

Jakub Gomułka

NIEINTUICYJNA CZY NIEKONSEKWENTNA TEORIA INTENCJONALNOŚCI? SPÓR MIĘDZY DENNETTEM I SEARLEM

STRESZCZENIE

W artykule zestawiam dwie koncepcje intencjonalności: Johna Searle'a oraz Daniela Dennetta. Wyjaśniam związek pomiędzy teoriami Searle'a w dziedzinie filozofii języka a jego stanowiskiem w kwestii intencjonalności, następnie prezentuję jego ideę trzech rodzajów intencjonalności. W dalszej kolejności przedstawiam Dennetta teorię intencjonalności jako cechy adaptacyjnej, a także jego koncepcję nastawienia intencjonalnego. Podkreślam fakt, że według Dennetta – używając terminologii Searle'a – żadna istota bądź wytwór techniki nie posiada intencjonalności wewnętrznej, a jedynie pochodną, tj. taką, którą przypisujemy danemu bytowi w celu utworzenia lepszych przewidywań zachowania tego bytu. Konsekwencją tej tezy jest nieintuicyjna: Dennett musi przyznać (i faktycznie to robi), że także intencjonalność obserwatora jest intencjonalnością tylko przypisywaną mu przez jakichś innych obserwatorów w celu tworzenia lepszych przewidywań jego zachowania. Searle odrzuca to twierdzenie jako jawnie absurdalne. Staram się pokazać, że stanowisko Searle'a jest zakorzenione w kartezjanizmie i wskazuję na podobieństwa pomiędzy jego krytyką Dennetta i Husserlowską krytyką naturalizmu. Biorąc pod uwagę fakt, że Searle uważa się za naturalistę, należy uznać jego teorię za niekonsekwentną.

Słowa kluczowe: intencjonalność, redukcjonizm, naturalizm biologiczny, John Searle, Daniel Dennett

WSTĘP

Współczesna filozofia umysłu jest areną wielowątkowego sporu między zwolennikami szeroko lub wąsko rozumianego materializmu i redukcjonizmu a przedstawicielami emergentyzmu, dualizmu własności, czy innych prób nieredukcyjnego ujęcia problematyki związanej z zagadnieniem relacji ciało-umysł. Do tego obszaru należy między innymi problem intencjonalności, a także problem *qualiów*. Sądzę, że wszystkie szczegółowe problemy

filozofii umysłu mają ze sobą wiele wspólnego, w szczególności uważam, że zagadnienie *qualiów* wykazuje pewne strukturalne podobieństwo do zagadnienia intencjonalności, choć oczywiście każde z nich posiada własną specyfikę. Ten artykuł dotyczy zagadnienia intencjonalności treści mentalnych, a konkretnie sporu o intencjonalność, który prowadzą ze sobą Daniel C. Dennett i John R. Searle.¹ Wybór adwersarzy uzasadniony jest faktem, że obaj autorzy bez wątpienia zaliczają się dziś do czołówki najbardziej wpływowych myślicieli zajmujących się filozofią umysłu, a jednocześnie ich stanowiska są całkowicie sobie przeciwstawne w pewnym bardzo istotnym punkcie. Analiza owych stanowisk, a także przebiegu dyskusji na temat intencjonalności, pozwoli mi odsłonić charakter i skalę problematyczności teorii Dennetta i Searle'a.

Pierwsza część artykułu poświęcona będzie propozycji Searle'a. Omówiona zostanie geneza jego teorii intencjonalności, struktura dziedziny intencjonalnej (tzw. Sieci i Tła), a także rozróżnienie trzech typów intencjonalności. W części drugiej zaprezentuję pewne istotne dla problemu intencjonalności koncepcje Dennetta: teorię trzech nastawień i ideę wieży generowania i testowania. W ostatniej części zestawię poglądy obu autorów, aby wskazać pewne ciekawe zbieżności, a także uwypuklić punkty sporne. W zakończeniu wspomnę o własnym stanowisku w kwestii intencjonalności.

INTENCJONALNOŚĆ WEDŁUG SEARLE'A

Specyfika podejścia Searle'a do problematyki intencjonalności bierze się z jego próby pogodzenia dwóch skonflikowanych stanowisk. Według pierwszego z nich współczesne problemy filozofii umysłu, poczucie tajemnicy czy niezwykłej trudności pewnych kluczowych zagadnień, są w dużej mierze iluzją związaną z naszym przywiązaniem do tradycyjnych kategorii, które odziedziczyliśmy po Kartezjuszu (Searle 1995, 12n). Tymczasem stany intencjonalne – jak wskazuje nasza aktualna wiedza o świecie – są, po pierwsze, przyczynowo warunkowane przez stany neurofizjologiczne, po drugie realizowane w naszych mózgach. Searle twierdzi więc, że tzw. problem umysł-ciało po prostu nie istnieje, podobnie jak nie istnieje problem żołądek-trawienie (Searle 1983, 15). Swoje stanowisko określa mianem naturalizmu biologicznego i uważa je za całkowicie oczywiste (Searle 1999a, 15).

¹ Ze względu na to ograniczenie tematyczne prezentacja dyskusji między Searlem i Dennettem będzie oczywiście niepełna: obu filozofów dzieli też podejście do zagadnienia *qualiów*. Niekiedy „wykrajanie” problemu intencjonalności będzie zabiegiem nieco sztucznym, ponieważ np. w argumentacji Searle'a przytoczonej w jego *The Mystery of Consciousness* problem *qualiów* wyraźnie interferuje z problemem intencjonalności (Searle 1997, 95–131). Trzeba jednak zastrzec, że Searle przejawia czasem pewną rezerwę wobec koncepcji *qualiów* (Searle 1997, 8n).

Zarazem jednak Searle odrzuca materializm i wszelkie formy redukcjonizmu utrzymując, że intencjonalność musi pozostać pojęciem pierwotnym. Píše:

Według mnie nie da się przeprowadzić analizy logicznej intencjonalności [sfery] umysłowej w terminach pojęć bardziej elementarnych, ponieważ intencjonalność jest, że tak powiem, własnością parterową umysłu, a nie cechą logicznie złożoną, zbudowaną z prostszych elementów. Nie istnieje stanowisko neutralne, z którego moglibyśmy badać relacje między stanami intencjonalnymi a światem, a następnie opisywać je w kategoriach nieintencjonalnych. Z tego powodu każde wyjaśnienie intencjonalności ma miejsce wewnątrz kręgu pojęć intencjonalnych. (Searle 1983, 26)²

Powyższy cytat wymaga pewnego wyjaśnienia. Searle – jak sam pisze (Searle 1999a, 9n) – zainteresował się problematyką intencjonalności ze względu na swoje wcześniejsze badania związane z teorią czynności mowy. W jego mniemaniu bowiem teoria ta pozostałaby niepełna, gdyby nie została podbudowana jakąś koncepcją intencjonalności. Filozofia języka jest według niego dziedziną filozofii umysłu, a sam język jest pochodną struktury intencjonalnej świadomości. Zarazem jednak narzędzie badawcze, jakie filozof stosuje do badań intencjonalności, jest takie samo, jak w przypadku badań języka: jest nim nieformalna analiza logiczna, metoda przejęta od Gilberta Ryle'a, Johna L. Austina, Petera Strawsona, częściowo także inspirowana późnym Ludwigiem Wittgensteinem.³ Analiza, o której mowa, nie polega na redukcji pojęć złożonych do pojęć pierwotnych, lecz jest najczęściej metodą odsłaniania wzajemnych związków między różnymi pojęciami; przyjmuje się bowiem, że to, co najlepiej znane nie musi być wcale logicznie podstawowe. Tak właśnie jest w przypadku Searle'owskich logicznych analiz intencjonalności: amerykański filozof prowadzi je wychodząc od czegoś lepiej znanego (wypowiedzi językowych) i dochodząc do czegoś bardziej pierwotnego (stanów intencjonalnych) ujawniając po drodze zależności i podobieństwa strukturalne między jednym a drugim.

Według Searle'a struktura stanów intencjonalnych odpowiada strukturze czynności mowy. Typowa czynność mowy zawiera dwa wyróżnione momenty: treść zdaniową (*propositional content*) i siłę illokucyjną (*illocutionary force*). Ta sama treść zdaniowa (np. *to, że wyjdiesz z pokoju*) może występować w wielu czynnościach różniących się między sobą siłą illokucyjną (np.

² Ten i inne cytaty z języka angielskiego w tłumaczeniu autora.

³ Na marginesie, warto zapoznać się z dyskusją na temat metody Searle'a pomiędzy nim a Hubertem L. Dreyfusem (Dreyfus 2002; Searle 2002; Dreyfus 1999). Dreyfus przez lata uważał Searle'a za fenomenologa. Swoją pomyłkę zauważył dopiero wówczas, gdy ten sam mu ją wytknął (w drugim z wyżej wskazanych artykułów). Można to potraktować jako dobre empiryczne świadectwo na rzecz tezy, że metoda fenomenologiczna w praktyce niewiele różni się od metody analizy języka potocznego. Do podobieństw między stanowiskiem Searle'a i Husserla wróć jeszcze pod koniec niniejszego artykułu.

w *oczekiwaniu*, że wyjdiesz z pokoju, *sugestii*, żebyś wyszedł z pokoju, *rozkazie*, że wyjdiesz z pokoju itp.). W przypadku stanu intencjonalnego mamy odpowiednio moment treści reprezentacyjnej (*representative content*) – zwanej również treścią intencjonalną (*intentional content*) – oraz trybu psychologicznego (*psychological mode*). Przykładowo, treścią reprezentacyjną (bądź intencjonalną) może być to, że wyjdiesz z pokoju, natomiast trybami psychologicznymi mogą być obawa, nadzieja, przekonanie itp. (Searle 1983, 5n).

Co prawda treść reprezentacyjna stanu intencjonalnego – jeśli w ogóle chcemy o nim mówić – musi być wyrażana w zdaniach bądź nazwach (np. gdy chodzi o oczekiwanie na Boże Narodzenie, albo miłość do kogoś), to jednak nie jest ona tym samym, co treść zdaniowa czynności mowy. Treść reprezentacyjna, jak wskazuje jej nazwa, coś reprezentuje. Tym czymś są warunki spełniania (*conditions of satisfaction*) danego stanu intencjonalnego, czyli fakty lub przedmioty (bądź osoby). Jeśli stan intencjonalny jest przekonaniem o czymś, np. o tym, że śnieg pada, to jego treścią reprezentacyjną jest fakt będący warunkiem spełniania tego przekonania, np. fakt, że śnieg pada (Searle 1983, 11n).

Treść zdaniowa czynności mowy nie reprezentuje, lecz *znaczy*. Znaczenie (*meaning*) jest cechą właściwą wypowiedziom i przydaje im to, czego nie mają one same z siebie, a mianowicie intencjonalnego odniesienia do warunków satysfakcji. Tu właśnie, według Searle'a, ujawnia się pochodność języka względem subiektywnych stanów mentalnych. Te ostatnie same z siebie posiadają wewnętrzną (*intrinsic*) intencjonalność, są więc wewnętrznie związane z rzeczywistością i nie należy ich rozumieć jako pewnych symboli wewnętrznego języka, które dopiero muszą być rozpoznane i zdekodowane przez wewnętrznego homunkulusa.⁴ Natomiast wypowiedzi – produkty czynności mowy – będące pewnymi mniej lub bardziej regularnymi układami fizycznych przedmiotów lub zjawisk, mogą mieć tylko pochodną intencjonalność, którą nadaje im umysł w akcie wypowiedzi bądź interpretacji. Searle pisze:

„Znaczenie” jest pojęciem, które ma literalne zastosowanie do wypowiedzi zdaniowych i czynności mowy, ale nie do stanów intencjonalnych. Przykładowo, ma sens pytanie co znaczy jakieś zdanie albo wypowiedź, lecz niedorzecznością jest pytać w tym sensie o to, co znaczy przekonanie lub pragnienie. [...] Znaczenie istnieje wyłącznie tam, gdzie istnieje różnica między treścią intencjonalną i formą jej uzewnętrznienia, a pytać o znaczenie to pytać o treść intencjonalną, która pojawia się wraz z ową formą uzewnętrznienia (Searle 1983, 28).

⁴ To by oczywiście generowało regres w nieskończoność, gdyż automatycznie powstawałoby pytanie o stany mentalne tegoż homunkulusa.

Widzimy zatem, że semantyka Searle'a jest zakorzeniona w subiektywnych stanach mentalnych: według niego znaczenie wypowiedzi językowych jest pewną relacją do rzeczywistości, lecz jest to relacja zapośredniczona w stanach intencjonalnych, ponieważ tylko one posiadają intencjonalność wewnętrzną. Powstaje jednak pytanie, w jaki sposób stany intencjonalne mogą posiadać ową wewnętrzną intencjonalność? Odpowiedzią Searle'a są idee Sieci (*Network*) i Tła (*Background*).

Koncepcja Sieci wynika z konstatacji, że nie ma sensu mówić o treści intencjonalnej jako o czymś atomowym i samodzielny. Gdy mówimy na przykład o kimś, kto chce startować w wyborach na prezydenta Stanów Zjednoczonych, musimy założyć, że ów pojedynczy stan intencjonalny musi z konieczności zakładać bardzo wiele innych stanów intencjonalnych, najczęściej o formie przekonań (np. o tym, że Stany Zjednoczone są republiką z ustrojem prezydenckim, że stanowisko prezydenta jest kadencyjne itd.). Tylko dzięki temu, że dany stan intencjonalny posiada swoje miejsce w Sieci innych stanów intencjonalnych, da się w ogóle zdeterminować jego warunki spełniania (Searle 1983, 141).⁵

Samo istnienie Sieci nie wystarcza jednak do tego, by stan intencjonalny mógł w ogóle posiadać intencjonalność, a więc związek z rzeczywistością. Drugim koniecznym elementem jest Tło, czyli zespół założeń nieintencjonalnych, a więc zdolności, umiejętności społecznych i nawyków. Searle naprowadza nas na koncepcję Tła wskazując, że śledząc uwarunkowania dowolnego węzła Sieci (np. wspomnianego stanu intencjonalnego chęci kandydowania na prezydenta) dochodzimy w końcu do sformułowań, o których trudno nam powiedzieć, że są przekonaniem (nawet nieświadomymi): że wybory dokonują się na powierzchni Ziemi, że jest ona twarda i nie zapada się, że karty do głosowania nie rozpadają się w powietrzu na chmurę cząstek itd. Ponadto, właściwie wszystkie nasze zamiary uwarunkowane są pewnym zespołem nabytych bądź wrodzonych możliwości działania: umiemy jeść, biegać, ale także jeździć na nartach, prowadzić samochód, wreszcie posługiwać się jednym bądź większą ilością języków (Searle 1983, 142n). Wszystko to wskazuje, że Sieć stanów intencjonalnych musi w pewnym momencie stopniowo przechodzić w nieintencjonalne Tło. Searle dowodzi:

Żałómy, że prawdziwe jest przeciwieństwo hipotezy Tła, to znaczy, że całe intencjonalne życie umysłowe i wszystkie zdolności poznawcze mogą być całkowicie zredukowane do reprezentacji [...]. Każda z tych reprezentacji może być wyrażana w postaci jasnej treści semantycznej [...], a procesy umysłowe polegają na przechodzeniu od jednej takiej treści do drugiej. Jednak ten obraz rodzi pewne trudności. Treści semantyczne, w które zaopatruje nas ta wizja, nie stosują się same. Nawet jeśli nimi dysponujemy, to wciąż musimy jeszcze

⁵ Znaczna część przekonań i innych rodzajów stanów intencjonalnych Sieci jest nieświadoma mniej więcej w sensie przedświadomości Freuda, to znaczy że nie utrzymujemy ich aktualnie w świadomości, ale w każdej chwili gotowi jesteśmy je podać (Searle 1983, 142).

wiedzieć, co z nimi *począć*, jak je zastosować, a ta wiedza nie może polegać na kolejnych treściach semantycznych, o ile nie chcemy popaść w *regressus ad infinitum*. [...] Reprezentacje faktycznie istnieją, niektóre z nich działają przyczynowo w wytwarzaniu naszych zachowań, lecz ostatecznie po łańcuchu reprezentacji osiągamy podłoże zdolności (Searle 1983, 152n).⁶

Amerykański filozof postuluje prostą geografie Tła, w której rozróżnia biologicznie uwarunkowane i wspólne całemu gatunkowi ludzkiemu „Tło głębokie” oraz nabyte w procesie wzrastania w określonej kulturze „Tło lokalne”. Czyni też rozróżnienie na zdolności Tła dotyczące tego, „jak się rzeczy mają” i dotyczące tego „jak coś zrobić” (Searle 1983, 143n).⁷

Na podbudowie swej semantyki Searle rozwija krytykę tzw. mocnego programu sztucznej inteligencji i związanego z nim klasycznego komputacjonizmu. Nie ma tu potrzeby prezentować znanego eksperymentu myślowego Chińskiego Pokoju, który autor przedstawił po raz pierwszy w artykule z 1980 roku (Searle 1980), a następnie powtórzył w kilku innych pracach. Wystarczy wskazać, że podstawowym celem argumentacji Searle’a jest sprowadzenie do absurdu idei, według której maszyny syntaktyczne – a są nimi wszystkie działające programy pracujące na bazie uniwersalnej maszyny Turinga, czyli wszystkie programy uruchamiane na współczesnych komputerach – mogą posiadać zdolności myślenia i rozumienia. Zdolności te – jak często powtarza filozof – zakładają semantykę i nie dają się zredukować do operacji na składni (Searle 1995, 33).

Searle broni tej tezy na różne sposoby, z których przedstawię tu, moim zdaniem, najsilniejszy. Filozof wyróżnia trzy rodzaje intencjonalności, z których dwa pierwsze były już wspomniane przy okazji omawiania relacji stanów intencjonalnych do języka: jest to intencjonalność wewnętrzna, intencjonalność pochodna i intencjonalność jak-gdyby. Pierwszą z nich możemy przypisywać wyłącznie istotom obdarzonym umysłowością, a więc posiadającym stany mentalne. To właśnie o nią chodzi w zdaniach wyrażających przekonania, pragnienia itp. (np. „Jestem głodny”). Druga przynależy znakom językowym i innym (np. tym, które opisują przekonania), które jednak posiadają ją tylko o tyle, o ile istnieją umysły zdolne je interpretować, a więc uznawać je za znaki z określonym znaczeniem. Trzecia właściwie w ogóle nie jest „prawdziwą” intencjonalnością i dotyczy wypowiedzi metaforycznych w rodzaju: „Te rośliny chcą pić, trzeba je podlać” (np. Searle 1999a, 113–116).

⁶ Searle odwołuje się tutaj do tzw. paradoksu reguł z *Dociekań filozoficznych* Wittgensteina (2000, 118–120, §198–202). Bardziej sugestywną ilustracją problemu samoaplikacji językowej reprezentacji reguł jest żartobliwa historyjka Lewisa Carrolla z 1895 roku zatytułowana *What the Tortoise Said to Achilles* (1995).

⁷ Nieco szersze omówienie koncepcji intencjonalności Searle’a i jej związków z teorią czynności mowy można znaleźć w artykule P. Kołodziejczyka (2002). Semantyka Searle’a nie będzie tu poddawana krytycznej analizie, chciałbym tylko zwrócić uwagę na fakt, że koncepcja Tła jest dość kontrowersyjna. Poważne zarzuty wobec niej wysunął m.in. Barry Stroud (1991).

Który z tych trzech rodzajów intencjonalności można przypisać maszynom syntaktycznym? Searle wskazuje, że trzeci, ponieważ komputer jest systemem, który może najwyżej zachowywać się tak, jakby był intencjonalny (Searle 1999a, 114n), a także drugi, ponieważ wszelkie zautomatyzowane operacje na symbolach mogą coś znaczyć wyłącznie dzięki temu, że to my – istoty obdarzone intencjonalnością wewnętrzną – możemy je interpretować jako znaczące (Searle 1999b, 151). W późniejszych pracach Searle wyostrza jeszcze swoje stanowisko mówiąc, że coś może być maszyną syntaktyczną tylko ze względu na obserwatora zdolnego zidentyfikować coś jako składnię. Tylko dzięki interpretacji pewne fizyczne zjawisko w obrębie stworzonego przez nas artefaktu jest „bitem informacji”. Gdybyśmy chcieli, moglibyśmy uznać za stan czy proces informacyjny każdy inny fakt, nie tylko stany termostatu, ale również na przykład to, że leży przed nami pióro (Searle 1999a, 272–278, Searle 1995, 32). Krótko mówiąc, bycie maszyną syntaktyczną (w sensie Searle’a) nie jest własnością wewnętrzną (Searle 1997, 14–16).

Searle zaznacza wyraźnie, że jego argumentacja nie podważa tezy, że mózg w jakimś sensie jest maszyną – tę tezę autor całkowicie akceptuje, a nawet nie wyklucza, że kiedyś w przyszłości będziemy w stanie zbudować sztuczną maszynę o podobnym działaniu. Możliwe, że nazwiemy ją komputerem. Nie będzie to jednak na pewno komputer cyfrowy w sensie znanych nam wszystkim maszyn syntaktycznych (Searle 1997, 13, 110n; Searle 1995, 32, 36).

INTENCJONALNOŚĆ WEDŁUG DENNETTA

Ujęcie zagadnienia intencjonalności zaproponowane przez Daniela Dennetta wydaje się nieco mniej skomplikowane niż propozycja Searle’a i budowane jest z perspektywy całkowicie jej przeciwstawnej. Dennett lubi bowiem podkreślać, że jego punkt widzenia zakłada starą i wielokrotnie wypróbowaną metodę badań naukowych, które wykorzystują dane gromadzone z perspektywy trzecioosobowej (Dennett 2003, 19). Podobnie jak Searle, nazywa siebie naturalistą, lecz w przeciwieństwie do niego akceptuje potrzebę wyjaśnienia redukcyjnego.

Swoją teorię intencjonalności Dennett rozwijał w ciągu ostatniego półwiecza, począwszy od książki *Content and Consciousness* (Dennett 1969), poprzez artykuł *Intentional Systems* (Dennett 1971⁸) książkę *The Intentional Stance* (Dennett 1987), skończywszy na *Consciousness Explained* (Dennett 1991) i *Darwin’s Dangerous Idea* (Dennett 1995) a także poświęconym nieco innemu tematowi *Odczarowaniu* (Dennett 2008a). Już w pierwszej z wymienionych książek pojawia się koncepcja intencjonalności oparta na

⁸ Artykuł ten został przedrukowany w: (Dennett 1981, 3–22). Powołując się dalej na *Intentional Systems*, odnoszę się do tego źródła.

teorii ewolucji oraz cybernetyce. Wątkiem przewodnim myśli Dennetta cały czas pozostaje teza o bliskiej analogii procesu rozwoju życia z jednej strony i rozwoju technologii, w tym przede wszystkim robotyki i informatyki z drugiej.⁹

Dennett wiązał początkowo swoją koncepcję intencjonalności z krytyką programu behawioryzmu. Behawioryzm był próbą wyeliminowania z psychologii odwołań do stanów wewnętrznych i wyjaśnienia zachowania wyłącznie w kategoriach bodźca i reakcji. Kierowało nim przekonanie, że pojęcia związane ze stanami wewnętrznymi generują poważne zobowiązania teoretyczne – a mianowicie założenie ich inteligentnego interpretatora – o których nie wiadomo, czy kiedykolwiek będą „spłacone” (Dennett 1981, 12). Filozof wskazywał jednak, że psychologia behawiorystyczna albo rezygnuje z idei celowości i wewnętrznego motywowania, ale wówczas nie jest w stanie zbudować satysfakcjonującej teorii zachowania (Dennett 1969, 63), albo faktycznie zakłada intencjonalność, tyle że stara się to ukryć (Dennett 1981, 14n). W tej sytuacji Dennett zaproponował podejście, w którym zaciąganie zobowiązań intencjonalnych jest dopuszczalne, o ile będziemy starali się je „spłacić” dzięki teoriom wyjaśniającym wewnętrzne uwarunkowania zachowań w kategoriach mechanistycznych regularności. Dobrym przykładem takiego podejścia jest, według niego, program rozwoju sztucznej inteligencji i program badań neurofizjologicznych (Dennett 1981, 15n). Innymi słowy, wbrew obawom behawiorystów, nie mamy w psychologii problemu z intencjonalnością, ponieważ istnieją programy badawcze obiecujące jej redukcję do prostszych, lepiej znanych struktur i prawidłowości. Co więcej, istnieją konteksty teoretyczne – w rodzaju teorii gier czy teorii ekonomicznych wyjaśniających funkcjonowanie zbiorowości jednostek ludzkich – w których taka redukcja nie jest nawet konieczna dla predykcyjnej efektywności (Dennett 1981, 16).

Jednocześnie Dennett, zwłaszcza w późniejszym okresie, podkreślał, że redukowalność stanów intencjonalnych nie znaczy, że są one nierealne. Swoje stanowisko określał mianem łagodnego realizmu (*mild realism*) i sytuował je pomiędzy mocniejszymi stanowiskami realistycznymi Jerry’ego Fodora i Donalda Davidsona a jeszcze słabszym realizmem Rorty’ego i eliminatywizmem Paula Churchlanda (Dennett 2008b, 303). Dla wielu krytyków filozofa – w tym również dla Searle’a – „łagodność” owego realizmu jest posunięta zbyt daleko.

Autor *The Intentional Stance* zaproponował teorię, wedle której intencjonalność jest redukowalna (co zapewnia jej status legalnego składnika nauki), a jednocześnie eksplanacyjnie autonomiczna i użyteczna. W teorii tej mamy do czynienia z trzema zasadniczymi nastawieniami (*stances*) poznaw-

⁹ Na marginesie można zaznaczyć, że bardzo podobne wątki znajdujemy w twórczości filozoficznej Stanisława Lema (np. Lem 2010).

czymi.¹⁰ Ze względu na skomplikowanie struktur, z jakimi ma do czynienia najprostsze jest nastawienie fizyczne (*physical stance*). Wyjaśnienia i predykcje formułowane w obrębie tego nastawienia dotyczą przedmiotów bądź układów poddanych prostym fizycznym oddziaływaniom, które zmieniają ich proste fizyczne własności. Przykładowo, możemy przewidzieć, co stanie się z garnkiem pełnym wody, kiedy pozostawimy go na gorącej płycie kuchenki: woda zagotuje się, po pewnym czasie całkowicie wyparuje a garnek się spali. Wiemy też, dlaczego Księżyc krąży wokół Ziemi po orbicie o takich a nie innych parametrach, bowiem znamy masy tych obiektów kosmicznych i prawa rządzące oddziaływaniami grawitacyjnymi. W środowisku, które nas otacza, znajdujemy bardzo wiele zjawisk dających się wyjaśnić w obrębie nastawienia fizycznego, w których chodzi o wzajemne relacje między parametrami, takimi jak masa, rozmiar, odległość, gęstość, ciepło, sprężystość itp.

Nasze otoczenie składa się jednak również z takich układów, wobec których wyjaśnienie fizyczne okazuje się niewystarczające, a mówiąc dokładniej jest ono zbyt skomplikowane, by mogło być użyteczne. Przykładem takiego układu może być kalkulator. Możemy oczywiście wyjaśniać działanie kalkulatora odwołując się wyłącznie do oddziaływań fizycznych między poszczególnymi jego częściami. Opis, jaki wówczas byśmy uzyskali, byłby jednak niezmiernie skomplikowany. Wobec tego znacznie praktyczniej jest przyjąć w odniesieniu do kalkulatora nastawienie projektowe (*design stance*), wedle którego rozpatrujemy już nie proste fizyczne cechy przedmiotu, lecz jego funkcjonalną strukturę i powiązania między oddziałującymi na siebie częściami mechanizmu. Dennett zauważa, że nastawienie projektowe cechuje zarówno nasze podejście do wytworów techniki, jak i do znacznej części wytworów ewolucji biologicznej: wiemy, że drzewo składa się z korzeni, które rozrastają się w ziemi, aby dostarczać organizmowi wodę i substancje odżywcze, z liści, w których dzięki energii słonecznej zachodzą procesy metaboliczne i z pnia, który transportuje wodę do gałęzi i liści, a także pozwala utrzymywać roślinie liście w miejscu lepiej nasłonecznionym. Wiemy, że prosty fototaksyczny organizm jednokomórkowy posiada receptory światła i przyczynowo powiązany z nimi układ efektorów ruchu. Fizykochemiczne wyjaśnienie funkcjonowania komórki czy układu wielokomórkowego również jest możliwe, lecz w wielu sytuacjach niepraktyczne i niestosowane do formułowania przewidywań zachowań tego rodzaju obiektów.¹¹

Jednak nie wszystkie organizmy sensownie jest traktować w kategoriach mechanizmów. Potoczny opis zachowania już tak drobnego zwierzęcia, ja-

¹⁰ Wyjaśniona poniżej idea nastawień przedstawiana była przez autora w wielu miejscach (Dennett 1981, 4–9; Dennett 1987, 15–19; Dennett 1995, 229–238).

¹¹ Dennett zauważa, że istnieją oczywiście sytuacje, w których organizm bądź mechanizm zostaje poddany prostym oddziaływaniom fizycznym i wówczas nastawienie fizyczne jest bardziej uzasadnione (Dennett 1981, 5).

kim jest pszczoła, zawiera elementy intencjonalne: mówimy na przykład, że pszczoła wie, jak trafić do swego ula, że jej działanie jest celowe itp. Im większe zwierzę, im bardziej jego działanie jest skomplikowane, tym chętniej korzystamy przy jego opisie z intencjonalnego słownika mówiąc o wiedzy, oczekiwaniu, chęci, czy radości. Ten trzeci sposób opisu i wyjaśnienia nazywany jest przez Dennetta nastawieniem intencjonalnym (*intentional stance*).

Nastawienie intencjonalne, czyli przypisywanie pewnym systemom posiadania informacji i celów, jest zasadniczo naszym domyślnym nastawieniem nie tylko wobec osób czy większych zwierząt, ale właściwie wobec wszystkiego, co się porusza. Mówimy np., że rośliny *nie chcą* kwitnąć w pierwszych cieplejszych dniach wiosny, bo są *ostrożne* (Dennett 1987, 22). Jest to po prostu użyteczna strategia poznawcza pozwalająca uzyskać przewidywania w przybliżeniu wystarczającym dla potrzeb naszych mniej cywilizowanych przodków. Co prawda dzisiaj, dzięki rozwojowi badań naukowych jesteśmy w stanie opisać większość zjawisk w kategoriach nastawienia fizycznego bądź projektowego, to jednak owa skłonność do opisu w kategoriach intencjonalnych wciąż pojawia się w na poły żartobliwych określeniach typu: „woda szuka swojego poziomu”, czy „piorun szuka najlepszej drogi ku ziemi” (Dennett, 2008, 151n).

Podobnie jak w przypadku nastawienia projektowego, fakt, że dany system wygodnie jest nam ujmować w kategoriach intencjonalnych (a więc, pragnień, przekonań, celów itp.), nie znaczy, że tego samego systemu nie da się opisać w kategoriach projektowych czy fizycznych. W ujęciu Dennetta stany mentalne nie są ani samodzielne substancjalnie (jak głosił klasyczny dualizm, np. kartezjański), ani nawet nie stanowią samodzielnej ontologicznie kategorii własności (jak głosi dualizm własności).

Czy zatem kwestia uznania czegoś za system intencjonalny jest arbitralna? Autor odpowiada na to pytanie przecząco. Jego zdaniem istnieje coś takiego, jak obiektywny sens przypisywania intencjonalności, a w konsekwencji istnieje różnica między rzeczywistym systemem intencjonalnym, a systemem, który tylko wydaje się niekiedy intencjonalny, nawet jeśli nie ma żadnej dającej się wskazać istotowej cechy intencjonalności. W swojej pierwszej książce Dennett przedstawia analogię z głosem argumentując, że on także jest czymś, czego ostatecznie nie możemy wskazać (Dennett 1969, 8–14). Brak odniesienia nie znaczy jednak, że fakt, iż ktoś coś powiedział, zależy całkowicie od naszej interpretacji. Różnica pomiędzy systemami potrafiącymi mówić a tymi, które mówić nie potrafią, jest obiektywna, a jej doświadczalnym fundamentem jest obserwowany efekt głosowego przekazywania informacji. Podobnie jest z intencjonalnością. Tu także oparciem dla obiektywności jest obserwowalna efektywność przewidywań. Jak mówi Dennett: „Decyzja, by przyjąć nastawienie intencjonalne jest wolna, ale fakty dotyczące tego, czy przyjęte przez kogoś nastawienie odniosło sukces bądź

zawiodło, są całkowicie obiektywne” (Dennett 1987, 24). Można to wyjaśnić na jego ulubionym przykładzie szachów. Gdy rozgrywamy partię, musimy przewidywać ruchy przeciwnika. W zasadzie moglibyśmy za podstawę naszych przewidywań przyjąć nastawienie projektowe (a nawet fizyczne), ponieważ nasz przeciwnik jest albo organizmem biologicznym, albo maszyną, ale byłoby ono kompletnie nieefektywne. W praktyce musimy oprzeć nasze analizy na założeniu, że oponent *zna* reguły, *wie*, jaka jest sytuacja na planszy, *potrafi przewidzieć*, co z niej wyniknie i *kieruje się celem*, jakim jest wygranie partii. Bez względu na to, czy oponentem tym jest człowiek czy komputer, musimy przyznać, że strategia intencjonalna jest praktycznie jedyną sensowną strategią, a przynajmniej jest strategią najbardziej efektywną, ponieważ pozwala przewidzieć posunięcia przeciwnika z wystarczająco dobrym przybliżeniem (Dennett 1987, 24). Według łagodnego realizmu należy w takich przypadkach stwierdzić, że obserwowana intencjonalność gracza jest rzeczywista.

Dennett proponuje więc model funkcjonowania pojęć intencjonalnych takich jak znaczenie, informacja, postrzeganie, przekonanie, działanie celowe, dzięki któremu możemy uniknąć zarówno relatywizmu pojęciowego, jak i uwikłania w kłopotliwe ontologiczne zobowiązanie. Istnieją obiektywne kryteria wyróżniania systemów intencjonalnych, a jednocześnie nie istnieje pierwotna własność intencjonalności, której nie dałoby się wyjaśnić poprzez odwołanie do prostszych mechanizmów (poziom projektowy), praw i procesów (poziom fizyczny). Wobec tego, z jednej strony Dennett może wyjaśniać intencjonalność w oparciu o teorię ewolucji, a z drugiej przewiduje (a właściwie już opisuje¹²) powstanie sztucznych systemów intencjonalnych jako efektu ewolucyjnego postępu technologii.

Intencjonalność, jak twierdzi autor *The Intentional Stance*, pojawiła się na świecie w podobny sposób, w jaki pojawiło się na nim życie: nie da się wskazać momentu, w którym jakaś bezrozumna struktura zyskała nagle możliwość żywienia przekonani i dążenia do celów. Był to raczej proces, w którym pierwotnie „bezmąsne” systemy dynamiczne stawały się stopniowo coraz bardziej inteligentne, a to znaczy: przejawiały coraz bardziej złożone zachowania. Początkowo systemy intencjonalne były raczej *quasi-inteligentne* i ich intencjonalność może być nazwana raczej *proto-intencjonalnością* – nie wykazywały wielkiego podobieństwa do organizmów zdolnych do podejmowania decyzji na bazie wiedzy, takich jak choćby pszczoły czy szczury. Jednak dokładnie tak samo było w przypadku życia –

¹² Dennett od dawna uważał, że odpowiednio oprogramowane komputery są systemami intencjonalnymi, ponieważ mogą spełniać wskazane przed chwilą kryterium – opis ich funkcjonowania w kategoriach intencjonalnych jest predykcyjnie bardziej efektywny niż opis w kategoriach projektowych. Można powiedzieć – zauważał on – że takie systemy są relatywnie prymitywne i bardzo różnią się od wysoce skomplikowanych i rozbudowanych systemów intencjonalnych, jakimi są wyższe zwierzęta (nie mówiąc już o ludziach), ale nie zmienia to faktu, że należą do tej samej klasy posiadaczy przekonani i istot działających celowo (Dennett 1987, 29).

jego pierwociny prawdopodobnie nie spełniały wielu definicyjnych własności życia, jakie znamy (Dennett 1995, 205).

Wedle wizji Dennetta, rozwijanej od 1975 roku (Dennett 1975¹³) rozwój życia obdarzonego umysłowością posiadał kilka istotnych etapów różniących się sposobem wcielania ogólnego prawa selekcji naturalnej. Na pierwszym etapie nie było żadnej intencjonalności: kolejne pokolenia prostych organizmów dziedziczyły w pełni swój repertuar zachowań. Dzięki mutacjom kodu genetycznego repertuar ten różnicował się jednak w obrębie populacji, a prawo selekcji premiowało osobniki lepiej przystosowane, wobec czego udział ich potomków w następnych pokoleniach wzrastał. Ten czysto darwinowski model akumulacji mechanizmów przystosowawczych miał jedną podstawową wadę – „przeprogramowanie” zachowania organizmów wymagało bardzo długiego czasu w stosunku do szybkości możliwych zmian środowiska, wobec czego „istoty darwinowskie” (określenie Dennetta) mogły rozwijać przystosowania wyłącznie w środowiskach stabilnych, co było oczywiście silnym ograniczeniem.

W toku ewolucji rozwinął się jednak pewien podtyp organizmów darwinowskich, który wyróżniała możliwość osobniczego dostosowania pewnej klasy zachowań do środowiska. Dostosowaniem zachowań pojedynczego organizmu tej klasy rządziło uchwycone przez behawiorystów prawo efektu mówiące, że działania nagradzane są powtarzane – jest to behawiorystyczny mechanizm uczenia, a więc przystosowania niewymagającego zmian w kodzie genetycznym. Dennett nazywa więc ten podtyp organizmów „istotami skinnerowskimi” i zauważa, że ową behawiorystyczną zasadę można potraktować jako transpozycję prawa selekcji naturalnej (Dennett 1981, 73).

Behawioryści nie dostrzegli jednak, że – analogicznie do prawa selekcji – także prawo efektu ulega transpozycji dla pewnej podklasy organizmów. Podklasa ta ma ciekawą własność interioryzacji ograniczeń środowiskowych, dzięki czemu adaptacja zachowań nie musi polegać na realnym testowaniu par bodziec-reakcja (które zawsze stanowi pewne zagrożenie dla życia organizmu), tylko na konfrontowaniu zinterioryzowanych par ze zinterioryzowanymi ograniczeniami. Oczywiście, ostatecznym selektorem zachowań jest środowisko, lecz jest ono w stanie niejako „delegować” swoją funkcję selekcyjną do czegoś we wnętrzu organizmu, podobnie jak na wcześniejszym etapie śmierć „wydelegowała” swoją funkcję selekcyjną na ból (Dennett 1981, 78). Zinterioryzowane ograniczenia środowiskowe mogą być wrodzone i dziedziczone, lecz jest to oczywiście dość nieefektywny wehikuł adaptacyjny, jeśli założymy, że środowisko podlega relatywnie szybkim przemianom. Lepiej, jeśli ograniczenia są przynajmniej częściowo kształtowane w trakcie życia pojedynczych organizmów i ich kontaktów ze światem. Na tym etapie

¹³ Artykuł został przedrukowany w: (Dennett 1981, 71–89). Powołując się na niego odnoszę się do tego źródła.

uczenie zyskuje sens inny niż behawiorystyczny; polega mianowicie na interioryzacji pozytywnych i negatywnych wzmocnień.

Zauważmy, że o ile w przypadku istot Skinnerowskich trudno było mówić o intencjonalności – była to raczej proto- czy *quasi*-intencjonalność – to w przypadku istot wyższego poziomu trzeba już dokonywać pewnego wysiłku, żeby słownika intencjonalnego nie używać. Czym są bowiem zinterioryzowane kopie ograniczeń środowiskowych, jeśli nie teoriami dotyczącymi świata? Właśnie z tego powodu Dennett nazywa podklasę organizmów zdolnych wytwarzać reprezentacje świata „istotami Popperowskimi” – nawiązując do sławnego powiedzenia Karla Poppera, że „nasze hipotezy umierają za nas”.

My, ludzie, należymy oczywiście do istot Popperowskich, lecz coś ważnego różni naszą umysłowość od umysłowości większości innych zwierząt. Aby działać i zdobywać informacje, posługujemy się narzędziami, to znaczy zaprojektowanymi przez nas częściami środowiska. Należymy więc (wraz z niektórymi gatunkami innych ssaków) do najbardziej zaawansowanej intelektualnie podklasy organizmów, które Dennett nazywa „istotami gregoriańskimi” – od nazwiska psychologa Richarda Gregory’ego, twórcy pojęcia inteligencji kinetycznej (Dennett 1995, 377). Prawdopodobnie tylko ludzie posiadają najbardziej wyrafinowane narzędzie, jakim jest język.¹⁴

Powyżej opisana wielostopniowa struktura została przez autora nazwana więżą generowania i testowania. Proces generowania i testowania organizmów, par bodziec-reakcja, a wreszcie teorii, stanowi dla Dennetta uniwersalny schemat ewolucji systemów żywych, a w szczególności systemów intencjonalnych.¹⁵ Pojęcie generowania i testowania pochodzi z informatyki i teorii sztucznej inteligencji i dotyczy jednej ze strategii algorytmicznego rozwiązywania problemów (Dennett 1981, 81). Jest to kolejne odwołanie do ogólnej tezy tego filozofa, że rozwój urządzeń przetwarzających informacje jest paralełą ewolucji biologicznej. Proste organizmy jednokomórkowe są z jego punktu widzenia po prostu naturalnymi samoreplikującymi się robotami a my jesteśmy ich bezpośrednimi potomkami (Dennett 1996, 22). W istocie, jak Dennett mówił w jednym ze swych internetowych wykładów, „jesteśmy robotami zrobionymi z robotów zrobionych z robotów zrobionych z robotów” (Dennett 2009).

Zarówno nastawienie projektowe jak i intencjonalne opisuje zatem roboty – naturalne bądź wytworzone przez człowieka – a więc pewne fizyczne systemy zdolne do samodzielnego działania. W przypadku części tych systemów możemy mówić o tym, że mają one cele i przekonania, a więc że posia-

¹⁴ Język jest narzędziem dość specyficznym, ponieważ nie jest właściwie elementem środowiska, lecz należy do dziedziny jego wewnętrznej reprezentacji. Można więc powiedzieć, że istoty gregoriańskie posługujące się językiem należą do swojej własnej, najwyższej podklasy organizmów inteligentnych. Dennett nie dokonuje jednak tego ostatniego rozróżnienia.

¹⁵ Do analogicznych wniosków doszedł np. Robert R. Traill (2008).

dają wewnętrzne reprezentacje rzeczywistości. To właśnie założenie, że system posiada takie reprezentacje jest istotą nastawienia intencjonalnego. Dennett zastrzega jednak, że

nie jest tak, że przypisujemy (bądź powinniśmy przypisywać) przekonania i pragnienia wyłącznie tym przedmiotom, w których odnajdujemy wewnętrzne reprezentacje, lecz raczej tak, że gdy odkrywamy przedmiot, w przypadku którego strategia intencjonalna działa, skłonni jesteśmy interpretować pewne jego stany bądź procesy wewnętrzne jako wewnętrzne reprezentacje. Ty, co czyni jakąś wewnętrzną własność przedmiotu reprezentacją, może być tylko jej rola w regulacji zachowań systemu intencjonalnego. (Dennett 1987, 32)

Autor *The Intentional Stance* twierdzi więc, że podobnie jak nie istnieje żadna tajemnicza istota systemu intencjonalnego, tak też nie istnieje nic, co w definitywny sposób wyróżniałoby wewnętrzną reprezentację, a więc treść intencjonalną. Uzasadnieniem praktyki uznawania czegoś za taką treść jest taki charakter działania całego systemu, który pozwala na efektywne stosowanie wobec niego nastawienia intencjonalnego. I nie ma żadnego znaczenia, czy system jest naturalny, czy też sztuczny.

ISTOTA SPORU SEARLE'A Z DENNETTEM

Jasne jest, że zapatrywania Searle'a i Dennetta na intencjonalność nie dają się ze sobą uzgodnić. Nie są one jednak całkowicie przeciwstawne. Pomimo tego, że obaj autorzy odnoszą się do siebie dość nieprzyjaźnie i nazywają wzajemnie swoje teorie „formami intelektualnej patologii” (Searle 1997, 112, 115), można w nich dostrzec zaskakująco wiele punktów wspólnych bądź przynajmniej przestrzeni możliwego konsensu. Zarówno Dennett jak i Searle deklarują się jako naturaliści, co samo w sobie niewiele jeszcze znaczy, ponieważ termin „naturalizm” jest obecnie stosowany w różnych znaczeniach terminu, czego najlepszym przykładem jest naturalizm drugiej natury Johna McDowella (McDowell 1996, 84). Dopowiedzieć jednak można, że dla obu filozofów naturalizm wiąże się z deklaratywnym odrzuceniem dualizmu. Co więcej, Dennett uważa za oczywiste, że można stworzyć sztuczny system intencjonalny, ale również Searle nie ma ostatecznie nic przeciw temu – różnica zdań dotyczy tylko szczegółów budowy takiego systemu.¹⁶

Na pozór pole konfliktu otwiera się wraz z Searle'owskim holizmem związanym z koncepcją Sieci (pojedynczy stan intencjonalny nie może istnieć bez powiązań z innymi stanami). Dennett nigdzie *explicite* nie formułuje wobec treści intencjonalnej wymogu bycia częścią systemu. Przeciwnie, podaje

¹⁶ Nawet i tu różnica zdań nie jest aż tak wielka. Dennett w jednym ze swoich tekstów broni tezy, że system oparty o znaną nam technologię komputera cyfrowego prawdopodobnie nigdy nie osiągnie inteligencji organicznego ludzkiego mózgu (Dennett 1987, 324–330).

przykład termostatu jako „zdegenerowanego przypadku systemu intencjonalnego”. Systemowi temu, jak twierdzi (choć z wyraźnym przymrużeniem oka), można w zasadzie przypisać „pół tuzina różnych przekonań i nieco mniej pragnień”, np. że bojler jest wyłączony, że dobrze byłoby, gdyby został włączony itp. (Dennett 1987, 29n). Searle oczywiście nie może zaakceptować termostatu jako systemu obdarzonego umysłowością, ale dla Dennetta również zasadniczo nie jest to prawdziwy system intencjonalny, ponieważ jest on zbyt prosty, by zamiana nastawienia projektowego na intencjonalne przyniosła zwiększenie efektywności predykcji. Rzeczywisty system intencjonalny, w myśl łagodnego realizmu Dennetta, to taki, w przypadku którego przyjęcie nastawienia intencjonalnego jest nieodzowne. Można się spodziewać, że treści intencjonalne takiego systemu muszą tworzyć całkiem rozległą sieć.

Jest również możliwe odnalezienie u Dennetta odpowiednika Searle'owskiej koncepcji Tła. Jeśli Tłem są nieintencjonalne dyspozycje organizmu do działania, to właściwie każdy mechanizm (biologiczny, czy techniczny) posiada takowe. Nawet wspomniany termostat, ponieważ bez odpowiednich efektorów i bez oddziaływania systemu ze środowiskiem nie sposób byłoby mu – w myśl koncepcji Dennetta – przypisać stanów intencjonalnych.

W kontekście problematyki związku systemu ze swym środowiskiem Dennett stara się wykazać, że idea „maszyny syntaktycznej” pełni w argumentacji Searle'a funkcję „słomianej kukły”. O ile bowiem zarówno formalny opis funkcjonowania maszyny cyfrowej, czy dowolny napisany na nią program, można uznać za „czysto syntaktyczny”, o tyle już implementacja tego programu – realne działanie komputera, który czymkolwiek steruje, choćby podłączonym do niego monitorem – ma „wejścia” i „wyjścia”, jest więc dwukierunkowo związany z rzeczywistością. Każdy program komputerowy i działanie każdej maszyny można opisywać abstrahując od tego związku – uzyskamy wówczas pewną formalną strukturę, ale działająca maszyna i działający program zazwyczaj nie jest w ścisłym sensie „maszyną syntaktyczną”.¹⁷

Kiedy jednak Searle mówi o maszynach syntaktycznych, chodzi mu właśnie o działające komputery cyfrowe sterowane programem. Uznaje je za „syntaktyczne” nie z tego powodu, że nie mają związków z rzeczywistością, ale dlatego, że same z siebie nie posiadają „semantyki”. Ma na myśli to, że intencjonalność operacji, które komputery wykonują – nawet jeśli operacje te mają jakiś wpływ na rzeczywistość – jest tylko pochodna. Notabene w późniejszym wzmocnieniu swojej argumentacji dodaje, że pochodna jest również struktura syntaktyczna takich maszyn. I tak dochodzimy do istoty jego sporu z Dennettem, to znaczy do Searle'owskiej idei trzech rodzajów intencjonalności: wewnętrznej, pochodnej i jak-gdyby.

¹⁷ W podobny sposób atakuje stanowisko Searle'a również m.in. William Rapaport (2000, 474).

Dennett wcale nie twierdzi, że intencjonalność, jaką przypisuje maszynom, jest wewnętrzna. Tu także zgadza się ze swoim adwersarzem: intencjonalność komputerów jest – zależnie od sposobu ujęcia – intencjonalnością pochodną, albo nawet intencjonalnością jak-gdyby (co nie zmienia faktu, że jest ona w jego ujęciu rzeczywista). Nie ma czegoś takiego, jak ostateczna czy pierwotna funkcja tych urządzeń, nic nie jest komputerem mocą swej „istoty” – wszystko zależy od sposobu interpretacji działania i budowy. Punktem niezgody jest istnienie intencjonalności wewnętrznej; Dennett argumentuje, że po prostu czegoś takiego nie ma. Znów działa tu analogia rozwoju technologii i ewolucji biologicznej: podobnie jak nie ma żadnego wewnętrznego powodu, by jakaś funkcja pewnego fizycznego układu cyfrowego była dodawaniem, tak nie ma żadnego wewnętrznego powodu, by określonego typu działanie żywej komórki było syntezą białek w oparciu o kod genetyczny – to my, obserwatorzy, nadajemy taką interpretację obserwowanemu zjawisku. Mało tego, tak jak nie ma wewnętrznego powodu, by coś było nazwane funkcją bądź cyfrowym układem fizycznym, tak też nie ma wewnętrznego powodu, by nazywać coś komórką albo wyodrębnić typy jej działań. Analogia z twórcami techniki narzuca nam myślenie o intencji twórcy: coś jest syntezą białka, ponieważ tak zdecydował projektant, czyli Matka Natura. Lecz oczywiście Matka Natura nie jest żadną osobą, nie podejmuje decyzji, niczego nie projektuje; to tylko metaforyczne ujęcie zespołu praw przyrody regulujących proces ewolucji biologicznej. Wracamy więc do obserwatorów – ostatecznym źródłem wszelkiej klasyfikacji, wszelkiego wyodrębniania funkcji i systemów, jest nasza intencjonalna decyzja (co jeszcze nie świadczy przeciw temu, że owe klasy i funkcje są rzeczywiste).

Problem w tym, że także my jesteśmy twórcami ewolucji biologicznej, a zatem również nasze działania zawierają się w zbiorze zjawisk, których klasyfikacja wymaga zewnętrznych kryteriów. Okazuje się więc, że nasza sytuacja nie różni się zasadniczo od sytuacji maszyny cyfrowej – ludzka intencjonalność też jest intencjonalnością tylko w sensie pochodnym (jeśli rozpatrujemy ją ze względu na pierwotną intencjonalności Matki Natury) lub jak-gdyby intencjonalnością (jeśli przyjmujemy ostatecznie, że Matka Natura to tylko metafora).

Łagodny realizm Dennetta pozwala mu w tym punkcie utrzymywać, że nie jest to żaden problem: co prawda intencjonalność obserwatora nie jest wewnętrzna, ale nie znaczy to, że nie jest rzeczywista. Krytycy wskazują, że ta nieproblematyczność osiągnięta jest dzięki pewnej sztuczce – Dennett zręcznie balansuje między realistyczną a relatywistyczną wykładnią racjonalności, wykorzystując raz jedną, raz drugą perspektywę w zależności od potrzeb. Zręczność ta jest jednak tylko retoryczna, a pozycja Dennetta próbującego pogodzić dwie przeciwstawne tendencje jest na dłuższą metę niestabilna (np. Heitner 2000).

Uwzględniając tego rodzaju krytykę można spytać, w jaki sposób Dennett zamierza „spłacić” zobowiązanie zaciągnięte w założeniu rzeczywistego charakteru intencjonalności obserwatora. Odpowiedź jest ewolucjonistyczna: zdolność przyjmowania przez nas nastawienia intencjonalnego jest pewną własnością przystosowawczą. To jednak znaczy, że również ona nie jest zdeterminowana wewnętrznie. Rozpoznajemy siebie jako istoty posługujące się nastawieniem intencjonalnym, ponieważ stosujemy nastawienie intencjonalne wobec nas samych. Czyli, mówiąc językiem Searle'a: *jak-gdyby intencjonalne* nastawienie intencjonalne jest uznawane za rzeczywiste nastawienie intencjonalne w świetle *jak-gdyby intencjonalnego* nastawienia intencjonalnego.

Przedstawiona konstrukcja – o ile nie jest po prostu paradoksalna – jest silnie nieintuicyjna. Zaryzykowałbym tezę, że teoria Dennetta jest najbardziej nieintuicyjną ze wszystkich sensownych teorii, jakie jesteśmy sobie w stanie wyobrazić, oczywiście o ile w ogóle uznamy ją za sensowną. Searle tego nie robi, uważając, że Dennett ostatecznie wikła się w paradoks. Możemy mówić – argumentuje – że padamy ofiarą iluzji co do tego czy owego faktu, ale nie możemy powiedzieć, że padamy ofiarą iluzji co do faktu, że padamy ofiarą iluzji (Searle 1997, 111n).¹⁸

Łatwo zauważyć, że korzenie tego argumentu Searle'a sięgają Kartezjusza. Koncepcja intencjonalności wewnętrznej związana jest ściśle z ideą asymetrii epistemicznej: intencjonalność taka jest stwierdzalna wyłącznie w doświadczeniu subiektywnym, co Searle przyznaje kreśląc wizję nauki obejmującej zarówno sferę obiektywną jak i subiektywną (Searle 1997, 114). Takie założenie prowadzi jednak do niepożądanego solipsystycznego wniosku: jestem w stanie odnaleźć wewnętrzną intencjonalność wyłącznie w sobie samym. Wniosek ten Searle próbuje co prawda osłabiać, dowodząc, że skoro wiemy, iż nasze mózgi posiadają wystarczającą moc przyczynową, by wytworzyć intencjonalność, a jednocześnie wiemy, że inni ludzie mają mózgi, to wiemy tym samym, że również inni ludzie posiadają wewnętrzną intencjonalność (Searle 1999a, 41–44, 107–110). Dennett wytyka mu jednak, że pierwsza przesłanka tego wniosku – nasza rzekoma wiedza o przyczynowym związku mózgu i intencjonalności – nie może być w żaden sposób uzasadniona empirycznie (Dennett 1987, 334).

Ostatecznie koncepcja Searle'a okazuje się bliska podstawowym założeniom fenomenologii Husserla: sens językowy zależy od intencjonalności dostępnej w prywatnym doświadczeniu. Także konstrukcja argumentu przeciw naturalizacji intencjonalności proponowanej przez Dennetta jest bardzo podobna do konstrukcji argumentu Husserla przeciw naturalizmowi (np.

¹⁸ Wielu innych autorów również odrzuca stanowisko Dennetta jako nie do przyjęcia (np. Beiseker 2002).

Husserl 1992, 23).¹⁹ Argument ten wydaje się dziś co najmniej poważnie osłabiony przede wszystkim analizami Wittgensteina, który starał się wykazać, że doświadczenia pojedynczego podmiotu nie mają mocy normatywnej, wobec czego źródłem tej ostatniej może być tylko wspólnota (Wittgenstein 2000, 134–136, §§ 258–265).²⁰ Z analiz tych korzysta Dennett, którego teoria opiera się na ich naturalistycznej i redukcjonistycznej interpretacji.

Searle może oczywiście – podobnie jak wielu innych – nadal uznawać moc argumentu naturalistycznego. Problem w tym, że filozof ten chce się jednocześnie uważać za biologicznego naturalistę. Jak wskazywałem wcześniej, termin „naturalizm” ma dziś tak wiele znaczeń, że któreś z nich mogłoby być może objąć także stanowisko Husserla, jednak sam Searle dookreśla swoje rozumienie naturalizmu mówiąc, że intencjonalność ma przyczyny neurofizjologiczne. Stwierdza, że mózg ludzki jest pewnego rodzaju maszyną i nie ma zasadniczo nic przeciw tezie, że dzięki postępowi nauki będziemy kiedyś w stanie w pełni zrozumieć działanie tej maszyny i wytworzyć artefakt generujący intencjonalność.

Naturalizm Searle’a jest więc przekonaniem, że intencjonalność znajdzie kiedyś pełne wyjaśnienie w ramach uprawianej przez nas nauki. I tu właśnie, jak sądzę, tkwi problem. Uważam bowiem, że David Chalmers wyraża pewną fundamentalną prawdę, gdy pisze, że nauka empiryczna na dowolnym etapie zaawansowania nie zaoferuje nam wyjaśnienia innego niż w kategoriach struktur i dynamiki procesów (Chalmers 2010, 209n), a skoro intencjonalność wewnętrzna Searle’a ma być czymś prostym, nieredukowalnym i danym w doświadczeniu subiektywnym (Searle 1999b, 96n; Searle 1983, 26), to żadna wyobrazalna nauka wychodząca od danych obiektywnych („trzecioosobowych”) nie wytworzy konstrukcji, która mogłaby być z ową intencjonalnością utożsamiona. Nauka może jedynie wytworzyć konstrukcje, które będą dawały się utożsamiać z intencjonalnością jak-gdyby.²¹

Zasadniczo cała „obudowa teoretyczna” Searle’owskiej idei intencjonalności, w której mowa o pewnych dynamicznych procesach i strukturach, a więc koncepcja relacji reprezentacji, Sieci i Tła, pozwala się dość dobrze dopasować do wizji Dennetta. Tym, co nie daje się do niej dopasować, jest

¹⁹ Nawiasem mówiąc, pogłębione analizy założeń fenomenologii mogą prowadzić do zakwestionowania dość popularnych we współczesnych naukach kognitywnych prób jej naturalizacji, na co wskazywała np. Barbara Trybulec (2015).

²⁰ Co ciekawe, Searle częściowo uznaje te analizy (zob. przypis 6), ale tam, gdzie ich konsekwencje godzą w jego stanowisko, stara się ignorować. Przykładem jest traktowanie §580 *Dociekań filozoficznych*, w którym czytamy, że „Zjawisko wewnętrzne» wymaga zewnętrznych kryteriów” (Wittgenstein 2000, 215) – Searle określa go mianem „zagadkowego” (Searle 1999a, 27). Nawiasem mówiąc, związek antynaturalistycznej argumentacji Husserla z obserwacjami Wittgensteina dotyczącymi problemu reguł oraz tzw. języka prywatnego jest ciekawym zagadnieniem badawczym wartym co najmniej osobnego artykułu.

²¹ Chalmers ma więc rację twierdząc, że koncepcja Searle’a jest rodzajem dualizmu własności (Chalmers 2010, 227). Ten ostatni, co prawda, deklaratywnie odrzuca dualizm własności jako jeszcze jedną wariację na temat kartezyjanizmu (Searle 1999a, 16n; Searle 1997, 135–163), lecz w *The Mystery of Consciousness* sam pisze wprost o ontologicznej autonomii subiektywności (Searle 1997, 98).

krypto-kartezjańska teza o oczywistym subiektywnym dostępie do wewnętrznej intencjonalności. Za odrzuceniem kartezjanizmu w filozofii umysłu przemawia bardzo wiele poważnych argumentów – można je przeczytać między innymi u Dennetta, ale można je również znaleźć w dziełach Searle'a, co znaczy, że jego propozycja jest po prostu niespójna.

ZAKOŃCZENIE

W artykule starałem się przybliżyć sposób, w jaki zagadnienie intencjonalności treści mentalnych rozpatrywane jest we współczesnej filozofii umysłu. Zaprezentowałem zestawienie dwóch przeciwstawnych stanowisk sformułowanych przez Daniela Dennetta i Johna Searle'a. Analiza tych stanowisk wykazała, że o ile pierwsze jest spójne, choć wysoce nieintuicyjne, o tyle drugie, choć rozwijane w zgodzie z naszymi podstawowymi intuicjami dotyczącymi nas samych, zawiera wewnętrzne sprzeczności. Nieintuicyjność koncepcji Dennetta wynika z odrzucenia intencjonalności wewnętrznej. Stanowisko Searle'a jest sprzeczne, ponieważ autor ten przyjmuje intencjonalność wewnętrzną, a jednocześnie stara się utrzymać naturalizm.

Uznając, że stanowisko Dennetta nie jest sprzeczne, staję w tym sporze zasadniczo po jego stronie, choć za krytykami w rodzaju Reese'a Heitnera uznaję, że łagodny realizm jest stanowiskiem niestabilnym. Ponadto, wbrew Dennettowi, sądzę, że nieintuicyjność jego teorii stanowi nieusuwalny problem i źródło filozoficznego niepokoju. Gdzie indziej zwracałem uwagę, że konsekwencją uznania nieredukowalności tego niepokoju może być postawa, którą nazwałem „metamisterianizmem”. Postawa ta nie zawiera żadnej specyficznej teorii intencjonalności. Wynika ona z jednej strony z rezygnacji z pomysłu budowania naukowego obrazu podmiotu przy założeniu nieredukowalności intencjonalności, a z drugiej na uznaniu, że odpowiedź redukcjonistyczna zawsze będzie niezadowolająca. Z intencjonalnością jest bowiem tak, jak kiedyś – w kontekście nieco innych problemów – zauważył Wittgenstein: „Czujemy, że gdyby nawet rozwiązano wszelkie *możliwe* zagadnienia naukowe, to nasze problemy życiowe nie zostałyby jeszcze nawet tknięte. Co prawda, nie byłoby już wtedy żadnych pytań; i to jest właśnie odpowiedź” (Wittgenstein 1997, 82, 6.52).

BIBLIOGRAFIA

- D. Beisecker, *Dennett and the Quest for Real Meaning: In Defense of a Myth*, *Philosophy in the Contemporary World*, 9, 1, 2002, 11–18.
L. Carroll, *What the Tortoise Said to Achilles*, *Mind*, 104 (416), 1995, 691–693.
D. Chalmers, *Świadomy umysł*, przeł. M. Miłkowski, WN PWN, Warszawa 2002.
H. L. Dreyfus, *Alchemy and Artificial Intelligence*, RAND Corporation, Santa Monica (CA) 1965.
_____, *What Computers Can't Do*, MIT Press, New York 1972.

- _____, *The Primacy of Phenomenology Over Logical Analysis*, *Philosophical Topics*, 27, 2, 1999, 3–24.
- _____, *Phenomenological Description Versus Rational Reconstruction*, *Revue internationale de philosophie*, 216, 1971, 180–196.
- D. C. Dennett, *Content and Consciousness*, Routledge & Kegan Paul, New York 1969.
- _____, *Intentional Systems*, *Journal of Philosophy*, LXVII, 4, 1971, 87–106.
- _____, *Why the Law of Effect Will Not Go Away*, *Journal for the Theory of Social Behaviour*, 5, 2, 1975, 169–187.
- _____, *Brainstorms: Philosophical Essays on Mind and Psychology*, MIT Press, Cambridge (MA) 1981.
- _____, *The Intentional Stance*, MIT Press, Cambridge (MA) 1987.
- _____, *Consciousness Explained*, Back Bay Books & Little, Brown and Company, New York 1991.
- _____, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Penguin Press, London 1995.
- _____, *Kinds of Minds: Towards an Understanding of Consciousness*, Basic Books, New York 1996.
- _____, *Who's on First? Heterophenomenology Explained*, *Journal of Consciousness Studies*, 10, 9–10, 2003, 19–30.
- _____, *Odczarowanie. Religia jako zjawisko naturalne*, przeł. B. Stanosz, PIW, Warszawa 2008a.
- _____, *Rzeczywiste wzorce*, przeł. M. Miłkowski, w: M. Miłkowski, R. Poczobut (red.), *Analityczna metafizyka umysłu. Najnowsze kontrowersje*, Wyd. IFiS PAN, Warszawa 2008b, 299–325.
- _____, *How People Are Like Robots*, „Big Thing”, <http://http://bigthink.com/videos/how-people-are-like-robots> (dostęp: 10.09.2015) 2009.
- O. J. Flanagan, *The Science of the Mind*, MIT Press, Cambridge (MA) 1991.
- R. M. Heitner, *Is Design Relative or Real? Dennett on Intentional Relativism and Physical Realism*, *Minds and Machines*, 10, 2, 2000, 267–283.
- E. Husserl, *Filozofia jako nauka ścisła*, przeł. W. Galewicz, Aletheia, Warszawa 1992.
- P. Kołodziejczyk, *Między fenomenologią a cognitive science. Johna Searle'a teoria intencjonalności*, 2002 <http://www.kognitywistyka.net/artykuly/pk-mfacs.pdf>, (dostęp: 15.09.2015).
- S. Lem, *Summa technologiae*, *Dziela*. Tom XXVIII, Biblioteka Gazety Wyborczej, Agora, Warszawa 2010.
- J. McDowell, *Mind and World*, Harvard University Press, Cambridge (MA) 1996.
- W. J. Rapaport, *How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition*, *Journal of Logic, Language, and Information*, 9, 4, 2000, 467–490.
- J. R. Searle, *Minds, Brains and Programs*, *Behavioral and Brain Sciences*, 3, 1980, 417–457.
- _____, *Intentionality. An Essay in the Philosophy of Mind*, Cambridge University Press, Cambridge 1983.
- _____, *Umysł, mózg i nauka*, przeł. J. Bobryk, PWN, Warszawa 1995.
- _____, *The Mystery of Consciousness*, Granta Books, London 1997.
- _____, *Umysł na nowo odkryty*, przeł. T. Baszniak, PIW, Warszawa 1999a.
- _____, *Umysł, język, społeczeństwo: filozofia i rzeczywistość*, przeł. D. Cieślak, CIS i WAB, Warszawa 1999b.
- _____, *Neither Phenomenological Description Nor Rational Reconstruction: Reply to Hubert Dreyfus*, *Revue internationale de philosophie*, 216, 2002, 277–284.
- B. Stroud, *The Background of Thought*, w: E. Lepore, R. Van Gulick (red.), *John Searle and his Critics*, Blackwell, Cambridge 1991, 245–258.
- R. R. Traill, *Thinking by Molecule, Synapse, or Both? – From Piaget's Schema to the Selecting/Editing of ncRNA*, *Ondwelle*, Melbourne 2008 <http://www.ondwelle.com/OSMo2.pdf> (dostęp: 15.09.2015)
- B. Trybulec, *Fenomenologia a kognitywistyka – dwie metody analizy podmiotu poznania. Perspektywa współpracy i problemy*, *Filozofia i Nauka*, 3, 2015, 281–299.
- L. Wittgenstein, *Tractatus logico-philosophicus*, przeł. B. Wolniewicz, PWN, Warszawa 1997.
- _____, *Dociekania filozoficzne*, przeł. B. Wolniewicz, PWN, Warszawa 2000.

**COUNTERINTUITIVE OR INCONSISTENT THEORY
OF INTENTIONALITY?
A DISPUTE BETWEEN DENNETT AND SEARLE**

ABSTRACT

The paper compares the conceptions of intentionality held by the two leading figures in the contemporary philosophy of mind: John Searle and Daniel Dennett. I explain the relation between Searle's theories in philosophy of language and his theory of intentionality and his idea of three kinds of intentionality. Then I present Dennett's theory of intentionality as an adaptive feature and his idea of intentional stances. I underline the fact that, using Searle's vocabulary, according to Dennett no creature or artifact possesses intrinsic intentionality — all we deal with is just derived intentionality: the intentionality the observer attributes to a being in order to form a better prediction of the being's behavior. The consequence of this standpoint is highly counterintuitive: Dennett must (and in fact does) accept a claim that the observer's own intentionality is also only attributed to one by some other observers in order to form better prediction of one's behavior. This claim is rejected by Searle as an obvious absurd. I argue that Searle's standpoint is founded on the Cartesian grounds and point to the similarity between his critique of Dennett and Husserl's critique of naturalism. Since Searle claims his conception to be naturalistic, his theory is inconsistent.

Keywords: intentionality, naturalism, artificial intelligence, syntactic machine, John R. Searle, Daniel C. Dennett.

O AUTORZE — dr filozofii, adiunkt, afiliacja: Katedra Filozofii Boga na Wydziale Filozoficznym Uniwersytetu Papieskiego Jana Pawła II, ul. Kanonicza 9, 31-002 Kraków.

Email: jakub.gomulka@upjp2.edu.pl