

Arkadiusz Chrudzinski

Wilfrida Sellarsa krytyka „mitu danych”

1. Mit danych

Wilfrid Sellars przeszedł do historii jako niezmordowany tropiciel i krytyk wszelkich postaci tzw. „mitu danych”. Mit ten, który skutecznie zniewala filozofujące umysły od czasów Kartezjusza, to spłot trzech idei: idei sfery epistemicznie uprzywilejowanego dostępu; idei fundamentu wiedzy; oraz idei znaczenia jako bytu dostępnego w uprzywilejowany sposób, stanowiącego zasadę podziału na to, co analityczne i syntetyczne. Każda z tych idei mogłaby być wprowadzie potraktowana oddzielnie, jednak łączą się one w naturalny sposób i w tej postaci tworzą silną doktrynę, która jeszcze do niedawna zasługiwała na miano współczesnej ortodoksji.

Idea sfery uprzywilejowanego dostępu ma różne źródła, najnaturalniej występuje jednak w powiązaniu z koncepcją fundamentu wiedzy. Jedną z typowych filozoficznych gier jest gra w „jak możesz uzasadnić, że wiesz, że...”. Grę tę uprawiali już starożytni sceptycy, lecz jej kanoniczną formę ustalił Kartezjusz. On też przekazał przyszłym pokoleniom pomysł klasycznej „strategii wygrywającej”. W grze trzeba po prostu cofnąć się aż do *fundamentu*, do czegoś, co dane nam jest w taki sposób, że pytanie „jak możesz to uzasadnić?” traci sens. Wszystko inne, co jest uzasadnione, może być uzasadnione tylko przez określone relacje do tego fundamentu. Fundamentem tym stały się na długie lata *wrażenia*, byty o charakterze prywatnym, dostępne podmiotowi poprzez nieomylną (bądź „prawie nieomylną”) introspekcję.

Zaraz potem pojawia się następny element: idea znaczenia oraz podziału na to, co analityczne i syntetyczne. Już Kartezjusz analizował ideę Boga i udawał na tej podstawie wiele ciekawych rzeczy. U podstaw tej procedury leży założenie, że, o ile możemy się mylić co do istnienia przedmiotów naszych idei, o tyle co do kwestii tego, jakie idee posiadamy oraz jaka dokładnie jest ich zawartość, pomyłka jest wykluczona. Po Kartezjuszu przychodzi zaś Hume, który powie całkiem wyraźnie, że jedyne, co możemy wiedzieć, oprócz faktów dotyczących naszej sfery uprzywilejowanego dostępu, to relacje, jakie między elementami tam występującymi zachodzą. W ten sposób Hume’owskie *relacje pomiędzy ideami*

stały się pierwowzorem współczesnych prawd analitycznych. Niezbyt jasne pojęcie idei zostało naturalnie dopracowane. Jego miejsce zajął Fregowski *sens*, rozumiany jako abstrakcyjny byt pośredniczący w naszym odniesieniu intencjonalnym. Podstawowe rysy koncepcji Kartezjusza-Hume'a zostały jednak zachowane. Sensy (znaczenia) należą do obszaru tego, co dane; zaś zastosowane do nich procedury analityczne, zapewniają wyniki bliskie nieomyślności.

Najpełniejsza forma mitu danych miałaby postać następującej teorii: (1) dla każdego podmiotu istnieje sfera epistemicznie uprzywilejowanego, „bezpośredniego” dostępu, do której należą *wrażenia* oraz *znaczenia*; (2) *znaczenia* są bytami abstrakcyjnymi, które, jako identycznie te same, mogą być ujmowane przez wiele podmiotów; status ontologiczny wrażenia pozostaje z reguły nieokreślony; (3) wszelkie odniesienie intencjonalne poza sferę uprzywilejowanego dostępu odbywa się poprzez *ujęcie* pewnego *znaczenia*; *znaczenie* to specyfikuje cechy (ewentualnego) przedmiotu odniesienia; (4) tego typu odniesienie intencjonalne może być *uzasadnione* tylko poprzez określoną relację ujętego *znaczenia* do *wrażeń*, które podmiot aktualnie posiada; (5) to, jakie *znaczenia* podmiot ujmuje oraz jakie *wrażenia* posiada, może być mu znane w sposób „bezpośredni”, tzn. inny niż opisany w punktach (3) i (4), nie zapośredniczony przez inne *znaczenia* i *wrażenia*, i z tego względu szczególnie „niezawodny”; żaden podmiot nie może jednak posiadać tego typu bezpośredniej wiedzy o *innym* podmiocie.

2. Konkurencyjna teoria języka

Zarysowana wyżej teoria, inteligentnie zastosowana, jest w stanie wyjaśnić rzeczywiście imponująco dużo. A jednak jest ona obecnie gwałtownie atakowana. Twierdzi się, że cena, jaką przychodzi zapłacić za owo wyjaśnienie, jest zbyt wysoka. Pierwszym aspektem tej ceny ma być z gruntu fałszywy obraz języka, jakim zostajemy uraczeni.

Jak wiadomo, intencjonalność została przez Fregego wyjaśniona bez odwołania się do pojęcia języka, w terminach mentalnego ujęcia odpowiedniej treści abstrakcyjnej (Fregowskiego *sensu*)¹. W związku z tym koncepcja języka, jaka się naturalnie nasuwa, to koncepcja *instrumentalna*, ujmująca język jako *narzędzie komunikacji*, pozwalające przekazać nasze *myśli* innym. Język posiadałby zatem jedynie intencjonalność „nadaną” w ten sposób, że słowa zostają celowo użyte dla wywołania określonego skutku w słuchaczu. W uproszczeniu, jeżeli podmiot *S* mówi (używając wyrażenia *W*), że *p*, to muszą mieć miejsce następujące fakty: (1) Podmiot *S* mentalnie ujmuje *znaczenie*, że *p*; (2) podmiot *S* chce komuś zakomunikować, że *p*; oraz (3) podmiot *S* celowo używa wyrażenia *W* do zakomunikowania, że *p*.

¹ Jest to oczywiście znana teoria z rozprawy *Sens i znaczenie*, w: G. Frege, *Pisma semantyczne*, PWN, Warszawa 1977, s. 60-88.

Ważnym ogólnym aspektem rozważanego podejścia jest ujęcie aktywności myślowej jako zasadniczo niezależnej od zachowań językowych. Aby móc powiedzieć, że *p*, trzeba wcześniej móc pomyśleć, że *p*; jednak dla pomyślenia, że *p*, kwestia jego językowego wyrażenia jest absolutnie nieistotna. Ten właśnie punkt uważa Sellars za nie do przyjęcia. Język *jest* oczywiście narzędziem komunikacji, *po pierwsze* jednak i *przede wszystkim* jest on *medium*, w którym myślimy. Wszelkie myślenie jest językowe, tę ideę proponuje on przyjąć za filozoficzny punkt wyjścia. Jak zobaczymy, ruch ten prowadzi do kompletnej destrukcji tradycyjnej koncepcji.

Aby ruszyć z miejsca machinę zniszczenia, musimy zacząć konstruować konkurencyjną teorię. Konstrukcja ta postępować będzie według linii ogólnie pojętego behawioryzmu językowego. Zamiast obrazu użycia słów w intencji powiedzenia czegoś, skupmy się, proponuje Sellars, na obrazie *głośnego myślenia*.

Mamy ponury, jesienny wieczór. Ostatni autobus przejeżdża obok przystanku bez zatrzymania. Zmoknięty podróżny wykrzykuje spontanicznie: „Dlaczego się nie zatrzymał?!” Czy chce on kogoś o coś zapytać lub zakomunikować komuś swoje oburzenie? Na pewno nie *przede wszystkim*. Przede wszystkim, sugeruje Sellars, *pomyślał* on to, co powiedział, w naturalny, *jawny* sposób. Podobna sytuacja: Na przystanku czeka dwóch podróżnych. Autobus przejeżdża bez zatrzymania i obaj chcą wykrzyknąć to samo, lecz, jako że nauczyli się „zachowywać swoje myśli dla siebie”, nie mówią nic. Jednak ich postawa, zdumiony i oburzony wyraz twarzy, mimowolne ruchy szczęk, wskazują, iż mają to „na końcu języka”.

Obraz ten nie ma zbyt wielkiej siły, jeśli nie powiąże się go z odpowiednią teorią nabywania języka. Chodzi oczywiście o teorię behawiorystyczną zarysowaną przez Skinnera, a na filozoficzne salony wprowadzoną przez Quine’a².

W psychologii — pisze Quine — możemy być, lub nie być behawiorystami, w teorii języka nie mamy jednak takiego wyboru, behawioryzm jest tu opcją jedyną, a decyduje o tym natura przedmiotu, który badamy — języka³. Aby to zrozumieć, wystarczy rozważyć, jak się języka uczymy. Dziać się to może jedynie poprzez selektywne wznacnianie werbalnych reakcji dziecka na określone sytuacje bodźcowe. W procesie uczenia podsuwa się dziecku jednowyrazowe zdania (np. „Mama”) w określonych sytuacjach percepcyjnych (w tym przypadku w sytuacjach pojawienia się matki). Po pewnym czasie dziecko zaczyna powtarzać te zdania, na początku prawdopodobnie niezależnie od sytuacji. Zadaniem nauczyciela jest teraz odpowiednie wznacnianie jego werbalnych zachowań. Wymówienia zdania „Mama” są nagradzane, o ile matka znajduje się w takim położeniu, iż jest przez dziecko percypowana. Jeśli dziecko osiągnęło taki stopień wprawy, że niemal nieomylnie reaguje na pojawienie się matki odpowiednimi słowami, nauczyciel uznaje, że nauczyło się ono zwrotu „Mama”.

² Por. Przede wszystkim W. V. O. Quine, *Word and Object*, Cambridge, Mass. 1960.

³ Por. W. V. O. Quine, *Pursuit of Truth*, Cambridge, Mass. 1990, s. 38.

A zatem w procedurze uczenia zakłada się powszechnie, że dziecko *rozumie* wyrażenie, jeśli używa go systematycznie w odpowiedniej sytuacji bodźcowej. Jednak *rozumieć* wyrażenie, to to samo, co *znać jego znaczenie*. W problematyce znaczenia nie tkwi zatem nic więcej, niż to, co opisano powyżej.

Widać tu pewien uderzający fakt. Zgodnie z omawianą teorią język posiada intencjonalność *absolutnie niezależną od jakichkolwiek zdarzeń mentalnych*⁴. Istotą sprawy jest wyrobienie pewnych społecznie kontrolowanych regularności, wiążących wejście pobudzeniowe z wyjściem werbalnym. Jeśli mamy to, mamy tym samym intencjonalność i nie potrzebujemy niczego więcej.

Jeśli jednak jest tak w istocie, powrócić możemy do naszych znajomych z przyśtanku. Rozważmy tym razem tych „małomównych”. W obecnym stadium analizy przypadek ten opisać możemy następująco: Dzięki standardowemu treningowi językowemu, jaki obaj przeszli, mieli oni *dyspozycję* do wypowiedzenia w tej sytuacji zdania: „Dlaczego się nie zatrzymał?!” Dyspozycje z natury swej dążą do swego uzewnętrznienia (jest to prawdziwe definicyjnie), jednak mogą wystąpić czynniki hamujące, które spowodują, iż dyspozycje nie uzewnętrzną się. W naszym przypadku czynnikiem tym była inna dyspozycja o formie ogólnego imperatywu: „Zachowaj swe myśli dla siebie!” Ważnym morałem z tej analizy jest to, iż niewypowiadania swoich myśli musimy się uczyć tak samo jak kłamstwa. Podstawowym fenomenem teorii jest zaś szczere i otwarte głośne myślenie, uwarunkowane standardowym treningiem językowym, który wszyscy przechodzimy⁵. Takie głośne myślenie, jeśli ma być *medium* wszelkiej intencjonalności, nie może, pod groźbą nieskończonego regresu, być rozumiane jako celowe działanie stosujące się do reguł⁶. Musi być więc zinterpretowane naturalistycznie, jako wyuczona regularność zachowania, *opisywalna* w kategoriach reguł, nie będąca jednak świadomym stosowaniem się do tych reguł.

Jeśli jednak nauczyliśmy się już milczeć i kłamać, potrzebne staje się pojęcie niewyrażonej myśli. Podstawowe takie pojęcie powstaje, zdaniem Sellarsa, przez analogiczne rozszerzenie opisów głośnego myślenia. Myślenie nie wyrażone to, w pierwszym przybliżeniu, z jakichś względów zablokowane głośne myślenie. Myślenie nie wyrażone, to słowa, które mieliśmy „na końcu języka”.

W stosunku do tradycyjnego Fregowskiego obrazu mamy tu zatem całkowite odwrócenie porządku. To myśl ma intencjonalność pochodną w stosunku do intencjonalności języka. Trzeba jednak uważać, aby nie dezinterpretować teorii Sel-

⁴ Por. W. Sellars, *Naturalism and Ontology*. (The John Dewey Lectures for 1973-74), Reseda, California 1979 (dalej jako *NO*), s. 74, 117.

⁵ Por. *NO*, s. 75.

⁶ Zrozumienie reguł wymagałoby użycia metajęzyka, którego reguł z kolei trzeba by się nauczyć używając meta-metajęzyka itd. Por. W. Sellars, *Science, Perception and Reality*, London 1963 (dalej jako *SPR*), s. 321; W. Sellars, *Science and Metaphysics: Variations on Kantian Themes*. (The John Locke Lectures for 1965-66), London 1967 (dalej jako *SM*), s. 74-76; oraz *NO*, s. 76n, 80.

larsa zgodnie z instrumentalnymi schematami. Pochodność oznacza tu bowiem coś zupełnie innego. Myśl nie jest pochodnie intencjonalna dlatego, że aby uzyskać intencjonalność musi zostać „skojarzona” z odpowiednim jawnym zachowaniem językowym, lecz dlatego, iż myśl w tym znaczeniu po prostu *jest* (aczkolwiek w sensie analogicznym) pewnym zachowaniem językowym. Semantyczny wymiar jawnych zachowań językowych stał się bowiem *modelem* dla teorii semantycznego wymiaru myśli⁷.

Po tym pierwszym przybliżeniu do teorii myśli następuje jednak jej dalszy, ciekawy rozwój. Otóż behawioralna teoria języka nie rości sobie bynajmniej pretensji do bycia bezzałożeniową filozofią pierwszą, i z tego też względu nie ma żadnych oporów przed użyciem dorobku nauk przyrodniczych. W zastosowaniu do naszego problemu daje to jeden ważny wynik. Otóż społecznie wyrobione i kontrolowane dyspozycje do zachowań językowych, które stanowią jądro intencjonalności języka i myśli, są z punktu widzenia teorii fizycznej pewnymi strukturami zrealizowanymi w mózgu. Jeśli jednak tak, to możemy stworzyć dojrzałą teorię myśli, ujmującą je jako elementy funkcjonalne w ramach naszego układu nerwowego, pośredniczące między wejściem pobudzeniowym a wyjściem działaniowym.

Taka „drobnoziarnista” teoria myśli stoi do wyjściowej „gruboziarnistej” teorii, posługującej się prostą analogią ze zdaniem publicznego języka, w podobnym stosunku, w jakim stoi molekularna teoria fizyczna do teorii operującej zdroworozsądkowym pojęciem makroskopowego ciała fizycznego. Teoria taka może uwzględnić o wiele więcej niż teoria wyjściowa. Może ona włączyć elementy funkcjonalne odpowiadające myślom rozumianym jako odpowiedniki niewypowiedzianych i wypowiedzianych zdań w szerszą strukturę pozwalającą powiązać je z innymi elementami funkcjonalnymi, pojawiającymi się chociażby w związku z działaniami bodźców fizycznych na powierzchnie naszych receptorów. Sellars twierdzi, że właśnie to ostatnie pojęcie zająć musi w ramach obecnego obrazu miejsce Kartezjańskiego wrażenia.

⁷ Por. *NO*, s. 117. To powiązanie między myślą a zachowaniem językowym (między tym, co ukryte, a tym, co jawne, uzewnętrznione), mogło być dla filozofów przeszłości – pisze Sellars – odkryciem. Dla filozofa współczesnego jest to jednak coś w rodzaju prawdy logicznej. W istocie zaś, jest to przypadek *syntetyczności a priori*. Por. *SM*, s. 68n, 162n. Ta „syntetyczność” stosunku między myślami rozumianymi jako epizody wewnętrzne a ich językowym wyrażeniem ma być istotą odróżnienia pozycji Sellarsa od pozycji *logicznego behawioryzmu*, który utrzymuje, że myśli po prostu *definitywnie* są dyspozycjami do odpowiednich zachowań językowych. Stosunek tych dwu pojęć byłby więc analityczny. Por. *SM*, s. 73. Na temat Sellarsowskiej koncepcji syntetycznego *a priori*, por. jego: *Is there a synthetic a priori*, w: *SPR*, s. 298-320.

3. Naturalistyczne pojęcie wrażenia

Warto porównać to pojęcie wrażenia z różnymi wcieleniami teorii danych zmysłowych, jakie formułowali filozofowie „mitu danych”. Można wyróżnić zasadniczo dwie odmiany takiej teorii: wersję „przedmiotową”, gdzie dane występują jako swego rodzaju obiekty w sposób uprzywilejowany dostępne „oku umysłu”, oraz teorię „przysłówkową”, gdzie zamiast o „czerwonych danych zmysłowych” mówi się o „jawieniu się czerwono”, jednak owo jawienie się, jako że stanowić ma epistemiczny fundament, jest również w sposób uprzywilejowany dostępne podmiotowi. Pierwszy typ to oczywiście klasyczna postać teorii, tak jak sformułowana została przez Moore’a i Russella, drugi znaleźć można współcześnie u Chisholma⁸.

To, co odróżnia koncepcję Sellarsa od obu tych odmian, to fakt, iż jego wrażenia *nie są*, jako takie, dostępne oku umysłu. Wrażenia w sensie Sellarsa należą do „gramatyki głębokiej” myśli; są to *były postulowane teorii* tak, jak elektrony czy geny. Wywołują one wprawdzie przyczynowo odpowiednie przekonania, zwykle jednak nie są obiektem przekonań. Nigdy zaś nie są obiektem zwróconej nań aktywności kognitywnej o charakterze *quasi-percepcyjnym*.

Ten rys omawianej koncepcji każe nam zwrócić się ku teoriom, które przyjmują, iż wrażenia, choć w pewnym sensie dostarczają „materiału” poznania, to jednak normalnie nie występują w charakterze jego obiektu⁹. Tego rodzaju koncepcję głosił Kant i rzeczywiście Sellars widzi wiele prawdziwych myśli w jego filozofii. Przede wszystkim Sellars uważa za słuszną tezę Kanta dotyczącą „zjawiskowego” charakteru świata zdroworozsądkowego. Naukę, zdaniem Sellarsa, należy brać poważnie, a to znaczy, że należy traktować ją nie jako „instrument działania”, lecz jako coś, co dostarcza nam, jeśli nie *prawdziwego* obrazu świata, to przynajmniej obrazu *prawdziwszego* niż nasza wiedza potoczna. Nauka mówi nam jednak, że przedmioty fizyczne nie są i nie mogą być kolorowe. Przypisywanie im barw wygląda wręcz na pomyłkę kategorialną. Paradoksalna teza, iż barwy istnieją tylko w naszych wrażeniach, okazuje się, w pewnym swym znaczeniu, prawdziwa.

⁸ Por. B. Russell, *Problemy filozofii*, Warszawa 1995; R. Chisholm, *Teoria poznania*, Lublin 1994.

⁹ Z drugiej strony wrażenia *mogą* stać się obiektami naszego poznania i to w sposób dużo bardziej „bezpośredni” niż inne przedmioty. Zarówno przedmioty świata zewnętrznego, jak i tak rozumiane epizody wewnętrzne wywołują pewne stany kognitywne w sposób przyczynowy. W przypadku przedmiotów świata zewnętrznego jest jednak bardzo wiele pośredniczących ogniw, których brak w przypadku fizycznie rozumianego wrażenia. Tą bliskością tłumaczy się po części bardzo duże prawdopodobieństwo prawdziwości sądów w rodzaju „Teraz czuję ból”. Por. Sellars, *The Metaphysics of Epistemology. Lectures by Wilfrid Sellars* edited by Pedro V. Amaral, Ridgeview Publishing Company, Reseda, California 1989, s. 114.

Otóż wrażenia, rozumiane jako elementy funkcjonalne, stanowiące w naszym układzie nerwowym najbliższe odpowiedniki bodźców fizycznych, posiadać muszą pewną strukturę, *na podstawie której* powstają odpowiednie przekonania. W bardzo uproszczonym zdroworozsądkowym modelu wyglądałoby to następująco:

- (1) Podmiot S stoi naprzeciw czerwonego obiektu, światło odbite od jego powierzchni pada na siatkówkę oka powodując pewien stan neuronalny, który ze względu na jego funkcję nazwać możemy „wrażeniem czerwieni” (zaś tradycyjni filozofowie danych zmysłowych mówiliby wręcz o „czerwonym wrażeniu”). Funkcjonowanie układu nerwowego podmiotu S powoduje, iż w typowych okolicznościach na podstawie posiadania wrażenia czerwieni uformowana zostanie myśl (przekonanie) o treści „przede mną znajduje się czerwony obiekt”. Przekonanie to jest w tym przypadku *prawdziwe*.

Fizyka mówi nam jednak, że, ściśle biorąc, nie ma żadnych czerwonych obiektów. To, co w języku zdroworozsądkowym opisaliśmy jako czerwony obiekt, to po prostu określona struktura mikrocząsteczkowa, której można oczywiście przypisać pewne własności, jednak na pewno nie własność bycia czerwoną. A zatem naszą sytuację opisać trzeba nieco inaczej:

- (2) Podmiot S stoi naprzeciw pewnej struktury mikrocząsteczkowej posiadającej cechę C_{fz} , ze względu na tę cechę strumień fotonów „odbity” od tej struktury i padający na siatkówkę oka podmiotu S , ma taki charakter, iż powoduje stan neuronalny o własności C_{wr} który ze względu na jego funkcję nazwać możemy „wrażeniem czerwieni”. Funkcjonowanie układu nerwowego podmiotu S powoduje, iż w typowych okolicznościach na podstawie posiadania wrażenia czerwieni uformowana zostaje myśl (przekonanie) o treści „przede mną znajduje się czerwony obiekt” („przede mną znajduje się obiekt, który jest C_{fen} ”).

W naszym (1) odnosiliśmy się tylko do *jednej* własności. Przedmiot był *czerwony*, wywołane zostało wrażenie *czerwieni* („czerwone” wrażenie) i przekonanie podmiotu dotyczyło tego, że przedmiot jest *czerwony*. Teraz jednak widzimy, że słowo „czerwony” nie mogło być tam użyte jednoznacznie. W naszym przeformułowaniu (2) pojawiają się bowiem zamiast tego *trzy* własności. Przedmiot percepcji jest czerwony w sensie fizycznym (C_{fz}), wrażenie, które wywołuje, jest czerwone w sensie wrażeniowym (C_{wr}), natomiast przekonanie podmiotu, jako że sformułowane jest w terminach zdroworozsądkowych, stwierdza, że w polu widzenia znajduje się przedmiot czerwony „w sensie zwykłym”, który nazwać możemy sensem fenomenalnym (C_{fen}). Jasne jest, że głównym pytaniem teorii stać się musi kwestia stosunku, jaki zachodzi pomiędzy tymi trzema pojęciami czerwieni. W szczególności zaś, jako że czerwień fizyczna nie jest na pewno czerwiecią fenomenalną, powstaje pytanie, czy konkluzją naszego (2) stać się musi teza, iż przekonanie podmiotu jest przekonaniem *falszywym*? A jeśli tak, to co odróżnia

rodzaj jego fałszywości od rodzaju fałszywości, które miałyoby przekonanie, że obiekt jest *zielony*; przekonanie klasyfikowane w potocznym dyskursie jako przypisujące przedmiotowi fałszywą barwę?

Odpowiedź zaczyna się od stwierdzenia, że nasze trzy czerwienie (C_{fiz} , C_{wr} oraz C_{fen}) należeć muszą do trzech rodzin własności („barw”), z których każda uporządkowana jest pewnymi relacjami. Najłatwiej zacząć od barw fenomenalnych, czyli tych barw, które przypisujemy przedmiotom ujmowanym w kategoriach zdroworozsądkowych. Mamy różne barwy fenomenalne: czerwony, niebieski, zielony itd... Barwy te wykluczają się wzajemnie (jeśli coś jest czerwone, to tym samym nie jest ono niebieskie, zielone ani żółte) i dają się uporządkować zgodnie z porządkiem, w jakim następują po sobie w spektrum barw.

Jak pamiętamy, przekonania dotyczące przedmiotów barwnych w sensie fenomenalnym wywoływane były przez pewne stany neuronalne, które nazwalismy wrażeniami. Rozsądną hipotezą jest przyjęcie, iż wrażenia takie muszą posiadać pewne cechy (typy struktur rozkładu ładunków w odpowiednim obszarze kory mózgowej), które odpowiadają za to, jaką barwę fenomenalną przypisujemy przedmiotom spostrzeganym. Te cechy konfiguracji pobudzeń neuronalnych nazwijmy barwami wrażeniowymi i uporządkujmy je w zależności od tego, jakie barwy fenomenalne odpowiadają im w naszych przekonaniach.

Przejdźmy teraz do przedmiotów opisywanych fizykalnie. To, co decyduje, iż strumienie fotonów „odbijają” się od „powierzchni” przedmiotów w taki a taki sposób i wywołują takie a takie wrażenia, to nic innego, jak pewne cechy konfiguracji atomowych, tworzących te przedmioty. Te cechy uporządkować możemy odpowiednio do barw wrażeniowych, jakie wywołują, i nazwać z kolei barwami fizykalnymi.

Jak widzimy, barwy fizykalne, wrażeniowe i fenomenalne to całkiem różne własności. Przypisywanie przedmiotom rozumianym fizykalnie barw fenomenalnych byłoby podobną pomyłką kategorialną, jak przypisywanie ich liczbom. A zatem, jeśli uznajemy opis fizykalny świata za opis prawdziwy (lub przynajmniej „kategorialnie” bliższy prawdy niż opis zdroworozsądkowy) będziemy musieli twierdzić, że nasze zdroworozsądkowe przekonania są, mówiąc językiem Ryle’a, kategorialnie fałszywe.

Jednak barwy fizykalne, wrażeniowe i fenomenalne stanowią *analogiczne* rodziny własności. W każdej z tych rodzin mamy elementy (konkretne barwy) oraz pewne relacje porządkujące. Istotnym rysem tego, co nazywamy percepcją, jest fakt, że elementy jednej rodziny odwzorowywane są na elementy innych rodzin przy zachowaniu porządkujących relacji (czyli izomorficznie ze względu na tę relację). W związku z tym jest sens twierdzić, że pod pewnym względem przedmiot fizykalnie czerwony jest bardziej podobny do przedmiotu fizykalnie pomarańczowego niż do przedmiotu fizykalnie zielonego. Przekonanie to jest prawdziwe i odpowiada ono w pewnym sensie przekonaniu o tym, że przedmiot fenomenalnie

czerwony jest pod pewnym względem bardziej podobny do przedmiotu fenomenalnie pomarańczowego niż do przedmiotu fenomenalnie zielonego.

Dzięki temu, mimo iż kategorialnie fałszywie klasyfikujemy pewien przedmiot, możemy na tej podstawie *prawdziwie* twierdzić, że jest on bardziej podobny do tego przedmiotu niż do tamtego (i to pomimo faktu, że również te dwa ostatnie przedmioty klasyfikujemy kategorialnie fałszywie). W ten sposób zrozumieć możemy, jak nasze kategorialnie fałszywe przekonania zdroworozsądkowe mogą być mimo to „strukturalnie” prawdziwe i w jakim sensie można powiedzieć, iż barwy istnieją „tylko w naszych wrażeniach”, choć w sensie ścisłym mówienie o (fenomenalnie) czerwonym wrażeniu jest podobnym absurdem, jak mówienie o kwaśnych zbiorach.

4. Konkurencyjna ontologia znaczenia

Aby zarysowana powyżej behawiorystyczna destrukcja idei sfery uprzywilejowanego dostępu (świadomości, umysłu), jako elementu istotnego dla teorii języka i spostrzeżenia, miała szanse powodzenia, trzeba dotknąć pewnych kwestii ontologicznych. Idea takiej sfery łączy się bowiem w sposób naturalny z ideą znaczenia, jako bytu o charakterze platońskim, pośredniczącego w odniesieniu intencjonalnym. Aby zatem destrukcja była pełna, należy zaatakować również ten aspekt. Pozostawienie go *ad acta* grozi bowiem tym, iż spotkamy się z następującą reakcją:

To wszystko piękne, co powiedziałaś o procesie nabywania języka. Być może jest w tym wiele rzeczy prawdziwych, nie może być to jednak *cała prawda*, a powód tego jest bardzo prosty. Można skonstruować „mówiącą maszynę”, która wykazywać będzie odpowiednie „zachowanie językowe”, która jednak nie będzie *rozumiała*, co wypowiedane przez nią słowa znaczą¹⁰. Istotne dla autentycznego użycia języka jest więc pojęcie znaczenia. Musimy umieć wyjaśnić nie tylko nasze „zachowania językowe”, lecz także to, iż wiemy takie rzeczy, jak to, że niemieckie słowo „Hund” oznacza *pies*. Obawiamy się poważnie, że aby wyjaśnienie takie było możliwe, wprowadzić trzeba będzie na powrót znaczenia, rozumiane jako pewne byty, do których posiadać będziemy pewien specyficzny sposób dostępu.

Aby zablokować możliwość tego rodzaju krytyki, behawiorysta powinien zbudować jakąś „antyplatońską” ontologię znaczenia. Aby jednak to zrobić, należy najpierw dokładnie zlokalizować platońskie intuicje rządzące klasyczną teorią znaczenia. Sellars twierdzi, że intuicje te mają swe źródło przede wszystkim w typowych kontekstach podających znaczenie. Właśnie te konteksty stanowią punkt wyjścia dialektyki prowadzącej do idei *rozumienia* czy *uchwycenia* znaczenia. Chodzi o konteksty typu:

¹⁰ Por. J. R. Searle, *Umysły, mózgi i programy*, w: B. Chwedeńczuk (red.), *Filozofia umysłu*, Warszawa 1995, s. 301-324.

(3) (Niemieckie) zdanie „Es regnet” znaczy *pada deszcz*.

Pierwsze, co się narzuca, to relacyjna forma powierzchniowa tego zdania. W związku z tym mamy skłonność sądzić, że zdania takie mówią o *relacji* pomiędzy pewnym typem wyrażenia, a jego znaczeniem. Jeśli rozważymy poza tym konteksty stwierdzające *tożsamość znaczenia*, czyli zdania typu:

(4) (Niemieckie) „Es regnet” znaczy to samo, co (polskie) „Pada deszcz”,

zauważymy, że znaczenie nie jest ograniczone do *jednego* języka. Wiele różnych wyrażen, należących do różnych języków może znaczyć *to samo*. Łatwo będzie nam więc zaakceptować klasyczną teorię, uznającą znaczenia za byty *pozajęzykowe*, które ze względów ekonomicznych będziemy z kolei skłonni uznać za te same byty, do których odnosimy się mówiąc o identyczności cech czy faktów. W tym ujęciu znaczeniami nazw stają się pojęcia indywidualne, znaczeniami predykatów odpowiednie własności i relacje, znaczeniami zaś zdań stany rzeczy (lub *propositions*)¹¹.

Jeśli zwrócimy poza tym uwagę, że, na nieszczęście dla filozofów, możemy mówić różne rzeczy o Centaurach, Pegazach i kwadratowych kołach, jasne stanie się, że będziemy musieli wprowadzić pojęcia *niezrealizowane*, własności i relacje *nieegzemplifikowane* oraz stany rzeczy *niezachodzące*. W ten sposób byty pełniące funkcję znaczeń muszą znaleźć się poza zasięgiem tego, co faktyczne, a teoretyk znaczenia nie zatrzyma się wcześniej niż w platońskim niebie.

Sellars proponuje jednak, aby ukrócić naszą platońską fantazję i o zdaniach typu (3) myśleć w inny sposób. To, co dokonuje się w takim zdaniu, to przede wszystkim *przekład* zdania pewnego języka na odpowiednie zdanie języka, o którym zakłada się, że jest dla słuchacza zrozumiały. To, co się tu zatem w istocie stwierdza, to nie zachodzenie relacji pomiędzy zdaniem a jakimś bytem pozajęzykowym, lecz raczej to, że odpowiednie zdanie pewnego języka (tu niemieckiego) pełni w ramach tego języka *taką samą funkcję*, jaką pełni podane na końcu zdanie w naszym języku.

W naszym (3) zarówno „Es regnet”, jak i „pada deszcz” występują w sposób bardzo specyficzny. Cudzysłowy przy „Es regnet” sygnalizują nie to, że mówi się o słowach, jak np. w zdaniu: „‘Koń’ ma trzy litery”, lecz ogniskują uwagę słuchacza na *funkcji*, jakie zwrot ten pełni w języku niemieckim. Cudzysłowy są tu więc pewnego rodzaju klasyfikacją funkcjonalną. Użycie słów „pada deszcz” jest nie mniej dziwaczne. Słowa te w zasadzie również powinny znaleźć się, jak „Es regnet” w cudzysłowie, gdyż i tu chodzi w istocie o ich funkcję. To, że tak nie jest, tłumaczy się zaś faktem, że występują one w specyficznej roli *funkcjonalnej klasyfikacji ilustrującej*. Zwróćmy bowiem uwagę, że zdanie (3) nie mówi nam tego samego, co zdanie (4). To ostatnie stwierdza po prostu *identyczność znaczenia*, ale nie *podaje* go. Zdanie (3) natomiast *podaje* znaczenie, a dzieje się to

¹¹ Por. teorię *intensji* w: R. Carnap, *Meaning and Necessity*, Chicago 1960, s. 19-28.

w ten sposób, że w zdaniu tym *użyte* jest wyrażenie, o którym zakłada się, że *jest ono znane rozmówcy*¹².

W zasadzie wszystko, co zdanie (3) przekazuje w taki ilustracyjny sposób, można by przekazać za pomocą *wyraźnego opisu funkcjonalnego*. Można by po prostu *explicitie* podać wszystkie reguły, które rządzą użyciem słów „Es regnet” w ramach niemieckiej gry językowej¹³. Taki opis byłby jednak z oczywistych względów niezwykle skomplikowany i dlatego praktycznie dogodniejsza jest klasyfikacja ilustracyjna, podsuwająca po prostu uczestnikowi pewnej (tu polskiej) gry językowej określony element *jego* gry i mówiąca: „Patrz, niemieckie ‘Es regnet’ używane jest właśnie tak”¹⁴.

A zatem, zgodnie z Sellarsowską analizą, *znaczyć* nie jest żadną relacją, zachodzącą pomiędzy wyrażeniem a jego *znaczeniem*, lecz jedynie pewną wyspecjalizowaną formą spójki. Zdanie (3) mówi po prostu, że, jeśli wolno się tak wyrazić,

(5) „Es-regnet”-y *są* pada-deszczami,

gdzie „*Es-regnet*”-y to określone funkcjonalne elementy językowej gry niemieckiej, a *pada-deszcze* to elementy funkcjonalne językowej gry polskiej, które funkcjonują tak samo. Zdanie (5) jest więc podobne do zdania mówiącego, że Jan gra w szachy w ten sposób, iż monety pełnią funkcję pionów. Jednak *być* pionem to tyle, co pełnić określoną funkcję w grze w szachy, a zatem można po prostu powiedzieć, że u Jana monety *są* pionami. Podobnie w języku niemieckim „Es-regnet”-y po prostu *są* pada-deszczami. Widzimy tu wyraźnie, że, wbrew swej gramatyce powierzchniowej, słowo „znaczy” jest w istocie pewną wyspecjalizowaną odmianą słowa „jest”, ograniczoną do stwierdzania *identyczności funkcjonalnej elementów językowych*.

Aby uwidocznic lepiej charakter funkcjonalnej klasyfikacji ilustrującej, jaki ma występujące w zdaniu (5) wyrażenie „pada-deszcz” wprowadza Sellars specjalną notację *cudzysłówów kropkowych*¹⁵. Zdanie (5) wyglądałoby zatem następująco:

(6) „Es-regnet”-y *są* •pada-deszcz•ami

gdzie zapis •pada-deszcz• sygnalizować ma, iż chodzi o pewien element funkcjonujący *tak a tak*, niezależnie od tego, do jakiego języka by należał. Czyni ona

¹² Por. *SPR*, s. 161-164.

¹³ Por. *NO*, s. 92; *SM*, s. 79.

¹⁴ Por. *NO*, s. 91.

¹⁵ Por. *SM*, s. 79-82; oraz W. Sellars, *Towards a Theory of Predication*, w: J. Bogen, J. E. McGuire (eds), *How Things are*, Dordrecht 1985 (s. 285-322), s. 291-295 (dalej jako *TTP*).

również jasnym, iż powód, dla którego między kropkami wystąpiło wyrażenie polskie, a nie np. angielskie lub rosyjskie, jest wyłącznie ilustracyjnej natury¹⁶.

Ta analiza znaczenia pozwala Sellarsowi uwolnić się również od innych kontekstów, które wydają się wymuszać platońską ontologię. Podstawowymi takimi kontekstami są użycia abstrakcyjnych terminów jednostkowych w pozycji podmiotu. Skoro pozycja ta nadaje się do kwantyfikowania, to zgodnie z tzw. *kryterium Quine'a*, głoszącym, że *być to być wartością zmiennej kwantyfikowalnej*, zobowiązani jesteśmy do platońskiej ontologii. Sellars jednak zapatruje się sceptycznie na to kryterium. Twierdzi on, że nie ma oczywistych powodów, dlaczego współczesna teoria kwantyfikacji miałaby rozstrzygać kwestie ontologiczne¹⁷. Zamiast kryterium Quine'a proponuje on zatem przyjąć raczej *zasadę tolerancji Carnapa*, pozwalającą każdemu stosować taką notację, jaka wydaje mu się wygodna, bez rozstrzygania kwestii ontologicznych za pomocą magii słów „zmien-na” i „kwantyfikator”¹⁸.

Wszelkie środki *nominalizacji* pozwalające z wyrażen takich, jak „jest czerwony”, „jest krewnym” czy „pada deszcz” utworzyć odpowiednie jednostkowe terminy abstrakcyjne: „czerwoność”, „pokrewieństwo” czy „to, że pada deszcz”, są w istocie tym samym specyficznym typem *kontekstów cudzysłowowych*, które tworzą metajęzykowe klasyfikacje funkcjonalne¹⁹, z którymi mieliśmy już do czynienia przy okazji pseudo-relacji: *znaczy*, a które Sellars zapisuje jako cudzysłowy kropkowe²⁰. A zatem zgodnie z ujęciem Sellarsa, jeśli powiemy, że:

(7) Centaurowatość jest nieegzemplifikowana,

mówimy w istocie coś w rodzaju:

(8) Nie istnieje żaden element funkcjonalny typu nazwy indywidualowej, który połączony z •jest-centaurem• (czyli z elementem, który pełni funkcję taką, jaką w naszym języku pełni wyrażenie „jest centaurem”), dałby zdanie prawdziwe.

Cudzysłowy kropkowe mogą być zaś, jak widzieliśmy, wyeksplikowane w kategoriach dopuszczalnych konfiguracji napisów i wypowiedzi w ramach skomplikowanych typów zachowań, które przez analogię nazwać możemy „grami języ-

¹⁶ To ujęcie znaczenia jest równoważne twierdzeniu, że podać znaczenie wyrażenia to tyle, co sklasyfikować je (funkcjonalnie). Że jednak kryteria takiej klasyfikacji są plastyczne, przeto *tożsamość* znaczenia staje się przypadkiem granicznym *podobieństwa* znaczenia. Kryteria te są poza tym mocno zależne od kontekstu. I tak, niemieckie „nicht” czasami będzie prawidłowo klasyfikowalne jako •nie• rozumiane klasycznie, czasami zaś, jako •nie• intuicjonistyczne. Por. *TTP*, s. 296.

¹⁷ Por. *NO*, s. 32, 59.

¹⁸ Por. *NO*, s. 24n.

¹⁹ Por. *NO*, s. 94; *SM*, s. 102-107.

²⁰ Tak więc, o ile platonik odwołuje się w swej semantyce i ontologii do bytów *poza-językowych*, Sellars, używając swej idei funkcjonalnej klasyfikacji, mówi raczej o bytach *między-językowych*. Por. *SM*, s. 100.

kowymi”²¹. Wydaje się więc, że byty platońskie stopniowo znikają z naszego ontologicznego horyzontu.

W sformułowaniu (8) widzimy jednak dwie ciekawe rzeczy. Po pierwsze kwantyfikuje się tu po *elementach funkcjonalnych*, po drugie zaś, w sposób istotny występuje pojęcie *prawdy*. Obie te kwestie wymagają omówienia.

Jeśli chodzi o pierwszą, to jest to standardowy zarzut pod adresem teorii Sellarsa. Jak można być usatysfakcjonowanym z uwolnienia się od takich bytów abstrakcyjnych, jak własności czy stany rzeczy — można zapytać — jeśli w zamian za to przyjmuje się równie abstrakcyjne, a na pierwszy rzut oka dużo bardziej dziwaczne, jako że dużo gorzej zadomowione w filozoficznym dyskursie, byty nazywane *funkcjami*, *rolami* czy *figurami* gier językowych?²²

Sellars odpowiada na ten zarzut nieco zagadkowo. Po pierwsze twierdzi, że z żadnych bytów abstrakcyjnych nie zrezygnował, lecz wyjaśnił jedynie ich prawdziwą (czyli *metajęzykową*) naturę. Po drugie zaś, proponuje użycie analogicznej strategii eliminacyjnej do pojęcia funkcji czy roli elementu w języku²³.

Pojawia się zresztą natychmiast dalszy problem. Chodzi o problem *nieegzemplifikowanych* funkcji językowych. Czy nie możemy mówić o takich elementach funkcjonalnych języka, które miały to nieszczęście, że nie są zrealizowane w żadnym języku? Czy nie to dokładnie zrobiliśmy w ostatnim zdaniu? Taka nieegzemplifikowana funkcja zaczyna bardzo niebezpiecznie przypominać nieegzemplifikowaną centaurowość.

Sellars powołuje się, jak widzieliśmy, na możliwość wyeksplikowania pojęć funkcjonalnych w terminach pojęć należących do *dyskursu normatywnego*, mówiącego o tym, co dozwolone, a co nie, w danej grze językowej. Jest to prawdą, zamiast mówić o funkcjach możemy mówić o *dopuszczalnych* strukturach zachowań językowych. Ta możliwość oznacza jednak jedynie to, że zamiast bytów abstrakcyjnych zaakceptować musimy pewien *dyskurs modalny*. To, że możemy tak zrobić (możemy zredukować byty abstrakcyjne na rzecz pewnych pojęć modalnych albo odwrotnie) jest dość oczywiste, lecz właśnie dlatego niekoniecznie filozoficznie interesujące²⁴.

W dyskusji naszej dotknęliśmy z konieczności pewnego bardzo przygnębiającego rysu wszelkich argumentacji filozoficznych związanych z problemem powszechników. Rys ten polega na tym, że, gdy tylko rozważymy nieco poważniej jakąkolwiek próbę redukcji bytów ogólnych, wszystkie problemy, których roz-

²¹ Por. *SM*, s. 107.

²² Por. *NO*, s. 112.

²³ Por. *NO*, s. 166.

²⁴ Por. na ten temat: *Correspondence with Michael Loux*, w: *NO*, s. 149-180. Interesujące jest, iż w kwestii bytów abstrakcyjnych Sellars zdaje się iść w kierunku dokładnie przeciwnym niż Quine. Quine, o ile tylko może, redukuje pojęcia modalne na rzecz kwantyfikacji zmiennych wyższego poziomu, co jest dla niego równoznaczne z akceptacją odpowiednich przedmiotów abstrakcyjnych.

wiązaniem miało być wprowadzenie uniwersaliów, zdają się nieuchronnie powracać na wyższym poziomie. Redukcja przez odwołanie się do pojęcia podobieństwa wywołuje natychmiast pytanie o ontologiczny charakter podobieństwa. Redukcja konceptualna, jak w przypadku teorii Sellarsa, prowadzi zaś do trudnych kwestii dotyczących funkcji nieegzemplifikowanych, wielu egzemplarzy tej samej funkcji (tego samego pojęcia) oraz kryterium ich identyczności²⁵.

Teoria Sellarsa nie daje klasycznie rozumianej *odpowiedzi* na te pytania. Przez sugestię, że do uniwersaliów metajęzykowych należy zastosować tę samą strategię, co do uniwersaliów z poziomu przedmiotowego, zdaje się postulować ideę nieskończonej hierarchii metajęzyków w wyjaśnianiu filozoficznym. Idea taka, dodajmy, może być bardzo frustrująca, nie jest jednak bynajmniej absurdalna. Zasadnicza „niezamkniętość” problemów filozoficznych, udokumentowana dostatecznie przez dzieje tej dyscypliny, dostarcza idei tej dość silnego poparcia.

To, o co Sellarsowi w istocie chodzi, to mówiąc językiem Wittgensteina, *zmiana obrazu*, jaki leży u podstaw naszej filozoficznej dialektyki²⁶. Krytykowane idee sfery uprzywilejowanego dostępu, znaczeń jako bytów pośredniczących, tradycyjnej koncepcji danych zmysłowych oraz instrumentalnej teorii języka (jako narzędzia komunikacji a nie medium myślenia), to swego rodzaju „pojęciowe kołki”, na których od wielu pokoleń „zawieszamy” nasze argumentacje i które wyznaczają granice naszego filozoficznego dyskursu. To, co osiągnął Sellars przez swą destrukcję tych filozoficznych struktur, to nie jakieś ostateczne wykazanie ich fałszywości, lecz raczej poważna próba ich sproblematyzowania; sproblematyzowania z punktu widzenia (czegoż by innego?) innych takich kołków. Kołki Sellarsa to idea języka, jako podstawowego medium intencjonalności, behawiorystyczna koncepcja jego nabywania i funkcjonalistyczna technika podejścia do problemów znaczenia oraz powszechników. Czy ta zmiana obrazu przyczyniła się pozytywnie do uplastycznienia naszego filozoficznego widzenia świata, osądzić musimy sami.

5. Gry językowe

Jeśli chodzi o pojęcie prawdy, to Sellars definiuje prawdziwość zdania (w języku L) jako jego poprawną stwierdzalność, zgodnie z semantycznymi regułami owego języka. Mówiąc krótko: Zdanie *p* (języka L) jest prawdziwe =Df. *zdanie p jest (w języku L) semantycznie stwierdzalne* (S-stwierdzalne)²⁷. Pojęcia tej po-

²⁵ Dobrym przeglądem wszystkich tych problemów jest D. M. Armstrong, *Universals and scientific realism*, Cambridge 1978, tom I.

²⁶ Por. *NO*, s. 112.

²⁷ Por. *SM*, s. 101. Prawdziwość jest więc relatywna do reguł semantycznych odpowiedniego języka, por. *SM*, s. 132n. Sellarsowska prawdziwość w pewnym sensie nie dopuszcza stopni. Albo coś jest S-stwierdzalne w pewnym języku, albo nie. W innym jednak sensie mogą być języki (struktury pojęciowe) mniej lub bardziej adekwatne. W tym ostatnim sensie

prawności oraz *reguły* wyjaśniane zaś są w modelu języka rozumianego, w duchu późnego Wittgensteina, jako gra.

Gra językowa, która stanowić może model ułatwiający zrozumienie funkcjonowania naszego języka, zawierać musi zdaniem Sellarsa trzy rodzaje „ruchów”.²⁸ (a) *wejście*, czyli pewne mechanizmy przetwarzania świat-język, które odpowiadają za ustalenie „pozycji wyjściowych”, z których rozpoczyna się gra. Chodzi tu o mechanizmy pozwalające na uzyskanie tzw. „empirycznej informacji”, które odpowiadają omówionemu pojęciu wrażenia. Sellars porównuje to często z początkowym ustawieniem figur w szachach, jednak ze względu na zmienny i „kontyngentny” charakter pozycji wyjściowej lepsza byłaby chyba metafora „rozdania kart”; (b) *ruchy wewnątrzjęzykowe*, które polegają na wszelkich przekształcaniach formuł, które nazywamy zwykle *wnioskowaniami*; oraz (c) *wyjście*, które wiąże pewne pozycje w grze językowej z określonym *zachowaniem*. I tak na przykład pozycja w grze: „przyjmuję: teraz podnoszę rękę” kulminuje w sposób naturalny w podniesieniu ręki.

Każdy z tych trzech typów ruchów rządzony jest odpowiednimi regułami. Wszystkie te reguły wzięte łącznie nazywa Sellars *regułami semantycznymi* języka²⁹. Jeśli więc Sellars mówi nam, iż „prawdziwe” znaczy po prostu: „prawidłowo stwierdzalne ze względu na reguły semantyczne odpowiedniego języka”, nie chodzi mu po prostu o wyrażenie akceptacji dla podstawowych idei semantyki logicznej, nawiązującej do definicji Tarskiego. Sellars akceptuje wprawdzie te idee, lecz raczej jako coś prawdziwego trywialnie i w tym sensie filozoficznie nieinteresującego. Odwołując się do pojęcia reguł semantycznych chce natomiast powiedzieć, że zdanie prawdziwe, to coś w rodzaju *właściwej pozycji w grze*, rozpatrywanej ze względu na holistycznie pojęte jej reguły (i być może z dodatkiem: w nieskończonym przebiegu)³⁰. Ten filozoficzny obraz zastąpić ma obraz *korespondencji* między zdaniem a światem, który wykształcił się ze względu na mylącą relacyjną formę powierzchniową zdań podających znaczenie.

Aby nie było jednak nieporozumień, Sellarsowski język myśli nie jest „monadą bez okien”. Relacje ze światem istnieją, ale nie są to już teraz *pseudo-relacje intencjonalne*, lecz *autentyczne relacje przyczynowe*. Chodzi oczywiście o związki przyczynowe uwikłane w regularności dotyczące ruchów typu (a) i (c). Świat kauzalnie oddziałuje na podmioty, a podmioty kauzalnie oddziałują na świat. Tylko w ten sposób powstawać mogą teorie, które zasługują na miano em-

możemy zdaniem Sellarsa mówić o mniej lub bardziej adekwatnych prawdach. Por. *SM*, s. 134.

²⁸ Por. *SPR*, s. 329; *SM*, s. 114.

²⁹ Tak więc semantyka Sellarsa zawiera w istocie syntaktykę i pragmatykę Morrisa.

³⁰ Por. nawiązania do Peirce’owskiej idei definicji prawdy jako tego, co akceptowałoby idealne społeczeństwo badaczy. Sellars ujmuje to w terminach idealnie adekwatnego języka (struktury pojęciowej), który nazywa „Peirceish”. Por. *SM*, s. 140-142.

pirycznych i tylko dlatego istnieją języki ucieleśniające takie teorie, czyli właśnie takie języki, których używamy.

Przyczynowo powiązane ze światem wejście języka opisuje Sellars w terminach „obrazkowej” teorii języka z *Traktatu* Wittgensteina³¹. Istnieją zdania atomowe, które stanowią obrazy przedmiotów świata, zaś zasada obrazowania to właśnie skomplikowana struktura przyczynowa, decydująca o regularnościach wejścia i wyjścia, czyli ustalająca reguły rządzące ruchami typu (a) i (c). Sellars akceptuje też Wittgensteinowską koncepcję języka nie zawierającego wyrażen innych niż nazwy indywiduów³², zaś funkcję, jaką pełnią predykaty, powierzającego *konfiguracjom*, w jakich występują nazwy. Idea takiej notacji w powiązaniu z wyżej omówionym podejściem do ontologii znaczenia, pozwala Sellarsowi realizować program *filozofii naturalistycznej*, która chce być zarazem *filozofią nominalistyczną*.

7. Introspekcja

Jest jednak jeszcze jeden aspekt, który ciągle wydaje się stanowić silny argument dla Kartezjańskiej idei sfery uprzywilejowanego dostępu. Chodzi o naszą introspekcyjną wiedzę na temat własnych niewypowiedzianych myśli oraz doznawanych wrażeń, którą każdy z nas rozporządza, a którą wydaje się nie rozporządzać nikt inny. Wiedzy takiej towarzyszy poza tym tak wysoki stopień uzasadnienia, iż bardzo wielu wybitnych myślicieli było skłonnych do mówienia w tym kontekście o *nieomyślności*.

Sellars przyjmuje tę argumentację. Nawet jeśli odrzucimy ideę generalnej nieomyślności w stosunku do własnych stanów mentalnych, zjawisko *niezwykle wysokiej wiarygodności* każdej szczerzej autodeskrypcji rzeczywiście ma miejsce. Da się ono jednak wyjaśnić w ramach behawiorystycznej teorii języka zarysowanej powyżej.

Przypomnijmy, że według behawiorystycznego modelu myśl jest przede wszystkim *głośnym* myśleniem, zaś myśl nie wyrażona głośno jest *dyspozycją* do takiego głośnego myślenia, wyjaśnialną z kolei, jak chcemy wierzyć, w ramach przyszłej teorii neurofizjologicznej. Jak jednak pogodzić to z faktem, że wiedza dotycząca własnych stanów mentalnych jest w normalnych warunkach wiedzą bezpośrednią (nie-inferencyjną). Aby wiedzieć, że jestem obecnie zdenerwowany, nie muszę (1) obserwować systematycznie przez lata swojego zachowania, aby ustalić indukcyjnie, iż ilekroć myślę głośno, że jestem zdenerwowany, pocą mi się

³¹ Por. *SM*, s. 118-128, 136n; *TTP*, s. 313-321.

³² To, czy Wittgensteinowskie nazwy powinny być rzeczywiście rozumiane jako nazwy *indywiduów*, jest sprawą sporną. Sellars w każdym razie argumentuje za taką interpretacją. Powiązania pomiędzy nazwami indywiduowymi a przedmiotami, do których się one odnoszą, ujmując Sellars jako „psycho-socjo-historyczne relacje”, co zbliża jego ujęcie do znanej teorii Kripkego. Por. *TTP*, s. 318; S. Kripke, *Nazywanie a konieczność*, Warszawa 1988, s. 92.

ręce; (2) zaobserwować, że teraz pocą mi się ręce; oraz wnioskować z tego, że (3) w takim razie prawdopodobnie mam obecnie dyspozycję do głośnego wypowiedzenia słów „jestem zdenerwowany”. Jeśli jestem zdenerwowany, to z reguły wiem o tym w podobnie bezpośredni sposób, jak np. o tym, że znajduje się przede mną jabłko. Właśnie te fenomenologiczne cechy samowiedzy generują *quasi-percepcyjną* metaforykę „oka umysłu”, która do dziś włada analizami stanów mentalnych.

A jednak model percepcyjny nie jest tu potrzebny. Przypomnijmy sobie obraz treningu językowego, który leży u podstaw teorii Sellarsa. Fundament stanowią proste zdania obserwacyjne takie jak „Mama” czy „Piłka”, których systematyczna korelacja z odpowiednią stymulacją bodźcową decyduje o tym, że funkcjonują one później jako „uzasadnione bezpośrednio” przekonania. Jeśli założymy, że Jan przeszedł standardowy trening językowy, to w przypadku, kiedy szczerze i spontanicznie mówi on: „O! Piłka.”, mamy bardzo dobre podstawy, aby sądzić, że rzeczywiście znajduje się przed nim piłka, nawet jeśli sami tego nie widzimy.

Istotnym elementem treningu językowego jest opanowanie zdolności mentalnej autodeskrypcji. Zastanówmy się, jak może się to dziać? Mały Jasio upada i rozbija sobie kolano, siada na ziemi i głośno płacze. Mama, usiłując uspokoić malca mówi: „Oh! Boli cię!”, Jasio, który nie zna jeszcze tego zwrotu, powtarza „Bo-li, Bo-li”. Jeśli następnego dnia upadnie znowu, wypowie być może te same słowa, a mama, jeśli będzie w pobliżu, potwierdzi: „Tak! Boli cię. Ale nie płacz.” W ten sposób Jaś opanowuje stopniowo jeden z ważnych elementów autodeskrypcji.

Jeśli behawioryzm językowy ma być konsekwentny, twierdzić trzeba, iż wszelka samowiedza zasadza się na tego typu warunkowaniu. Na początku nauki nie istnieje więc żadna sfera uprzywilejowanego dostępu. Rodzice, którzy zwykle dość dobrze znają swoje dzieci, mogą, na podstawie ich jawnych zachowań w sposób bardzo wiarygodny przewidywać ich stany mentalne (zwłaszcza, że w tym stadium nie mogą być one jeszcze zbyt skomplikowane) i to *ich* wiedza jest motorem nauki. Na tej podstawie mogą oni wytrenować dziecko do wiarygodnego werbalnego odpowiadania na własne dyspozycje do głośnego myślenia „*p*” w formie: „Myślę, że *p*”³³. Późniejsza historia, wprowadzająca do programu nauki imperatywy kładące zachowywać swoje „stany wewnętrzne” dla siebie, wprowadza dwa rodzaje bloków. Po pierwsze (na mocy maksymy: „zachowaj swe myśli dla siebie”), blokowane są głośne wypowiedzenia o formie „Myślę, że...”; po drugie (na mocy maksymy: „nie okazuj swych uczuć”), blokowane są również zachowania niewerbalne, które stanowią naturalny wyraz odpowiednich stanów dyspozycyjnych. Jeśli więc Jaś poddany zostanie zwykłym naciskom społecznym, po pewnym czasie nikt nie będzie w stanie odkryć, że coś go boli, choć on sam będzie to wiedział w sposób bezpośredni.

³³ Por. Sellars, *The Structure of Knowledge*. (The Machette Foundation Lectures for 1971 at the University of Texas), w: Hector-Neri Castaneda (ed.), *Action, Knowledge and Reality, Studies in Honor of Wilfrid Sellars*, Indianapolis 1975 (s. 295-347), s. 326n.

Dodajmy na koniec, że to, co dokonuje się w ramach treningu językowego związanego z autodeskrypcją, daleko wykracza poza to, co nazwać moglibyśmy nauką „nazywania”, czy nawet „rozpoznawania” własnych stanów mentalnych. Tak wygląda to jedynie z punktu widzenia nauczyciela, który interpretuje „życie wewnętrzne” dziecka w analogii do własnego. Z punktu widzenia ucznia mamy tu jednak coś w rodzaju *tworzenia nowej warstwy bytu*. Dziecko, dopóki nie było dyspozycji do werbalnego odpowiadania na własne „stany wewnętrzne”, po prostu *nie ma takich dyspozycji*, to jednak oznacza ni mniej ni więcej, tylko to, że *nie ma ono stanów wewnętrznych* w tym znaczeniu, w jakim chcemy nazywać je *stanami mentalnymi*. Posiadanie dyspozycji do odpowiedniego jawnego zachowania werbalnego *należy bowiem do pojęcia stanu wewnętrznego, rozumianego jako stan mentalny*. Stany mentalne nie są więc czymś, co „istnieje w sobie” i co musimy się tylko nauczyć opisywać czy wyrażać. Warstwa stanów mentalnych to raczej coś, co dopiero tworzy się w samym procesie nauki ich wyrażania. W tym sensie nie ma filozoficznego „problemu innych umysłów”. Kryterium przypisywania stanu mentalnego wbudowane jest rzeczywiście, tak jak tego chciał Wittgenstein, w samo jądro logiki języka mentalistycznego.

Wilfrid Sellars' Critique of the Myth of the Given

Wilfrid Sellars is famous for his brilliant critique of the 'myth of the given'. Viewed by many as demolishing, it is by no means a piece of internal criticism. On the contrary, it is rooted in his behavioristic theory of language, and in his ideas about science, the mind and ontology. Without this background it is hardly understandable. In the article the author tries to make this point in more detail. The myth of the given consists principally of three main ideas: (i) that there is a sphere of epistemologically privileged access; (ii) that knowledge is based on foundations, and (iii) that the classical theory of meaning and the related synthetic/analytic distinction are not to be challenged. The force of Sellars' arguments lied in the bold refutation of these assumptions. He proposes an alternative picture in which (i) language is conceived as a behavioristically describable language-game, (ii) thought is equated with 'silent speech', and (iii) theory of meaning is presented in functionalist terms. The final result of this criticism is a forceful rejection of the Cartesian concept of mind.