

Stanisław Judycki

### Hipoteza ‘języka myśli’ i jej krytyka\*

Na pewno trudno byłoby przedstawić jakąś bezdyskusyjną listę rankingową dorobku tzw. filozofii umysłu (*philosophy of mind*) po drugiej wojnie światowej, to znaczy powiedzieć, które z nowo wprowadzonych koncepcji i pojęć są najważniejsze, które pozostaną trwałym dorobkiem filozofii na następne dziesięciolecia, a być może stulecia, a które po prostu ulegną zapomnieniu. Trudność ta wynika z banalnego powodu zbyt małej odległości czasowej, co powoduje, że nowości ostatnich czasów wydają się nam być prawie że ostatnim słowem możliwym do powiedzenia na dany temat. Mimo to jest kilka pojęć i koncepcji, które, jak się wydaje, mogą przetrwać próbę czasu: teoria identyczności, funkcjonalizm, superweniencja koneksjonizm, czy sformułowana przez J. Fodora hipoteza ‘języka myśli’ (*language of thought*) (Fodor 1975). Należy ona już do stałego repertuaru konceptualnego dzisiejszej filozofii umysłu, każdy filozof o niej przynajmniej słyszał, jest ciągle żywo dyskutowana i - co ciekawsze - mimo wielorakich krytyk nie ma właściwie konkurentek (oprócz, być może, podejścia koneksjonistycznego).

Celem tego artykułu jest przedstawienie treści tej hipotezy i jej krytyki. Mimo że w krajach angielskojęzycznych teoria ta i jej krytyka stały się standardowym fragmentem edukacji studentów filozofii zajmujących się naturą umysłu, sądzę jednak że warto jeszcze raz te kwestie przedstawić, zwłaszcza na gruncie polskiej literatury filozoficznej i psychologicznej, gdyż w ostatnim dziesięcioleciu zaczynamy dopiero asymilować duże fragmenty dyskusji filozoficznych, do których mieliśmy bardzo słaby dostęp. Drugim powodem jest to, że atrakcyjnie wyglądające nowe idee filozoficzne wymagają ciągle nowych przemyśleń kierowanych pytaniem, czy rzeczywi-

---

\* Poniższy tekst został opublikowany w: A. Klawiter/W. Dziarnowska (red.), *Studia z kognitywistyki i filozofii umysłu*, t. 2, Poznań 2005, s. 257-274.

ście da się w filozofii powiedzieć coś nowego. Fodorowski ‘język myśli’ budzi bowiem podejrzenia, że jest to coś zupełnie niemożliwego i że radykalna krytyka tej hipotezy może ukazać, iż jej atrakcyjność bierze się raczej z powierzchownego uchwycenia dyskutowanych problemów. Najpierw zreferuję treść hipotezy sformułowanej przez Fodora, następnie krótko przedstawię główne zarzuty, natomiast centralną częścią artykułu będzie przedstawienie pewnej radykalnej krytyki symbolicznej teorii myślenia, a w tym też krytyki koncepcji ‘języka myśli’.

Według Fodora dominujący w dawnej psychologii asocjacionizm nie był w stanie stworzyć hipotezy wyjaśniającej, jak racjonalność naszego umysłu może się realizować w oparciu o relacje przyczynowe między myślami. Relacje te miały wyrażać tzw. prawa asocjacji. Dopiero pojawienie się komputerów i powstała w oparciu o inspirację komputerową tzw. komputacyjna teoria umysłu umożliwiły sformułowanie tego rodzaju hipotezy.

Oto najprostszy zarys tego, jak to nowe ujęcie mogłoby się przedstawiać: łączycie przyczynowe własności symbolu *via* jego syntaksę. Syntaksa symbolu jest jedną z jego fizycznych własności wyższego rzędu. Pierwsze metaforyczne przybliżenie tej idei wygląda następująco: możemy myśleć o syntaktycznej strukturze symbolu jako o abstrakcyjnej własności jego kształtu. Ponieważ syntaksa redukuje się do kształtu, i ponieważ kształt symbolu jest potencjalnym determinantem jego roli przyczynowej, to dość łatwo jest dojrzeć, jak jest możliwe istnienie sytuacji, w których przyczynowa rola symbolu koreluje z jego syntaksą. Łatwo sobie wyobrazić egzemplarze (*tokens*) symbolu wzajemnie oddziaływujące na siebie w oparciu o ich struktury syntaktyczne. Syntaksa symbolu może determinować przyczyny i skutki jego egzemplarzy w ten sam sposób, jak geometria klucza określa, które zamki będzie otwierał (Fodor 1987, s. 18-19).

Koncepcja Fodora opiera się na następującym przypuszczeniu: introspekcyjnie dostępne stany umysłowe przyczynowo wywołują inne stany umysłowe oraz wywołują zachowanie ludzkie. Efektywność kausalna stanów mentalnych nie jest jednak osadzona w tych cechach, które tym stanom przypisuje introspekcja. Określone przekonanie jest efektywne kau-

zalnie nie dzięki treści danej w introspekcji, lecz dlatego, że w mózgu dokonuje się jego translacja w niezinterpretowany system znaków, którym z kolei odpowiadają pewne klasy zdarzeń fizycznych (fizjologicznych). Istnieje więc według Fodora równoległość pomiędzy tym, co semantyczne, syntaktyczne i fizyczne. Jeżeli na przykład ktoś chce  $P$  oraz jest przekonany, że nie- $P$  pod warunkiem, że  $Q$ , to zazwyczaj stara się zrealizować  $Q$  - czymkolwiek byłyby  $P$  i  $Q$ . Tego rodzaju uogólnienie - i oczywiście wiele innych uogólnień należących do 'skarbcza' tzw. psychologii potocznej (*folk psychology*) - odnoszą się do wielu indywidualów i sprawdzają się w życiu, tzn. są efektywne jako strategia wyjaśniająca zachowanie się ludzi dlatego, że te same treści (przedstawione symbolicznie jako  $P$  i  $Q$ ) powracają w kausalnych interakcjach między stanami mentalnymi. Strategia wyjaśniania zachowania się ludzi za pomocą tego typu uogólnień jest skuteczna dlatego, że relacje kausalne (często) respektują relacje semantyczne.

W przypadku komputerów przyporządkowanie treści mentalnych związkom syntaktycznym dokonuje się dzięki umysłowi ludzkiemu, który jest w stanie dostrzegać 'abstrakcyjne własności pewnych kształtów przestrzennych'. Na tej podstawie możliwe jest przyporządkowanie danych introspekcyjnie treści kształtom przestrzennym. Same te treści nie mają przecież żadnego kształtu przestrzennego. Hipoteza głoszona przez Fodora posiada szereg założeń: reprezentacjonistyczną teorię umysłu, funkcjonalizm, teorię identyczności psychofizycznej w jej wersji egzemplarycznej, kausalną teorię znaczenia. Umysł ludzki funkcjonuje jako system reprezentacji zrealizowanych w medium fizyczno-neuronalnym. Relacje między tymi reprezentacjami mają charakter czysto syntaktyczny, a dokładniej rzecz biorąc, posiadają one według Fodora charakter zdaniowy. Przyjmowanie zdań istniejących w naszych głowach przestaje się jednak wydawać dość dziwnym pomysłem, gdy weźmiemy pod uwagę, że zdania, o jakie tu chodzi, są rozumiane jako obiekty abstrakcyjne, których egzemplarze (*tokens*) mogą być zrealizowane na różne sposoby: jako formy fal dźwiękowych (w

wypadku języka mówionego), jako sekwencje kropek i kresek (kod Morse'a) lub jako sekwencje naładowanych elektrycznie cząstek (komputer). Według Fodora zdania te są zrealizowane w mózgu jako sekwencje egzemplarzy zdarzeń elektro-chemicznych w ten sposób, aby odzwierciedlały różnicowania syntaktyczne.

Jednym ze standardowych zarzutów wobec hipotezy 'języka myśli' jest to, że nie wszystkie istoty zdolne do spełniania aktów kognitywnych wyposażone są w język. Zwierzęta i małe dzieci są w stanie myśleć, lecz nie są w stanie mówić. Zwolennik koncepcji Fodora może próbować odpowiadać, że wprawdzie niektóre istoty nie są w stanie mówić, a niektórym brakuje języka publicznego, lecz z tego nie wynika, że nie posiadają one niemówionego, wewnętrznego języka mentalnego. Inny podstawowy zarzut dotyczy sposobu uczenia się pojęć. Akceptując hipotezę Fodora trzeba byłoby przyjąć, że uczenie się pojęć polega na przetwarzaniu i kombinacji stałego repertuaru reprezentacji mentalnych. Tego rodzaju kombinacja i przetwarzanie musiałoby móc wrażyć każdą treść, której kiedykolwiek bylibyśmy w stanie się nauczyć. Hipoteza 'języka myśli' łączy się więc z mocną tezą o wrodzoności naszego podstawowego repertuaru pojęciowego. Ta silna konsekwencja natywistyczna prowadzi z kolei do innego zarzutu, a mianowicie zarzutu atomizmu. Według Fodora dwie myśli pozostają do siebie w pewnej relacji wtedy, gdy mają wspólne elementy należące do 'języka myśli', a więc wspólne elementy należące do wewnętrznego kodu mózgowego, lecz same te elementy pozostają w stosunku do siebie w całkowitej izolacji - podobnie jak słowa w języku publicznym, według którego hipoteza 'języka myśli' jest modelowana. Semantycznie patrząc, elementy 'języka myśli' nie mają nic ze sobą wspólnego, nie należą do żadnej wspólnej 'przestrzeni semantycznej' (Maloney/Clark 1994, s. 401-412).

Teraz przechodzę do przedstawienia krytyki symbolicznej teorii myślenia, którą w ostatniej swojej książce sformułował L. Bonjour (Bonjour 1998, s. 162-186). Krytykę tę uważam za bardzo ważną, gdyż w sposób

jasny i radykalny kwestionuje podstawowe założenia dominującego dzisiaj naturalistycznego podejścia do natury umysłu ludzkiego. Głównym przedmiotem analiz BonJoura nie jest jednak metafizyczna natura umysłu (filozofia umysłu), lecz kwestia tradycyjnie uważana za epistemologiczną, tzn. kwestia możliwości wiedzy koniecznej (wiedzy *a priori*). W tym aspekcie BonJour argumentuje przeciw dominującemu współcześnie empiryzmowi i jednocześnie na rzecz możliwości racjonalnego wglądu (*rational insight*) w konieczne stany rzeczy. Zgodnie z tradycją pogląd uznający istnienie tego rodzaju wglądu nazwa racjonalizmem. Przeciwstawia go przede wszystkim różnym odmianom empiryzmu, ale także aprioryzmowi Kanta. Stanowiska empirystyczne, odrzucające możliwość wglądu w konieczne stany rzeczy, muszą się według BonJoura zawsze skończyć jakąś wersją sceptycyzmu. Jednocześnie jednak stanowisko Kanta BonJour uważa za niewystarczające, gdyż relatywizuje ono wiedzę konieczną (w terminologii Kanta wiedzę syntetyczną *a priori*) do faktycznej struktury umysłu ludzkiego. Wiedza konieczna według BonJoura nie pociąga za sobą konieczności przyjmowania istnienia przed-doświadczeniowych form naoczności i kategorii, lecz oparta jest na szczególnego rodzaju władzy intuicyjnego ujmowania konieczności. Fragmentem argumentacji BonJoura na rzecz tak rozumianego racjonalizmu jest krytyka symbolicznej natury myślenia. Gdyby bowiem symboliczna teoria myślenia była prawdziwa, to uzasadniana przez BonJoura możliwość wglądu w konieczne stany rzeczy musiałaby zostać uznana za całkowicie pozorną. Umysł przetwarzający wyłącznie symbole fizyczne nie jest w stanie dokonywać żadnych wglądów w konieczne stany rzeczy. W kwestii metafizycznego źródła znaczenia mentalnego BonJour ostatecznie preferuje rozwiązania Arystotelesa i św. Tomasza z Akwinu, według których źródło znaczenia nie opiera się na symbolach, lecz na procesie egzemplifikacji własności rzeczy w umyśle.

Symboliczna teoria umysłu jest uważana przez dużą część współczesnych filozofów za jedyną poprawną teorię, a nawet za jedyną teo-

rię dającą się w ogóle pomyśleć. Według tej teorii myślenie jest procesem o charakterze symbolicznym (lingwistycznym). Proces myślenia polega na operowaniu systemem reprezentacji, który jest przynajmniej analogiczny do języka naturalnego. Pogląd tego rodzaju nie tylko jest wpływowym stanowiskiem w dzisiejszej filozofii umysłu, ale można go też uznać za stanowisko charakterystyczne dla całego nurtu filozoficznego nazywanego filozofią analityczną.

Według tradycyjnego (racjonalistycznego) podejścia, gdy jakaś osoba wypowiada na przykład sąd, że czerwień i zieleń nie mogą jednocześnie występować na tej samej powierzchni, to świadomie ujmuje obie te jakości i na tej podstawie ‘widzi’, iż nie mogą one być w tym wypadku jednocześnie zrealizowane. Jakości czerwieni i zieleni znajdują się bezpośrednio w zasięgu jej ‘widzenia’ w ten sposób, że da się uchwycić ich niekompatybilność. Inaczej jest jednak według teorii symbolicznej. Tym, co się dzieje, gdy myślę czerwieni i zieleni, jest tylko to, że (polskie) słowa ‘czerwień’ i ‘zieleń’ (lub bliskie ich analogie) pojawiają się w moim ‘języku myśli’. Patrząc od strony jego implementacji, język ten mógłby być jakimś rodzajem kodu, pewnego rodzaju inskrypcją w postaci sekwencji powtarzających się wyładowań elektrycznych w mózgu. Współczesny stan badań neurofizjologicznych nie daje wprawdzie żadnych podstaw do przyjmowania, że tego rodzaju hipoteza filozoficzna jest empirycznie prawdopodobna, jest ona jednak atrakcyjna teoretycznie, gdyż pokazuje, jak materialistycznie można wyjaśnić sposób funkcjonowania naszego umysłu.

BonJour podkreśla, że według tego poglądu lingwistyczne (czy quasi-lingwistyczne) elementy nie stanowią tylko jakiejś ingrediencji w procesie myślenia, lecz że elementy tego rodzaju miałyby wyczerpywać cały subiektywny (wewnętrzny) aspekt myśli. Oczywiście, zwolennicy teorii symbolicznej muszą być w stanie pokazać, w jaki sposób niezinterpretowane symbole ‘języka myśli’ uzyskują to znaczenie (treść), które faktycznie posiadają. Muszą pokazać, jak to się dzieje, że nie są to tylko niezinterpreto-

wane symbole. Faworyzowanym w tym wypadku stanowiskiem jest tzw. kauzalna teoria znaczenia. Cokolwiek jednak kauzalna teoria znaczenia miałaby do powiedzenia na ten temat, to nie może ona sugerować istnienia jakiegoś elementu o charakterze treściowym (znaczeniowym), żadnego rodzaju wewnętrznej świadomości znaczenia (świadomości treści), które byłyby skojarzone z symbolami, ale od nich różne. Przyjęcie takiego myślowego, asymbolicznego elementu równałoby się odrzuceniu samego rdzenia stanowiska symbolicznego. Tego rodzaju dodatkowa świadomość tworzyłaby niesymboliczną odmianę myślenia, a ten rodzaj myślenia hipoteza nazywana hipotezą *'language of thought'* programowo odrzuca.

BonJour podkreśla, iż niezaprzeczalnym faktem jest to, że gdy myślimy o czymś, to treść naszych myśli - przynajmniej w pewnym znaczącym stopniu - jest dla nas widoczna, uświadamiamy ją sobie od wewnątrz. Jeżeli teoria symboliczna nie mogłaby wyjaśnić, jak możliwa jest ta wewnętrzna świadomość tego, o czym myślimy, to trzeba byłoby ją uznać za całkowicie niewiarygodną. Na rzecz takiej właśnie tezy o niewiarygodności argumentuje BonJour.

W jaki sposób jestem świadomy takich przedmiotów, własności i stanów rzeczy, jak na przykład: koty, telefony, piłki do koszykówki, drzewa wiśniowe w ogrodzie, Immanuel Kant, Bill Clinton, czerwień, trójkątność? W jaki sposób uświadamiam sobie to, że jedna rzecz jest dłuższa od drugiej, czy też to, że Napoleon przegrał bitwę pod Waterloo? BonJour zwraca uwagę, że teoria symboliczna koncentruje się zazwyczaj nie na wymienionych wyżej przykładach obiektów, własności i stanów rzeczy, lecz na przekonaniach i na praktykach przypisywania komuś przekonań. Jest to jednak według niego błąd, gdyż tego rodzaju praktyki w dużej mierze zależą od czynników kontekstualnych i znaczeniowych, a te czynniki są z kolei istotnie niezależne od tego, co rzeczywiście dzieje się w umyśle osoby, która je sobie uświadamia. W związku z tym BonJour preferuje przykłady dotyczące

obiektów, własności i stanów rzeczy, gdyż w ich przypadku występuje mniejsza zależność od czynników kontekstualnych.

Jak to jest możliwe, że możemy myśleć na przykład o ‘trójkątności’? Teoria symboliczna odpowiada, że jesteśmy zdolni myśleć o trójkątności dlatego, że posiadamy e g z e m p l a r z (*token*) słowa ‘trójkątny’ w naszym umyśle (mózgu), albo też dlatego, że dysponujemy jakąś jego translacją w języku, który wykorzystują nasze operacje kognitywne. To właśnie egzemplarz słowa ‘trójkątność’ stanowiłby wehikuł naszych myśli o trójkątności. Według niektórych interpretacji język ten jest identyczny z językiem naturalnym (lub jest jemu bliski), podczas gdy inne podejścia uznają, że jest to specjalny język, właśnie ‘język myśli’ jakoś zakodowany w naszym mózgu. To jednak według BonJoura nie ma tu rozstrzygającego znaczenia. Bardzo ważne jest natomiast to, że tego rodzaju teorie przyjmują, iż egzemplarze są symbolami w taki sam sposób, w jaki zwykle słowo ‘trójkątność’ jest symbolem, a więc że ich zdolność do reprezentowania nie jest wyznaczana przez ich wewnętrzną naturę. Ta zdolność do reprezentowania jest natomiast wyznaczona przez relacje zewnętrzne, w które wchodzi dany symbol (jego egzemplarz). Teoria symboliczna w tym znaczeniu stanowi odmianę tego, co Fodor nazywa ‘reprezentacjonistyczną teorią umysłu’. Według Fodora tzw. postawy propozycjonalne (*propositional attitudes*), a także inne stany umysłu reprezentujące coś, są relacjami do wewnętrznych reprezentacji, tzn. relacjami do odpowiednich lingwistycznych lub quasi-lingwistycznych egzemplarzy symboli (Fodor 1981).

W tym miejscu BonJour wprowadza odróżnienie pomiędzy (a) reprezentacjonistyczną teorią umysłu *simpliciter* a (b) reprezentacjonistyczno-symboliczną teorią umysłu. W wypadku reprezentacjonistycznej teorii umysłu *simpliciter* reprezentacje mentalne mają wewnętrzną treść niezależną od ich relacji zewnętrznych. Tego rodzaju teoria satysfakcjonowałaby następujące wymagania stawiane przez Fodora każdej ogólnej teorii umysłu. Dostarczałaby mechanizmu pośredniczącego pomiędzy umysłem a

abstrakcyjnymi stanami rzeczy. Wyjaśniałaby zjawisko nieprzejrzystości semantycznej (*opacity*). Nadawałby reprezentacjom formę logiczną. Wyjaśniałaby strukturalny paralelizm między myśleniem a językiem. Reprezentacjonistyczna teoria umysłu *simpliciter* znajduje się jednak według BonJoura w konflikcie z drugą doniosłą tezą Fodora (i innych), a mianowicie z tezą, że procesy psychiczne mają charakter czysto komputacyjny, a więc że tylko formalno-syntaktyczne cechy reprezentacji symbolicznych odgrywają w nich rolę. Ta teza jest więc w mniejszym lub większym stopniu równoważna z tezą o symbolicznym charakterze myślenia.

BonJour stwierdza, że istnieją dwa podstawowe problemy, z którymi musi się skonfrontować każda z wersji teorii symbolicznej. Po pierwsze, jeżeli treść reprezentująca pewnego symbolu myślowego nie pochodzi z jego własności wewnętrznych, to przez co treść ta jest determinowana? Co sprawia, że słowo 'trójkątność' reprezentuje tę a nie inną własność, a nie reprezentuje nieokreślonej liczby jakichkolwiek innych rzeczy, które mogłoby reprezentować (albo w ogóle nic nie reprezentuje). Po drugie, w jaki sposób osoba myśląca uzyskuje bezpośredni dostęp do treści reprezentujących związanych z symbolami, które posiada w umyśle? Jak ta osoba uświadamia sobie te treści? Te problemy nie pojawiają się w wypadku zwykłych symboli językowych. Tutaj oczywistym - chociaż dzisiaj niemożliwym - wyjaśnieniem jest to, że treść symbolom nadają odnoszące się do nich myśli. To rozwiązanie w oczywisty sposób nie wchodzi w rachubę dla teorii symbolicznej, gdyż według niej to same myśli są symbolami (np. inskrypcjami w mózgu). Jeśli zaś chodzi o drugi problem, problem bezpośredniego dostępu, to gdy myślimy na przykład o martwym krabie na plaży, to mamy tę treść przed oczyma, możemy ją odróżnić od innych treści, a ten fenomen jest według BonJoura z pewnością różny od mentalnego posiadania czystych symboli, do których nie mamy żadnego bezpośredniego dostępu.

Na pytanie pierwsze, co sprawia, że na przykład słowo ‘trójkątność’ reprezentuje tę a nie inną własność, istnieją współcześnie dwie strategie odpowiedzi. Pierwszą nazywa się kauzalno-historyczną teorią znaczenia (treści) reprezentacji mentalnych. Drugą określa się mianem semantyki ról konceptualnych.

Teorie kauzalno-historyczne uznają, że pewien określony symbol posiada treść reprezentującą (znaczenie) dlatego, że egzemplarze tego symbolu znajdują się w relacji przyczynowej do tego, co reprezentują. Na przykład egzemplarze ‘języka myśli’, które reprezentują ‘trójkątność’, posiadałyby tę właśnie treść dlatego, że znajdowałyby się w zewnętrznej, przyczynowej relacji do rzeczy trójkątnych (ewentualnie w relacji przyczynowej do abstrakcyjnej i ogólnej własności trójkątności). W stosunku do takiego stanowiska Bonjour wysuwa jednak bardzo podstawowy zarzut. Zgodnie z tą teorią osoba myśląca o trójkątności musiałaby pozostawać niejako ‘uwięziona’ po jednej stronie relacji przyczynowej, a mianowicie miałaby ona dostęp tylko o do skutku, więc dostęp tylko do egzemplarzy danego symbolu. Nie miałaby więc żadnych szans stania się świadomą, co reprezentuje ten symbol. Nie miałaby żadnego pojęcia, o czym myśli, mając do dyspozycji tylko obecność samego egzemplarza symbolu w swoim umyśle. Trzeba tu natychmiast zastrzec, że według teorii symbolicznej powiązanie między relacją kauzalno-historyczną a samą treścią nie jest znowu jakimś rodzajem relacji kauzalnej. Pogląd tego rodzaju nie głosi, że relacja przyczynowa niejako dokonuje ‘infuzji’ treści do surowego symbolu, która to treść może potem stać się dostępna dla świadomego podmiotu. Teza jest tu raczej taka (i wydaje się, że nie może być inna), iż relacja kauzalno-historyczna konstytuuje posiadanie przez egzemplarz symbolu pewnej treści w ten sposób, że treść ta realizuje się wtedy, gdy realizuje się cała relacja, nie zaś gdy realizuje się tylko któryś z jej składników sam dla siebie. Z tego musi wynikać, że podmiot miałby dostęp do treści swoich myśli tylko wtedy, gdyby miał dostęp do całej sytuacji kauzalno-historycznej. Wydaje się

jednak, że taki rezultat jest absurdalny. Bonjour stwierdza z pewną ironią, że jednak wiemy, o czym myślimy, żadne więc odwołanie się do samych relacji przyczynowych nie jest w stanie wyjaśnić myślenia. Samo istnienie relacji przyczynowych pomiędzy nami a pewnym przedmiotem lub własnością nie jest wystarczające, aby umieścić ten przedmiot lub własność ‘przed naszym umysłem’ w ten sposób, abyśmy mogli mieć o nim myśli.

Podobna trudność według Bonjoura dotyczy drugiego stanowiska próbującego wyjaśnić treść mentalną (znaczenie) za pomocą odwołania się do relacji zewnętrznych pomiędzy umysłem a światem. W tym wypadku treść symboli umysłowych, treść ich egzemplarzy, miałyby być ukonstytuowana przez praktyki i konwencje wspólnoty językowej, w której dana osoba się znajduje (lub przez praktyki językowe ekspertów w danej dziedzinie). Jednakże i przy tej próbie rozwiązania ten, kto używałby egzemplarza danego symbolu, nie miałby dostępu do całej społecznej złożoności tego użycia (a nawet do jego większej części), a stąd treść jego myśli nie mogłaby być dla niego świadomie i bezpośrednio dostępna. Tak więc gdybyśmy zaakceptowali najbardziej dzisiaj obiegową teorię symboliczną myślenia, to musielibyśmy nieodwołalnie dojść do wniosku, że nie posiadamy żadnej świadomości naszych myśli, żadnego wewnętrznego dostępu w ogóle do tego, o czym myślimy.

BonJour sugeruje w tym miejscu, że zwolennik teorii symbolicznej dysponuje dwiema strategiami obronnymi w stosunku do przedstawionych zarzutów. Po pierwsze, może próbować zanegować, że treści reprezentujące (znaczenia) są wewnątrznie i bezpośrednio dostępne. Po drugie, może też stracić się wykazać, że źródłem treści myśli (znaczenia myśli) nie są relacje zewnętrzne umysłu do świata, lecz że treści te powstają w wyniku funkcjonowania całego systemu semantycznego danej osoby. W takim wypadku zwolennik teorii symbolicznej będzie odwoływał się do teorii nazywanej ‘semantyką ról konceptualnych’. Gdyby semantyka ról konceptualnych mogła wyjaśnić, jak egzemplarze symboli otrzymują swoje znac-

nie, to wewnętrzny i bezpośredni dostęp do naszych myśli nie byłby nam potrzebny.

Pierwsza strategia obronna symbolizmu polegać musi na jakimś argumente na rzecz tego, że nie mamy bezpośredniego dostępu do treści (znaczenia) naszych myśli. Istnieje taki argument, a jego hasłem wywoławczym jest stwierdzenie Putnama, że ‘znaczenie lub treść nie są czymś, co istnieje w umyśle (*‘content or meaning just ain’t in the head’*)’ (Putnam 1975, s. 227). Bonjour zauważa, że przeciwnik symbolizmu, który chciałby odrzucić teorię znaczenia Putnama, nie musi jednak twierdzić, że wszystko, co on powiedział na temat znaczenia, jest fałszywe, albo że cała jego teoria jest tylko pewnym pomysłowym trickiem filozoficznym. Wystarczy zwrócić uwagę, że poglądy Putnama w tej kwestii są przesadzone i że wynikająca z nich dystynkcja pomiędzy wąską a szeroką treścią (znaczeniem w sensie wąskim i znaczeniem w sensie szerokim) da się utrzymać, ale że jednocześnie można twierdzić, iż to właśnie wąska treść jest nam wewnętrznie i bezpośrednio dostępna. Jeżeli jednak wąska treść jest wewnętrznie i bezpośrednio dostępna, to symboliczna teoria umysłu pozostaje nadal fałszywa.

BonJour referuje eksperyment myślowy Putnama, według którego istnieje fikcyjna bliźniacza Ziemia wraz z mieszkańcami, którzy co do ostatniej molekuly są tacy sami jak my. Zachodzi tylko jedna różnica: podczas gdy my dzięki badaniom przyrodniczym wiemy, że mokra i przezroczysta substancja, którą nazywamy wodą, jest  $H_2O$ , to wyniku badań przyrodniczych na Ziemi Bliźniaczej okazuje się, że mokra i przezroczysta substancja, zewnątrznie nieodróżnialna od naszej wody, jest w rzeczywistości XYZ. Gdy więc, powiedzmy, Hilary<sup>1</sup> na Ziemi i Hilary<sup>2</sup> na Ziemi Bliźniaczej myślą o wodzie przed dokonaniem badań naukowych, to treści ich myśli są identyczne. Według Putnama okazuje się jednak, że prawdziwym znaczeniem słowa woda na Ziemi jest  $H_2O$ , natomiast na Ziemi Bliźniaczej tym prawdziwym znaczeniem jest XYZ. Stąd Putnam wnosi, że znaczenie

(treść) nie miałyby być czymś, co istnieje w naszych głowach, tzn. czymś, co jest nam bezpośrednio dostępne.

W trakcie dyskusji nad tym eksperymentem myślowym zwrócono jednak uwagę, że przed dokonaniem badań przyrodniczych bezpośrednio dostępne dla Hilarego<sup>1</sup>, jak i dla Hilarego<sup>2</sup> są wąskie treści, natomiast różnią się oni co do niedostępnej im bezpośrednio przed badaniami treści szerokiej, którą stanowią odpowiednio H<sub>2</sub>O i XYZ. Spostrzeżenie Putnama na temat pojęcia treści wydaje się jednak sprowadzać tylko do tego, że treść naszych myśli, która jest nam bezpośrednio dostępna, nie determinuje treści szerokiej, czyli składu chemicznego. Aby jednak móc obronić teorię symboliczną, potrzebna byłaby jednak według BonJoura teza mocniejsza, a mianowicie, że nie jesteśmy świadomi treści (znaczenia) w żadnym stopniu. Szeroka treść dotyczy relacji semantycznej pomiędzy reprezentacją a światem i jest ona funkcją tego, co rzeczywiście świat zawiera oraz funkcją innych czynników kontekstualnych. Tylko wąska treść jest nam bezpośrednio dostępna (choć bardzo trudno jest określić granicę między tymi odmianami treści). Jeżeli jest jednak dostępna, to nie może być funkcją samych 'surowych symboli'.

BonJour wyciąga stąd wniosek, że relacje kauzalne nie mogą wyjaśnić treści myśli, a jeżeli przynajmniej wąska treść jest niewątpliwie nam dostępna, to powstaje pytanie, jakie jeszcze rozwiązanie pozostaje zwolennikowi teorii symbolicznej. Wydaje się, że pozostaje mu mocny argument odwołujący się do tzw. semantyki ról konceptualnych. Jest to próba wyjaśnienia treści myśli za pomocą odwołania się do wewnętrznych relacji pomiędzy samymi symbolami. Jako reprezentatywne dla tego stanowiska BonJour wymienia poglądy N. Blocka (Block 1986, s. 615-678). Potrzebne są jednak dwa zastrzeżenia. Po pierwsze, zwolennicy semantyki ról konceptualnych głoszą tę teorię wraz z kauzalną teorią znaczenia. Kauzalna teoria znaczenia została omówiona i skrytykowana poprzednio. Po drugie, zwolennicy semantyki ról konceptualnych nie mogą się odwołać do

relacji między symbolami (między ich egzemplarzami) a innymi wewnątrznie dostępnymi stanami umysłu, np. percepcjami zmysłowymi czy wyobrażeniami (obrazami). Wtedy bowiem musieliby przyjąć, iż te inne stany posiadają wewnętrzną (istotnie z nimi związaną) treść (znaczenie), że nie są tylko zwykłymi egzemplarzami symboli. To jednak byłoby ukryte odrzucenie teorii, którą głoszą, czyli teorii symbolicznej.

Podstawowym argumentem ‘semantyki ról konceptualnych’ jest to, że znaczenie lub treść jakiegoś zdania w ‘języku myśli’ wywodzą się z roli, jaką zdania tego typu odgrywają w różnych rodzajach wnioskowań dedukcyjnych i probabilistycznych. Dla każdego zdania z ‘języka myśli’ istniałyby określone struktury inferencyjne zawierające inne zdania należące do języka myśli, struktury, w których zdanie to raz występowałoby jako wniosek, innym razem zaś jako przesłanka. Gdybyśmy mogli tego rodzaju struktury w pełni rozpoznać, to określiłyby one dokładnie, na czym polega wąska treść tego zdania. Pojedyncze symbole, potraktowane osobno, mogłyby wprawdzie znaczyć coś całkiem innego niż faktycznie znaczą, ale wzięte wraz z całym kontekstem potencjalnych struktur inferencyjnych miałyby całkowicie zdeterminowaną treść. ‘Posiadać treść’, to według tej teorii ‘istnieć w specyficznej sieci relacji inferencyjnych’.

Dla lepszego wyjaśnienia tej koncepcji Bonjour referuje następujące porównanie zaproponowane przez Fodora (Fodor 1987, s. 78-79). Abstrakcyjne treści myśli spełnianych z asercją, a więc sądy w sensie logicznym (*propositions*), tworzyłyby swoistego rodzaju sieć logiczną, powiazaną relacjami dedukcyjnymi i probabilistycznymi. Z drugiej strony znajdowałyby się zdania należące do ‘języka myśli’. Zdania te pozostawałyby do siebie w relacjach przyczynowo-inferencyjnych. Chodzi tu o te wnioskowania, których użytkownik ‘języka myśli’ faktycznie dokonuje lub co do których znajduje się dyspozycji do ich wypowiedzenia. Zdania te tworzyłyby sieć kauzalno-inferencyjną. Według Fodora możemy teraz założyć, że sieć logiczna i sieć kauzalno-inferencyjna są ze sobą skorelowane

(są izomorficzne). Wtedy teza zwolennika symbolicznej teorii myślenia brzmiałaby następująco: określony element należący do ‘języka myśli’ (określony symbol) posiada jako swoją treść tę treść abstrakcyjną z sieci logicznej, z którą jest skorelowany. Bonjour podkreśla w tym miejscu, że zwolennik symbolizmu, odwołujący się do semantyki ról konceptualnych, nie może zakładać treści, które chce wyjaśnić. Inaczej mówiąc, relacje inferencyjne pomiędzy zdaniem należącymi do ‘języka myśli’ powinny dać się zdefiniować w terminach czysto kauzalnych (nie-intencjonalnych, nie-semantycznych). Osoba, która znajduje się w dyspozycji do uznania pewnej konkluzji, robi to wtedy, gdy znajduje się w dyspozycji do uznania przesłanek. Tylko do takiej sytuacji semantyka ról konceptualnych może się odwołać. Inne elementy należące do języka myśli, z którymi dany element jest powiązany, też muszą uzyskiwać swoje znaczenie tylko przez to, że można je umieścić w całej kauzalno-inferencyjnej sieci powiązań.

BonJour formułuje następujące trudności, z którymi semantyka ról konceptualnych powinna dać sobie radę. Jednocześnie jednak stwierdza, że najbardziej podstawowe z tych trudności zupełnie dyskwalifikują tę teorię jako możliwą teorię wąskiej, bezpośrednio dostępnej świadomości treści. Oto najważniejsze z nich.

**Problem informacji pobocznej.** Zwolennik semantyki ról konceptualnych musi przyjąć, że każde dodatkowe przekonanie, które ktoś nabywa, musi wpłynąć na znacznie danej reprezentacji. Gdyby tak jednak było, to należałoby uznać, że dwie osoby posiadające istotnie odmienne zbiory przekonań, nie mogłyby mieć żadnego wspólnego przekonania.

**Problem niedopasowania.** Relacje wynikania zachodzące między treściami abstrakcyjnymi (*propositions*) są o wiele bogatsze i bardziej skomplikowane niż relacje kauzalno-inferencyjne pomiędzy symbolami-egzemplarzami. Postulowana przez Fodora izomorfia mogłaby być więc co najwyżej tylko częściowa. Być może istnieje wiele częściowych izomorfii pomiędzy siecią logiczną a siecią kauzalno-inferencyjną. Gdy

weźmiemy na przykład dwa słowa należące do ‘języka myśli’, KOT i LEW (duże litery są tu użyte dla podkreślenia, że chodzi o mózgowie inskrypcje należące do hipotetycznego języka myśli), to być może dla każdego zdania zawierającego symbol KOT istnieje strukturalnie izomorficzne zdanie zawierające symbol LEW, wtedy jednak trzeba byłoby stwierdzić, że sama struktura nie daje żadnych wskazówek, które ze zdań ‘języka myśli’ należałoby skorelować z abstrakcyjnymi treściami na temat kotów. Każdy bowiem faktycznie istniejący język zawiera terminy odpowiadające stosunkowo małej liczbie własności, rodzajów i gatunków.

**Problem warunku wystarczającego.** Nawet jeśli założymy, że izomorfia pomiędzy relacjami inferencyjnymi w ‘języku myśli’ a abstrakcyjnymi relacjami wynikania w sieci logicznej jest ścisła i ma charakter unikalny, to nadal pozostaje problem, czy z tego wynika, że zdania należące do ‘języka myśli’ mają tę samą wąską treść, co skorelowane z nimi zdania dotyczące treści abstrakcyjnych. Dlaczego nie mogłoby być tak, że pomimo istnienia unikalnej korelacji zdania ‘języka myśli’ nie miałyby żadnego w ogóle znaczenia. Posiadanie odpowiedniej struktury może wprawdzie być - ale nie musi - warunkiem koniecznym posiadania treści, ale posiadanie odpowiedniej struktury nie jest z pewnością warunkiem wystarczającym posiadania treści. Nie ma żadnych powodów, aby twierdzić, że izomorfizm, o którym mowa, stanowi warunek wystarczający posiadania treści.

**Problem bezpośredniego dostępu.** Gdy myślę na przykład o krabach, to jestem w stanie pomyśleć również o różnicy pomiędzy krabami a homarami. Ta zdolność do myślenia o jednych treściach na podstawie innych treści nie może jednak, jak chce zwolennik teorii ról konceptualnych, polegać na uchwyceniu straszliwie skomplikowanej struktury relacji kauzalno-inferencyjnych pomiędzy słowem KRAB a słowem HOMAR. Normalnie jest tak, że najpierw intuicyjnie rozumiemy pewną treść, a potem dopiero, mając ją do dyspozycji, dokonujemy na jej podstawie takich lub innych wnioskowań. Według semantyki ról konceptualnych

musiałyby być jednak odwrotnie, a mianowicie treści musiałyby powstawać dzięki wnioskowaniom. Jeżeli jednak by tak było, to skąd istnieje w nas silne przeświadczenie, że najpierw mamy dane intuicyjnie (wąskie) treści, a potem dopiero możemy przechodzić do procesu wnioskowania.

Problem wyjaśnienia bezpośredniego dostępu. Jeżeli nasza świadomość treści nie pochodzi ze świadomego uchwycenia ogromnie skomplikowanej struktury kauzalno-inferencyjnej, jak musiałby twierdzić zwolennik teorii ról konceptualnych, to wtedy potrzebujemy jakiegoś niezależnego wyjaśnienia, jak rzeczywiście dochodzimy do uchwycenia tych treści, rzekomo ukonstytuowanych przez tę skomplikowaną strukturę. Wyjaśnienie to nie może się odwoływać, jak widzieliśmy, do symboli i relacji między nimi. Wtedy jednak, jak się wydaje, dla teorii symbolicznej nie istnieje już żaden inny ruch do wykonania.

Bonjour dodaje na koniec następującą pointę. Żadna czysto relacyjna teoria myślenia nie jest w stanie wyjaśnić możliwości dojścia do skutku myślenia refleksyjnego, które operuje bezpośrednio dostępnymi świadomości treściami.<sup>1</sup> Gdy mimo to ktoś obstaje przy teorii symbolicznej, to musimy albo dojść do sceptycznego wniosku, że po prostu nie znamy swojego umysłu (nie wiemy, o czym myślimy), albo ten rezultat należy potraktować jako formę *reductio ad absurdum* relacyjnych wyjaśnień myślenia. Symboliczną teorię myślenia trzeba więc uznać za stanowisko całkowicie nie do utrzymania.

## BIBLIOGRAFIA

- BLOCK N. 1986: *An Advertisement for a Semantics for Psychology*, w: „Midwest Studies in Philosophy” 10, s. 615-678.
- BOGHOSSIAN, P. 1989: *Content and Self-Knowledge*, w: „Philosophical Topics” 17:1989, s. 5-26.

---

<sup>1</sup> Według Bonjour w całej literaturze dotyczącej tej problematyki tylko P. Boghossian dochodzi do podobnych wniosków (Boghossian 1989, s. 5-26).

- BONJOUR L. 1998: *Defence of Pure Reason. A Rationalist Account of A Priori Justification*, Cambridge: Cambridge University Press.
- FODOR J.A. 1975: *The Language of Thought*, New York: Thomas Y. Crowell.
- FODOR J.A. 1981: *Representations*, Cambridge MA: MIT Press.
- FODOR J.A. 1987: *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*, Cambridge MA: MIT Press.
- MALONEY J. CH. / CLARK A. 1994: *Language of Thought*, w: S. Guttenpalm (ed.), *A Companion to the Philosophy of Mind*, Cambridge MA: Blackwell, s. 401-412.
- PUTNAM H. 1975: *The Meaning of 'Meaning'*, w: H. Putnam, *Mind, Language and Reality*, Cambridge: Cambridge University Press.