marciszewski. Wydanie 2 zmienione, Wrocław – Warszawa – Kraków – Gdańsk – Łódź, Zakład Narodowy im. Ossolińskich.

MORGAN HENRY LEWIS

[1887] *Spoleczeństwo pierwotne, czyli badania kolei ludzkiego postępu od dzikości poprzez barbarzyństwo do cywilizacji*. Przekład Aleksandry Bąkowskiej. Objąsnienie i uzupełnienie według Marksa i Engelsa przez Ludwika Krzywického, Warszawa, Redakcja „Prawdy”.

PALUCH ANDRZEJ K.

[1981] *Malinowski*, Warszawa, Wiedza Powszechna, „Myśli i Ludzie”.

POSERN-ZIELIŃSKI ALEKSANDER

[1987a] *ewolucja kulturowa*, [w:] *Słownik etnologiczny. Terminy ogólne*, pod redakcją Zofii Staszczak, Warszawa – Poznań, PWN, s.113-116.

[1987b] *ewolucjonizm*, [w:] tamże, s. 116-118.

[1987c] *neowolucjonizm*, [w:] tamże, s. 254-256.

[1987d] *przeżytek*, [w:] tamże, s. 299-300.

TYLOR EDWARD BURNETT

[1889] *O metodzie badania rozwoju instytucji w zastosowaniu do praw małżeństwa i pochodzenia*, „Wisła”, t. 3, s. 614-641.

[1896] *Cywilizacja pierwotna. Badania rozwoju mitologii, filozofii, wiary, mowy sztuki i zwyczajów*. Przekład (z trzeciego wydania angielskiego z roku 1891) Z. A. Kowerskiej, ze wstępem i dodatkami, dotyczącymi rzeczy słowiańskich, a zwłaszcza polskich, oraz życiorysem autora przez Jana Karłowicza, t. 1-2, Warszawa, Wydawnictwo „Głosu”.

[1923] *Antropologia. Wstęp do badania człowieka i cywilizacji*. Przetłumaczyła z angielskiego Aleksandra Bąkowska. Z przedmową Ludwika Krzywického, Warszawa.

WALIGÓRSKI ANDRZEJ

[1973] *Antropologiczna koncepcja człowieka*, Warszawa, PWN.

## **Analogia jako źródło innowacji w naukach społecznych i filozofii**

Analogia wiąże się z typem rozumowania, które wymyka się jednoznacznym klasyfikacjom. Z jednej strony wnioskowanie przez analogię można uznać za “ułamną” wersję rozumowania logicznego. Takie wnioskowanie jest rozpowszechnione w myśleniu potocznym i ma znikome znaczenie w dyskursie naukowym, przy tym jest bardzo trudne do formalizacji. Z drugiej strony można wskazać na wszechobecność analogii w myśleniu potocznym i związany z nią fenomen psychologicznej, subiektywnej pewności, tak mocnej, jak w przypadku rezultatów we wnioskowaniu logicznym. “Myślenie przez analogię stanowi jedną z najbardziej zagadkowych form aktywności poznawczej człowieka” – tak otwiera swą książkę poświęconą podstawom wnioskowania przez analogię psycholog (Biela [1981] s. 5). Wskazuje to zarazem na konieczność interdyscyplinarnego podejścia do badania wnioskowania przez analogię – należy się odwołać do ustaleń logiki, filozofii nauki i języka, lingwistyki, psychologii i antropologii kultury.

Szczególnie interesujące jest przy tym znaczenie myślenia przez analogię dla przełamania utartych, kanonicznych form myślenia i opisu rzeczywistości, począwszy od przewyższania prostych oczywistości aż po rewolucje naukowe. Ta innowacyjna funkcja analogii daje się odnaleźć w historii wszystkich nauk szczegółowych, zarówno przyrodniczych, jak i społecznych. Właśnie temu wymiarowi myślenia przez analogię poświęcony jest niniejszy szkic. Celem artykułu nie jest formalizacja rozumowania przez analogię, ani opis znaczenia analogii - i rozumowania na niej opartego - dla przełomów w historii nauki. Chodzi tu raczej o modelowe ujęcie innowacyjnej funkcji analogii w naukach społecznych i filozofii.

### ***1. Analogia w logice***

Z punktu widzenia logiki możemy wyróżnić dwa podstawowe znaczenia nazwy “analogia”:

1. **Analogia jako proporcja.** Jest to wąskie rozumienie analogii. Analogia zachodzi między parą przedmiotów (zdarzeń, stanów rzeczy)  $S_1$  i  $S_2$  oraz parą przedmiotów (zdarzeń,

stanów rzeczy)  $S_3$  i  $S_4$  zawsze i tylko, wtedy gdy  $S_1$  pozostaje do  $S_2$  w takim samym stosunku, co  $S_3$  do  $S_4$  (Krajewski [1996] s. 16). Symbolicznie:

$$(1) \quad (S_1 \text{ i } S_2) \text{ a } (S_3 \text{ i } S_4) \equiv R(S_1, S_2) \cong R'(S_3, S_4),$$

gdzie  $S_1 \dots S_4$  – dowolne przedmioty, zdarzenia lub stany rzeczy,  $R, R'$  – dowolne relacje,  $a$  – relacja analogii, symbol “ $\cong$ ” oznacza “bycie podobnym”. (Np. klucz ma się tak do zamka jak hasło do komputerowego programu zabezpieczającego dane, tzn. zarówno klucz, jak i hasło “otwiera”, “daje dostęp”, “umożliwia wejście do wydzielonego obszaru”, etc.)

Zdefiniowanie analogii jako proporcji zakłada, że analogia oparta jest na równoważności relacji w obu członach. Jest to klasyczne rozumienie analogii, występujące w pracach Platona (np. *Państwo*, 534 a b) i Arystotelesa (np. *Analitiki II*, 74 a 18). Traktowanie analogii jako proporcji przetrwało aż do połowy XIX wieku.

**2. Analogia jako podobieństwo.** Przedmioty (stany rzeczy, zdarzenia)  $S_1$  i  $S_2$  są analogiczne zawsze i tylko, wtedy gdy cechy  $W_1, \dots, W_n$  przedmiotu (stanu rzeczy, zdarzenia)  $S_1$  są do siebie w podobnej relacji, jak cechy  $V_1, \dots, V_n$  przedmiotu (stanu rzeczy, zdarzenia)  $S_2$  (Krajewski [1996] s. 16). Symbolicznie:

$$(2) \quad S_1 \text{ a } S_2 \equiv R(W_1, \dots, W_n) \cong R'(V_1, \dots, V_n),$$

gdzie  $W_1, \dots, W_n$  i  $V_1, \dots, V_n$  – cechy przedmiotu, zdarzenia lub stany rzeczy,  $R, R'$  – dowolne relacje,  $S_1, S_2$  – dowolne przedmioty, zdarzenia lub stany rzeczy,  $a$  – relacja analogii, symbol “ $\cong$ ” oznacza “bycie podobnym”. (Np. średniowieczne miasto jest podobne do mrowiska za względu na połączone funkcje mieszkalne, reprodukcyjne i obronne.)

W oparciu o analogię prowadzi się różnorodne rozumowania. Ściśle rzecz ujmując, wnioskowanie z analogii można zaliczyć do wnioskowań uprawdopodobniających. Występuje ono w dwóch głównych odmianach (Ziemiński [1996] s. 191-192):

1. Wiedząc, że  $S_1$  jest  $P$ ,  $S_2$  jest  $P$ , ...,  $S_n$  jest  $P$ , skłaniamy się do sądu, że kolejne przedmioty rodzaju  $S$  ( $S_{n+1}$ ), które napotkamy, będą miały własność  $P$ . W przypadku indukcji wniosek byłby mocniejszy: każde  $S$  jest  $P$ . Jeśli we wnioskowaniu tym przesłanki będą prawdziwe, to może się okazać, że wniosek jest fałszywy. Czyli:

$$(3) \quad S_1 \in P$$



$$\begin{array}{l}
S_2 \in P \\
S_3 \in P \\
\text{.....} \\
S_n \in P \\
\hline
S_{n+1} \in P
\end{array}$$

lub inaczej:

$$\begin{array}{l}
(4) \quad (S_1 \in P) \wedge (S_2 \in P) \wedge (S_3 \in P) \wedge \dots \wedge (S_n \in P), \\
S_1 \cong S_2 \cong S_3 \cong \dots \cong S_n \cong S_{n+1}, \\
\hline
S_{n+1} \in P
\end{array}$$

2. Znając przedmiot  $S_1$ , który ma cechy A, B, C i jest P, wnioskujemy, że przedmiot  $S_2$ , który także ma cechy A, B, C, również jest P. Podstawą tego typu wnioskowania przez analogię jest założenie, że między cechami A, B, C i P zachodzi związek rzeczowy. Czyli:

$$\begin{array}{l}
(5) \quad S_1 \in A \text{ i } B \text{ i } C \text{ oraz } P \\
S_2 \in A \text{ i } B \text{ i } C \\
\hline
S_2 \in P
\end{array}$$

lub inaczej:

$$\begin{array}{l}
(6) \quad (S_1 \in A) \wedge (S_1 \in B) \wedge (S_1 \in C) \wedge (S_1 \in P), \\
(S_2 \in A) \wedge (S_2 \in B) \wedge (S_2 \in C), \\
S_1 \cong S_2 \\
\hline
S_2 \in P
\end{array}$$

W obu przypadkach rozumowania przez analogię między przesłankami a wnioskiem nie zachodzi stosunek wynikania.

Kontrowersje wokół obecności analogii w dyskursie naukowym dotyczą zatem nie tyle stwierdzania występowania analogii, ile raczej - zastosowania wnioskowania przez analogię. Dotyczy to zwłaszcza wnioskowania pierwszego typu (schematy 3 i 4), które jest jakby “wybrakowaną” indukcją, tzn. w pewnych kontekstach może sugerować, że sąd na niej oparty ma moc prawa (każde S jest P).

Generalnie należy zauważyć, że wnioskowanie przez analogię nie jest nastawione na generowanie uniwersalnych praw naukowych:

“Wnioskowanie przez analogię przebiega «od szczegółu do szczegółu», co każe go traktować jako odrębne zarówno od wnioskowania indukcyjnego jak i od dedukcyjnego; treścią wniosku analogicznego jest bowiem orzekanie zawsze o jakiejś pojedynczej, szczegółowej cesze, zdarzeniu lub stanie rzeczy  $P_{n+1}$  w danej sytuacji lub też orzekanie o występowaniu danej cechy, zdarzenia lub stanu rzeczy w nowej, kolejnej sytuacji  $S_{n+1}$ ” (Biela [1981] s. 11).

Problemy zaczynają się wtedy gdy autorzy tekstów naukowych o tym zapominają.

## ***2. Psychologia o myśleniu przez analogię***

Myślenie przez analogię jest przedmiotem badań psychologicznych od wielu dziesięcioleci. Sposób ujmowania analogii zależy przy tym od przyjętego paradygmatu psychologii. Psychologia zorientowana behawioralnie skupia się na mechanizmach uczenia się ujmowania związku analogii, czyniąc z wniosku analogicznego tzw. reakcję antycypacyjną spełniającą. Psychologia psychoanalityczna akcentuje podświadome mechanizmy kierujące rozumowaniem analogicznym. Psychologia poznawcza traktuje człowieka myślącego przez analogię jako układ informacyjno-decyzyjny (Biela [1981] s. 6-7).

Psychologowie rozumieją analogię znacznie szerzej niż logicy. Inaczej definiują także część z tych kategorii, którymi posługują się ci drudzy. Dotyczy to na przykład *pojęcia relacji*. Z punktu widzenia logicznego istnienie relacji uzależnione jest od istnienia jej członów, w aspekcie psychologicznym zaś można mówić o poznawczym ujmowaniu relacji bez konieczności jednoczesnego ujmowania jej członów (Biela [1981] s. 30).

Psychologowie posługują się testami analogii m.in. do badania rozwoju umysłowego człowieka. Schemat formalny testu wygląda następująco (Biela [1981] s. 22):

$$(7) \quad a R b :: c R' d,$$

co oznacza: a pozostaje w relacji R do b, podobnie jak c pozostaje w relacji R' do d. Jak widać testy analogii oparte są na rozumieniu analogii jako proporcji (por. schemat 1).

Badania laboratoryjne wskazują, że podstawową operacją umysłową przy rozwiązywaniu zadań analogii jest ujmowanie związku analogii (podobieństwa) między terminami testu. W przypadku testu, w którym znane są trzy pierwsze terminy, a nieznanym jest czwarty (w schemacie 7 – d), pierwotnym psychologicznie etapem rozwiązania jest ujęcie relacji R między terminami pierwszej pary i “przeniesienie” jej na terminy drugiej pary (Biela [1981] s. 67). (Np. samochód – silnik :: bryczka – x.)

Generalnie, współczesne badania nad analogią można sprowadzić do dwóch podstawowych perspektyw (Chlewiński [1999] s. 260 i n): strukturalnej (formalnej, syntaktycznej) i funkcjonalnej (pragmatycznej). Pierwsza polega na szczegółowej analizie wewnętrznego odwzorowania relacji między analogicznymi zdarzeniami, a druga – na poszukiwaniu wewnętrznego odwzorowania w zakresie relacji i własności porównywanych przedmiotów, związanych przyczynowo z osiągnięciem zadanych celów.

Adam Biela na podstawie analizy logiczno-psychologicznej wyodrębnił następujące etapy we wnioskowaniu przez analogię (Biela [1981] s. 124 i n.): (1) dostrzeżenie i ujęcie związku analogii (ujęcie związku analogii między obiektami, wyodrębnianie elementów lub faktów pozostających w pewnych relacjach), (2) sformułowanie przesłanek wnioskowania (obiektywizacja procesu dostrzeżenia analogii z fazy pierwszej: sformułowanie podstaw związku analogii; zidentyfikowanie pewnej relacji w jednym z obiektów lub zjawisk), (3) uświadomienie sobie charakteru implikacji analogicznej (świadomość probabilistycznego charakteru implikacji analogicznej lub traktowanie jej jako wynikania logicznego), (4) wyciągnięcie wniosku analogicznego (stopień kategoryczności sądu zależy od wyników etapu trzeciego).

Strategia analogiczna jako pewna wspomagająca strategia myślowa

“nie jest więc stereotypowym sposobem organizowania procesów umysłowych, lecz oznacza wysoki stopień aktywności podmiotu w zakresie przystosowania znanych wzorów i schematów postępowania do rozwiązywania problemów w nowych warunkach” (Biela [1981] s. 79).

Myślenie przez analogię można traktować zatem jako skuteczny mechanizm adaptacyjny organizmu ludzkiego wobec sytuacji deficytu informacji oraz wobec ograniczeń obliczeniowych mózgu ludzkiego.

Z punktu widzenia psychologii wyciąganie wniosku analogicznego ma zarazem charakter procesu poznawczego, i problemu decyzyjnego (konflikt motywów) (Biela [1981] s. 159). W relacjach interpersonalnych na myślenie przez analogię w istotny sposób wpływa postawa osoby wnioskującej wobec osoby, o której coś się orzeka (Biela [1981] s. 154). Nie bez znaczenia jest także to, o jakich relacjach wypowiada się wnioskujący, zwłaszcza gdy chodzi o koniunkcje relacji (np. różne ujmowanie związku w przypadku koniunkcji: pozytywna emocja i negatywne działanie vs. negatywna emocja i pozytywne działanie). Nie jest to zatem czysta czynność poznawcza.

Niewątpliwie w związku z tym psychologia zwraca uwagę przede wszystkim na pragmatyczny wymiar myślenia przez analogię w różnych kontekstach społecznych. Dotyczy to zarówno wymiaru indywidualnych procesów poznawczych i decyzyjnych, jak i funkcji użytkowych analogii w różnych narracjach (np. w edukacji, w tym edukacji moralnej - por. np. Klus-Stańska [2002] s. 197).

Wnioskowanie przez analogię w wymiarze psychicznym zbliża się bardzo do myślenia indukcyjnego. Główny wpływ na to ma probabilistyczny charakter myślenia w obu przypadkach (Biela [1981] s. 167). Nieuświadomienie sobie charakteru implikacji analogicznej w myśleniu przez analogię prowadzi do nieuprawnionych uogólnień. Nie należy jednak obu typów myślenia utożsamiać. Zasadnicza różnica polega na tym, że – jak podkreślono wyżej – wnioskowanie przez analogię przebiega “od szczegółu do szczegółu”, podczas gdy w indukcji dąży się do wydania sądów ogólnych. Po drugie zaś, w przypadku myślenia indukcyjnego chodzi o proces oceny prawdopodobieństwa, zaś w przypadku myślenia przez analogię - o proces “ważenia relacji”.

“«Ważenie relacji» stanowi więc istotną operację myślową w kształtowaniu się postawy krytycznej wobec wnioskowania przez analogię, gdyż od wyniku tej operacji zależy, jakie relacje zostaną uznane za dostateczną podstawę dla uformowania się związku analogii, czyli jakie relacje będą stanowiły tzw. relacje bazowe analogii” (Biela [1981] s. 127; por. także s. 166-167).

Choć myślenie przez analogię nie jest czystym myśleniem probabilistycznym, to jednak jego zawodny charakter i odwoływanie się przez wnioskującego w pewnych kontekstach do prawdopodobnych rozwiązań sprawia, że otwiera się ciekawa perspektywa teoretyczna

ujmowania analogii przez pryzmat rozumowania opartego na modelach umysłowych (*mental models*)<sup>108</sup>.

### 3. *Metafora a analogia*

Metafory są nie tylko rozpowszechnione w języku, ale także - zdaniem niektórych - stanowią genezę większości pojęć naszego języka (Lakoff, Johnson, [1988]). Jest wiele definicji metafory. Na potrzeby tego szkicu przez metaforę będziemy rozumieli figurę poetycką polegającą na nowym zestawieniu semantycznym pewnego zespołu wyrazów, odmiennym od tego zestawienia, jakie wynikałoby ze znaczeń poszczególnych wyrazów (Krajewski [1996] s. 122).

O metaforze mówi się zwykle, że “daje do zrozumienia” coś, czego dosłownie nie znaczy. Zatem - jak podkreśla Jan Srzednicki - należy “*uchwycić* tę analogię, symboliczne znaczenie zawarte w takich wypowiedziach” (Srzednicki [1993] s. 195). Przywołany autor wskazuje dalej, że:

“1. Dopóki nie domyślamy się, co mogłoby być właściwym tłumaczeniem metaforycznego stwierdzenia, dopóty nie rozumiemy go.

2. W rozumieniu metaforycznego zdania pomaga, jak się wydaje, to, że związek między elementami jest tak daleki. Pomaga to *uchwycić* analogię.

3. Posmak metafory nadaje zdaniu właśnie to, że trzeba *uchwycić* dalekie związki” (Srzednicki [1993] s. 197).

“Chwytnie” dalekich związków jako podstawa rozumienia metafory sprawia, że można traktować metaforę jako wyrażenie oparte na analogii. Wydaje się to trafną decyzją ze względu na rozważany w tym artykule problem.

Należy w związku z tym odróżnić metaforę “*martwą*”, która jest utartym zestawieniem słownym i nie wymaga poszukiwania bazowej analogii (ani w ogóle nie wymaga szczególnych operacji intelektualnych), od metafory żywej, która rzeczywiście nie jest w pierwszej chwili zrozumiała (Chlewiński.[1999] s. 270). Innym zjawiskiem językowym jest

---

<sup>108</sup> Np. w oparciu o schematy zarysowane w: (Johnson-Laird [1999]). “Główną tezę niniejszej pracy jest twierdzenie, że wiedza ogólna oraz przekonania, wspólnie z opisami sytuacji, prowadzą do konstruowania modeli umysłowych, które są stosowane do szacowania prawdopodobieństw” (s. 238).

także katachreza, która oznacza przesunięcie znaczenia w sytuacji, gdy wyraz o ustalonym znaczeniu stosuje się do nazwania nowych, bezimiennych przedmiotów (procesów, zjawisk).

Oparcie metafory na analogii sprawia, że dopiero uchwycenie tej drugiej umożliwia zrozumienie metafory:

“Metafora powstaje, wtedy gdy pewne użycie słów, które z powodu czy to swej własnej budowy, czy kontekstu, jest pozornie niezrozumiałe, zawiera jednak, jak się okazuje, pewien element podobieństwa do czegoś, co jest już wprost zrozumiałe. To podobieństwo, aczkolwiek dalekie, jest dostatecznie określone na to, by umożliwić uchwycenie jego ważności, można je zatem traktować jako wystarczającą wskazówkę intencji mówiącego” (Srzednicki [1993] s. 198).

Działanie metafory można także określić jako dodawanie do hierarchii pojęciowej nowych związków, które mogą powodować “zacieranie się” różnic między kategoriami (Chlewiński.[1999] s. 278). Celem metafory jest z tej perspektywy ukazanie bardziej abstrakcyjnych powiązań między pojęciami (a nie precyzacja pojęć). Nośnik i temat metafory łączy przeto pojęcie bardziej ogólne, znajdujące się wyżej w hierarchii pojęciowej (Chlewiński.[1999] s. 280). (Innymi słowy – upraszczając - metafora pozwala na odnalezienie “wspólnego mianownika” obu pojęć.)

Generalnie sędzi się, że metafora nie jest naturalnym punktem wyjścia do poszerzenia znaczenia wyrażenia (por. np. Srzednicki [1993] s. 198, 205). Często pozorne rozszerzenie znaczenia sprowadza się jedynie do utworzenia homonimów, które funkcjonują w języku równolegle. Jednak nie sposób nie dostrzegać innowacyjnego znaczenia metafory w historii nauk, zarówno humanistycznych, jak i przyrodniczych (aby trzymać się tego konwencjonalnego podziału). Po tych wprowadzających uwagach spróbujmy zatem wreszcie zmierzyć się z tytułowym problemem.

#### ***4. Historyczne przykłady zastosowania analogii***

Spójrzmy na kilka przykładów zastosowania analogii i wnioskowania na niej opartego w historii nauk, głównie humanistycznych. Prawdopodobnie, myślenie przez analogię rozpowszechnione było w czasach przednaukowych. Sposób wyzwiania się z pierwotnych systemów np. magicznych nie polegał jednak na całkowitym zerwaniu z wyobrazeniami teoretyków takich jak Paracelsus, których idee stanowiły źródło inspiracji intelektualnej elity aż do końca XVII wieku (por. np. Webster [1992] s. 15).

Ważną rolę odgrywała analogia w systemie Tomasza z Akwinu. Uważał on, że nazwy, które nadajemy Bogu (imiona, przymioty), mają się tak do samego Boga, jak skutek do przyczyny - ten sposób mówienia o Bogu (“niezupełnie wieloznacznie”) jest w rozumieniu św. Tomasza analogią (Gilson [1960] s. 157). Analogia (proporcja) może występować w dwóch głównych przypadkach: analogii między objawem i przyczyną (np. “zdrowa krew” - “zdrowe lekarstwo”), kiedy to kilka rzeczy pozostaje w jakimś stosunku do innej rzeczy, choć nie jest to stosunek jednakowy, oraz między przyczyną a skutkiem (“zdrowe lekarstwo” - “zdrowy człowiek”), kiedy to analogia zachodzi z racji istniejącego między rzeczami stosunku (Gilson [1960] s. 158). Przedmiot poznania w filozofii Tomasza (Bóg) sprawia, że szczegółowe analizy analogii mają w naszym przypadku nikłe zastosowanie, ale podstawowa charakterystyka i podział analogii są w pełni przydatne.

Analogia w systemie św. Tomasza pełniła specyficzną funkcję (rozważania nad analogią kontynuowali spadkobiercy teoretyczni Tomasza z Akwinu). Inne znaczenie miała analogia w odniesieniu do “ziemskich” przedmiotów poznania. W tym wypadku lepiej mówić o tworzeniu modeli<sup>109</sup> złożonych systemów. Szczególnie płodne poznawczo było stosowanie analogii z organizmami żywymi i systemami ekologicznymi:

“Niektórzy myśliciele, stykając się z systemami o zorganizowanej złożoności, zwracali uwagę na systemy żyjące jako na modele. Wiedza o systemach żyjących była poprzez *analogię* zastosowana do innych systemów, aby zdobyć pewien punkt oparcia w rozważaniu ich złożoności” (Weinberg [1979] s. 48).

Oczywiście takie podejście obciążone było pewnymi błędami. Gdy modelem był człowiek, popełniano często grzech antropocentryzmu. Na przykład Lamarck tłumaczył dostosowania ewolucyjne roślin i zwierząt pewnego rodzaju “inteligencją” – organizmy “działały” tak, by

---

<sup>109</sup> Jest to bardzo wieloznaczny termin i trudno podać jego definicję. Peter Abell rozpoczyna swoją książkę *Modele w socjologii* (1975) następującą uwagą: “Termin «model» stosowany jest w socjologii w wielu różnych znaczeniach, co dało początek dyskusji nad zaletami poszczególnych znaczeń. Autorowi dyskusja ta wydaje się przede wszystkim stratą czasu. Termin model rozumiemy tu możliwie szeroko: zarówno jako zbiór istotnych zmiennych, jak i jako określone związki funkcjonalne między nimi” (s. 11). Można ją uzupełnić uwagami Stefana Nowaka (1985): “[...] budowa modeli to po prostu pewien określony sposób uprawiania teorii. [...] Modelem pewnego zjawiska (przedmiotu, klasy zjawisk lub przedmiotów) ze względu na pewną rozważaną własność czy zbiór własności tego zjawiska jest więc jego opis w terminach pewnych elementów składowych i związków między nimi lub układu sprzężeń i zależności między własnościami tego przedmiotu, spełniający wymóg adekwatnego wyjaśnienia badanej własności czy ich zespołu” (s.444).

zwiększyć swoje korzyści. Błędy tego rodzaju mogły także zniewalać wyobraźnię kolejnych pokoleń badaczy.

Człowiek bardzo często był modelem dla opisu przez analogię struktury różnych społeczeństw. Wystarczy przywołać bajki Ezopa. Tą drogą poszedł także Hobbes, traktując państwo jako Lewiatana (“ciało polityczne”), zaś jego instytucje jako jego organy:

“Gdy to się stanie, wielość ludzi, zjednoczona w jedną osobę, nazywa się *państwem*, po łacinie *civitas*. I tak powstaje ten wielki Lewiatan, a raczej (mówiąc z większym szacunkiem) ten *bóg śmiertelny*, któremu, pod władztwem *Boga Nieśmiertelnego*, zawdzięczamy nasz pokój i naszą obronę” (Hobbes [1954] s. 151).

Oczywiście Hobbes nie był niewolnikiem przyjętej analogii i zbudowanej na niej metafory. Z tego wypływał m.in. jego sukces.

Analogia ma wpływ nie tylko na najwyższy szczebel teorii (na drodze ogólnego wyobrażenia struktury teorii), ale także na rezultaty żmudnych badań, ciągłego zmagania się z empirią. Nie każdy teoretyk ma jednak pełną świadomość tego faktu. Co więcej, proces ten nie jest ograniczony do nauk humanistycznych. Mało który fizyk ma taką samoświadomość teoretyczną, jak Einstein, który zauważył:

“Nauka nie jest niczym innym, jak wyrafinowaną formą codziennego myślenia. Z tego powodu krytyczne rozważania fizyka nie mogą ograniczać się do analizy pojęć z jego dziedziny. Fizyk nie może posunąć się dalej bez uprzedniego krytycznego zbadania znacznie trudniejszego problemu, to znaczy bez przeanalizowania natury naszego codziennego myślenia” (Einstein [1997] s. 43).

A, jak wiadomo z ustaleń psychologów (przywołanych wyżej), podstawowym instrumentem potocznego myślenia i poznawania jest analogia.

Przykłady można by mnożyć. Można zaryzykować tezę, że nie ma dzieła z nauk humanistycznych, które nie zawierałoby choć jednej analogii lub wyniku wnioskowania przez analogię. Skutki stosowania analogii były bardzo różne. Niewątpliwie trzeba pamiętać o tym, że “granica między wyjaśniającą analogią a upodobnianiem na siłę jest bardzo cienka” (Wacquant [2001] s. 11), szczególnie, gdy brakuje nam bardziej obiektywnych narzędzi do zweryfikowania wyników takiego myślenia. Z czego zatem wynika sukces myślenia przez analogię, skoro – co oczywiste – nie z samego przywołania analogii? Spójrzmy na ten problem w sposób bardziej systematyczny.



## 5. *Innowacyjna funkcja analogii*

Do tej pory mówiliśmy o dwóch generalnych sposobach ujmowania analogii. Pierwszy sposób charakterystyczny jest dla logiki i nauk formalnych. Traktuje on wnioskowanie przez analogię jako specyficzny, zawodny proces poznawczy. Drugi sposób, omówiony w oparciu o ustalenia psychologii, traktuje myślenie przez analogię jako złożony proces poznawczy i decyzyjny zarazem, którego wyniki uzależnione są od kontekstu i postawy badacza. W dalszej części interesować nas będzie przede wszystkim poznawczy aspekt myślenia przez analogię, co nie oznacza rezygnacji z przywoływania wymiaru decyzyjnego, bez którego nie sposób nakreślić adekwatnego obrazu myślenia przez analogię w naukach humanistycznych.

Przede wszystkim spójrzmy na wyróżnione dwa typy analogii z perspektywy ich “innowacyjnego potencjału”.

1. **Analogia jako proporcja.** To wąskie rozumienie analogii ma ograniczone znaczenie innowacyjne na polu nauk społecznych. Kilka słów na ten temat wypowiedziano na marginesie analizy miejsca analogii w systemie Tomasza z Akwinu. I tak, odwołując się do przykładu zastosowania analogii w rozumieniu Tomasza, przestępczość, samobójstwa i ubóstwo można uznać za objawy “chorego społeczeństwa” (lub ich nikłe odsetki za objaw “zdrowego społeczeństwa”), a np. niesprawiedliwą redystrybucję dochodów - za przyczynę tego stanu rzeczy. Wtedy zachodzi analogia (pierwszego typu) między przestępczością i niesprawiedliwą redystrybucją dochodów, gdyż oba człony proporcji pozostają w pewnym (różnym) stosunku do tej samej rzeczy - społeczeństwa (“chorego społeczeństwa”). Analogia drugiego typu zachodzi zaś między niesprawiedliwą redystrybucją dochodów (“chorą redystrybucją”) a społeczeństwem (“chorym społeczeństwem”).

Analogie obu typów mają praktyczne zastosowanie na tych obszarach nauki, gdzie przedmiot poznania jest “skryty”, poznawany przez relacje pośrednie (postrzegamy x, o którym wiemy, że powstaje na skutek działania y, zatem wnioskujemy o pojawieniu się y, którego badamy), czyli np. cząstki elementarne w fizyce kwantowej lub struny w kosmologicznej teorii strun. Analogie te mają także zastosowanie we wnioskowaniu o stanie psychicznym innej osoby (Biela [1981] s. 152). W praktyce mogą zaś być przydatne np. w oswojaniu świata poznawanego bez pewnych zmysłów, np. u niewidzących.

Analogia jako proporcja ma ograniczone znaczenie heurystyczne, choćby ze względu na to, że znacznie ogranicza wnioskowanie z zastosowaniem modeli systemów społecznych, w którym to wnioskowaniu posługujemy się podobieństwem.

2. **Analogia jako podobieństwo.** Ograniczenia analogii jako proporcji przewyższa analogia jako podobieństwo. Należy w tym miejscu zaznaczyć, że problematyka podobieństwa wiąże się bezpośrednio z kategoryzacją zdarzeń. Powstaje tu groźba błędnego koła (co jest podstawą kategoryzacji? – podobieństwo; czemu przedmioty są podobne? – bo należą do jednej kategorii), na co zwracali w uwagę G. L. Murphy, D. L. Medin i L. Brooks (Por. Chlewiński.[1999] s. 264). Wedle tych autorów, dostrzeżenie podobieństwa jest funkcją wiedzy o tym, że porównywane przedmioty należą do tej samej klasy (Chlewiński.[1999] s. 264). Pamiętając o tych problemach, w dalszej części pozostaniemy jednak przy klasycznym rozumieniu podobieństwa, jako wielkości względnie stałej, wyznaczonej przez syntaktyczne reguły korespondencji porównywanych zdań<sup>110</sup>.

Analogie omówione do tej pory opierały się na - choćby pobieżnej - znajomości związku rzeczowego między wybranymi cechami wyróżnionych obiektów. Innowacyjność analogii na polu nauk społecznych może opierać się zaś na zestawieniu dwóch przedmiotów (stanów rzeczy), których nie jesteśmy w stanie od razu uporządkować ze względu na relacje przyczynowo-skutkowe lub inne prawa rzeczowe. Np. wyrażenie “społeczeństwo jest mrowiskiem” można potraktować jako metaforę, która opiera się na analogii między strukturą społeczną interakcji i praw nimi rządzących w ramach społeczeństwa mrówek i społeczeństwa ludzkiego.

Generalny schemat takiego rozumowania może nakreślić w następujący sposób (będę dalej używał wygodnego pojęcia “system”<sup>111</sup>). Podstawą rozumowania przez analogię jest dostrzeżenie i ujęcie związku analogii, na co wskazują ustalenia psychologów. Bez tego trudno mówić o procesie myślenia przez analogię. Drugim etapem jest wybór znanego systemu (rodziny systemów)  $S_1$ . Ponieważ sądzimy, że jest to system, który pomoże nam w opisie badanego systemu, tworzymy jego model ( $M_1$ ), który będzie podstawą dalszych rozumowań. Na tym etapie czynimy szereg założeń dotyczących relacji, na których oparty jest związek analogii. Trzeci etap polega na przekształceniu modelu  $M_1$  w model badanego systemu ( $M_2$ ). Przekształcenie polega na wieloaspektowym dostosowaniu modelu do wymagań nowego systemu. Jest to najbardziej twórczy etap myślenia przez analogię. Generalnie schemat “innowacyjnego” rozumowania przez analogię będzie wyglądał następująco:

---

<sup>110</sup> Por. Chlewiński.[1999] s. 267. Jest to założenie charakterystyczne dla strukturalnej perspektywy analizy analogii.

<sup>111</sup> System można traktować jako układ elementów pozostających w interakcji (por. Bertalanffy [1984] s. 48, 68, 86-87, 115-117).

1. Dostrzeżenie i ujęcie związku analogii (przyjęcie, że system  $S_1$  jest podobny do systemu  $S_2$ ),
2. Skonstruowanie modelu  $M_1$  znanego systemu  $S_1$ ,
3. Przekształcenie  $M_1$  w  $M_2$ .

Mówiąc o schemacie rozumowania zakładamy wysoki poziom samoświadomości metodologicznej wnioskującej osoby. Nie chodzi zatem, jak w psychologii, o stworzenie modelu rozumowania przez analogię, który adekwatnie opisywałby potoczne czynności poznawcze wszystkich (normalnych) ludzi, lecz o zarysowanie schematu, który wymaga świadomej, kontrolowanej implementacji w procesie badawczym. Postulat samoświadomości wiąże się między innymi z koniecznością świadomości ułomności (zawodności) implikacji analogicznej.

Zastosujmy opisany schemat do analizy metafory “społeczeństwo jest mrowiskiem”. Po pierwsze dostrzegamy, że istnieje analogia między społecznością mrówek a społeczeństwem ludzkim. Załóżmy, że naszym celem badawczym jest stworzenie ogólnej teorii wyjaśniającej zachowania społeczne. Zdanie sobie sprawy z występowania analogii może opierać się na bardzo różnych przesłankach, a nawet wrażeniach. To “poczucie” jest jednak konieczne dla zainicjowania wnioskowania przez analogię.

Następnie, w drugim etapie, tworzymy model społeczeństwa mrówek, które jest dla nas systemem znanym (lub takim, który uważamy za znany). Konstruowanie tego modelu będzie już pośrednio sterowane naszym celem badawczym, dlatego np. skupimy się raczej na wspólnych działaniach owadów, wzajemnym porozumiewaniu się i efektach grupowych, niż na anatomii mrówek. Tworząc model możemy także dostrzegać kolejne analogie z innymi społecznościami owadów, np. z osami społecznymi czy termitami. Ostatecznie zatem możemy stworzyć model o wiele szerszej kategorii niż wyjściowa, tzn. model owadów społecznych jako takich (por. Wilson [1979]).

Ponieważ założyliśmy, że “struktura społeczna” mrowiska jest podobna pod pewnymi względami do struktury społeczeństwa ludzkiego, to w trzecim etapie uprawnieni jesteśmy do przekształcenia modelu społeczności mrówek (owadów społecznych) w model społeczeństwa ludzkiego. Dzięki temu uzyskamy możliwość zbudowania “socjobiologicznej” teorii wyjaśniającej np. altruizm w społeczeństwie ludzkim.

Nie wnikając w rzeczywistą wartość teoretyczną powyższej operacji, ani tym bardziej - nie polemizując z socjobiologią, której status teoretyczny jest nadzwyczaj dyskusyjny, spójrzmy, jakie generalne błędy wiążą się z tak rozumianym myśleniem przez analogię. Po

pierwsze, może się okazać, że nasza wiedza o wybranym systemie  $S_1$  wcale nie jest tak szeroka, jak sądziliśmy. Wskutek tego być może skonstruowaliśmy błędny model  $M_1$ . W związku z tym poszukiwanie podobieństw może prowadzić do błędnych rezultatów, choć oczywiście - nie musi. Paradoksalnie błędny model wyjściowy może prowadzić do nowych odkryć teoretycznych, które rzeczywiście poszerzą naszą wiedzę. Po drugie, może się okazać, że wbrew naszym intuicjom  $S_1$  nie jest (ze względu na istotne cechy) podobne do  $S_2$ . I tak w naszym przykładzie, może się okazać, że zupełnie inny jest ontologiczny status indywiduum biologicznego w zbiorowości owadów i w zbiorowości ludzi, w związku z czym wyróżnione relacje, które “przenosimy” z jednego systemu na drugi, mogą przebiegać między zupełnie różnymi obiektami (obiektami z różnych poziomów). Wreszcie, po trzecie, można błędnie przekształcić  $M_1$  w  $M_2$ , tzn. możemy skonstruować adekwatny model  $M_1$  i mieć do czynienia z rzeczywistym podobieństwem obu systemów, ale błędnie “przenosić” relacje z  $M_1$  na  $M_2$  (np. między funkcjonalnie odmiennymi elementami).

Zarysowany schemat wykorzystuje kategorię “model”, ale nie wyklucza to porównywania wybranych, pojedynczych relacji. Wskazanie na model jest jednak ważne, gdyż wybierając pewne relacje (czy zbiory relacji) czynimy na ich temat szereg założeń, a ponadto celem wnioskowania przez analogię jest stworzenie teorii (modelu).

Widać z tego, że aby wnioskowanie przez analogię uzyskało pełnoprawne obywatelstwo w królestwie metod naukowych, musi być zaślubione z którąś ze szlachetnych metod rozumowania. Zatem w zarysowanym schemacie rozumowania przez analogię konieczny jest czwarty etap, który polega na wyprowadzeniu modelu  $M_2$  – uzyskanego na podstawie myślenia przez analogię, z materiału empirycznego lub z dowiedzionych praw danej dyscypliny.

Nie pozbawia to jednak analogii jej mocy innowacyjnej.

“W istocie niekiedy możemy odnieść korzyści, budując modele systemów słabo znanych. Pewien świeży punkt widzenia może być pożyteczny, jeśli tylko *cokolwiek* wiemy o danym analogu. Analog co najmniej pobudza nasz umysł, a przecież w niebiesiech wiadomo, że umysł ludzki wymaga nieco bodźców” (Weinberg [1979] s. 49).

Jak zauważył G. M. Weinberg, “łańcuch rozumowania może stanowić sto ogniwi logicznych, albo też jeden skok analogii”. Nie znaczy to oczywiście, że po wirtuozerskim skoku można zrezygnować ze żmudnego łączenia kolejnych ogniwi logicznych, które świadczą o naukowej spistości i sile teorii. Stawia to analogię po stronie metod heurystycznych. Jednak analogia

jako metoda – jak się wydaje z racji predyspozycji gatunkowych – ma duży potencjał innowacyjno-adaptacyjny.

Trzeba odróżnić innowacyjność analogii, która ogranicza się do ograniczonych zakresowo zmian w ramach systemu czy paradygmatu teoretycznego, od innowacji, która w konsekwencji prowadzi do podważenia wyjściowego paradygmatu. Wedle T. Kuhna:

“Zasadniczo istnieją tylko trzy rodzaje zjawisk, na gruncie których tworzyć można nową teorię. Po pierwsze, mogą to być zjawiska uprzednio już dobrze wytłumaczone przez istniejące paradygmaty; rzadko kiedy jednak są one motywem czy też punktem wyjścia do konstruowania nowej teorii. Gdy jednak tak się dzieje [...] uzyskane w rezultacie teorii nie znajdują zazwyczaj uznania, brak bowiem dostatecznych racji, by rozstrzygnąć o ich słuszności. Po drugie, mogą to być zjawiska, których naturę określa istniejący paradygmat, lecz których szczegóły rozumiane być mogą tylko w wyniku dalszego uszczegółowienia teorii. Są to zjawiska, których badaniu uczoney poświęca większość swego czasu. Jego celem jednakże jest tu raczej uściślenie istniejących paradygmatów niż zastąpienie ich innymi. Dopiero, wtedy gdy tego rodzaju próby uściślenia zawiodą, uczoney ma do czynienia z trzecim rodzajem zjawisk – z rozpoznanymi anomaliami, które charakteryzuje to, iż uporczywie opierają się ujęciu za pomocą istniejących paradygmatów. Ten rodzaj zjawisk sam daje początek nowym teoriom” (Kuhn [2001] s. 175).

Przy zmaganiu się z anomaliami mamy do czynienia z *sytuacją analogiczną*, jak w przypadku sytuacji deficytu informacji w świecie wobec konieczności podjęcia działania. I tak jak analogia ułatwia rozwiązanie problemu decyzyjnego w życiu, tak wspomaga proces poznawczy w świecie nauki. Porównanie to może się wydawać kontrowersyjne, ale koresponduje ono z ustaleniami podejścia pragmatycznego w dyskusji o racjonalności człowieka. Wskazuje ono na adaptacyjny charakter ludzkich procesów rozumowania (por. np. Lewicka [2001]).

Nie każde ufundowanie nowego modelu w oparciu o analogię prowadzi do rewolucji naukowej w sensie kuhnowskim. Czasami jest to pozorna rewolucja, zwłaszcza, wtedy gdy mamy do czynienia z zastosowaniem paradygmatu jednej dziedziny na polu drugiej dziedziny, bez zmian własnych systemów kategorii badacza (Weinberg [1979] s. 53). Co oczywiste, analogia nie gwarantuje innowacji, choć może do niej prowadzić.

W związku z tym można stwierdzić, że dobór metafor w tekście naukowym przebiega na dwa sposoby. Po pierwsze, metafory wplata się w dyskurs specjalistyczny jako retoryczne ozdobniki, ułatwiające percepcję zapisanego wniosku, lub zdające sprawę z niemożności ścisłego opisanie danego problemu (wtedy analogia jest bardziej strawnym dla

ambitnego autora zapisaniem, że dany sąd nie uzyskał potwierdzenia w badaniach lub dedukcji). Po drugie, metafora może być elementem rozumowania uprawdopodobniającego lub dyskusji, które nie kończy się jednoznaczną konkluzją. Wtedy metafora i analogia, na której jest oparta, wskazują na “kierunek” rozumowania, który rekomenduje autor, choć nie jest w stanie zakończyć go ścisłym, logicznym rozumowaniem.

Paradoksalnie, wnioskowanie przez analogię uwieńczone pełnym sukcesem, tzn. zakończone czwartym etapem potwierdzenia dedukcyjnego lub indukcyjnego, maskuje swoje analogiczno-metaforyczne korzenie. Badacz wiedziony nicią analogii, dochodząc do wyjścia z labiryntu, wytycza oświetlony trakt dowodu naukowego, który po nici nie zostawia ani śladu.

Należy jeszcze powrócić do problemu dwóch wymiarów wnioskowania przez analogię: poznawczego i decyzyjnego. Uwzględnienie wymiaru decyzyjnego wcale nie musi oznaczać “merkantylizacji” poznania naukowego: motywy mogą mieć bowiem charakter czysto autonomiczny (poznawczy). Dotyczy to społecznego kontekstu odkrycia naukowego, np. kryzysu, który badacz za wszelką cenę chce ogarnąć poznawczo, a uniemożliwia mu to dominujący paradygmat. Kryzys ma szczególną rolę innowacyjną, zwłaszcza w filozofii, do której najpełniej stosuje się teoria rewolucji paradygmatów Kuhna (por. Ochocki [2001] s. 21 i n).

Podsumowując, analogia ma istotną funkcję innowacyjną w naukach społecznych i filozofii. Pozwala na przewycięzanie kryzysów poznawczych, gdy standardowe metody nie dają zadowalających rezultatów. Wynika to prawdopodobnie z gatunkowych predyspozycji człowieka, który posługuje się analogią w życiu codziennym, zwłaszcza w rozwiązywaniu problemów decyzyjnych. Analogia jest wyróżnioną metodą heurystyczną, która – gdy badacz świadomy jest jej ograniczeń – prowadzić może do cennych innowacji, a nawet głębokich zmian paradygmatów. Rezultaty tego procesu muszą jednak znaleźć potwierdzenie w indukcji lub dedukcji.

## BIBLIOGRAFIA

ABELL P.

[1975] *Modele w socjologii*, Warszawa.

ARYSTOTELES

*Analityki II.*

BIELA A.

[1981] *Psychologiczne podstawy wnioskowania przez analogię*, Warszawa.

BERTALANFFY L. von

[1984] *Ogólna teoria systemów. Podstawy, rozwój, zastosowania*, Warszawa.

CHLEWIŃSKI Z.

[1999] *Umysł. Dynamiczna organizacja pojęć. Analiza psychologiczna*, Warszawa.

EINSTEIN A.

[1997] *Teoria względności i inne eseje*, tłum. P. Amsterdamski, Warszawa.

GILSON E.

[1960] *Tomizm. Wprowadzenie do filozofii św. Tomasza z Akwinu*, Warszawa.

HOBBS T.

[1954] *Lewiatan*, tłum. Cz. Znamierowski, Warszawa.

JOHNSON-LAIRD P.N.

[1999] *Modele umysłowe a myślenie probabilistyczne*, tłum. J. Radzicki, w: Z.

CHLEWIŃSKI (red.), *Modele umysłu. Zbiór tekstów*, Warszawa.

KLUS-STAŃSKA D.

[2002] *Narracje w szkole*, w: J. Trzebiński, *Narracja jako sposób rozumienia świata*, Gdańsk.

KRAJEWSKI W.

(red.) [1996] *Słownik pojęć filozoficznych*, Warszawa.

KUHN T. S.

[2001] *Struktura rewolucji naukowych*, tłum. H. Ostromecka, Warszawa.

LAKOFF G., JOHNSON M.

[1988] *Metafory w naszym życiu*, Warszawa.

LEWICKA M.

[2001] *Czy jesteśmy racjonalni?*, w: M. Kofta, T. Szustrowa (red.), *Złudzenia, które pozwalają żyć*, Warszawa.

NOWAK S.

[1985] *Metodologia badań społecznych*, Warszawa.

OCHOCKI A.

[2001] *Filozofia i burze dziejowe*, Warszawa.

PLATON

*Państwo.*

SRZEDNICKI J.

[1993] *Kłopoty pojęciowe*, Warszawa.

WEBSTER CH.

[1992] *Od Paracelsusa do Newtona. Magia i powstanie nowożytnej nauki*, tłum. K. Kocińska, A. Zapałowski, Warszawa.

WACQUANT L.

[2001] *Wprowadzenie*, w: P. Bourdieu, L. Wacquant, *Zaproszenie do socjologii refleksyjnej*, przeł. A. Sawisz, Warszawa.

WEINBERG G.M.

[1979] *Myślenie systemowe*, tłum. Cz. Berman, R. Wielburski, Warszawa.

WILSON E. O.

[1979] *Spoleczeństwa owadów*, tłum. D. Tymowska, Warszawa.

ZIEMBIŃSKI Z.

[1996] *Logika praktyczna*, Warszawa.



## Realizm i konstruktywizm w historii a język historycznego dyskursu

Przedstawienie, a przede wszystkim objaśnianie rzeczywistości historycznej, związane jest z zasadniczym problemem jej dyskursu: w jego ramach występują albo pojęcia, którym odpowiadają dobrze określone desygnaty (będące jednak konglomeratami składników, a nie pojedynczymi przedmiotami), i pojęcia ogólne - o często niejasnym związku między danymi zdarzeniami a istotą, pojęcia mające cechy *sui generis* interpretacji, albo pojęcia rozumiane jako bezsprzeczne *fakty historyczne*. Przykładem tego ostatniego może być upadek Rzymu; jego stwierdzenie jest bezsporne (nikt nie może twierdzić, iż cesarstwo rzymskie przetrwało), z drugiej strony trwają jednak nieustanne spory o cechy i istotę stwierdzanego faktu<sup>112</sup>, a więc o wszystko, to co nie jest samym stwierdzeniem. W pewnym sensie zatem upadek Rzymu dla wszystkich historyków istnieje tylko jako fakt zupełnie nagi, gdyż to, iż miał miejsce, byłoby może jedyną rzeczą, z którą wszyscy mogliby się zgodzić. Przykładem faktu rodzaju pierwszego może być zdobycie miasta Rzym przez Alaryka 24 VIII 410 r.<sup>113</sup>, jednak tylko pojmowane jako zespół zdarzeń, które zaszły tamtego dnia, a nie jako jego historyczne znaczenie, które istnieje tylko dzięki osadzeniu w szerszej historycznej całości<sup>114</sup>, w tym przypadku – w kontekście upadku cesarstwa rzymskiego.

Obecne są również pojęcia pośrednie, które w zależności od użycia mogą mieć charakter przede wszystkim przedmiotowy, lub przede wszystkim - interpretacyjny - gdy postulują inną *istotę* danego zjawiska historycznego. Typowym przykładem może być tu pojęcie rewolucji; kanonicznie wręcz mówi się o „rewolucji francuskiej”. Z drugiej strony, na obszarze badań nad antykiem, mówienie o rewolucji jest prezentyzmem, który jednak objaśnia nam *w nowy sposób* ciąg prostych faktów historycznych, a więc w pewnym sensie zmienia ich istotę, poprzez automatyczne proponowanie innych niż dotychczas powiązań

---

<sup>112</sup> pojęcie „faktu” jest dla mnie w tej pracy równoważne z faktem historycznym, czyli najogólniej mówiąc, zjawiskiem uznanym przez *communis opinio* badaczy, które może być wywiedzione w oparciu o źródła historyczne.

<sup>113</sup> materia historyczną, w której osadzone zostaną rozważania będzie świat antyku, przede wszystkim późnego, jako najbliższy autorowi

<sup>114</sup> Norman A. P. [1998], s.158.

przyczynowo-skutkowych między zdarzeniami. Uznany polski bizantynista, Kazimierz Zakrzewski, postąpił w ten sposób mówiąc o „rewolucji Odoakra”<sup>115</sup> w odniesieniu do wydarzeń z 476 r., zaś wybitny angielski historyk antyku, Sir Ronald Syme poddaje nam przykład niemal paradygmaticzny – w swoim słynnym dziele „The Roman Revolution”<sup>116</sup>, opisującym objęcie i sprawowanie władzy przez pierwszego cesarza rzymskiego, Oktawiana Augusta.

Historycy najchętniej umieszczają swą praktykę w poznawaniu rzeczywistości rozumianym jako znajdowanie tylko tego, co już – może nienazwane lub niezauważone – ale jednak - było. Niemniej, trudno zaprzeczyć istnieniu pewnego konstrukcyjnego aspektu tego poznania, obecności pojęć, których sama przeszłość w żaden oczywisty sposób nie przywołuje. Niektórzy (np. przedstawiciele pewnych koncepcji narratywistycznych<sup>117</sup>) powiedzieliby nawet, iż - skoro ów element konstrukcyjny stanowi immanentną część poznania, to *de facto*, niczym transcendentalne kategorie Kanta, tę historyczną rzeczywistość *ustanawia*. Moim celem jest próba przyjrzenia się (przede wszystkim z pozycji praktykującego historyka) zagadnieniu realizmu i konstruktywizmu w historii - właśnie w odniesieniu do owych ujęć najogólniejszych, które są wątpliwe ontologicznie, a - jednocześnie niezbędne dla wiedzy historycznej. To przecież wokół nich ogniskuje się od zawsze duża część dociekań badaczy; przy czym analizowanym przez nas przykładem będzie wspomniany już upadek Rzymu.

Koncepcja faktów historycznych o charakterze ogólnym, w rodzaju wspomnianego upadku Rzymu, sytuuje samo ujęcie historyczne w sferze już pewnej nadrzędnej konstrukcji pojęciowej, jaką bezsprzecznie jest ów kres Rzymu. Niezbędność dla dyskursu i wyjaśniania historycznego tego rodzaju metakoncepcji historycznych wydaje się być jednym z poważniejszych argumentów przeciw obecnie dominującej narratywistycznej teorii w filozofii historii. Nie zwraca ona bowiem uwagi na arbitralność, która musi się kryć chyba w każdym ogólnym fakcie historycznym: oto pewien określony zbiór zdarzeń z V w. nazwany został „upadkiem Rzymu”, niemniej nie był on uwarunkowany żadnymi koniecznościami zewnętrznymi. Był to akt jednocześnie nazwania i zrozumienia; stworzenia pewnej szerszej historycznej całości (taka nigdy nie istnieje na poziomie samych zdarzeń), rzecz, moim zdaniem, typowa właśnie dla wyjaśniania w obrębie historii. Dla narratywistów owo nazwanie i zrozumienie pewnego elementu procesu historycznego jest uwarunkowane

---

<sup>115</sup> Zakrzewski K. [1933].

<sup>116</sup> Syme R. [1939].

<sup>117</sup> np. Carr D. [1998].

strukturą narracji. Owa narracyjna konieczność określa nie tylko elementy formalne dyskursu (np. konieczność istnienia wątku), na co można by jeszcze przystać, ale - charakteryzuje je pod względem estetyczno-etycznym.

Jest to główna teza (później szeroko powielana i rozwijana) jednego z fundamentalnych i paradygmatycznych dzieł dla tego sposobu myślenia, jakim była *Metahistoria* Haydena White'a. White uważa dyskurs historyczny za jednocześnie wtórny i tym samym strukturyzowany poprzez cztery typy narracji fabularno-literackiej: romans, tragedię, komedię i satyrę<sup>118</sup>. Wedle tej koncepcji konstruktywizm istnieje w historii jako nadająca kształt „opowieści historycznej” struktura narracji. Ponieważ jednak narracja ta jest koniecznością niektórzy próbują już mówić po prostu o zrelatywizowanym do każdego szczególnego przypadku „narracyjnym realizmie” (mającym jednak swe źródło w tezie o charakterze idealistycznym, głoszącej, iż przedmiot historii jest konstytuowany dopiero przez język), który *de facto* staje się jedyną dostępną dla nas historyczną rzeczywistością. Pojedynczym zdarzeniom jako takim (nie włączonym w żadną „opowieść”) odbiera on „historyczny” charakter i nadaje im status właściwie Kantowskiego *noumenu*. Ta teza została jednoznacznie sformułowana przez Franklina Ankersmita, który, w ramach „linguistic turn” stał się jednym z prekursorów narratywizmu: głosząc, że autonomiczność rozważanych ogólnych pojęć i wyjaśnień historycznych, pod postacią specjalnie wprowadzonej kategorii „substancji narracyjnych” (dla której strukturalnym wzorem mają być monady Leibniza) jako autonomii samego tekstu historycznego<sup>119</sup>.

Proponowania samoreferencyjności tekstów historycznych unikają częściowo te koncepcje narratywistyczne, które wysuwają twierdzenie o tym, iż przeszłość jest, zanim zaczniemy o niej opowiadać, już narracyjnie ustrukturyzowana. Powołując się na analizy fenomenologiczne stwierdza się, iż życie i jego zdarzenia są dla nas spójne i zrozumiałe *przed* pisaniem historii, zatem historyczna narracja jest prawdziwa, *odzwierciedlając*, to co już właściwie w gotowej narracyjnej (historycznej) formie istnieje<sup>120</sup>. Podstawowym elementem analizy jest tu akcja, której różnorakimi wymiarami się zongluje. Poprzez samą akcję ma być również ukonstytuowana ludzka tożsamość, bo to właśnie pojedyncze „ja” działa i doświadcza. Analogicznie traktuje się czas i tożsamość „społeczną”<sup>121</sup>. Jednak, to, skupienie się na wymiarze w gruncie rzeczy pojedynczej akcji i jej wykonawcy, nie pozwala tej

---

<sup>118</sup> White H. [1973].

<sup>119</sup> Ankersmit F.R. [1983].

<sup>120</sup> Norman A. P. [1998], s.156.

<sup>121</sup> Carr D. [1998], s.146-147.

koncepcji na analizowanie pojęć historycznych, które są w jakiś sposób ogólne - i zarazem proponowane; wtórne i autonomiczne wobec doświadczenia (jak choćby przytaczany upadek Rzymu, które to pojęcie zaproponował dopiero Marcellinus Komes na pocz. VI w.), i mają zawartość interpretacyjną.

Wydaje się, że narratywiści popełnili jednak zasadniczy błąd, jakim jest pominięcie swoistej struktury ontologicznej historii (pomijając już brak uzasadnienia przesłanek dla ich koncepcji), która zawsze jest tylko pewną spełnioną możliwością. Trudno, żeby – w sytuacji, kiedy niemal każde zdarzenie historyczne ma charakter *przygodny* (mogło zajść lub nie) – epistemologia historii nadawała, tak jak chcą tego narratywiści, (mówiąc o koniecznościach wymuszanych przez narrację) pewien obiektywny, nadrzędny charakter naszemu poznaniu historycznemu. Nie sposób bowiem wskazać elementy, które mogłyby albo jednoznacznie determinować np. tragiczne aspekty upadku Rzymu (wedle wspomnianej koncepcji White'a, jako specyfikacja poetyki narracji) albo jednoznacznie o nim świadczyć.

Przykładowo, jeśli wkroczenie Alaryka do Rzymu miałyby być poczytane jako jego (tj. Rzymu) upadek lub też (wedle narratywistów) stanowić wyznacznik ujęcia tej opowieści jako tragedii, to znaczyłoby to, że gdyby Alaryk nie wkroczył do Rzymu, to pojęcie upadku Rzymu by nie powstało, a dana narracja by nie zaistniała. Prawdopodobnie jednak o upadku Rzymu nadal by mówiono, gdyż skłaniało do tego wiele innych wydarzeń. Żadne jednak nie skłaniało do tego w sposób zasadniczy, żadne nie mogło nadać sensu tym całościom. Być może koncepcja upadku Rzymu w ogóle by się – wobec nie wkroczenia Alaryka do Rzymu – nie pojawiła. Podobnie, w każdym innym historycznym momencie nie jesteśmy w stanie wskazać zdarzenia, które uprawniałoby nas do osadzenia w nim szerszej historycznej całości. Być może zdobycie Rzymu przez Alaryka miało zasadniczy charakter dla pojawienia się myśli o upadku Rzymu. Być może jednak to i każde inne zdarzenie nie mogło być determinującą przyczyną dla pojawienia się takiej myśli. Jeżeli tak jest, to mówienie o koniecznościach w narracji historycznej zawsze jest fałszem, gdyż nie sposób wskazać takiego wydarzenia (nawet jeśli wydaje się ono kluczowym punktem), które zmuszałoby nas do danego ujęcia pewnych historycznych zdarzeń i narzucałoby dany typ narracji (lub czyniłoby niemożliwym inny). Zatem każda narracja i każde szersze wyjaśnienie historyczne ma charakter arbitralny, gdyż w historii nie istnieją (lub nie jesteśmy w stanie ich wskazać) żadne obiektywne przyczyny.

Powiązany jest z tym bardzo ciekawy problem „historii niebyłej”<sup>122</sup>, - zdarzeń, które nigdy nie zaistniały, lecz stać się mogły. Z tego punktu widzenia, to co zaistniało, jest pod wszystkimi innymi względami tylko pewnym wariantem owych „historii niebyłych”, zrealizowaną możliwością, która zachowała jednak wszystkie cechy możliwości, gdyż to, że zaistniała, w żaden sposób nie przydaje jej jakiegokolwiek – wbrew pozorom - konieczności ani uprzywilejowania (przykładowo, *implicite* wyrażanej kiedy pytamy: dlaczego właśnie to się zdarzyło?). Myślenie o historii w kategoriach kauzalności wyłącznie między stwierdzonymi faktami historycznymi<sup>123</sup>, jak czyni większość narratywistów i, niestety, sporo historyków, jest zatem zawsze błędem.

O samej arbitralności i historycznym zrelatywizowaniu pojęcia „upadku Rzymu” świadczyć może istnienie perspektyw, które lansowały ideę wręcz przeciwną. Pojmując Rzym jako pewien zespół idei religijno – politycznych, zarówno bizantyjscy bazyleusi, jak i rzymscy papieże, mówili o nieprzerwanej kontynuacji, dla której przejęcie politycznej władzy przez Germanów (którzy i tak zresztą rządili przez zgoła kilka dziesiątków lat w ramach rzymskich struktur władzy) było tylko jednym z wielu przewrotów politycznych, jakich cesarstwo doświadczyło w ciągu setek lat swej historii. Jeśli zaś mówić o cywilizacji rzymskiej, traktując cesarstwo przede wszystkim jako pewien rodzaj i typ kultury, to upadek stanie się jeszcze mniej oczywisty, i są historycy, którzy w takim ujęciu wolą podkreślać właśnie ciągłość. Przykładem może tu służyć praca znakomitego angielskiego uczonego, Petera Browna, który w swoim dziele „Świat późnego antyku”<sup>124</sup> daje wyraz hipotezie, iż kultura antyczna, ewoluując ciągle i niemal niedostrzegalnie, nagle, niczym wąż zmieniający skórę, odnalazła się w nowej postaci – w żadnym jednak wypadku to przejście nie miało charakteru nagłego załamania.

Wydaje się zatem, iż te ogólne pojęcia historyczne (upadek Rzymu czy rewolucja francuska), mające charakter zarówno szerokiej konstrukcji zdarzeń historycznych, jak i całościowej konstrukcji interpretacyjnej, są pewnego rodzaju propozycjami

---

<sup>122</sup> ten problem próbuje rozważyć niemiecki historyk Alexander Demandt w swojej pracy, przełożonej na język polski, pt: „Historia niebyła” [1999].

<sup>123</sup> Jest to charakterystyczne dla typowego sposobu myślenia o historii, który często bazuje na analitycznej idei właściwego podziału materiału, co wyraża się w nacisku na określenie właściwych cezur, zaś samo objaśnianie historii dokonać się ma poprzez wskazywanie (w obrębie wyznaczonych cezurami podziałów) odpowiednich związków przyczynowo-skutkowych między analitycznie określonymi zdarzeniami. Często spór o nie właśnie jest dla dzisiejszej historii sporem o prawdę

<sup>124</sup> Brown P. [1991].

historycystycznymi<sup>125</sup> - zważywszy i to, iż sens czy wyjaśnienie historyczne odwołuje się bardzo często do ogólnego ujęcia danego czasu (owej propozycji), które nie jest jednak obecne w samym procesie historycznym *in se*. Sens dziejów jest zatem funkcją relacji do wartości, które zapośrednicza i wskazuje ludzki *wybór*<sup>126</sup>. Ta decyzja, będąc wyborem – jak dowodzą wyżej – arbitralnym, jest zawsze w jakiś sposób relatywizacją danego kawałka historii, tworząc pewną zamkniętą całość, jako że dane zjawisko historyczne zostaje *zindywidualizowane* (uznana zostaje jego historyczna niepowtarzalność) i ujęte jako ewoluujący jednak tylko w ramach proponowanej koncepcji (upadek Rzymu to rozważanie jego genezy i przebiegu, ale nie np. początków średniowiecza) - proces historyczny<sup>127</sup>.

Historycyzm może być również propozycją samej procedury badawczej w historii, która postuluje operowanie szerokim kontekstem historycznym, zarówno interpretacji, jak i samych zdarzeń, by znaleźć wyjaśnienie pojedynczych, oderwanych od siebie wzmianek w źródłach. Poszukiwane znaczenie danej wzmianki powinno zatem być elementem pewnej szerszej całości historycznej, wpasowywać się w nią, a nie - w kilka tylko zdarzeń, wąsko powiązanych wzajemnymi odniesieniami często niewiele mówiącymi o badanym zdarzeniu. Jest to procedura przydatna przede wszystkim do tworzenia roboczych hipotez, wtedy gdy mamy stosunkowo niewiele wzajemnie powiązanych źródeł historycznych, a z drugiej strony

---

<sup>125</sup> Autor chciałby w tym miejscu wyraźnie zaznaczyć, że jego przedmiotem zainteresowania jest historycyzm, jaki można znaleźć w zwykłej literaturze historycznej, nie zaś spekulatywne ujęcia całości dziejów w rodzaju czynionych przez Hegla, Spenglera czy Toynbeego, i to już choćby dlatego, że żaden historyk w ścisłym tego słowa znaczeniu nie podejmie się pisania historii w kategoriach globalnego i czasowego uniwersalizmu.

Literatura przedmiotu niekiedy wprowadza tu rozróżnienie na historycyzm (ten pierwszy) i historycyzm (ten drugi), nie jest one jednak zbyt powszechnie przyjęte (szczególnie na obszarze anglosaskim), przy czym historycyzm jest terminem o wiele powszechniejszym, dlatego by nie wprowadzać dodatkowych komplikacji, z powyższego rozróżnienia zrezygnowałem.

Dla celów praktycznego pisania historii historycyzm można ująć jako praktykę ujmowania historii w ramach możliwie szerokich cezur i obszarów (perspektywa holistyczna, integrująca fakty), przy czym te propozycje ram, jak i wszystkie fakty w ich obrębie muszą się poddawać skutecznej weryfikacji prawdziwości lub fałszu (perspektywa analityczna). Generalnie, praktyka powinna wykluczać propozycje perspektywiefalsyfikowalnych, stąd wspomniana wcześniej propozycja „rewolucji Odoakra” jest poprawna, ale „rola rewolucji w dziejach Rzymu” grzeszy już zbyt dużą spekulacją.

<sup>126</sup> Rossi P. [1975], s.24.

<sup>127</sup> Tamże, s.16.

zachodzi potrzeba interpretowania każdej wzmianki – (a więc szczególnie w badaniach nad antykiem)<sup>128</sup>.

Podobną ocenę można spotkać w tym niewielkim nurcie narratywistycznej teorii historiografii, który nie próbuje narracji historycznej wiązać z jej literackim odpowiednikiem. Idea - iż zawsze to narracja bardziej spekulatywna i bardziej odważna (zakładamy jednak, iż wszystkie twierdzenia o składających się na nią zdarzeniach są prawdziwe) będzie bardziej owocną w badaniach historycznych - ma być nawet w pewnym sensie paralelna do twierdzenia Poppera, że to teorie o największym stopniu falsyfikowalności ostatecznie dostarczą nam najwięcej nowych danych i stwierdzeń<sup>129</sup>. Argumentem wysuwany z kolei przez narratywizm za niepowtarzalnością historycznych fenomenów jest intensjonalny charakter języka historycznych tekstów<sup>130</sup>: żadne streszczenie czy parafraza „Upadku Rzymu” Edwarda Gibbona nie zastąpi samego dzieła, nie odzwierciedli bez różnic ani jego znaczeń, ani - stylu czy metody. Gdyby było inaczej, wystarczyłoby

Intensjonalność w rozumieniu Ankersmita jest zatem tą właściwością tekstu historycznego, która określa nam kryterium *tożsamości tylko jego znaczenia*:

„Jeśli jednak zawsze będą różnice między znaczeniem (meaning) tekstu pierwotnego a jego streszczeniami, to które z różnic w znaczeniu wykraczają poza kryteria tożsamości tego znaczenia, a które nie (tak byśmy byli wystarczająco usprawiedliwieni, by mówić o tożsamości znaczenia)? Oczywiście, ten problem jest dokładnie tym, czym pisanie historii (lub, w tym przypadku, pisanie historii pisarstwa historycznego) jest w zupełności.” (Ankersmit [1990], s.164); a nie kryterium *tożsamości wartości logicznej* zdania złożonego, dla którego wartość logiczna zdań składowych zależy od ich treści<sup>131</sup>. Zatem zdanie intensjonalne w ostatnim sensie może nie mieć cech tej swoistej narracyjnej intensjonalności Ankersmita. Jakkolwiek, trzeba zauważyć, że jego teza broni indywidualizacji historycznych konstrukcji ściśle jako tworów narracji (stąd wcześniej historyczne „fenomeny”, a nie zjawiska czy zdarzenia), nie zaś – jak historycyzm – jako koncepcji, które mają jednak same mówić coś o świecie, a zatem o realistycznym odniesieniu.

Mimo postulowania zarówno realistycznego odniesienia, jak i uzasadnienia każdej historycyzycznej koncepcji, nie mamy tu do czynienia z *regułami przekładu*. Stanowiło by to zaprzeczenie jednej z głównych tez historycyzmu, mówiącej, że rozważane zjawisko

---

<sup>128</sup> West D. [2000].

<sup>129</sup> Ankersmit F.R.[1983], s.3, 85.

<sup>130</sup> Ankersmit F.R. [1990], s164.

<sup>131</sup> Pelc [1982], s.257.

historyczne jest całkowicie niepowtarzalne, osadzone i rozumiane tylko w danym unikalnym procesie historycznej zmiany. Próba przypisania historii pewnych określonych reguł translacji czy zasad „przepisywania” badanej rzeczywistości w historyczną wiedzę o niej, ma podstawowe znaczenie we wszelkich próbach czynienia z historii nauki społecznej, zorientowanej na wzorec stworzony przez nauki przyrodnicze. Próby takowe nie są jednak w stanie wyjść poza podstawową metodologię badań historycznych, czyli to - co sami historycy najczęściej określają mianem krytyki zewnętrznej źródeł. Traktowanie historii jako pewnego zestawu określonych procedur, jakie są dokonywane wobec przedmiotu badań, nie pozwala nam bowiem z sukcesem wyjaśnić szerokiej konstrukcyjnej warstwy historii, jej elementów interpretacyjnych i narracyjnych.

Z drugiej strony, problem *przekładu* we współczesnej teorii historii można ująć z innej perspektywy (wywołanej częściowo przez akcentowanie i uprawomocnienie przez postmodernizm pluralizmu interpretacyjnego i perspektywizmu): perspektywy translacji na pewien dany, konkretny język (dyskurs) współczesności. Wszystkie te konceptualizacje właściwe są naszej perspektywie, nie zaś perspektywie czasu przeszłego *in se*. Mamy tu do czynienia z przeciwieństwem hermeneutycznego podejścia w historii. Współczesna mnogość owych perspektyw (zwróćmy uwagę na niemałą liczbę grup wykluczonych, z których pozycji można napisać historię) owocuje tak naprawdę poważnymi trudnościami, jeśli nie niemożliwością, w proponowaniu jakiegokolwiek interpretacji historycznej, o którą można by się sensownie spierać (nie negując od razu jako punktu widzenia naznaczonego ideologią). Wbrew pozorom, powiększanie liczby dyskursów, w ramach których można opisać dane zdarzenie, nie służy naszej wiedzy o tyle, o ile brak im wspólnej, syntetyzującej perspektywy, czy dyskursu, który próbował by je wzajemnie uprawomocnić i zintersubiektywizować. Wobec jego nieobecności, mamy do czynienia w pewnym sensie z nową formą historii antykwarycznej, gdzie kolekcjonuje się już nie tyle same zdarzenia, co ich klasyfikacje.

Tak silne odzwierciedlanie się intelektualnych prądów współczesności w historii stanowi chyba o silniejszym niż kiedykolwiek przedtem warunkowaniu danego przekładu osobą odbiorcy. Nie chodzi tu nawet tak bardzo o to, iż owa publiczność wyznacza perspektywy, lecz o to, iż historia ma być *przedstawiana* tak, by była dla publiczności zrozumiała. Ma to miejsce często nie tylko w potocznym ujmowaniu pewnych koncepcji historycznych, ale i wobec wielości proponowanych perspektyw badawczych i ich formalnego wyrafinowania – także w naukowych próbach ujmowania przeszłości. Ów postulat prowadzi do zastępowania nazw własnych faktów historycznych poprzez użycie swego rodzaju przekładów tych nazw, które mają przekazać pewne ich wybrane znaczenia i



sensy; oto, załóżmy, w antycznym dramacie pada nazwa pewnego państwa greckiego, które ma przywołać określony kontekst polityczny. Bez dłuższych wyjaśnień słuchacz lub czytelnik nie może tego kontekstu pojąć, czy zatem nie lepiej zastąpić od razu nazwy tego państwa nazwą państwa współczesnego, z pewnością słuchaczom znanego, a które – wedle tłumacza – przywoła równoważny kontekst?<sup>132</sup> Czy zatem mamy w tym momencie równoległe dwie równoważne i prawidłowe odpowiedzi, dwie prawdy, tę- przekładu literalnego, i tę - uwspółcześioną ? Czy taka narracja ma prawo nazywać się jeszcze historyczną, czy też nie? Przykład jest może trochę prymitywny, lecz wskazuje na istotę problemu: przeszłość może być poznawana - z jej perspektywy, którą to badający stara się przyjąć lub - poprzez perspektywę badacza. Przyjęcie tej ostatniej w pewnym sensie neguje jednak obiektywną konieczność istnienia przeszłości jako przedmiotu badań; historia we współczesnej perspektywie to bowiem historia w coraz to nowej perspektywie, historia pisana na coraz to nowy sposób, będąca już tylko przekazem *ahistorycznych* sensów. Można by zatem zapytać, czy – jeszcze jest to historia ? Niekończąca się liczba translacji, postulowana przez ten model, w wypadku jego autonomii, musi ostatecznie zagubić pierwotne zdarzenie i jego pierwotny sens. Pozostaje tylko aktualność sensu, ciągle odnawianego; jest to zatem już pamięć, która w przeciwieństwie do historii, jest właśnie zawsze tym tylko, co wobec upływu czasu wciąż żywe, w sposób nieprecyzyjny jednak, bo wybiórczo i subiektywnie odtwarzane.

Niemniej, te problemy wywołane przez postmodernistyczne pomysły kształtujące dziś sporą część teorii historii i samych badań historycznych, ogniskując się wokół pytań o tożsamość perspektywy historycznej, jednocześnie wprowadzają na nowo do historii - usunięty niegdyś przez Weberowski postulat „etycznej neutralności” - problem wartości. Historia nie jest nauką tylko o faktach. Oprócz swojej części opisowej zawiera w sobie także sądy o wartościach. Co prawda, są one najczęściej ukryte w wyborze historycznych konceptualizacji i perspektyw historycznych, które jako takie zawsze są jednak wyborami normatywnymi.

Jak pokazał Wittgenstein, wiedza o znaczeniu pojęcia zakłada zdolność do poprawnego posługiwania się tym pojęciem; a to z kolei zakłada umiejętność wskazania typu rzeczy, do których to pojęcie się odnosi, a także warunków, dla których można o nim powiedzieć, iż jest prawdziwe. Podobnie jak składowe stwierdzenia deskryptywne - (w

---

<sup>132</sup> przykład pochodzi z internetowej listy dyskusyjnej (X.2001) kwartalnika *History and Theory*: rozważano tam problem, czy zdanie „Chciałbym cię widzieć panem spraw Greków, i wodzem wszystkich jej hoplitów...” w celu uczynienia zawartej w nim aluzji jasną bez długich, dodatkowych komentarzy, można zastąpić zdaniem „Chciałbym cię widzieć, jak dowodzisz armiami Afaganistanu...” ?

historii uznają je za zdania o faktach stwierdzanych wyłącznie na podstawie dobrze określonych desygnatów) ustanawiają zatem znaczenie danego stwierdzenia historycznego jako kryterium jego prawdziwości<sup>133</sup> (ale teź nie weryfikują: korespondencja między danym faktem historycznym a jego znaczeniem nie może nam posłużyć jako warunek wystarczający do kontroli prawdy) - tak - stwierdzenia normatywne zakładają pewną teorię moralności, czy szerzej - wartości, która jednak najczęściej funkcjonuje jako bezproblemowe tło. W ten sposób twierdzenie o upadku Rzymu w starciu z Germanami (użyte w sensie wspomnianego na wstępie faktu historycznego o charakterze interpretacyjnym) - wydawałoby się - dyktowane przez samą historyczną naoczność (gdy patrzymy na liczne wojny końca imperium), zawiera w sobie pewien wybór na korzyść Rzymian, a zatem - ukrywa normatywną definicję cywilizacji rzymskiej jako tej „prawdziwej”. To zaś upoważnia nas z kolei na przykład do twierdzenia o późniejszym regresie kultury. Moglibyśmy przecież również mówić o powstaniu nowych, twórczych społeczności i kultur – „świecie narodów”, jak się go niekiedy nazywa (z kolei „kultura narodowa” definiowana jest jako ta „prawdziwa”) - lub o skutkach niepełnej integracji kultur (w tym przypadku problem dziedzictwa i misji cywilizacyjnej Rzymu (nie zaś wojen) staje się wyznacznikiem wartościowania sytuacji historycznej).

Niemniej, nie ma tu miejsca, postulowane przez postmodernistów, rozdzielanie narracji historycznej od dowodu, w postaci niemożności wypowiedzenia stwierdzeń niedyskusyjnych, które by redukowały problem „co w istocie się zdarzyło” mimo możliwości zaproponowania właściwie nieskończonej liczby innych, konkurencyjnych narracji<sup>134</sup>. Warstwy: normatywna i deskryptywna, wzajemnie się warunkując i uzupełniając, są bowiem ściśle ze sobą powiązane. Można tu również przytoczyć oczywiście klasyczny argument przeciw radykalnemu perspektywizmowi, który głosząc swoją tezę, nie jest w stanie ustanowić sam siebie.

Język historii byłby zatem zdolny do *równoczesnego* wypełniania funkcji normatywnych i reprezentowania; jak już wcześniej wskazywano, tak charakterystycznego

---

<sup>133</sup> Na przykład: wzmianki o danej wojnie w źródłach historycznych są pewnym faktem, którego znaczeniem może być ekspansja terytorialna, konflikt interesów handlowych lub zobowiązania wypływające z traktatów sojuszniczych (lub wszystko razem). Odpowiednio, kryterium prawdziwości (falszu) dla każdej z tych propozycji będzie obecność (brak) innych faktów o tym samym znaczeniu. Stąd – w przypadku epok o relatywnie niewielkiej ilości faktów historycznych (antykw) – często szczególnie trudno nawet w miarę jednoznacznie zweryfikować prawdziwość danego znaczenia (relacji).

<sup>134</sup> White [1993]

dla historii jednoczesnego *nazwania i zrozumienia*. Stwierdzenia deskryptywne, jak na przykład „Domicjan był kolejnym cesarzem, który oparł swe rządy na armii”, „Podział najwyższej władzy w cesarstwie między czterech rządzących, przeprowadzony przez Dioklecjana, miał w dużej mierze służyć lepszej obronie państwa.”, mogą być także odczytane jako stwierdzenia normatywne. Rozróżnienie na sądy o faktach i sądy o wartościach nie może być zatem dłużej przyjmowane na obszarze dyskursu historycznego za dane, i służyć do zawężania pola historii przez jej nieustanne relatywizowanie.

Ten fakt wydaje się być uwarunkowany w dużej mierze, postulowanym również przez historycyzm, holistycznym charakterem wyjaśniania w historii, które to wyjaśnienie warunkowane jest czymś, co dodatkowo określa ramy rozważań, a mianowicie - „horyzontem oczekiwań”<sup>135</sup>, który można próbować przybliżyć jako daną tradycję historiograficzną (na przykład, konserwatywną lub marksistowską). Może ona też służyć jako rodzaj metajęzyka dla wspomnianych wyżej faktów historycznych typu interpretacyjnego. Określa ona najbardziej ogólną formułę danej, napisanej, historii; do jego zaistnienia przyczynia się osadzenie dyskursu historycznego w sferze językowej pragmatyki - żadne dzieło historyczne nie pojawia się przecież w próżni społecznej. Jak już wspominaliśmy, w tle - bardziej lub mniej, ale obecni są konkretni słuchacze, z kolei normatywna warstwa dyskursu ogranicza możliwe do uwzględnienia czynniki. Stąd być może nieobecność w historii konsensu i nieustanny pluralizm stanowisk; również bardzo rzadko operuje się na jej obszarze pojęciem „racjonalnego argumentu”, takiego, który zmuszeni by byli przyjąć wszyscy; o tym, że nie jest to na polu historii możliwe, świadczy właśnie owa znaczna liczba współistniejących stanowisk, niemożliwa do utrzymania na obszarze nauk przyrodniczych.

Relacja między językiem a rzeczywistością bezsprzecznie nie jest zatem w historii „przejrzysta”, niemniej - nie musi to wcale prowadzić do zaprzeczenia wszelkiemu odniesieniu do rzeczywistości, wystarczy konkluzja chyba o wiele bardziej umiarkowana i „realistyczna”: referencyjność i korespondencja muszą być tu ograniczone do specyficznych ram pojęciowych, z których czerpią swe znaczenia; tym samym istnieją w tych ramach jako względne i, niejako, wewnętrzne. Pytanie „co jest rzeczywiste?” jest zatem zawsze zależne od specyfikacji językowego obszaru, na którym jest formułowane, gdzie ów kontekst ma charakter pewnych propozycji deskryptywno-normatywnych, a zarazem historycystycznych, proponują one bowiem pewne całościowe ujęcie danej rzeczywistości historycznej. Tym

---

<sup>135</sup> Lorenz [1998], s.364.

samym, pewne ogólne koncepcje typu „antyk” czy „renesans” wcale nie muszą być uważane za mniej realne od faktów typu „Oktawian zmarł w 14 r. n.e.”, które są zrelacjonowane przez świadków. Wiedza historyczna zachowuje swój praktyczny charakter *rekonstrukcji*, i nie staje się tylko *konstrukcją*<sup>136</sup>. Lokuje się bowiem na obszarze powyższego, Putnamowskiego „wewnętrznego realizmu”<sup>137</sup>, którego immanentny charakter decyduje jednak o jego konstrukcyjnym uwarunkowaniu.

## *BIBLIOGRAFIA:*

Ankersmit F.R.

[1983] *A Semantic Analysis of the Historian's Language*, Hague.

[1990] *Reply to Professor Zagorin*, *History and Theory* 10/1990, Wesleyan Station, Middletown, Connecticut.

Brown P.

[1991] *Świat późnego antyku*, przeł. A. Podzielna, Warszawa.

Carr D.

[1998] *Narrative and the Real World: An Argument for Continuity*, „*History and Theory*” eds. B.Fay, Ph. Pomper, R. Vann, Oxford.

Demandt A.

[1999] *Historia niebyła*, przeł. M. Skalska, Warszawa.

Lorenz Ch.

[1998] *Historical Knowledge and Historical Reality: A Plea for „Internal Realism”*, *History and Theory* eds. B.Fay, Ph. Pomper, R. Vann, Oxford.

Norman A.P.

[1998] *Telling It Like It Was: Historical Narratives on Their Own Terms*, *History and Theory* eds. B.Fay, Ph. Pomper, R. Vann, Oxford.

Pelc J.

[1982] *Wstęp do semiotyki*, Warszawa.

Putnam H.

---

<sup>136</sup> Tamże, s.351.

<sup>137</sup> Putnam [1981].

[1981] *Reason, Truth and History*, Cambridge.

Rossi P.

[1975] *The Ideological Valences of Twentieth-Century Historicism*, History and Theory, Beiheft 14 (1975), Wesleyan Station, Middletown, Connecticut.

Syme R.

[1939] *The Roman Revolution*, Oxford.

West D.

[2000] *Speculative Historicism*, Histos 4 (2000), wydanie internetowe: [www.dur.ac.uk/Classics/histos/2000/west.html](http://www.dur.ac.uk/Classics/histos/2000/west.html)

White H.

[1973] *Metahistory*, Baltimore.

[1993] *The Fact of Modernism: The Fading of the Historical Event*, manuskrypt cyt. w Lorenz Ch.[1998]

Zakrzewski K.

[1933] *Rewolucja Odoakra*, Kraków.

## **Użyteczność pojęcia „gra językowa” dla nauk społecznych**

Poniżej zamieszczona analiza pojęcia gry językowej odbiega zapewne dość daleko od klasycznego już poglądu na tak zwaną późną twórczość Ludwika. Nie można wykluczyć, że wykorzystanie terminu, skonstruowanego przez Wittgensteina jako metafora funkcjonowania języka – z całą swobodą właściwą metaforze, a niedopuszczalną w filozofii nauki – odbiega daleko od Wittgensteinowskiego pomysłu na język. Czytając *O pewności, Dociekania filozoficzne czy Kartki*, odnosi się wrażenie, że główną ideą Autora było wyzwolenie języka z więzów analizy, ukazanie jego złożoności i żywiołowości. Narzędzia pojęciowe są wobec języka bezużyteczne, próby opisu są pomyłką lub nadużyciem kompetencji, które nam sam język przyznaje. „Mam tu skłonność do walki z wiatrakami, gdyż wciąż nie mogę powiedzieć tego, co naprawdę chcę powiedzieć” (Wittgenstein 1993, s. 80) – pisze Wittgenstein, dając wyraz ludzkiej bezsilności względem języka. Jest to tyleż idea estetyczna i etyczna, ile uwaga na temat metody filozoficznej i naukowej. Wittgenstein demistyfikuje filozoficzne nadużycia w analizie języka i wzywa do powrotu do języka naturalnego – nieanalizowalnego.

Argumenty Wittgensteina są na ogół przekonujące, szczerść i świeżość jego myśli budzą sympatię. Znacznie mniej przekonujące i sympatyczne są argumenty niektórych jego komentatorów, zwłaszcza w dziedzinie nauk społecznych, którą szczególnie będę się tu zajmować.

Spróbujmy krótko zrekonstruować przykładowe rozumowanie powołujące się na swoiście pojęty autorytet Wittgensteina. Ponieważ twierdzi on, że język to instrument służący radzeniu sobie z rzeczywistością, na język składają się gry językowe, a gry te są nieporównywalne pod względem racjonalności – otrzymujemy wniosek, że ludzkie sposoby postępowania w świecie przyrody nie są porównywalne w kategoriach racjonalności. Prowadzi to w prosty sposób do zniesienia wszelkich obiektywnych standardów oceny postępowania, możliwości porównań rozmaitych strategii „radzenia sobie w świecie” oraz kryteriów wyboru strategii. Proponowana przez neopragmatystów zasada większej użyteczności, pomijając nawet problemy z samym jej sformułowaniem, wydaje się niemożliwa do obrony w roli ewentualnego uniwersalnego punktu odniesienia. Jej

konsekwencją jest, albo atomizm w postrzeganiu społeczeństwa, albo też negacja wolnej woli, tożsamej w danym razie z możliwością wytworzenia przekonania nieredukowalnego do aktualnego stanu wiedzy społecznej. Powołując się przede wszystkim na tezy z *Dociekań filozoficznych* i inne, pokrewne im koncepcje filozoficzne, można więc stworzyć wizję świata pozbawionego elementów dotąd uznawanych za nader istotne, jak „prawda” czy „użyteczność”. Autorzy tacy, jak Richard Rorty, dekonstruuja kolejne fragmenty dekoracji „kartezjańskiego teatru” i wypychają rzeczywistość poza nawias sensownego dyskursu (Rorty w: *Habermas...* [1996] s. 58-59). Niektórzy spośród tych autorów są wnikliwymi czytelnikami, błyskotliwymi pisarzami i retorami, co w znaczący sposób przydaje ich tezom uroku. Nic w tym złego, naturalnie, zwłaszcza że na ogół nie roszczą oni sobie pretensji do uniwersalnej ważności swoich poglądów, twierdząc, że są one kolejnym przypadkiem metody dyskursywnej czy jedną z wielu alternatywnych opowieści o społeczeństwie.

Opowieść jednak nie chce się skończyć. Dopóki bowiem pozostajemy w obrębie analizy dyskursu potocznego, nie dzieje się nic szczególnego. Najbardziej nawet opętany filozof (jak choćby Platon zdaniem Rorty’ego) nie twierdziłby, że ludzie zawsze są racjonalni czy zawsze dążą do prawdy – twierdziłby natomiast, że jedynie pewni ludzie - w pewnych okolicznościach. Owymi pewnymi ludźmi w naszej kulturze byli we własnej opinii najpierw filozofowie, potem zaś - uczeni. Wcześni twórcy socjologii wiedzy czy marksiści nie przeczyli możliwości naukowego myślenia.

Wielki szturm demistyfikatorów na naukę symbolicznie rozpoczął się, jak wiadomo, wraz z wydaniem prac Thomasa Kuhna. Rozszerzenie pola zainteresowań socjologii wiedzy na obszary wiedzy potocznej zaowocowało rozwojem konstruktywizmu w naukach społecznych. Obecnie mamy już do czynienia ze skrajnymi wariantami konstruktywistycznej socjologii wiedzy, której tezy (ze słynnym Luhmannowskim „wszelka wiedza jest tylko wiedzą wiedzy”) negują nie tylko przydatność, ale i możliwość tradycyjnej epistemologii i filozofii nauki.

Dorobek krytyków tradycyjnej, uprawianej w duchu quasi-pozytywistycznym filozofii nauki zainspirował krytykę tej dziedziny - nawiązującą nie do społecznego zastosowania wyników badania naukowego, ale do procesów społecznych (zwłaszcza zaś procesów myślenia grupowego) zachodzących w samym kolektywie naukowym. Krytyka ta w znacznym stopniu opierała się na analizie zjawisk językowych zachodzących z udziałem członków świata nauki – czyli gry językowej uczonych. Operując kategoriami Rorty’ego – spory w kolektywie naukowym w fazie zmiany paradygmatu polegają na dyskusji pojęć – dyskusja ta prowadzi tylko od jednej gry językowej do innej, nigdy poza język (Rorty [1996])

s. 110). Tym samym nauka weszła w krainę języka, stając się jednym z obszarów chronicznej nieciągłości.

Nie jest moim zamiarem oddawać się dyskusji z neopragmatyzmem czy postmodernizmem. Nie mam również zamiaru bronić Wittgensteina przed stawianymi mu niekiedy zarzutami relatywizmu czy radykalnego sceptycyzmu. Chciałabym jedynie spróbować wykazać, że mylą się ci, którzy na gruncie filozofii *Dociekań* próbują uzasadnić tezę o absolutnej nieciągłości form życia społeczeństw ludzkich i braku kategorii dla opisu tychże (Nielsen w: *Mocny program*, [1993] s. 363) - ci zatem, którzy utrzymują, że Kuhn, dezawuuując wysiłki dziedzin opisowych, powiedział w filozofii nauki ostatnie słowo.

Zacznijmy od kwestii najbardziej fundamentalnej. Wychodząc od Wittgensteinowskiej analizy wpływu gramatyki języka na obraz rzeczywistości (większość istotnych rozważań na ten temat zawarta jest w pracy *Gramatyka filozoficzna*) niektórzy autorzy formułują przekonanie o bezużyteczności teoretycznej założenia istnienia intersubiektywnej i obiektywnej rzeczywistości (Nielsen w: *Mocny Program* [1993] s. 365). Założenie to jest moim zdaniem niezbędne właśnie ze względu na dokonane przez Wittgensteina powiązanie języka z jego kontekstem, ze względu na instrumentalny charakter wyrażeń języka i antypsychologizm w opisie zjawisk językowych. Opis języka jako zespołu spokrewnionych gier językowych musi informować o odniesieniu języka do zjawisk pozajęzykowych – nie o samych zjawiskach oczywiście, nic jednak nie stoi na przeszkodzie próbom formalnego ujęcia samego sposobu odniesienia. O ile opis każdej gry językowej jest hermetyczny o tyle, że nie wychodzi poza samą grę – nie wyjaśnia, bo gry językowej wyjaśnić nie można, o tyle opis całości złożonej ze spokrewnionych gier językowych (to jest języka) powinien mieć moc wyjaśniającą, czy przynajmniej proponować możliwe czynniki wyjaśniające dla zjawisk językowych. Nie można – z definicji gry językowej (por. dalej) – wyjaśnić języka, ale można za jej pomocą wyjaśnić i uzasadnić inne, spokrewnione z nią gry językowe. Uzasadnienie to powinno nawiązywać do innych niż definicyjne elementów wspólnych każdemu zjawisku językowemu. Wspólność tego rodzaju dostrzegam w sferze realizacji funkcji gry językowej w społeczności graczy. Użyteczność filozofii Wittgensteina dla nauk społecznych staje się oczywista, gdy uświadomimy sobie, jak silny kładzie ona nacisk na społeczny i kulturowy kontekst funkcjonowania języka.

Dla uzasadnienia tej tezy pozwolę sobie krótko zaprezentować moje wyobrażenie Wittgensteinowskiej teorii gier językowych. Prezentacja ta jest o tyle niezbędna, że postaram się omówić wszystkie zagadnienia, które uważam za istotne dla nakreślenia socjologicznej przede wszystkim użyteczności pojęcia „gry językowej”. Zacznę od kwestii definicyjnych.



Wittgensteinowskie pojęcie „gry językowej” zrobiło od czasu publikacji *Dociekań filozoficznych* w 1953 roku niebywałą wręcz karierę. Wyrażenie to pojawia się po raz pierwszy w późnych pracach filozofa. Powszechnie przyjmuje się, iż rozwinął on koncepcję gry językowej, aby dać wyraz nowej teorii znaczenia. Nie wnikając w tym miejscu zbyt głęboko w istotę zagadnienia, można w skrócie powiedzieć, że „późny” Wittgenstein odrzucał pogląd „wczesnego”, jakoby język był „zwierciadłem rzeczywistości”. Rezygnując z metafory języka-zwierciadła Wittgenstein wprowadził zarazem nową, trafniejszą jego zdaniem metaforę języka-gry.

Kariera późnego Wittgensteina, w przeciwieństwie do wczesnego, rozwijała się nie tylko w dziedzinie filozofii, do której takie dzieła, jak *Dociekania*, przynależą na mocy wprost wyrażonej woli autora. Nie do przecenienia okazał się jego wpływ na socjologię, zwłaszcza na tak istotne kierunki współczesnej myśli społecznej, jak socjologia wiedzy, etnometodologia, teoria ról społecznych czy analiza dyskursów. O ile *Traktat logiczno-filozoficzny* stał się swego rodzaju biblią adeptów filozofii analitycznej, o tyle *Dociekania* – jak słusznie, acz złośliwie zauważa Bogusław Wolniewicz we *Wstępie* do ostatniego polskiego wydania - świetnie się prezentują w towarzystwie freudyzmu, New Age i postmodernizmu (Wittgenstein [2000] s. XVI), czyli stosunkowo daleko od systematycznej filozofii, a blisko współczesnej psychologii, socjologii i antropologii kulturowej. Nie miejsce tu na szczegółowe wywody dotyczące różnic pomiędzy „pierwszym” i „drugim” Wittgensteinem ani też na zarzuty pod adresem *Dociekań*. Interesujące wydaje się jednak zagadnienie, dlaczego właśnie tak krytykowana późna filozofia stała się jedną z podwalin współczesnej refleksji na temat wszystkiego, której koroną i zwieńczeniem jest postmodernizm.

*Cui prodest?* Wittgenstein ofiarował naukom o rzeczywistości społecznej niezwykle użyteczny, pojemny i elastyczny schemat pojęciowy, one zaś ofiarowały mu ceną dla filozofa możliwość zastosowania wyników jego procesów myślowych do czegokolwiek poza samymi owymi procesami (Bloor [1993] s. 5).

Przyjrzyjmy się, jak Wittgenstein konstruuje obraz gry językowej. Na początku *Dociekań filozoficznych*, w punkcie trzecim znajdujemy opis idealnej współpracy dwóch pracowników branży budowlanej. Jeden z nich wykrzykuje nazwy rozmaitych przedmiotów, drugi zaś dostarcza mu te przedmioty po ich „wywołaniu”. Sytuacja właściwie doskonale poddaje się analizie w kategoriach komunikatu, nadawcy i odbiorcy. Następnie otrzymujemy jeszcze kilka przykładów gier językowych, które mają nas oswoić z problemem, po czym dowiadujemy się, o co autorowi chodziło. Zatem :

"Możemy też sobie pomyśleć, że cały proces używania języka w (2) jest jedną z owych gier, za pomocą których dziecko uczy się swego ojczystego języka. Gry te nazywać będę „grami językowymi”, stosując też czasem to określenie do języka prymitywnego" (Wittgenstein [1999] s. 11).

Wydaje się to zupełnie do przyjęcia. Autor sugeruje, że nauczanie dziecka języka ojczystego może się odbywać – i w istocie odbywa się zwykle - poprzez grę w język. W świetle rozmaitych znanych teorii socjalizacji, począwszy od teorii faz zabawy i gry, nie wydaje się to bynajmniej zaskakujące. Sam Wittgenstein kilka razy powraca w Dociekaniach do wątku tresury dziecka – w przeciwieństwie do odrzucanego przezeń modelu uczenia słów z użyciem definicji ostensywnych - przyuczania do używania języka danej społeczności. Uczenie słów przez wskazywanie wytwarza w nowym członku społeczności więź asocjacyjną pomiędzy słowem i rzeczą (Wittgenstein [2000] s. 10).

Trudno w pierwszej chwili docenić wagę tego na pozór niezbyt wyrafinowanego obrazu dla całej koncepcji Wittgensteina. Wiąże się to dociekaniem co do ewentualnej architektury Dociekań. Bogusław Wolniewicz, podejmując próbę jej odtworzenia, przedstawił próbę analizy kompozycyjnej części pierwszej, dzieląc ją na dwie części, zatytułowane: Znaczenie słów i Rozumienie słów. Dwie te części zajmują się, jak widać, odpowiednio semantyką i pragmatyką. Część pierwsza części semantycznej - punkty od 1 do 37 dotyczyć mają gry językowej i nazywania (Wittgenstein [2000] s. XII).

Niewątpliwie inspiracją dla takiego podziału stało się umieszczenie przez Wittgensteina w części pierwszej słynnego „objaśnienia” znaczenia nazwy jako jej użycia w języku (Wittgenstein [2000] s. 34). Takie określenie znaczenia zwrotu pojawia się zresztą już w Niebieskim Zeszycie (Wittgenstein [1998] s. 60). Początek pierwszej części Dociekań, ze schematycznym opisem sposobu użycia języka, przygotowuje grunt dla głównej tezy Wittgensteinowskiej semantyki. Nie zapominajmy jednak, że utożsamienie znaczenia i użycia wyrażenia to element konstytutywny metafory gry językowej (Gozzi [1998]). Z części pierwszej dowiadujemy się, że gra polegająca na nazywaniu jest najprostszą i najbardziej prymitywną z możliwych gier. Nie dowiadujemy się niestety – acz jest to kusząca perspektywa – czy jest to zarazem gra pierwotna, czyli – inaczej mówiąc – czy funkcja opisowa języka jest zdaniem autora Dociekań pierwotna względem innych (Hintikka [1992] s. 471). Przyjmując, że znaczeniem nazwy jest coś poza samym językiem, czujemy się w obowiązku wyjaśnić powstanie tego rodzaju intuicji znaczenia (o ile odrzucimy tezę, że

intuicja taka, jako *najprostsza, nie wymaga wyjaśnienia*). Poczucie obowiązku pcha nas w objęcia mentalizmu, który Wittgenstein krytykuje, wychodząc z pryncypialnego założenia, które ewidentnie zachował czasów Traktatu, że znaczenie musi być funkcją (w sensie matematycznym) słowa, zaś obraz mentalny taką funkcją nie jest i być nie może (Wittgenstein [2000] s. XXII). Argumenty antymentalistyczne są w Dociekaniach liczne i trafne – nie miejsce tu, by się nimi szczegółowo zajmować, warto jednak zauważyć, że pozwalają one zrozumieć potrzebę poszukiwania poza językiem czegoś, co da się opisać funkcją matematyczną. Również z logicznego punktu widzenia trudno takiemu postulatowi odmówić zasadności. Wittgenstein bowiem najwyraźniej chce filozofować, mimo że odżegnuje się od dotychczasowej filozofii i pozostaje, jak się wydaje, przy swoim młodzieńczym poglądzie, że koniec filozofii oznaczać będzie początek bardziej systematycznej ery w dziejach ludzkości. Chęć uczciwego filozofowania musi zaś oprzeć się na przyzwoitej podstawie prakseologicznej. Liczne w historii próby filozofowania na temat powszechnej niemożności pouczają nas, że jest to zajęcie na dłuższą metę niemożliwe.

Główny ideologiczny zwrot w myśli Wittgensteina polegał, jak się powszechnie przyjmuje, na przejściu od teoretycznego do praktycznego podejścia do języka, czy też – jak kto woli, na porzuceniu filozofii języka idealnego na rzecz filozofii języka naturalnego. Oto Wittgenstein postanowił przestać mówić ludziom, jak wygląda porządne zdanie. Zdecydował, że od tej pory wszystkie zdania będą porządne, a on – niegdyś miłośnik porządku – będzie tylko rejestrował formy porządku, jakie co bardziej pomysłowym użytkownikom języka zechcą przyjść do głowy. Trudno tę zmianę stanowiska uznać za racjonalną, - jeszcze trudniej za wynik przemyśleń czy obserwacji. Mamy więc pierwszą nieuprawnioną przesłankę teorii Wittgensteina, przesłankę, która wchodzi zarazem do definicji pojęcia znaczenia i pojęcia gry językowej. Otrzymujemy koło: znaczenie jest użyciem słowa w grze językowej (Wittgenstein [2000] s. XXI), zaś jedynym – co później postaram się szerzej omówić – kryterium jedności gry językowej, zapewniającym możliwość wytyczenia jej granic w chaosie języka jest jedność sposobu użycia słów.

Zainteresowanie językiem naturalnym kazało Wittgensteinowi przyjrzeć się ludziom w trakcie interakcji językowej. Jego filozofia wiąże ściśle język z kontekstem jego użycia. Podkreśla płynność i zmienność okoliczności, w których język może się znaleźć, wielość jego możliwych zastosowań. Koncepcja gry językowej dopuszcza elastyczność i dowolność w tworzeniu reguł, wygodę i oryginalność w wykonywaniu ruchów w grze (Gozzi [1998]). Pozwala w ten sposób od razu sformułować kluczowe, moim zdaniem, pytanie: czy metafora

gry dopuszcza rzeczywiście swobodę gracza, czy też tylko umożliwia postrzeganie jego zachowania w pewien sposób, w danym razie jako swobodnego?

Fakt, że kategorię „gry” uważa się za metaforyczną, zdaje się świadczyć, że silniejsza jest tendencja do uznawania jej, zgodnie z duchem późnej filozofii Wittgensteina, za narzędzie opisu raczej niż adekwatne odzwierciedlenie rzeczywistości. Również zdecydowana niechęć Wittgensteina do przenoszenia analizy gry językowej na płaszczyznę psychologiczną skłania ku opinii, że on sam traktował – w każdym razie początkowo – grę jako metaforę stosunków międzyludzkich w ich warstwie językowej, nie sugerując bynajmniej, że najbardziej trafną figurą jednostki ludzkiej staje się w związku z tym „gracz”. Być może obserwacja nie była dlań tak istotna, jak zdaje się nam sugerować, przytaczając wiele quasi-empirycznych ilustracji swoich tez. Prawdopodobnie jednak to rzeczywiście analiza sytuacji komunikacyjnych kazała mu stwierdzić, że jedyną pewną rzeczą, jaką możemy powiedzieć na temat słowa, jest opis jego użycia przez uczestników interakcji językowej. Tworząc pojęcie gry Wittgenstein stworzył zatem przede wszystkim metodę. Studiowanie zjawisk językowych powinno się zaczynać od ich wersji prymitywnych, w których cel i funkcjonowanie wyrazów leżą jak na dłoni (Wittgenstein [2000] s. 10). Trudno, po raz kolejny, odmówić dużej dozy zdrowego rozsądku założeniu, że sposób funkcjonowania wyrazów nie zmienia się zależnie od stopnia komplikacji kontekstu ich występowania – że więc słowo jest instrumentem dopuszczającym znaczny stopień dowolności użycia, ale nie uniwersalnym. Trudno jednak, z drugiej strony, uznać to założenie za w pełni uprawnione. Wittgenstein podaje liczne przykłady dowodzące, że sposób użycia danego słowa może różnicować się niemal w nieskończoność, brak natomiast w jego dziele podstaw dla określenia granic tego procesu. Brak tego, co nietrudno znaleźć w dziełach współczesnych postmodernistów – konsekwentnego przekonania, że język jest z natury swojej amorficzny, dlatego że słowo nie posiada żadnych cech, których nie nadawałoby mu użycie. Zauważmy więc mimochodem, że Wittgenstein widzi jednak możliwość podania reprezentatywnego przykładu gry językowej – pomimo wszystkich zastrzeżeń, którymi obudowuje tę możliwość.

Przed nami jednak jeszcze jedno znaczenie terminu „gra językowa”, które dało asumpt do najrozmaitszych kulturalistycznych wybryków pod szyldem *Dociekań*:

„Grą językową” nazywać też będę całość złożoną z języka i czynności, w które jest on wpleciony” (Wittgenstein [2000] s. 12).

Dwie tezy wyrażone w powyższym zdaniu zasługują na uwagę. Po pierwsze, język i czynność, w którą jest wpleciony, stanowią pewną całość. Nie bardzo wiadomo, jaka miałyby być natura tej całości, najprawdopodobniej jednak należy wstępnie założyć, że związek języka z daną sytuacją nie jest przypadkowy. Istnieje więc pomiędzy sytuacją i pojawiającym się w jej kontekście językiem. Po drugie, język „wplata się” w czynność. Wyrażenie co najmniej niejasne, choć intuicyjnie rozumiemy, że jeśli wpleciony, to niełatwo go wyplątać.

Wittgenstein podaje wiele - niekiedy dość zastanawiających – przykładów gier językowych. Pozwala nam to stwierdzić, że gra językowa w żadnym razie nie jest tożsama z żadnym ze znanych typów zachowań językowych – jest to pojęcie o zdecydowanie szerszym zakresie. Być może rzuca to nieco światła na naturę gry, jako całości złożonej nie tylko z języka, ale także z aktywności człowieka. W tym sensie filozofia Wittgensteina jest – o ile dotyczy gier językowych – także filozofią życia.

Sformułowanie metafory gry językowej przez Wittgensteina dalekie jest od wyczerpania tematu i dopuszcza wiele interpretacji, rozszerzających zakres zastosowania kategorii gry na coraz to nowe obszary, zwłaszcza w dziedzinę nauk społecznych (Gozzi [1998]). Metafora gry ma bowiem dwa oblicza. Z jednej strony, pozwala ona ujmować znaczenie słowa jako jego użycie – i to jest rozumienie, które można określić jako „filozoficzne”. Z drugiej jednak strony, powiedzieć, że ktoś „gra w grę” czy, nie daj Boże, „w gierki” („nie pogrywaj ze mną!”, „stop playing your games with me!”) znaczy przywołać bezpośrednio dostępne rozumienie tego słowa, ukształtowane przez jego użycie w języku życia codziennego – które nazywać będę „zdroworozsądkowym”. Gracz – to nie brzmi sympatycznie. Nauki społeczne dawno już jednak – mniej więcej w czasach Hobbesa – rozpoczęły trudny proces wydobywania się z zawinionej dziecięcej ufności w sympatyczną w gruncie rzeczy naturę człowieka. Człowiek nie jest ani łagodnym dzikusem, ani zimno kalkulującym *homo oeconomicus*. Człowiek jest istotą indywidualną żyjącą w społeczności innych istot sobie podobnych. Podkreślane w wielu teoriach socjologicznych i psychologicznych rozdwojenie natury ludzkiej na część społeczną i część indywidualną można uznać za próbę metaforycznego zdania sprawy z faktu, że człowiek, z natury bynajmniej niespołeczny (o czym świadczą zachowania egoistyczne, czy wręcz aspołeczne), zarazem z różnych względów nie może i nie chce żyć samotnie. Socjologia wciąż poszukuje normatywnego rozwiązania kwestii: ile społeczeństwa w społeczeństwie, zważywszy że, jak to celnie zaobserwował Durkheim, zarówno zbyt mały, jak i zbyt duży udział pierwiastka społecznego w życiu jednostki kończy się nieuchronnie jej psychiczną i fizyczną dezintegracją. Teorie form uspołecznienia zajmują się albo opisem funkcjonowania

społeczeństwa jako takiego, albo zagadnieniami w gruncie rzeczy ekonomicznymi, albo też - psychologią. Każdej z nich zarzucić można niepełne, wybiórcze i abstrakcyjne traktowanie współżycia realnych jednostek – zarzucano im oderwanie opisu form interakcji społecznej od faktycznego funkcjonowania jednostek w tych interakcjach. Stąd też nowe szkoły socjologiczne, wywodzące się z nurtu interakcjonistycznego, łatwo przyswoiły filozofię *Dociekań* i zinternalizowały Wittgensteina niemal całkowicie. Metafora gry okazała się niezwykle płodna dla analiz różnych typów społecznych interakcji. Stało się tak prawdopodobnie, dlatego że dostarcza ona rzadkiej okazji, by zrozumieć oba powyżej wymienione sposoby istnienia człowieka w świecie – indywidualny i społeczny - zarazem i w formie takiej, w jakiej sądzimy, że występują w rzeczywistości – jako jednolitą całość.

Wszelkie nurty socjologicznego badania człowieka-w-swiecie bazują na podstawowej obserwacji – tak to chyba mimo wszystko należy określić – filozofii idealistycznej, zwłaszcza późniejszej, fenomenologicznej wersji idealizmu.

Przede wszystkim, każe nam ona uznać, że człowiek z konieczności jest twórczy, oraz że jego wytwory podlegają reifikacji i obiektywizacji \*. Muszę od razu zastrzec, że nie widzę konieczności t r w a ł e g o obiektywizowania się danego wytworu w momencie transmisji dla skuteczności tejże. Ulega on oczywiście reifikacji o tyle, że traktowany jest jako przedmiot - w tym sensie wszelki przekaz - dotyczący obiektów nieprzekładalnych na język konkretów - nie może się obyć bez hipostazowania traktowanego abstraktu. Dodatkowym założeniem jest jednak, że człowiek akceptuje skutecznie tylko treści, które uznaje za obiektywne, co przyjmując traci się z oczu jedną z zasadniczych motywacji ludzkiego działania, a mianowicie pragnienie jego skuteczności. Treść przekazu kulturowego (czy językowego) dotyczy strategii radzenia sobie z rzeczywistością – jak przyjmuje wielu autorów, wśród nich Wittgenstein twierdzący, że język to „narzędzie funkcjonujące w przyrodzie” oraz instrument wpływu społecznego" (Wittgenstein [2000] s. 194-195). Jednostka ma przyjąć daną strategię na stałe - nie da się więc utrzymać wszystkich przyjętych założeń nie dodając kolejnego, a mianowicie że przekaz dotyczy tego, co okazało się już skuteczne (przy rozumieniu skuteczności w ramach danego dyskursu, rzecz jasna). W tym punkcie nasuwa się nie do końca, być może, uprawniona intuicja naturalistyczna – przekaz dotyczący sprawdzonych procedur, połączony z uczeniem się, spotyka się również wśród zwierząt – zarówno w warunkach laboratoryjnych (różne warianty eksperymentów na

---

\* Przez obiektywizację rozumiem tu wyodrębnienie się wytworów człowieka z jego świata i stopniowe ich uniezależnienie od procesów twórczych, które doprowadziły do ich powstania. Klasyczną analizę procesu obiektywizacji prezentują Berger i Luckmann w pow. cytowanym dziele na przykładzie instytucjonalizacji.

szympankach i bananach), jak i w naturze (do znudzenia powtarzane przykłady mew rozbijających muszle o skały). Nie jest do zaistnienia tego przekazu potrzebna ani kultura, ani jej teoria – potrzebne są tylko mechanizmy znane z prac Pawłowa. W tym sensie zatem można doskonale bronić czegoś w rodzaju minimalistycznego modelu socjalizacji, twierdząc, że fakt przyjęcia danej procedury ze względów pragmatycznych może zajść bez obiektywizacji przedmiotu przekazu (w jakimkolwiek innym sensie, niż chwilowa i wynikająca z natury komunikacji obiektywizacja jej przedmiotu). Zatem do skutecznego radzenia sobie z rzeczywistością kultura nie jest konieczna – o ile konsekwentnie traktujemy człowieka jako działającego celowo. W tym momencie dziękujemy socjologii humanistycznej – Wittgenstein się broni. Oczywiście, mówiąc o „celowości” i „skuteczności działania”, należy przypomnieć sobie niewdzięczne pojęcie „dyskursu” i ostatecznie pożegnać się z jakimikolwiek, w szczególności zaś z naszymi własnymi, kryteriami celowości lub skuteczności.

Dwa rozumienia metafory gry, filozoficzne i zdroworozsądkowe, obrazują stan ducha człowieka nieświadomego filozoficznych zawłości. Staje on – nieuleczalny aktywista - wobec konieczności działania, które może realizować dwojako. W świecie, który uznaje za obiektywny i pozbawiony zdolności manipulowania człowiekiem, aktor określa cele i wybiera sposób działania, dokonując bilansu dostępnych środków, przy czym nie ma zwykle możliwości zmiany zasobu środków i obniżenia kosztów. Mając dane reguły i stan początkowy, dokonuje ruchu – w tym przejawia się indywidualistyczny aspekt ludzkiej aktywności, jest to zarazem jej element ujmowany przez znaczenie filozoficzne. W interakcji międzyludzkiej natomiast – przy analogicznej, w gruncie rzeczy indywidualistycznej, strukturze działania – rozkład kart po rozdaniu nie determinuje wyniku gry, bowiem inni jej uczestnicy swoim działaniem mogą zmieniać stan naszych aktywów. Skłaniamy ich do tego za pomocą rozmaitych środków perswazji – tutaj wkracza opis zdroworozsądkowy. W momencie, gdy zaczynamy manipulować otoczeniem, któremu przyznajemy zdolność opierania się manipulacji – stajemy się zwierzęciem społecznym. Czysta postać działania pierwszego typu odpowiadałaby grom typu „oczko” lub „wojna” – przebieg gry daje się wyczerpująco opisać w kategoriach rachunku prawdopodobieństwa, w efekcie doskonale można grać w wojnę samotnie, drugie działanie przypomina natomiast partię brydża – z maksymalnie jednym dziadkiem. Nie znaczy to oczywiście, że nie można partii brydża opisać jako układu zdarzeń o określonym prawdopodobieństwie – nie jest to jednak opis, który narzucałby się jako najbardziej trafny. Za bardzo trafne natomiast większość brydżystów

uznałaby twierdzenie, że brydż jest grą komunikacyjną, w której umiejętność manipulacji przeciwnikiem odgrywa znaczącą rolę.

Powyższa skromna typologia działania człowieka w świecie nie pretenduje oczywiście do zupełności. Przede wszystkim, czysto subiektywne wydaje się postawienie granicy pomiędzy tym, co posiada i tym, co nie posiada zdolności manipulowania otoczeniem – czyli określenie, z kim gramy w oczko, a z kim w brydża. Z punktu widzenia wyznawcy religii animistycznej każdy element rzeczywistości wokół niego może się okazać potencjalnym źródłem bodźców manipulacyjnych (co należy chyba rozumieć jako zdolność samodzielnego działania celowego). Warto na marginesie zauważyć, że z punktu widzenia obserwatora zewnętrznego, który uznaje za jedyną przestrzeń działania człowieka świat kreowany wspólnie z innymi ludźmi, pierwszy typ działania albo nie istnieje, albo ma również charakter manipulacyjny. Innymi słowy, na gruncie radykalnego konstruktywizmu nie można obalić tezy o wyłącznie manipulacyjnym charakterze komunikacji językowej.

Porzucając meandryczne rozważania z zakresu ontologii społecznej, warto na koniec zaznaczyć, że dwa powyżej scharakteryzowane rozumienia pojęcia gry – filozoficzne i zdroworozsądkowe – nie wyczerpują prawdopodobnie metafory gry językowej (Gozzi [1998]). Wystarczą jednak w zupełności, by wytworzyć w badaczach przekonanie, że oto otrzymali narzędzie, które rozwiązuje podstawowy problem metodologiczny: jak pogodzić pozytywny i negatywny aspekt działania człowieka w interakcji z innymi ludźmi – inaczej mówiąc, jak zanalizować przebieg realizacji motywów pragmatycznych w konkretnych interakcjach.

Liczne teorie lansują pogląd, iż użycie języka może być wyłącznie manipulacyjne. Dotyczy to zarówno teorii powszechnie już akceptowanych, uznawanych nawet za klasyczne, jak choćby Goffmanowska teoria ról społecznych, jak i licznych wariantów konstruktywizmu społecznego. Oczywiście, nie da się tych stanowisk porównać, nie ustalwszy wcześniej znaczenia terminu "manipulacja", a bardzo trudno powiedzieć, co mógłby on oznaczać dla Wittgensteina. Utrzymuje on, że bez języka niemożliwie byłoby porozumiewanie się i wpływanie na innych (więc działanie). Zdaje się zatem przyznawać, że człowiek jest istotą manipulującą oraz, że warunkiem koniecznym manipulacji, jest używanie języka, traktuje to jednak jako immanentne i nie podlegające wartościowaniu cechy odpowiednio człowieka i pewnych jego gier. Za manipulację językiem w sensie pejoratywnym – za kłamstwo, uznałby on z pewnością przedstawianie dylematów gramatycznych jako rzeczywistych problemów – to jednak niewiele nam jeszcze wyjaśnia. Faktem jest jednak, że metafora „języka na wakacjach” stwarza szerokie pole dla dywagacji. Jak to z pasją podkreśla Harry Redner w



swojej pracy *Malign masters*, Wittgenstein lansuje w gruncie rzeczy *labour theory of language* (teorię języka pracującego?). Przykłady gier językowych, które podaje, przypominają sytuację współdziałania (nie są więc podobne do gier w jakimkolwiek znanym sensie). Również definiowanie znaczenia jako użycia słowa wymaga przyjęcia, że znaczeniem słowa jest jego praca, zatem słowo leniwe, takie jak słowo filozofa, traktowane jest z założenia negatywnie, powinno się je zmusić, by wzięło się do pracy i zajęło na przykład opisywaniem. Wittgenstein wierzy, zdaniem Rednera, w język potoczny i gorąco potępia lenistwo języka filozofii (Redner [1997] s. 77-79).

Pomijając ewentualne nieścisłości i ogólny ton wypowiedzi Rednera, zwraca on uwagę na istotny problem. Jak w ogóle zoperacjonalizować Wittgensteinowskie pojęcie gry językowej? Jeśli chcemy odrzucić jako niemerytoryczny zarzut, iż Wittgenstein, rozciągając język na cały obszar kultury, redukuje go zarazem, czy wręcz degraduje do poziomu gry, musimy ustosunkować się do tezy, że między grami, jakie opisuje Wittgenstein, a językiem istnieje oczywista nieciągłość (Redner [1997] s. 71). Rzeczywiście, metoda Wittgensteina polega w gruncie rzeczy na niezbyt dobrze się tłumaczącym uogólnieniu. Pokazuje się nam kilka prostych gier, które wszystkie Redner uważa za przykłady zastosowania prymitywnego systemu sygnałów – zatem, jego zdaniem, coś nieskończenie mniej bogatego, niż język. Cały pomysł gry językowej to z gruntu nieprawidłowe, uproszczone i trywialne przedstawienie natury ludzkiej komunikacji (Redner [1997] s. 99). Jednak poza obiekcjami natury oczywiście ideologicznej można wysunąć przeciw totalizmowi gier językowych także i bardziej zdroworoządkowe argumenty, jak to czyni Jaakko Hintikka. Zwraca on uwagę, że zdarzają się takie wystąpienia słowa, które nie mogą być uznane za ruch w jakiejkolwiek grze językowej. Najbardziej przekonujący z jego przykładów odwołuje się do typu wypowiedzi stanowiących relację w mowie niezależnej, typu „Hamlet powiedział: „Idź do klasztoru”. Nie mamy tu do czynienia z grą: „wysyłanie do klasztoru”, ale z grą „relacjonowanie wypowiedzi Hamleta”. Trudno niemniej zaprzeczyć, że użycie słowa „idź” w tym kontekście jest sensowne i informuje nas o pewnej regule gramatycznej. Zarazem jednak zgodzimy się, że słowa „idź do klasztoru” nie są ruchem w grze językowej „relacjonowanie wypowiedzi Hamleta”. Mamy zatem pokąźną klasę sytuacji, w których sensowne użycie słowa nie jest jeszcze ruchem w grze językowej (Hintikka [1992] s. 472-473). Podobnego przykładu dostarcza nam sam Wittgenstein, pisząc w *O pewności*: „Gdyby zdanie „Wiem, że to jest drzewo” wypowiedziano poza grą językową, to mogłoby też być cytatem (może z podręcznika języka polskiego)” (Wittgenstein [1993] s. 79). Słowa występujące w powyższych przykładach w cudzysłowach nie mogą być użyte w funkcji manipulacyjnej, jako

że nie są posunięciem w grze, a tylko wykonując ruch (bądź rezygnując z jego wykonania) możemy manipulować przeciwnikiem. Zwróćmy uwagę – gry językowe, w przeciwieństwie do większości gier – nazwijmy je tak – rekreacyjnych, mogą być grami o sumie niezerowej, co znaczy, że najprostsza intuicja gry nie będzie nam pomocna w zrozumieniu tego zjawiska. Przykładu gry „kooperacyjnej” o sumie niezerowej dostarcza Hintikka w postaci „gry dialogowej” (Hintikka [1992] s. 431 i dalej). Autor ten zwraca jednak uwagę na zupełnie inny aspekt problemu gier językowych, a mianowicie na fakt, że można je rozumieć jako sposób na powiązanie fragmentów języka z pewnymi aspektami świata. Relacja reprezentowania elementu świata przez znak językowy nie jest naturalna, nie można wywnioskować jej istnienia z obserwacji świata. Jediną możliwość powstania relacji reprezentowania stwarza poddane regułom działanie ludzkie – czyli gra językowa. Funkcją gier językowych jest tworzenie i podtrzymywanie relacji reprezentowania rzeczywistości poprzez język (Hintikka [1992] s. 445, 458).

Gra językowa pośredniczy więc między językiem a światem, zapewniając odniesienie w świecie wyrażeniom języka – co stawia nas przed kwestią zakresu reprezentacji świata w języku. Zamiast badać, czy każdemu elementowi języka odpowiada (w jakiejś grze językowej) jakiś element rzeczywistości – jak zdaje się proponuje Hintikka, można zastanowić się, czy każdemu elementowi rzeczywistości odpowiada (w jakiejś grze językowej) jakiś element języka. Jeśli bowiem można znaleźć fragment rzeczywistości, który nie byłby reprezentowany w języku, znaczyłoby to, że jest taka dziedzina aktywności ludzkiej – czy, bardziej górnolotnie, ludzkiego życia, w której ludzie nie są graczami – w każdym zaś razie nie grają w gry językowe. Należy przy tym pamiętać, że mowa tu o rzeczywistości w powyżej przedstawionym rozumieniu socjologii humanistycznej, nie o „twardej” rzeczywistości obiektywnej.

Pytanie o zasięg języka – w jego niebanalnej, socjologizującej wersji – to w gruncie rzeczy wyraz postawy tyleż liberalnej względem ludzkich sposobów używania języka, co bezsilnej w zakresie prób ich pojęciowego uporządkowania. Nie kładziemy żadnej tamy ludzkiej aktywności w dziedzinie tworzenia środków językowych. Pytamy natomiast, czy w każdą ludzką aktywność jest z konieczności wpleciony język – czy zatem każde działanie jest z konieczności, na mocy definicji Wittgensteina, grą językową. Kwestia ta sprowadza się do pytania, czy całość rzeczywistości jest objęta reprezentacją językową – czy zatem język jest właściwym przedmiotem badania wszystkich nauk społecznych.

Koncepcja języka jako gry pozwala w moim przekonaniu docenić aspekt społecznego użycia języka szczególnie istotny dla teorii decyzji, teorii gier właśnie oraz niektórych

kierunków w psychologii społecznej, tych mianowicie, które posługują się kategorią gry. Kierunki te spokrewnione są zresztą z filozofią Wittgensteina pod wieloma względami, z których na czoło wysuwa się antypsychologizm i behawioryzm. Co do behawioryzmu Wittgensteina, chciałabym podkreślić tylko jeden aspekt tego typu postawy – wszelkie wnioski na temat mechanizmów ludzkiego zachowania ma tu postać zwrotnego opisu stanu informacji podmiotu. Eksperymentalnie mogłoby to wyglądać następująco: podajemy uczestnikowi gry informację i obserwujemy, jak wpływa ona na jego ruchy w grze. Zakładamy więc związek, którego możemy oczywiście nie nazywać przyczynowym, a który ekstrapolujemy na dalsze zachowania uczestnika gry.

Behawioryzm w badaniu uczestnika gry językowej przydaje się na dwóch etapach analizy: po pierwsze, przy badaniu socjalizacji językowej, po drugie zaś, przy obserwacji zależności pomiędzy gramatyką – w sensie Wittgensteinowskim - użytowanego języka a zachowaniami podmiotów nim się posługujących. Metafora gry językowej pozwala więc zdać sobie sprawę z kapitalnej z punktu widzenia behawiorysty funkcji języka. Działa on mianowicie jako narzędzie upraszczania świata. Pojęcia języka służą do klasyfikacji i porządkowania rzeczywistości. Każda klasyfikacja jest jednak uproszczeniem potencjalnie nieskończenie złożonej rzeczywistości, zakłada bowiem stosowanie kryteriów właściwych dla danego typu zjawisk wobec każdego egzemplarza tego typu. Typologia musi opierać się na wyodrębnieniu cech istotnych. Z kolei, jako że wobec uprzedniego nabycia typologii właściwej danemu językowi w socjalizacji proces poznawczy jest przez tę typologię predeterminowany, klasyfikacja nie dopuszcza do powstania rozbieżności pomiędzy kryteriami języka a faktami. Stosowanie językowych typologii polega więc na adaptacji doświadczenia do posiadanego aparatu pojęciowego i automatycznej eliminacji, jako niewiarygodnych, danych niezgodnych z tym aparatem (Barnes w: *Mocny program* [1993] s. 118). Z punktu widzenia ekonomiki działania – czysty zysk. Tak w skrócie przedstawia się jedna z podstawowych funkcji języka w ujęciu socjologicznym. Jak widać, jest to podejście dość problematyczne. Nawet przyjąwszy wbrew zdrowemu rozsądkowi, że eliminacja szumów informacyjnych jest zawsze stuprocentowa, pozostajemy z dręczącym pytaniem o innowacyjność, czyli o sytuację, gdy zbiór aktualnie otrzymywanych danych jest nie sprzeczny, ale rozłączny ze zbiorem danych przewidzianych w siatce pojęciowej. Wittgenstein szczegółowo o tych kwestiach nie rozprawia, natomiast z pewnością można mu przypisać pogląd, że język jest skutecznym narzędziem upraszczającym działanie w świecie.

Raymond Gozzi w swoim cytowanym już artykule pod tytułem *Is Language a Game* relacjonuje, w jaki sposób doznał olśnienia co do tej funkcji języka. Czytał oto pewnego razu

artykuł na temat nowej generacji tak zwanych *role-playing games*. Wydawały się one doskonale – tworzyły bardzo złożone w porównaniu z wersją tradycyjną scenariusze, pełne, wielowymiarowe postacie, umożliwiały wielotorowy rozwój akcji. Twórcy owych nadzwyczajnych gier stanęli jednak twarzą w twarz z nieoczekiwanym problemem.

Okazało się, że jeśli dopuścimy zbyt dużą dozę interaktywności, jeśli pozwolimy użytkownikowi kontrolować zbyt dużą część akcji gry, tracimy kontrolę nad fabułą. Gra i jej bohaterowie wydają się ginać w chaosie możliwości. Zmuszeni więc jesteśmy uprościć historię i bohaterów, aby zapewnić satysfakcjonujące rozwiązanie. Wówczas jednak tracimy złożoność, która czyni zabawę atrakcyjną, zaś postacie biorące w niej udział - interesującymi. Doprowadziło to niektórych twórców do przekonania, że istnieje przepaść nie do przekroczenia pomiędzy „grą” (game) i „historią” (story). Ta ostatnia wymaga złożonych wewnętrznie bohaterów, wielości możliwych wyborów, wątków fabularnych i tym podobnie. Gra natomiast przeciwnie – wymaga jednowymiarowych postaci dążących do jasno określonych celów w schematycznie opisanych warunkach.

Rozróżnienie gry i historii można, zdaniem Gozziego, zastosować do określenia granic metafory gry językowej. Z pewnością nie ma wyraźnej granicy pomiędzy grą i historią, intuicyjne wyczucie różnic miałyby jednak w zupełności wystarczyć do skrytykowania pomysłu Wittgensteina. Gra – to sytuacja uproszczona. Na tym przede wszystkim polega urok gry: zapewnia ona jasne i ograniczone alternatywy. Daje niedwuznaczny wynik – inaczej niż tak zwany „świat realny”, w którym alternatywy są nie do końca jasne, a wynik działania – niepewny (Gozzi [1998]).

Wittgenstein w swej późnej filozofii porzucił, jak wiadomo, wizję języka jako lustro odbijającego rzeczywistość. Język stał się instrumentem działania w świecie. Potrzebujemy języka, by móc się odnaleźć w dwuznacznym (czyli potencjalnie niezrozumiałym) świecie naszego życia. Zdaniem Gozziego – upraszczanie życia poprzez upraszczanie użycia języka jest pewnym środkiem pomocniczym, o ograniczonej skuteczności. Wittgensteinowską opozycję lustro-gra chciałby on zastąpić nową parą: gra-historia - takie przeciwstawienie wydaje mu się bardziej produktywnie na dłuższą metę (Gozii [1998]). Wspominam o tym pomyśle krótko, wydaje mi się bowiem bardzo oryginalny (choć w pewnym sensie zbliżony do niektórych postmodernistycznych pomysłów na „narrację”). Czy jednak ludzie są zdolni do działania na zasadach właściwych historii? Przychodzi mi na myśl ilustracja literacka tego pytania: w jednej z powieści angielskiego pisarza Terry’ego Pratchetta bohaterowie (którzy w chwili, gdy to piszę, są od pewnego czasu także bohaterami komputerowej *role-playing game*), postanawiają – bo czemuż by nie – rzucić monetę.

- „Orzeł czy...coś w rodzaju ryby z nogami ? – zapytał Dwukwiat.
- Krawędź – odpowiedział Rincewind.”

*Na czym polega komizm tej wypowiedzi? Otóż w grze nigdy nie wypada krawędź.*

Radzenie sobie ze światem, funkcjonowanie w świecie odbywa się niejako „w trybie gry”, co Wittgensteinowskie ujęcie doskonale chwytą. Oprócz pozytywnego aspektu, związanego z pragmatyczną motywacją gracza, gra ma też aspekt negatywny – gracz gra w określoną grę o określonym typie wygranej. Nie wyłuszcza on swoich intencji, nie informuje, jaki jest rzeczywisty powód, dla którego gra. Stosując termin „gra” w rozumieniu zdroworozsądkowym, chcemy między innymi powiedzieć, że gracz gra, bo nie może osiągnąć swoich celów prostszą metodą. Uzyskujemy tym samym odpowiedź na pytanie o trafność zewnętrzną opisu interakcji międzyludzkiej w kategoriach gry. Ludzie – wbrew dobrej woli rozmaitych obserwatorów ich zachowań – są graczami. Działanie człowieka jest ciągłą walką z nadmiarowością otoczenia. Można to poprzeć argumentami biologizującymi, można jednak także filozoficznymi. Jeżeli przyjmiemy, że pierwotna motywacja człowieka jest pragmatyczna, musimy się zgodzić, że w ramach właściwej sobie racjonalizacji działania będzie on starał się odnaleźć taką metodę selekcji bodźców płynących z otoczenia, by zapewnić sobie możliwie prostą i pewną wygraną – czyli realizację celu. Tresura socjalizacyjna uczy go, jak się osiąga cele. Gra językowa to – z pewnego punktu widzenia – obszar możliwości realizacji celów za pomocą znanych środków. Aby jednak gra mogła się toczyć, aby zapewnić stały dopływ graczy, gra musi być pod każdym względem grą. Musi więc mieć uczestników, cel, ale musi też mieć reguły. Regułą gry w orła i reszkę jest, że jeżeli rzucimy monetę, chwycimy ją w locie i ułożymy płasko na dłoni, to po wykonaniu przez graczy ruchu – złożeniu deklaracji o treści „orzeł” lub „reszka”, gdy odsłonimy monetę, będziemy mogli z całą pewnością orzec, kto jest zwycięzcą. Jeśli mamy przypadkiem do czynienia z grą o sumie niezerowej (powiedzmy z grą na giełdzie, jak to przyznają wszyscy prócz antyglobalistów), w dalszym ciągu musimy być w stanie podać kryterium wygranej.

Oczywiście, rozmaite teorie systematycznie podważają metafizyczne przecucie, że kij ma dwa końce. Nie wpływa to jednak na naszą strategię działania w świecie. Nawet uznając, że dotychczasowe sukcesy w rutynowo rozgrywanych grach językowych są tylko wynikiem pomyślnych zbiegów okoliczności, nie przestajemy dalej grać, tak jak nam rutyna nakazuje. Witamy się ze znajomymi, na żądanie podajemy policjantowi dowód osobisty i odpowiadamy na pytanie „która godzina”. I naturalnie zdarza się, że znajomy mija nas nie odkłoniwszy się, policjant (jak się to ostatnio zdarzyło jednemu z moich przyjaciół w pewnym kraju ościennym) oddaje nasz dowód i odchodzi unosząc ukrytą w nim przez nas

przezornie znaczną sumę w gotówce, a miły młody człowiek, dowiedziawszy się, że jest za pięć szóstą, próbuje nas następnie przekonać, że Królestwo Boże jest znacznie bliższe, niż się nam za dziesięć szóstą wydawało. Wszelkie tego typu potknięcia nie podważają jednak naszego zdroworozsądkowego zaufania do reguł pewnych gier.

Warto też zauważyć, że zdaniem Wittgensteina powody, dla których w pewne gry gramy, inne zaś nie, nie mają nic wspólnego z takimi cechami ludzkiego działania, jak racjonalność. Wspomniałam już, że myśl późnego Wittgensteina wywarła wielki wpływ na postmodernistów. Otóż, ojcem postmodernizmu jest bez wątpienia relatywizm. Relatywistami teoriopoznawczymi – zdeklarowanymi bądź dorozumianymi – są niemal wszyscy pierwszoplanowi postmoderniści, z Richardem Rorty’em na czele. Ten ostatni, co nietrudno stwierdzić, czyta Wittgensteina przede wszystkim jako relatywistę (Soin [2001]).

Wittgenstein jest zatem relatywistą. Nie jest to jednak relatywizm dowolny, absolutny czy też – powiedzmy – spontaniczny. Relatywizm wymaga uzasadnienia, a koncepcja gry językowej takie uzasadnienie oferuje.

Pojęcie „gry językowej”, najbardziej popularne pojęcie „późnego” Wittgensteina, uważa się często za główny element jego domniemanego porzucenia prawdy. Z pojęciem tym wiąże się przede wszystkim sprzeciw wobec redukcji sensu zdania do obrazowania i potencjalnej prawdziwości, zatem sprzeciw wobec takiego ujęcia sensu zdania, jakie Wittgenstein rozwinął w *Traktacie*. Przypomnijmy krótko – zgodnie z tezą tego dzieła, tworzymy sobie obrazy faktów (*Tractatus logico-philosophicus* 2.1), które są modelami rzeczywistości (2.12), zgodnymi z nią lub niezgodnymi, trafnymi albo błędnymi, prawdziwymi wreszcie lub fałszywymi (2.21). Jak widać, jest to koncepcja pozytywistyczna, zakładająca jednolite odniesienie obrazu i możliwość jednoznacznego rozstrzygnięcia o jego prawdziwości lub fałszywości po porównaniu go z rzeczywistością (2.223). Skoro myśl jest także obrazem – „logicznym obrazem faktów” (3), zaś zdanie jest zmysłowo postrzegalnym wyrazem myśli (3.1), a myśl jest to zdanie sensowne (4) (Wittgenstein [1997] s. 9, 11, 12, 20) znaczy to, że każde zdanie sensowne musi być opisywalne w kategoriach prawdziwości i fałszywości.

Zgoła inaczej mają się sprawy w *Dociekaniach*. Przede wszystkim pojęcie „prawdziwości” schodzi na plan dalszy. Pojawiają się w związku z tym liczne problemy, jako że jedną z podstawowych cech „lustrzanej” teorii prawdy jest znaczne uproszczenie wszystkich kwestii semantycznych. Zamiast obiektywnego kryterium sensowności pojawia się więc kryterium, które nie ustanawia i nie wymaga żadnej relacji pomiędzy językiem a rzeczywistością, którą nazywamy obiektywną. W *Traktacie* język jest definiowany jako ogół

zdań – (4.001). W *Dociekaniach* język – to całość złożona z gier językowych. Opisuując rozmaite gry językowe, Wittgenstein nie używa terminu „prawdziwość”. Definicja znaczenia jako użycia wyklucza przyznanie zarazem prawdziwości podstawowej roli w określaniu sensowności zdania.

Maciej Soin, w swoim cytowanym już referacie, głosi, że zasadnicza różnica pomiędzy wczesnym a późnym Wittgensteinem nie leży w odrzuceniu lustrzanej teorii języka, ale w wyczerpaniu się programu analizy logicznej języka, wobec niemożności rozwiązania wspólnej Wittgensteinowi i Kołu Wiedeńskiemu kwestii zdań elementarnych.

"Sprawa [rzekomego odrzucenia pojęcia „prawdy”] wiąże się z jedną z najważniejszych różnic między „wczesnym” i „późnym” Wittgensteinem, mianowicie generalnym odwrotem od koncepcji formy logicznej na rzecz autonomicznej gramatyki. Nie możemy znów wchodzić w szczegóły tego zwrotu, ale wystarczy powiedzieć, że wraz z załamaniem się programu analizy logicznej, która miała doprowadzić do zdań elementarnych, Wittgenstein odrzucił również przekonanie, że możliwość obrazowania opiera się na nieziennej formie logicznej wspólnej obrazowi i temu, co obrazowane, zdaniu i rzeczywistości (por. np. TLP 2.18 i PG 163). Miejsce opróżnione przez formę logiczną zajęła gramatyka, a wraz z nią decydującego znaczenia nabrało przekonanie o autonomii czy arbitralności języka" (Soin, [2001]).

Gramatyka danego słowa to sposób jego używania w grach językowych. Wittgenstein, wychodząc z założenia, że znak pracuje tylko w systemie gramatycznym (Wittgenstein [1993] s. 132), twierdzi, że gramatyka to rachunek – czysty rachunek, nie jego zastosowanie do rzeczywistości, zaś sensem zdania jest jego rola w tym rachunku (Wittgenstein [1993] s. 312, 130). Przejęcie się problematyką nabywania kompetencji językowej prowadzi Wittgensteina do konkluzji, że o ile użycie zdania w szeregu myśli jest częścią rachunku i jest na gruncie jego reguł uzasadnione, to sam rachunek nie tylko nie jest, ale nie może być uzasadniony. Wracamy do paradoksów samoodniesienia – powód, dla którego używamy tego, a nie innego rachunku, postępując zgodnie z jego regułami, można podać tylko w tym samym rachunku. Otrzymujemy więc wniosek, że nasze przyjęcie danej gramatyki jest oczywiście „równie arbitralne, jak lęk przed ogniem” (Wittgenstein [1993] s. 110, 113) jednak sama gramatyka jest doskonale arbitralna. Nie rzeczywistość określa reguły użycia słów, lecz gra językowa. Kształt gry nie jest związany z rzeczywistością. Ostatecznie język jest charakteryzowany

przez czynności językowe, a nie przez ich odniesienie do świata pozajęzykowego (Wittgenstein [1993] s. 193).

Tej ostatniej konsekwencji nie dostrzegają przedstawiciele wielu nurtów Wittgensteinowsko zorientowanej socjologii, takich jak strukturalizm czy etnometodologia. Interpretatorzy owi usiłują jednak znaleźć jakieś *iunctim* pomiędzy językiem i praktyką społeczną angażującą język (Le Roy Finch w: *Metafizyka jako cień...* [1996] s. 102). Powrócę jeszcze do tego zagadnienia przy omówieniu problematyki form życia. Na razie wystarczy podkreślić, że w analizie języka nie można wyjść poza sam język (tej intuicji z *Traktatu* Wittgenstein pozostał wierny), a „prawda” i „fałsz” to tylko słowa, których znaczenie jest ustalane w toku gier językowych. Oczywiście, „gry w prawdę i fałsz” są grami specyficznymi.

"Można to wyrazić następująco: jak długo pozostajemy w dziedzinie gier w prawdę i fałsz, zmiana gramatyki może nas prowadzić tylko z jednej gry do drugiej ale nigdy od czegoś prawdziwego do czegoś fałszywego" (Wittgenstein [1993] s. 111) (przekł. MB).

Jest to jednak tylko reguła opisująca gry w prawdę i fałsz, nie zaś prawdę i fałsz jako takie. Gramatyczną konwencją jest w szczególności jedna z głównych tez *Traktatu*, zgodnie z którą zdanie jest prawdziwe ,wtedy gdy odpowiada mu fakt. Czytamy w *Dociekaniach*:

„Podanie zwrotu: „Jest tak a tak” jako ogólnej formy zdania jest w gruncie rzeczy tym samym, co wyjaśnienie: zdaniem jest wszystko, co może być prawdą lub fałszem. (...) Powiedzieć, że zdaniem jest wszystko, co może być prawdą lub fałszem, to tyle, co powiedzieć: zdaniem nazywamy to, do czego w *naszym języku* stosuje się rachunek funkcji prawdziwościowych” (Wittgenstein [2000] s. 79).

Pogląd taki radykalnie odcina problematykę prawdy od kwestii ontologicznych czy epistemologicznych i przesuwa ją na grunt pragmatyki. Zmienia się tu niepostrzeżenie cała koncepcja filozofii jako procesu rozwiązywania problemów, którą to rolę pełniła ona jako zwierciadło natury. Pojawia się inna filozofia, zwana niekiedy „terapeutyczną”, gdyż zamiast rozwiązywać problemy filozoficzne, leczy umysł z opętania językowego (Wittgenstein [2000] s. 11-12). Przejście do filozofii terapeutycznej uważa za odkrycie przesądające o znaczeniu Wittgensteina w historii myśli. O ile więc Rorty entuzjastycznie akceptuje tezy bliskie jego



własnym przekonaniom co do przygodności wszelkiego dyskursu, w tym i filozoficznego, Wolniewicz głosi, że *Dociekania* i cała niemal późna filozofia Wittgensteina to myśl jałowa, bo oderwana od jakiegokolwiek bazy naukowej i żywiąca się samą sobą. *Traktat* otwierał szeroką perspektywę przed myślą filozoficzną, *Dociekania* zaś usiłują tę drogę zamknąć albo kierują ją na manowce (Wittgenstein [2000] s. XII).

Arbitralność reguł gramatyki nie znaczy jednak, że nie istnieje nic poza regułami gramatyki, czego mogłyby one dotyczyć. Można tu przywołać jedną z ulubionych przez Wittgensteina gier językowych – mierzenie. Reguły gramatyki są arbitralne w tym samym sensie co wybór jednostki miary. Ale to może przecież znaczyć tylko, że są one niezależne od długości mierzonego przedmiotu i że wybór jednej jednostki nie jest „prawdziwy”, a innej „fałszywy” tak jak podanie długości jest prawdziwe lub fałszywe” (Wittgenstein [1993] s. 185). Wittgenstein zastrzega, że jest to tylko uwaga na temat gramatyki wyrażenia „jednostka długości”, niewątpliwie pojawia się tu jednak pewne ograniczenie arbitralności wyników naszych gier językowych.

Można twierdzić, że nawet jeśli Wittgenstein zrelatywizował prawdę do języka, czy gry językowej, to taka relatywizacja nie prowadzi do relatywizmu w kwestii prawdy. Z tego, że rzeczy nie dyktują nam, *jak* mamy o nich mówić, nie wynika bowiem, że nie możemy powiedzieć o nich nic prawdziwego (Soin [2001]). Oczywiście nie wynika. Nie możemy jednak powiedzieć nic prawdziwego w sensie, w jakim prawdę mówi fizyk twierdząc, że widmo światła dzieli się na określone barwy. Jest to tylko gra językowa fizyka – w języku, który nie przewiduje słów o odpowiedniej gramatyce, ta prawda nie może być wypowiedziana, z czego zdaniem na przykład konstruktivisty społecznego wynika, że nie jest ona dla użytkownika takiego języka prawdą.

Wywód powyższy dowodzi, że Wittgenstein nie był zwolennikiem wizji racjonalności jako zupełnej relatywizacji reguł uzasadniania twierdzeń do reguł gry językowej. W szczególności, nie przyjąłby na pewno postmodernistycznego przeciwstawienia prawdy sensowi, czyli jako jakości wytwarzanej przez jednostki bądź społeczności ludzkie. Jego filozofia pozwala uniknąć tak schematycznej dychotomii. Wciąż jednak nie możemy powiedzieć, że późny Wittgenstein widzi możliwość powiązania prawdziwości z faktami. Można tylko skonkludować – o czym i tak trudno byłoby wątpić – że Wittgenstein jest przekonany, że język (czy też system pojęć) nie stwarza świata, tak jak nie stwarza domu, w którym siedzi w Cambridge, acz można by go już śmiało winić za śmierć autora *Gramatyki filozoficznej*, gdyby dom ów nagle się zawalił.

Zdaniem Wittgensteina, nie musimy zajmować się prawdą, możemy się bez niej obyć – nie wszystkie gry językowe są grami w prawdę i fałsz. Wyznając teorię absolutyzmu w dziedzinie prawdy, zgłaszając akces do grupy, którą Rorty nazywa "platonikami" (Rorty w *Habermas...* [1998] s. 50), podejmujemy wysiłek beznadziejny – staramy się uprościć i zredukować ludzki sposób życia i sam język do jednego schematu. Niewyobrażalna wielość możliwych użyczeń języka, lekceważona przez radykalnych racjonalistów, jak choćby gruntownie skrytykowany przez Wittgensteina Frazer, czyni wszelką perspektywę z zewnątrz nie dość, że zbędną, to szkodliwą dla perspektyw naszego wczucia się w innych. „Muszę stale pograżać się w wodach zwątpienia” – pisze Wittgenstein i słowa te można uznać za najbardziej zwięzły wyraz jego metody filozoficznej. Nie można wszystkiego wyjaśnić, co więcej – to nie wyjaśnienie przynosi zadowolenie (Wittgenstein [1998] s. 11-13). Słowa te są wyrazem niechęci do pierwszego pomysłu jedynie słusznej gry językowej – gry w uniwersalną racjonalność, w jedność procedur eksplanacyjnych, w *adequatio rei et intellectus*.

Wittgenstein okazał się przodkiem ideologicznym współczesnej socjologii przynajmniej na dwóch polach. Po pierwsze, określenia przedmiotu zainteresowania – skierował bowiem uwagę socjologów na nierozzerwalny związek języka z ogółem działań człowieka i stworzył możliwość poszukiwania metody identyfikacji kultury jako całości charakteryzowanej przez zachodzące w jej obszarze zjawiska językowe. Po drugie zaś, Wittgenstein dostarczył metody badania człowieka jako istoty wewnętrznie spójnej, nie sprowadzając zarazem wyjaśnienia jego działań do prostej analizy czynnikowej. Perspektywa gry językowej pozwala umieścić działanie człowieka w szerokim kontekście społecznym, zdając równocześnie sprawę z jego indywidualnego charakteru. Pozwala to mieć nadzieję, że adaptacja filozofii Wittgensteina jest jedną z dróg ku niesprzecznej socjologii humanistycznej. Wychodząc od pojęcia „gry językowej” możemy ocalić rozsądek, atakowany przez postmodernizm, nie rezygnując z dostrzegania całej złożoności i swoistości świata społecznego. Możemy wrócić z wakacji i zająć się jakąś porządną robotą, nie opuszczając dziedziny urokliwie błogiej abstrakcji.

## BIBLIOGRAFIA:

Billing H.

[1980] *Wittgensteins Sprachspielskonzeption*, Bonn

Bloor D.

[1993] *Wittgenstein. A Social Theory of Knowledge*, Nowy Jork.

Engel M.

[1971] *Wittgenstein's Doctrine of the Tyranny of Language*, Haga.

Geier M.

[2000] *Gra językowa filozofów*, przeł. Jan Sidorek, Warszawa.

Gozzi R.

[1998] *Is Language a game, A Review of General Semantics*, t. 55 nr 2, baza danych: *Academic Search Elite*.

[1996] *Habermas, Rorty, Kołakowski. Stan filozofii współczesnej*, przeł. Józef Niżnik, Warszawa.

Haller R.

[1988] *Questions on Wittgenstein*, Londyn.

Hintikka J.

[1992] *Eseje logiczno-filozoficzne*, przeł. Adam Grobler, Warszawa.

Malcolm N.

[1989]. *Wittgenstein on Language and Rules*, *Philosophy* 64.

Chmielewski A., Orzechowski A.

[1996] *Metafizyka jako cień gramatyki. Późna filozofia Ludwika Wittgensteina*, Wrocław.

Mokrzycki E.

[1993] *Mocny program socjologii wiedzy*, przeł. Ziemowit Jankiewicz, Józef Niżnik, Waleria Szydłowska, Michał Tempczyk, Warszawa.

Pears D.

[1999] *Wittgenstein*, przeł. Katarzyna i Jacek Gurczyńscy, Warszawa.

Putnam H.

[1994] *Words and Life*, Cambridge.

Redner H.

[1997] *Malign Masters. Gentile, Heidegger, Lukacs, Wittgenstein*, Londyn.

Rorty R.

[1996] *Przygodność, ironia, solidarność*, przeł. Wacław Jan Popowski, Warszawa.

Soin M.

[2001] *Wittgenstein. Prawda i sens*, materiały seminarium *Fundamentalim vs antyfundamentalizm*, IFIS PAN.

[1998] *Wittgenstein w Polsce*, Warszawa.

*Wittgenstein*

[2000] *Dociekania filozoficzne*, przeł. Bogusław Wolniewicz, Warszawa.

[1999] *Kartki*, przeł. Sława Lisiecka, Warszawa.

[1998] *Niebieski i brązowy zeszyt*, przeł. przeł. Adam Lipszyc i Łukasz Sommer, Warszawa

.

[1993] *O pewności*, przeł. Małgorzata i Wojciech Sady, Warszawa .

[1993] *Philosophical Grammar*, przeł. Anthony Kenny, Oxford.

[1997] *Tractatus logico-philosophicus*, przeł. B. Wolniewicz, Warszawa 1997.

[1998] *Uwagi o „Złotej gałęzi” Frazera*, przeł. Andrzej Orzechowski, Warszawa.

Wolniewicz B.

[1968] *Rzecz i fakty. Wstęp do pierwszej filozofii Wittgensteina*, Warszawa.