

PIOTR WILKIN*

REPREZENTACJE ZNATURALIZOWANE – UŻYTECZNY CEL CZY UŻYTECZNA FIKCJA?

STRESZCZENIE: Jednym z kluczowych pojęć znaturalizowanej epistemologii, a także wyrastających z niej nauk kognitywistycznych jest znaturalizowane pojęcie reprezentacji poznawczej. W ramach tego znaturalizowanego pojęcia próbuje się jednolicie (dla świata ludzi oraz innych organizmów żywych) ująć pojęcie błędu reprezentacyjnego. W tekście pokazane są argumenty przeciwko adekwatności stosowania znaturalizowanego pojęcia błędu reprezentacyjnego oraz wątpliwości co do tak szerokiego programu naturalizacji zagadnień związanych z ludzką pojęciowością.

SŁOWA KLUCZOWE: reprezentacje poznawcze, błąd reprezentacyjny, naturalizacja

1. WSTĘP

W filozofii umysłu podejście naturalistyczne jest ostatnimi czasy coraz popularniejsze, a w kognitywistyce jest być może konstytutywne, przynajmniej dla pewnego rozumienia tej dziedziny. Jednym z fundamentalnych pojęć kognitywistyki, które często podlega próbom naturalizacji, jest pojęcie reprezentacji poznawczej. Jedną z najpopularniejszych propozycji naturalizacji pojęcia reprezentacji jest koncepcja Dretskego, wiążąca pojęcie reprezentacji z pewnymi funkcjami naturalnymi organizmu (Dretske 1986). Koncepcja Dretskego ma dwie zasadnicze zalety. Z filozoficznego punktu widzenia pozwa-

* Ośrodek Badań Filozoficznych. E-mail: ilintar@gmail.com

ła na rozwiązanie wielu problemów, z jakimi borykały się np. czysto informacyjne czy korelacyjne koncepcje reprezentacji poznawczych (zasadniczym problemem jest tu chociażby problem wszechobecności reprezentacji, z którym borykały się wspomniane koncepcje: dym jest często skorelowany z ogniem, ale czy dym jest reprezentacją poznawczą ognia? Czy ogień w związku z tym posiada reprezentację?). Natomiast z metodologicznego punktu widzenia koncepcja Dretskego pozwala na uniwersalne potraktowanie reprezentacji poznawczych u wielu klas organizmów, co pozwala na czerpanie empirycznych danych o reprezentacjach z badań nad prostymi organizmami zwierzęcymi, czy nawet nad bakteriami. Pozwala ona także na uznanie ludzkich mechanizmów poznawczych za proste przedłużenie analogicznych mechanizmów zwierzęcych – w szczególności jest mocnym ciosem we wszystkie dualistyczne koncepcje poznania ludzkiego.

Celem tego tekstu jest podważenie uniwersalności reprezentacji poznawczych – a w konsekwencji sprzeciw wobec naturalizacji pojęcia reprezentacji na modłę Dretskego. Co ważne, w tym tekście poruszamy się wyłącznie wewnątrz reprezentacjonistycznego stanowiska w kognitywistyce. Nie będziemy tutaj poruszać zagadnienia reprezentacjonizm-antyreprezentacjonizm i zastanawiać się, czy rzeczywiście reprezentacje poznawcze są pożądanym elementem krajobrazu kognitywistycznego. Taka dyskusja, chociaż niewątpliwie ciekawa i fundamentalna, wykracza poza zakres tego tekstu, a argumentacja na rzecz reprezentacjonizmu rozmyłaby główną argumentację zawartą w artykule. Kiedy zatem będziemy omawiać zalety i wady naturalistycznego podejścia do reprezentacji, należy pamiętać, że dyskusja dokonywana jest przy założeniu, że uwzględnienie reprezentacji poznawczych w kognitywistyce jest pożądane i istotne.

2. REPREZENTACJE ZNATURALIZOWANE I BŁĘDNA REPREZENTACJA

Ważną cechą, którą Dretske, a za nim wielu jego następców (m.in. Millikan), akcentuje w swojej koncepcji, jest to, że umożliwia ona skuteczne analizowanie reprezentacji z punktu widzenia ich poprawności – a więc w szczególności uznania ich za błędne. W ujęciu Dretskego reprezentacja własności X jest poprawna wtedy, kiedy (w normalnych warunkach) działa zgodnie ze swoją funkcją, którą to funkcją

jest wskazywanie obecności *X*. Kluczowe w tej koncepcji jest pojęcie funkcji – dokładniej rzecz ujmując, mówi się tu o pewnym specyficznym jej rodzaju: funkcji naturalnej. Jeśli chcemy naturalizować reprezentację, nie możemy pozostać przy ogólnym sformułowaniu funkcjonalnym, gdyż możliwych klas funkcji jest bardzo wiele i ryzykujemy podobny problem, co w wypadku korelacyjnych koncepcji reprezentacji – wszechobecność reprezentacji. W związku z tym ograniczamy się właśnie do klasy funkcji naturalnych, czyli takich, które gwarantują sukces ewolucyjny. Dretske przytacza chociażby przykład bakterii, która posiada magnetyczny detektor północy, co pozwala jej jednocześnie na kierowanie się ku wodom pozbawionym tlenu (nadmiar tlenu jest dla tej bakterii zabójczy). Ta sama bakteria, przeniesiona na półkulę północną, zginie, gdyż jej magnetyczny czujnik będzie błędnie kierował ją w kierunku przeciwnym do bieguna (zamiast ku biegunowi), ku zabójczym dla niej wodom bogatszym w tlen.

Warto zauważyć, że już sam Dretske, choć przykład bakterii wydaje się bardzo eleganckim znaturalizowanym przykładem błędnej reprezentacji, zauważa, że przypisywanie bakterii funkcji naturalnej reprezentowania wody pozbawionej tlenu (zamiast po prostu reprezentowania północy) może być zbyt pospieszne. Nawet jeśli przyjmimy ewolucyjne kryterium wyznaczania funkcji naturalnych, to trudno uznać, że mechanizm magnetyczny rzeczywiście reprezentuje obecność tlenu, a nie po prostu obecność odpowiedniego pola magnetycznego, co w środowisku, w którym żyje bakteria, wyznacza otoczenie niezbędne do przetrwania. Innymi słowy, przykład z bakterią jest nie tyle przykładem błędnej reprezentacji, ile przykładem anormalnego stanu świata. Co więcej, nawet jeśli założymy, że do wywołania danej reprezentacji wystarczy jeden z kilku niezależnych mechanizmów, nadal możemy uznać, że dana reprezentacja reprezentuje dysjunkcję cech prostych, a nie pożądaną przez nas własność (tzn. gdyby bakteria oprócz czujnika magnetycznego miała jeszcze czujnik światła, to nadal moglibyśmy uznać, że posiada reprezentację własności Światło-lub-Północ, a nie reprezentację własności Tlen). Dretske twierdzi, że dopiero organizmy posiadające zarówno zestaw niezależnych mechanizmów sterujących reprezentacją, jak i zdolność uczenia się, czyli modyfikowania owych niezależnych mechanizmów włączających reprezentacje w trakcie życia, posiadają zdolność do błędnego reprezentowania. Innymi słowy, dopiero gdy dany organizm posiada reprezentację wła-

sności X , która na różnym etapie nauki wywoływana jest przez różne bodźce pochodzące od X (ale wszystkie te bodźce mają wspólną cechę bycia wywoływanym przez X), można mówić o takim pojęciu reprezentacji, która ma zdolność błędnej reprezentacji.

Widzimy więc, że wbrew obiecującym początkom (bakteria mająca reprezentację), aby mówić o reprezentacjach w sensie Dretskego potrzebujemy bardziej skomplikowanych organizmów niż bakterie. Niemniej jednak nadal jest to koncepcja bardzo atrakcyjna z naukowego punktu widzenia – większość organizmów zwierzęcych, nawet tych bardzo prostych, posiada jakiegoś rodzaju zdolność adaptacyjną, moglibyśmy zatem czerpać wiele przykładów na reprezentacje poznawcze z bogatego świata zachowań zwierzęcych.

Co więcej, koncepcja Millikan (1995), która powstała jako odpowiedź na koncepcję Dretskego, radzi sobie z problemem reprezentowania nawet w tak prostym przypadku, jak bakterie. Millikan rozwiązuje bowiem problem nieokreśloności funkcji Dretskego, przyjmując, że dla danego organizmu funkcją właściwą danej reprezentacji (jak nazywa swoje rozwinięcie koncepcji funkcji naturalnej Dretskego, zob. (Millikan 1987)) jest wskazywanie takiej własności, która jest potrzebna temu organizmowi do przetrwania (osiągnięcia sukcesu ewolucyjnego). Innymi słowy, bakteria ma reprezentację własności Tlen, a nie własności Światło-lub-Północ, gdyż to od tej pierwszej zależy jej przetrwanie – druga nie jest jej potrzebna.

Wszystkie te koncepcje reprezentacji są bardzo dopracowane i obiecujące, jeśli chodzi o badanie zjawiska reprezentowania u zwierząt. Czy jednak rzeczywiście zbliżają nas do zrozumienia problemu reprezentowania u ludzi?

3. BŁĘDY LUDZKIE I FUNKCJA NATURALNA U CZŁOWIEKA

Rozpatrzmy paradygmatyczny przypadek błędu reprezentacji, który często występuje w ludzkim procesie uczenia się języka. Otóż dziecko, widząc wyjątkowo dojrzałe, czerwone jabłko, wyciąga ku niemu rękę i mówi „pomidor” – z wyraźną intencją zjedzenia tego jabłka jako pomidora. Nie nauczyło się bowiem do tej pory, że również inne, wyglądające nieco jak pomidor owoce bywają czerwone oraz posiadają rozmiar i kształt zbliżony do tego, jak wyglądały obserwowane (i jedzone) przez nie uprzednio pomidory.

Warto zauważyć, że mamy tutaj dwa możliwe wyjaśnienia sytuacji. Jedno to wyjaśnienie, zgodnie z którym dziecko ma błędną reprezentację pomidora. Innymi słowy, dziecko ma jakąś reprezentację, jednak nie jest to właściwa reprezentacja pomidora. Drugie wyjaśnienie to takie, że dziecko posiada reprezentację pomidora, która w tym przypadku błędnie zadziałała. Nazwiemy pierwsze rozumienie błędem ogólnym, a drugie – błędem jednostkowym. W obu przypadkach powstaje pytanie: w jaki sposób mielibyśmy naturalizować tak rozumianą reprezentację?

Zauważmy, że jeśli mówimy o znaturalizowanej koncepcji reprezentacji w jej ujęciu funkcjonalnym (zarówno w wydaniu Dretskego, jak i późniejszym jej rozwinięciu np. u Millikan), to zazwyczaj chcemy mówić o funkcji biologicznej – i to takiej, która pozwoli nam wspólnie ujmować reprezentacje zarówno u ludzi, jak u prostszych organizmów (choć, jak już widzieliśmy, Dretske wydaje się uważać, że do poprawnego określenia funkcji reprezentacyjnej wymagany jest pewien poziom skomplikowania organizmów). Z reguły funkcja ta ma być w jakiś sposób związana z ewolucyjnym (bądź bardziej bezpośrednio – reprodukcyjnym) sukcesem organizmu. Tutaj warto zauważyć, że Millikan wspomina o „reprodukowaniu reprezentacji”, a nie tylko o „reprodukcji organizmu”, co otwiera pole do rozumienia tego w sposób inny niż ewolucyjny. Ponieważ jednak większość rozważań Millikan dotyczy biologicznych mechanizmów reprodukcji, będziemy w poniższych fragmentach zakładać, że chodzi nam właśnie o reprodukcję biologiczną. Możliwość zastosowania szerszej interpretacji koncepcji Millikan omówimy oddzielnie w późniejszej części tekstu.

Zastanówmy się teraz, czy rzeczywiście potrafimy znaleźć taką funkcję biologiczną organizmu dziecka z naszej historyjki, w której właściwą reprezentacją pomidora byłby właśnie pomidor, a nie czerwone jabłko? Zauważmy, że nie chodzi nam tutaj o zwiększony poziom komplikacji. Gdyby różnica między reprezentacjami ludzkimi a reprezentacjami organizmów prostszych polegała tylko na różnicy stopnia, nie byłby to istotny zarzut przeciwko tak rozumianej koncepcji reprezentacji. Można bowiem sensownie wyznawać pogląd, zgodnie z którym poziom komplikacji funkcji naturalnej stojącej za określoną reprezentacją jest proporcjonalny do stopnia złożoności organizmu, który ową reprezentację posiada. W takim przypadku nie powinniśmy się dziwić, że ludzkie funkcje naturalne są dalece bardziej złożone niż funkcje na-

turalne ameby czy bakterii. Co więcej, przynajmniej dla Dretskego to bardziej zaleta niż wada – aby rzeczywiście mówić o funkcjach naturalnych i uniknąć dwuznaczności, musimy mieć dostatecznie skomplikowany system, który dokonuje pewnych wyborów na podstawie więcej niż jednego kryterium.

Założmy więc na potrzeby naszego wywodu, że rzeczywiście znaleźliśmy funkcję naturalną, która odpowiada reprezentacji pomidora u dziecka. Założmy też, że owa funkcja rzeczywiście tłumaczy, na czym polega błąd reprezentacyjny, który dziecko popełnia, kiedy na czerwone jabłko mówi „pomidor” (albo wyciąga po nie ręce celem zjedzenia go jako pomidora, żeby nie ograniczać się koniecznie do kryteriów językowych). Czy jednak naprawdę jakakolwiek taka funkcja może być rzeczywiście naturalna? Pomijamy tutaj sytuację, kiedy pojęcie „funkcji naturalnej” zostaje zdefiniowane tak szeroko, że traci swoje intuicyjne znaczenie. Pozostańmy przy takim rozumieniu funkcji naturalnej, które pozwala nam na rzeczywistą naturalizację reprezentacji – czyli odnosi skuteczność reprezentacji do skuteczności przeżycia organizmu w jego środowisku. Wówczas okaże się – jak spróbuję teraz pokazać – że tego typu analiza reprezentacji pomidora nie ma szans się obronić.

Zacznijmy od analizy najbardziej bezpośredniej. Otóż zwolennik koncepcji naturalistycznej mógłby twierdzić, że zdolność odróżniania jabłek od pomidorów ma bezpośredni wpływ na przeżycie. Wyobraźmy sobie np. przypadek dziecka, które jest śmiertelnie uczulone na jabłko (ale nie na pomidory). Wówczas błędne reprezentowanie przez nie jabłka jako pomidora może być bezpośrednią przyczyną jego śmierci (jeśli np. wyciągnie rękę po jabłko i je zje, zanim zdąży to zauważyć rodzice). Tego typu analiza może wydawać się atrakcyjna, odwołuje się bowiem do kryteriów czysto biologicznych. Co więcej, można podać mniej wydumane przypadki, w których rzeczywiście zdolność odróżniania jednych organizmów od innych organizmów ma kluczowe znaczenie w kontekście zatrucia. Przykład dobrze znany w naszym kręgu kulturowym to chociażby czubajka kania i muchomor sromotnikowy. Jest to przykład o tyle lepszy od przykładu z pomidorem i jabłkiem, że ma ważność ogólnogatunkową, a nie tylko jednostkową (muchomor sromotnikowy jest trujący dla ludzi jako gatunku, a nie tylko dla pojedynczych jego przedstawicieli), można więc obronić tezę o naturalności takiej funkcji jako korelującej z przetrwaniem gatunku.

Nasz język jest jednak na tyle bogaty, że nie wszystkie pojęcia poddają się tak prostej i bezpośredniej analizie, dlatego też ta droga jest prędzej czy później skazana na porażkę. Nie będziemy w stanie dla każdego pojęcia znaleźć jego funkcji bezpośrednio ewolucyjnej, chociaż oczywiście jesteśmy w stanie prawdopodobnie dla każdej pary obiektów wymyślić scenariusz, w którym posiadanie błędnej reprezentacji jednego jako drugiego prowadzi do śmierci organizmu. Tyle tylko, że tego typu wymyślanie scenariuszy to droga donikąd – równie dobrze możemy bowiem wymyślić scenariusz, w którym to właśnie nieumiejętność odróżnienia jednego od drugiego pozwala organizmowi przeżyć (np. scenariusz niemalże z baśni braci Grimm, w której nieświadomy różnicy między kanią a sromotnikiem Jaś przynosi sromotnika do domu i częstuje nim macochę, która właśnie miała w planie ugotowanie go w piecu). Dlatego, aby uzasadnić naturalizację reprezentacji, musimy mieć funkcję uniwersalną – taką, która daje się wyjaśnić ogólnogatunkowo, a nie tylko jednostkowo. W literaturze można znaleźć liczne dyskusje na temat tego, w jaki sposób opisywać poprawnie funkcje naturalne tak, aby rzeczywiście spełniały one kryterium naturalności (czyli mówiąc inaczej: poddawały się naturalizacji; dobrym przykładem są tutaj chociażby rozważania Millikan).

Nie będziemy tutaj kontynuować tego wątku – głównie z racji tego, że jakiegokolwiek zastrzeżenia wobec konkretnej koncepcji funkcji naturalnych nie będą stanowiły wiarygodnych zastrzeżeń wobec idei naturalizacji reprezentacji w ogólności. Nawet jeśli nie potrafimy wprost powiedzieć, jaka jest ewolucyjna przewaga z posiadania reprezentacji kabrioletu w odróżnieniu od posiadania reprezentacji kombi, to przecież sam fakt istnienia tego rozróżnienia sugeruje, że w jakiś sposób pozwala nam ono przetrwać. Zwolennik naturalizacji reprezentacji może stwierdzić, że sam fakt, że funkcja naturalna odpowiadająca bardziej złożonym pojęciom nie jest nam jeszcze znana, świadczy tylko o tym, że mamy pole do dalszych badań empirycznych, aby ją odkryć i opisać.

Dlatego też obiekcja wobec tak sformułowanego rozumienia reprezentacji musi mieć charakter bardziej fundamentalny. Na tę bardziej fundamentalną obiekcję nakieruje nas następujące pytanie: w jaki sposób orzekamy o błędach w reprezentacji w przypadku ludzi? Co powoduje, że stwierdzamy, że ktoś błędnie reprezentuje (w obu sensach: ogólnym i szczegółowym) jakąś klasę obiektów (np. pomidory

albo kabriolety)? I wreszcie: w jaki sposób uczymy się dokonywania odpowiednich dystynkcji? Odpowiedzi na te pytania – mam nadzieję – podadzą przynajmniej częściowo w wątpliwość rzetelność naturalistycznego wyjaśnienia reprezentacji u ludzi.

4. CZŁOWIEK A BŁĄD NATURALNY

Człowiek jest w całym świecie zwierzęcą istotą szczególną, gdyż istotna część jego reprezentacji ma podłoże społeczne – czyli powstaje i zmienia się nie tylko bezpośrednio wskutek specyficznych bodźców związanych bezpośrednio z reprezentowanym obiektem, ale również (a czasem nawet w większości) w procesie socjalizacji. Co więcej, nawet spośród zwierząt podlegających socjalizacji człowiek jest szczególny – bardzo duża część jego reprezentacji powstaje bowiem w sposób pośredniczony językowo. Nie chcę tutaj rozstrzygać, jaka jest dokładnie relacja między funkcjonowaniem społecznym a reprezentacjami poznawczymi – tego typu rozważania z pewnością wykraczałyby poza ramy tego skromnego artykułu. Wystarczy nam tutaj spostrzeżenie, że w istotnym stopniu nasze reprezentacje są skorelowane z pojęciami językowymi. Uwaga ta dotyczy w szczególności reprezentacji osobowych (w rozumieniu Dennetta (1969)) – tutaj argumentacja na rzecz owej korelacji jest dość oczywista: bardzo trudno nam będzie znaleźć reprezentacje, które sobie na poziomie osobowym uświadamiamy, a dla których nie mamy językowego odpowiednika. Dotyczy ona jednak także, co może być nawet ważniejsze dla dyskusji o kwestii naturalizacji, reprezentacji na poziomie subosobowym. Warto tutaj chociażby wymienić kryterium atrakcyjności – wydaje się, że reprezentacja potencjalnie atrakcyjnego partnera to coś, co zdecydowanie dzielimy ze światem zwierzęcym. Jednak chociażby prześledzenie historii ludzkości pokazuje, że kategoria atrakcyjności zmieniała się bardzo często – dużo częściej, niż dałoby się to rozsądnie wytłumaczyć z punktu widzenia czysto ewolucyjnego.

Nawet jeśli ograniczymy się do reprezentacji subosobowych, nie mamy gwarancji, że w ich powstawaniu nie brały udziału czynniki społeczne (chyba że mówimy o reprezentacjach wrodzonych – jak później będziemy argumentować, pochodzenie reprezentacji jest tutaj bardzo istotnym czynnikiem różnicującym). Co więcej, jeśli istotną cechą reprezentacji ma być ich trwałość, jest to mocny argument na rzecz spo-

łecznego, a nie ewolucyjnego pochodzenia wielu ludzkich reprezentacji – chociażby przytoczony powyżej przykład kanonów atrakcyjności pokazuje, że w przypadku człowieka interakcje społeczne w wielu przypadkach przeważają nad czynnikami ewolucyjnymi jako czynnik wyznaczający reprezentacje.

Wróćmy tutaj do przytoczonej na początku definicji błędu reprezentacyjnego i podstawmy zmienne. *Atanazy* ma niewłaściwą reprezentację *czubajki kani* w sytuacji, w której jego reprezentacja [*czubajka kania*] każe mu zebrać w lesie *muchomora sromotnikowego* (tak wygląda sytuacja błędu ogólnego, w przypadku błędu szczegółowego – błędnie rozpoznaje *czubajkę kanię* i zbiera *muchomora sromotnikowego*). Wydaje się, że ze względu na efekty biologiczne jest to paradygmatyczny przykład na możliwość naturalizacji pojęcia reprezentacji – wszak błąd reprezentacyjny prowadzi tutaj do śmierci osobnika. Powstaje jednak pytanie: czy jest to rzeczywiście błąd reprezentacyjny? Innymi słowy: czy rzeczywiście właściwa reprezentacja *czubajki kani* to ta, która wskazuje na to, co rozumiemy w języku przez pojęcie „*czubajki kani*”? Wszak równie dobrze możemy sobie wyobrazić, że reprezentacja *czubajki kani* nie ulega zmianie, a jedyne, co ulega zmianie, to skłonność do jej spożycia bądź nie. Oczywiście ten kontrargument można zbić argumentem, że posiadanie oddzielnych reprezentacji *czubajki kani* i *muchomora sromotnikowego* to scenariusz bardziej ewolucyjnie przydatny niż samo posiadanie reprezentacji *czubajki kani* jako grzyba niejadalnego (wyobraźmy sobie populację ludzi, których jedynym potencjalnym źródłem pożywienia jest las, w którym rosną same *kanie* i *sromotniki*). Równie dobrze możemy jednak sobie wyobrazić sytuację, w której dana populacja ma reprezentację *czubajki kani* oraz dodatkowe przekonania mówiące o tym, że te bardziej zielonkawe są trujące, a te mniej – nie (innymi słowy, przekonania przypisujące bycie jadalnym pewnemu stanowi organizmu, a nie innemu organizmowi – podobnie jak np. czynimy w odniesieniu do grzyba zwanego popularnie „*purchawką*”, którego wczesna forma jest jadalna).

Być może widać już tutaj pewną analogię do argumentu, który zapisał się na stałe w historii filozofii języka – argumentu Wittgensteina o braku „błędu prywatnego” (Wittgenstein 1953), który szczegółowo omawia Kripke (1982). Myślę, że analogia ta nie jest tylko dziełem przypadku. Moim zdaniem mówienie o błędzie reprezentacji w kon-

tekście naszych reprezentacji poznawczych (poza reprezentacjami wrodzonymi) w takim samym sensie, w jakim mówimy o błędzie reprezentacji prostszych organizmów w odniesieniu do ich funkcji naturalnych, jest ślepym zaułkiem.

Większość argumentów, które Wittgenstein (a z nim Kripke) odnosił do niemożliwości istnienia znaturalizowanego pojęcia błędu pojęciowego można z powodzeniem zastosować do opisu sytuacji błędu reprezentacyjnego. Powyższy przykład z muchomorem i kanią powinien być tutaj dobrą ilustracją. Nawet oryginalny przykład Wittgensteina (przypomnijmy – Wittgenstein, a za nim Kripke, argumentował, że nie jesteśmy w stanie ocenić, czy ktoś, mówiąc o operacji dodawania, czyli o znaku „plus”, ma rzeczywiście na myśli plus, a nie „kwus”. Kwus różni się od plusa regułami użycia w pewnych bardzo specyficznych okolicznościach, które akurat w tym przypadku nie zachodzą) można tutaj zastosować z powodzeniem (gdyby nie to, że dotyczy pojęć wysoce abstrakcyjnych, co do których raczej trudno mówić o prostych reprezentacjach poznawczych). Zauważmy bowiem, że *clou* argumentu jest takie samo. Wittgenstein (a za nim Kripke) mówi: nie jesteśmy w stanie wyłącznie za pomocą obiektywnych (rzeczowych) kryteriów rozstrzygnąć, który z dwóch opisów pojęcia jest tym właściwym – podobnie, jak nie jesteśmy w stanie rozstrzygnąć, który z dwóch opisów reprezentacji jest tym właściwym, poza dokonywaną *ad hoc* wsteczną racjonalizacją już istniejących reprezentacji („dziwna czubajka kania” vs. „czubajka kania / muchomor sromotnikowy”).

Co więcej, argument ten można zastosować dwa sposoby. Jeśli nasz błąd reprezentacji rozumiemy jako błąd ogólny, czyli posiadanie błędnej reprezentacji, to powstaje pytanie, co wyznacza poprawność owej reprezentacji (innymi słowy, jak wyznaczyć tę, a nie inną funkcję właściwą). Jeśli błąd reprezentacji rozumiemy jako błąd szczegółowy, czyli błędne zastosowanie reprezentacji w danym konkretnym momencie, to możemy przywołać argument Wittgensteina i zapytać: co przekonuje nas, że mamy do czynienia z błędem, a nie tylko wyjątkiem przewidzianym w regule?

Zauważmy jednak, że w stosunku do oponentów Wittgensteina z dyskusji o niemożliwości jednostkowego ustalenia znaczenia pojęć jesteśmy w lepszej sytuacji. Mamy bowiem cały czas nasze pojęcie funkcji naturalnej. Możemy zatem bronić naszych reprezentacji poprzez ich odniesienie do właściwych nam funkcji biologicznych. Tyle tylko, że ta

droga wydaje się ślepą uliczką. Nawet w przykładzie teoretycznie bardzo dobrze pasującym do znaturalizowanej wizji epistemologii trudno nam pokazać przewagę koncepcji dwóch reprezentacji (czubajki kani i muchomora sromotnikowego) nad koncepcją pojedynczej reprezentacji o zróżnicowanych stanach.

Do argumentów o charakterze fundamentalnym warto dodać jeszcze argument o charakterze empirycznym. Nawet jeśli zgodzilibyśmy się z tezą, że koncepcja błędnej reprezentacji w sensie naturalistycznym daje się obronić, to trudno będzie nam bronić tezy, że nasze reprezentacje poznawcze rzeczywiście powstają w ten sposób i że błędy w owych reprezentacjach diagnozujemy za pomocą ich skutków ewolucyjnych. Bogactwo naszego systemu pojęciowego, a co za tym idzie – systemu reprezentacyjnego (gdyż zauważyliśmy, że większość naszych reprezentacji ma swoje korelaty językowe), jest bowiem zbyt duże, a okres zmienności ewolucyjnej – zbyt krótki, aby takie wyjaśnienie rzeczywiście miało moc eksplanacyjną. Możemy tego typu teorii bronić w na przykład w odniesieniu do pszczół, których mechanizmy komunikacji mają rzeczywiście ewolucyjne podstawy – ale trudno jest w ten sposób bronić ludzkich reprezentacji, których historia jest zbyt krótka, aby kontekst ewolucyjny mógł mieć w ich wypadku istotne znaczenie.

Możemy oczywiście twardo obstawać za znaturalizowaną koncepcją reprezentacji także u ludzi, twierdząc, że poprawnymi reprezentacjami są te, które realizują pewne funkcje naturalne (bądź funkcje właściwe, jeśli wolimy używać terminologii Millikan), a społeczne uwarunkowania dotyczące zgody bądź niezgody na takie a nie inne stosowanie pojęć nie mają wpływu na poprawność reprezentacji. Tyle tylko, że tego typu podejście wymaga przyjęcia jednego z poniższych założeń – każde z nich wydaje się z różnych powodów problematyczne.

Po pierwsze, możemy założyć, że pojęciowość językowa i reprezentacje poznawcze nie są ze sobą bezpośrednio skorelowane – czyli, że pojęcia językowe nie są zakorzenione w reprezentacjach poznawczych. W tym tekście zakładamy, że takie zakorzenienie istnieje, można jednak sobie wyobrazić, że ktoś chciałby to zakorzenienie negować. Owa negacja może przyjąć różne poziomy. W wersji skrajnej (pojęcia językowe nie mają żadnego związku z reprezentacjami poznawczymi) jest kompletnie nie do obrony dla kogokolwiek, kto chce respektować naukowe podstawy kognitywistyki, w tym wyniki badań empirycznych chociażby z zakresu psychologii rozwojowej. Istnieje jednak możliwość

słabszej negacji tego zakorzenia – chociażby zaakceptowanie zakorzenia na poziomie typów (pojęcia językowe jako całość są zakorzenione w reprezentacjach poznawczych jako takich), ale nie na poziomie egzemplarzy (poszczególne pojęcia językowe nie są zakorzenione w poszczególnych reprezentacjach poznawczych). Trudno jednak dostrzec, w jaki sposób jakakolwiek taka negacja poza wersją skrajną pozwala uniknąć problemu opisanego powyżej.

Po drugie, możemy założyć, że aktualny stan języka nie jest adekwatnym miernikiem adekwatności reprezentacji poznawczych. Przy takim założeniu wnikamy się jednak w iście Leibniziańską koncepcję języka idealnego, który to miałby najlepiej służyć przetrwaniu gatunku i wedle którego mielibyśmy oceniać nasze reprezentacje poznawcze. Pomijając czysto metafizyczne wątpliwości dotyczące takiego rozwiązania, pozostaje wątpliwość fundamentalna: czy aby takie podejście rzeczywiście jest naturalistyczne, tzn. czy mamy naukowe podstawy do tego, aby orzekać, jakiego rodzaju język naturalny jest językiem idealnym, najlepiej sprzyjającym realizowaniu przez nas funkcji przetrwania gatunku.

Po trzecie, możemy zrezygnować z oceniania indywidualnych reprezentacji jako skorelowanych z indywidualnymi pojęciami językowymi i zamiast tego oceniać reprezentacje jako poprawne bądź błędne z punktu widzenia ich przynależności do systemu językowego, który lepiej służy przetrwaniu. Tyle tylko, że w tym miejscu wnikamy się w holizm, który jest próbą ukrycia istotnej części odpowiedzi na pytanie – że tak naprawdę pytanie o poprawność reprezentacji ma sens tylko w kontekście systemu językowego, w ramach którego owa reprezentacja powstała. Innymi słowy, znowu tracimy naturalizację – musimy bowiem zrezygnować z oceny błędności pojedynczych reprezentacji na rzecz oceny poprawności bądź błędności całego systemu reprezentacyjnego. Przywołując na moment alternatywę, o której wspominaliśmy we wstępie – takie rozwiązanie niebezpiecznie kieruje nas w kierunku rozwiązań antyrepresentacjonistycznych, których zdecydowaliśmy się w tym tekście nie rozważać.

Pomijając trzy wspomniane powyżej założenia, rozwiązanie ignorujące językową stronę reprezentacji poznawczych jest też słabo zakorzenione w danych empirycznych. Ignoruje bowiem rzeczywiste mechanizmy naszej oceny reprezentacji jako błędnych, kosztem pewnej wyidealizowanej koncepcji błędu, różnej od tego, co w języku potocz-

nym określamy tym mianem. Jako takie przypomina nieco koncepcje semantyczne, które, próbując uniknąć kontekstowej zależności dla niektórych wyrażeń, przypisują im literalną fałszywość niemal w każdych okolicznościach ewaluacji, tłumacząc ich zastosowanie obecnością prawdziwych implikatur. Takie rozwiązanie jest oczywiście formalnie możliwe, ale powstaje zasadnicze pytanie o jego eksplanacyjną przydatność.

5. REPREZENTACJE A SPOŁECZEŃSTWO

Skoro już podaliśmy argumenty przeciwko w pełni znaturalizowanej koncepcji reprezentacji, powstaje pytanie o pozytywną alternatywę: jakie rozwiązanie w kwestii reprezentacji ma być przeciwwagą dla pełnej naturalizacji? We wstępie odrzuciliśmy opcję antyreprezentacjonistyczną, pozostaje nam zatem sformułowanie jakiegoś pozytywnego stanowiska.

Z pomocą może nam przyjść analiza przypadków związana z językowymi korelatami reprezentacji. Kiedy powiemy, że ktoś posługuje się błędną reprezentacją? Wracając do naszego przykładu: kiedy uznamy, że Atanazy posługuje się błędną reprezentacją czubajki kani?

Powiedzieliśmy, że w jakimś stopniu ważny jest tutaj, jak u Wittgensteina, społeczny konsens – Atanazy posługuje się błędną reprezentacją kani, jeśli nie zgadza się ona z tym, co społeczeństwo wypracowało jako reprezentację. Czy jednak mamy przez to rozumieć, że błędność reprezentacji jest wyłącznie funkcją społecznej zgody, tzn. Atanazy ma poprawną reprezentację, jeśli społeczeństwo zgadza się, że jego reprezentacja jest poprawna?

Takie rozwiązanie z pewnością jest atrakcyjne i upraszcza wiele kwestii, jeśli chodzi o koncepcyjną stronę zagadnienia reprezentacji. Jest też wyraźnie antynaukowe, przynajmniej w sensie nauk empirycznych – na gruncie takiej koncepcji musielibyśmy porzucić wszelakie pretensje do redukcji reprezentacji poznawczych do przedmiotu tychże nauk. Samo to w sobie nie jest jednak jeszcze dostatecznym zarzutem – wszak wiele zjawisk socjologicznych uważamy za emergentne i nie postulujemy ich redukcji do poziomu biologicznego. Rozwiązanie to ma jednak inny, poważny problem – uniemożliwia pozytywne rozstrzygnięcie dylematu „jednostka vs. społeczeństwo”.

Rozpatrzmy historię opartą na popularnym, zwłaszcza w beletrystyce, motywie wybitnego samotnego naukowca. W tej narracji jed-

nostka, która wpadła na trop pewnego odkrycia naukowego, jest powszechnie odrzucana przez środowisko, jednak pod koniec opowieści okazuje się, że jednak miała ona rację. Rozważmy ten scenariusz w kategoriach reprezentacji, odwołując się np. do obecnej rzeczywistości w fizyce historii dyskusji dotyczącej natury czarnych dziur (abstrahując od sytuacji, czy owa dyskusja rzeczywistości miała charakter „jednostka vs. reszta społeczeństwa”, ponieważ taka sytuacja w przyrodzie praktycznie nie występuje). W sytuacji, gdy poprawność reprezentacji wyznaczana jest przez społeczny konsens, samotny naukowiec będzie w tej kwestii zawsze się mylił i nigdy nie będzie mógł udowodnić poprawności swojej reprezentacji. Zauważmy w szczególności, że zawodzi tutaj także opis społeczny, zgodnie z którym naukowiec stopniowo przekonywałby innych naukowców o swoim odkryciu – jeśli bowiem przyjmiemy społeczny opis poprawności reprezentacji, to przy takim opisie ów naukowiec nadal jest w błędzie przez cały czas, kiedy przekonuje społeczność do swoich racji i zaczyna poprawnie używać reprezentacji dopiero wtedy, kiedy ową społeczność do swoich racji przekona. Ewidentnie taki opis sytuacji wydaje się podejrzany – wolelibyśmy uważać, że ów naukowiec cały czas miał rację i cały czas poprawnie używał reprezentacji czarnych dziur, a to reszta społeczeństwa naukowego cały czas posiadała błędną reprezentację. W jaki sposób możemy uratować taki obraz reprezentacji?

Otóż wydaje się, że można pogodzić funkcjonalistyczną interpretację reprezentacji poznawczych, za którą stoi niewątpliwy dorobek intelektualny i precyzja pojęciowa, ze społeczną koncepcją reprezentacji. Aby to zrobić, musimy jedynie... zrezygnować z naturalizacji. Koncepcja taka dopuszcza bowiem nadal możliwość, że posiadanie pewnej reprezentacji to realizowanie pewnej funkcji – to, co się zmienia, to charakter owej funkcji. Przestaje być to bowiem funkcja naturalna – a staje się to funkcja narzucona przez społeczne reguły, w takim sensie, w jakim Wittgenstein rozumiał regulowanie przez społeczeństwo reguł znaczeniowych – znaczeniem słowa jest to, co akurat dana społeczność językowa wymusza jako jego znaczenie.

W takim opisie naukowiec badający samotnie czarne dziury może mieć rację, o ile jego funkcja reprezentacyjna zgadza się z regułami znaczeniowymi, które są narzucane przez społeczeństwo – nawet mimo tego, że różni się w opiniach z resztą społeczeństwa co do tego, co np. stanowi konkretnie obiekt tej reprezentacji (tak jak mogę np. zga-

dzać się ze znajomymi, że „najszybszy człowiek świata” znaczy „ten, kto przed chwilą najszybciej przebiegł 100 m na Igrzyskach Olimpijskich”, ale wskutek braku informacji być przekonanym, że jest to Justin Gallin, gdyż tak podały pierwsze komunikaty prasowe, podczas gdy późniejsza analiza fotokomórki wykazała, że szybszy był Usain Bolt). Ta charakterystyka naszego rozwiązania przypomina nieco uwagi Kripkego o prawdach koniecznych znanych *a posteriori* – z tego, że społeczeństwo zgadza się (*explicite* bądź *implicite*) co do reguł znaczeniowych dotyczących pojęcia językowego (a zatem, w naszym rozwiązaniu, również co do funkcji reprezentacyjnej, która realizuje owo pojęcie) nie wynika jeszcze, że posiada wiedzę dotyczącą wszystkich sądów, których składnikiem jest owo pojęcie, bo część owych sądów można poznać tylko na drodze doświadczenia empirycznego, a nie poprzez analizę pojęciową. Nasz przykładowy naukowiec może mieć rację w sprawie czarnych dziur w tym sensie, że może zgadzać się z innymi naukowcami co do ostensywnej definicji „czarne dziury = to, co znajduje się w centrach galaktyk”, a nie zgadzać się z nimi co do (esencjalnych) właściwości fizycznych tych obiektów (np. tego, że pochłaniają one światło czy zakrzywiają pole grawitacyjne).

Jeżeli ktoś jeszcze nie został przekonany powyższą argumentacją, że naturalistyczne pojęcie do ludzkich reprezentacji poznawczych jest trudne do obrony, oto alternatywna argumentacja – genetyczna, odwołująca się do wyników nauk kognitywnych (jej echa pobrzmiwają także w powyższej argumentacji). Zastanówmy się, co jest obiektem badania kognitywistyki, jeśli chodzi o reprezentacje u zwierząt i innych organizmów żywych, a co – u ludzi. Zakładając, że możemy w ten sam sposób ująć same reprezentacje, podając wspólną metafizyczną charakterystykę (np. funkcjonalną), powstaje jeszcze pytanie o sposób powstawania tych reprezentacji. Tutaj nie ulega wątpliwości, że o ile u zwierząt wiele, jeśli nie większość reprezentacji ma charakter wrodzony i źródło ewolucyjne, to u człowieka znacząca część reprezentacji (choćby licząc ich korelaty językowe) ma charakter przygodny, społeczny. Próbując zatem podać wspólną charakterystykę reprezentacji u ludzi i zwierząt za pomocą funkcji naturalnej, przesuwamy ciężar w kierunku łatwości badania, kosztem wiarygodności (bo łatwo nam jest dostarczyć wiarygodny opis empiryczny wraz z wiarygodnym uzasadnieniem ewolucyjnym wielu mechanizmów u prostych organizmów zwierzęcych, trudno jest natomiast zrobić to dla człowieka). In-

nymi słowy, z pary łatwość/wiarygodność eksplanacyjna wybieramy to pierwsze – co jest oczywiście atrakcyjnym rozwiązaniem, jeśli celem jest znalezienie opisu mechanizmu, który zmieści się w kilkunastostro-
nicowym artykule, natomiast budzi poważne wątpliwości metodolo-
giczne.

Wróćmy teraz na chwilę do porzuconego wcześniej w tekście roz-
wiązania Millikan i zastanówmy się jeszcze, czy istnieje możliwość ura-
towania podejścia naturalistycznego do reprezentacji poprzez szersze
rozumienie pojęcia reprodukcji reprezentacji. Czy możemy rozumieć
reprodukcję tutaj w sensie społecznym i uznać, że reprezentacje są
trwale wtedy, kiedy są trwale reprodukowane społecznie? Oczywiście
możemy, tylko że – jak się wydaje – dla naturalisty takie zwycięstwo
byłoby zwycięstwem pyrrusowym. O ile bowiem nie ulega wątpliwości,
że nauką odpowiadającą za opis mechanizmów ewolucyjnie stabilnych
jest biologia, o tyle trudno jest uznać, że biologia jest adekwatną nauką
do opisu mechanizmów społecznie stabilnych (takich jak chociażby re-
produkcja pojęć językowych, systemów językowych czy systemów kul-
turowych). Możemy tutaj oczywiście odwoływać się do socjologii, psy-
chologii bądź ekonomii jako nauk odpowiedzialnych za opis zjawisk
społecznych, jednak trudno będzie bronić nam tezy, że nadal mamy
do czynienia z naturalizacją. Naturalizacja zakłada bowiem standardo-
wo odwołanie się do nauk przyrodniczych – albo zatem przyjmujemy
niestandardowo szeroki zakres pojęcia „naturalizacja” (co jest sprzecz-
ne z zakresem przyjmowanym zazwyczaj w literaturze), albo zakła-
damy redukcję nauk takich jak socjologia czy ekonomia do biologii
i nauk przyrodniczych (co byłoby obroną naturalizacji reprezentacji za
pomocą dużo mocniejszej i jeszcze bardziej wątpliwej tezy).

Nie ulega więc wątpliwości, że jeśli chcemy uratować podejście
funkcjonalistyczne do reprezentacji, to będziemy musieli zmodyfiko-
wać dużą część założeń związanych z tym podejściem. Musimy bowiem
zrezygnować z „naturalnej” teleologii związanej z funkcjami ewolucyj-
nymi i z powiązanych z nią rozważań dotyczących ewolucyjnych kryte-
riów wyznaczania funkcji (które w znacznym stopniu widać chociażby
u Millikan). Z powyższego wywodu można wyciągnąć nawet dalej idą-
cy wniosek – musimy zrezygnować z teleologii w ogóle. Analiza różnic
podejścia do reprezentacji u ludzi i u organizmów prostszych wyda-
je się bowiem podpowiadać, że sensowniejsze byłoby rozpatrywanie
funkcji właściwych ze względu na przyczynę ich powstania, a nie

jej cel. Owo rozróżnienie wyjaśniałoby również, skąd teleologiczne podejście do funkcji reprezentacyjnych w dotychczasowych badaniach – w perspektywie ewolucyjnej przyczyna i cel są bowiem bardzo ściśle ze sobą związane i ich odróżnienie od siebie jest funkcjonalnie niemal niewykonalne („posiada taką-a-taką funkcję, gdyż jego geny przeżyły” vs. „posiada taką-a-taką funkcję, aby jego geny przeżyły” z perspektywy badacza, który ma zazwyczaj dostępne tylko dane historyczne jest praktycznie nieodróżnialne). Te dwie kategorie są jednak bardzo różne, jeśli chodzi o podejście do ludzkich reprezentacji – pozwalają bowiem mówić sensownie o reprezentacjach, które mają podłoże językowe, w odróżnieniu od tych, które mają podłoże ewolucyjne (zauważmy, że taka koncepcja nie przesądza też, że u ludzi nie istnieją reprezentacje mające podłoże ewolucyjne – znów, znając wyniki nauk kognitywnych, takie rozstrzygnięcie mogłoby być mocno kontrowersyjne, w szczególności część reprezentacji subosobowych, zwłaszcza prostych reprezentacji percepcyjnych u ludzi, wydaje się mieć charakter ewolucyjny).

6. PODSUMOWANIE

W świetle powyższej argumentacji całkowita naturalizacja reprezentacji poznawczych człowieka wydaje się rozwiązaniem trudnym do obrony. Przeciwno niej przemawiają zarówno argumenty natury czyisto filozoficznej, jak i argumenty metodologiczne, odwołujące się do pochodzenia reprezentacji u człowieka. Naszkicowana tutaj hybrydowa, społeczno-funkcjonalna koncepcja reprezentacji, odwołująca się nie do celu funkcji reprezentacyjnych, a do jej źródła, dużo lepiej zdaje sprawę z istotnych różnic pomiędzy reprezentacjami poznawczymi u człowieka i u prostszych organizmów oraz lepiej tłumaczy pojęcie błędu reprezentacyjnego. Otwarte pozostaje pytanie, do jakiego stopnia dostępne już wyniki badań nad funkcjonalizmem w odniesieniu do reprezentacji można zastosować do tak zmodyfikowanej koncepcji reprezentacji – uważam jednak, że tego typu mariaż mógłby być owocny, a przy okazji pomógłby zbliżyć do siebie konkurencyjne (naturalistyczne i antynaturalistyczne) podejścia do reprezentacjonizmu.

BIBLIOGRAFIA

- Dennett, D. (1969), *Content and Consciousness*, London: Routledge.
- Dretske, F.I. (1986), *Misrepresentation*, w: R. Bogdan (ed.), *Belief: Form, Content, and Function*, Oxford: Clarendon Press, s. 17–37.
- Saul, K. (1982), *Wittgenstein on Rules and Private Language*, Cambridge, Mass.: Harvard University Press.
- Millikan, R.G. (1989), *Biosemantics*, „The Journal of Philosophy” 86(6), s. 281–297.
- Millikan, R.G. (1984), *Language, Thought, and Other Biological Categories. New Foundations for Realism*, Cambridge, Mass.: MIT Press.
- Millikan, R.G. (1995), *Pushmi-Pullyu Representations*, „Philosophical Perspectives” 9, s. 185–200.
- Wittgenstein, L. (1953), *Philosophical Investigations*, Oxford: Blackwell Publishing.

NATURALIZED REPRESENTATIONS – A USEFUL GOAL OR
A USEFUL FICTION?

SUMMARY: One of the key concepts of naturalized epistemology as well as the cognitive sciences that stem from it is the naturalized concept of mental representation. Within this naturalized concept, many attempts have been made to unify (for humans as well as for other living organisms) the notion of representation error. This text makes an attempt to argue against the adequacy of using a naturalized concept of representation error as well as casts doubt on the wide program of naturalizing concepts related to human conceptuality.

KEYWORDS: mental representations, representation error, naturalization