

T o m a s z K o r b a k

Apercepcja transcendentalna w kantowskim modelu epigenezy czystego rozumu

Słowa kluczowe: *I. Kant, kognitywistyka, transcendentalna filozofia umysłu, apercepcja transcendentalna, epigeneza, samoorganizacja, naturalizm, reprezentacyjna teoria umysłu, funkcjonalizm*

Wstęp

Zadanie historycznej rekonstrukcji koncepcji filozoficznej zawsze obarczone będzie pewną sprzecznością intencji: konfliktem między dążeniem do adekwatności hermeneutycznej i adekwatności empirycznej. Recepcja idealizmu transcendentального Immanuela Kanta jest jaskrawym przykładem tego problemu: *Krytyka czystego rozumu* w ciągu przeszło dwóch wieków obrosła całymi bibliotekami neokantyzmów, rzadko kiedy motywowanych bezinteresowną, hermeneutyczną ciekawością. Chciałoby się powiedzieć, że filozofowie znajdują w trzech *Krytykach* to, co sami *a priori* w nie wkładają. I tak, wśród psychologów i filozofów umysłu: Helmholtz odczytywał czyste formy naoczności oraz kategorie jako ograniczenia nakładane przez fizjologię zmysłów (pomysł ten podchwycili Uexküll, Lorenz, a współcześnie przedstawiciele językoznawstwa kognitywnego), gestaltyści czuli się kantystami dzięki tezie o aktywnym charakterze percepcji, Freud postulując niedostępną introspekcyjnie nieświadomość u podłoża całego życia psychicznego, Jaspers zaś rozwijając kantowskie pojęcie rozumu do pojęcia osobowej egzystencji. W tradycji analitycznej znalazła nowych entuzjastów koncepcja częściowo apriorycznych mechanizmów poznania – w osobach Chomsky’ego (w odnie-

sieniu do kompetencji językowej), Jackendoffa (w odniesieniu do kategoryzacji pojęciowej, zob. Jackendoff 1999) i Fodora (jego wskrzeszenie koncepcji schematu transcendentznego w teorii pojęć w: Fodor 2011). Należy też powiedzieć o całej tradycji jawnie inspirowanej Kantem epistemologii i filozofii umysłu, zaczynającej się u Sellarsa, a kontynuowanej przez Putnama (1998), Davidsona i Brandoma. Wreszcie warto wspomnieć o enaktywistycznych fascynacjach trzecią *Krytyką* (Weber, Varela 2002). Wszystkie te różnorodne, nierzadko w dużej części sprzeczne podejścia łączy jakiegoś rodzaju roszczenie do miana transcendentnej filozofii umysłu.

I ja nie zamierzam podejmować się przedsięwzięcia szukania Kanta samego w sobie. W istocie bardziej filozoficznie interesujące jest śledzenie meandrów myśli kantowskiej we współczesnej kognitywistyce i filozofii umysłu. Spróbuję zdystansować się od dominującego w analitycznej filozofii umysłu odczytania Kanta jako funkcjonalisty i pozwolić mu pokazać nieco inne oblicze. Takie, w którym swojego patrona mogłyby znaleźć także nowe nurty współczesnej kognitywistyki¹.

Na czym polega transcendentna filozofia umysłu?

Opracowanie systematycznego opisu architektury ludzkiego umysłu nie należało do głównych celów filozofii krytycznej: bardziej niż znalezienie psychologicznych zasad jego funkcjonowania (to, co Kant nazywa dedukcją metafizyczną) interesowało ją znalezienie warunków koniecznych, jakie muszą być spełnione, aby cokolwiek mogło tym psychologicznym zasadom uczynić zadość (to zadanie nosi nazwę dedukcji transcendentnej). Widać to jasno, gdy Kant stwierdza w przedmowie do wydania A pierwszej *Krytyki*, że „jakkolwiek [dedukcja metafizyczna] ma wielkie znaczenie dla [jego] badania, to nie należy doń w sposób istotny” (xvi–xvii)², a główny cel *Krytyki* – uzasadnienie przedmiotowej ważności poznania naukowego – powiódłby się i bez niej.

¹ Jak zauważa Chemero (2011), aby nowe nurty kognitywistyki nie zasadały swojej tożsamości na prostej negacji podejścia klasycznego, muszą odkryć swoje własne korzenie, co w pewnej mierze zmusza do przemyślenia i napisania na nowo historii filozofii nowożytnej. Podobnie motywowanym i wartym przywołania przedsięwzięciem jest próba zrekonstruowania genealogii pojęcia nawyku (eksplanansu zdobywającego popularność wśród przedstawicieli poznania ucieleśnionego) w historii filozofii zachodniej i wschodniej, podjęta przez Barandiarana i Di Paolo (2014).

² Teksty Kanta cytuję podając dodatkowo datę wydania oryginalnego. Wyjątkiem jest *Krytyka czystego rozumu* (Kant, 1781/1986), którą cytuję z samym podaniem wydania (A, B) i numerem strony w tym wydaniu.

Transcendentalna filozofia umysłu bada więc umysł drogą okrężną: z pewnych niepowątpiewalnych, ale niekoniecznie usystematyzowanych faktów o Ja empirycznym (np. że uprawia naukę i para się spekulacjami metafizycznymi) wywodzi podstawową strukturę konieczną dla ich możliwości – Ja transcendentalne, cechujące się jednością i spontanicznością – a następnie bada, jakie inne własności będzie manifestowała ta struktura w każdym możliwym skończonym umyśle, a więc w szczególnym przypadku także w ludzkim, empirycznym Ja. Taki schemat rozumowania nie zapewnia, że transcendentalna struktura głęboka Ja empirycznego wyczerpuje możliwą wiedzę o nim; przeciwnie, różne Ja empiryczne mogą się mienić całą paletą psychologicznej różnorodności, która jednak leży poza zakresem zainteresowania transcendentalnej filozofii umysłu. Jakie są więc transcendentalne niezmienniki skończonych umysłów? To, co w różnych miejscach Kant mówi na ten temat, daje się właściwie sprowadzić do dwóch pojęć: syntezy i apercepcji transcendentalnej.

Spory o to, od czego wychodzą, co chcą udowodnić i czy rzeczywiście im się udaje oraz jaką strukturę mają dedukcje transcendentalne obu wydań, trwają od przeszło dwustu lat i zajmują dziś całe biblioteki. Spróbuję zarysować teraz w dużym uproszczeniu możliwie niekontrowersyjną wykładnię transcendentalnej architektury umysłu według dedukcji kategorii z wydania B jako punkt wyjścia, a dopiero później odnieść się do wybranych problemów interpretacyjnych. W doświadczeniu, powiada Kant, świat ukazuje się nam jako powiązana ze sobą całość i jest tak w sposób konieczny (a nie tylko empirycznie przypadkowy), a ponieważ zasady tej konieczności nie znajdziemy w samym doświadczeniu, to *musi* ona być transcendentalnym warunkiem doświadczenia. Tak dochodzi Kant do pojęcia syntezy transcendentalnej, spontanicznej czynności intelektu, która wprowadza jedność w różnorodność zmysłowości. Wprowadzanie jedności w różnorodność zakłada jednak jakiś zewnętrzny punkt odniesienia, który jest wzorcem jedności, co prowadzi do kolejnego transcendentalnego inwariantu: czystego Ja. „Tylko dlatego, że mogę objąć przedstawienia jedną świadomością, wszystkie one są moimi przedstawieniami; inaczej miałbym jaźń tak różnobarwną, jak wiele jest przedstawień, które posiadam” (B 134). Tym, co jednoczy różnorodność przedstawień, jest więc to, że wszystkie są dostępne dla świadomości Ja, co wyraża się w sloganowym już *passusie*, że „«Myślę» musi móc towarzyszyć wszystkim moim przedstawieniom” (B 131–132). Tak na arenie pojawia się moment, w którym Ja transcendentalne samo przygląda się swojemu myśleniu (syntezowaniu): ten akt nakierowania na swoją aktywność nazywa Kant właśnie apercepcją transcendentalną. Choć słowo „apercepcja” dziedziczy chwiejność znaczeniową leibnizjańskiego modelu świadomości, to widać, że Kant traktuje ją jako formę samowiedzy, która w pewien niebezpośredni sposób zawarta jest w każdej wiedzy przedmiotowej. Apercepcja transcendentálna jest przy tym

starannie odróżniona od empirycznej introspekcji (zmysłu wewnętrznego); ta ostatnia sięga co najwyżej Ja empirycznego, nie ma więc charakteru wiedzy koniecznej ani konstytutywnego charakteru³. To właśnie jedność pierwotnego aktu apercpepcji ustanawia apercpepcyjną jedność świadomości, która umożliwia jedność doświadczenia.

Mam nadzieję, że zarysowany przeze mnie szkic pozwala dostrzec pewne napięcie między pojęciem syntezy i apercpepcji transcendentalnej. Relacja pomiędzy nimi nie jest jasna: z jednej strony uznanie apercpepcji transcendentalnej za konieczny warunek jedności syntezy wydaje się kluczowym krokiem dedukcji transcendentalnej, z drugiej jednak są w *Krytyce* ustępy sugerujące odwrotną zależność:

(...) stała tożsamość apercpepcji pewnej różnorodności danej naocznie zawiera syntezę przedstawień i jest możliwa tylko dzięki świadomości tej syntezy (B 133).

Apercpepcję transcendentalną można więc rozumieć dwójako: funkcjonalnie, jako kompetencję Ja transcendentalnego do myślenia o swoich przedstawieniach, albo procesualnie, jako korelat procesu syntezy; w pierwszym przypadku jako pierwotną obiera się apercpepcję jako akt, z którego biorą się wszystkie inne akty świadomości, w drugim zaś pierwotna jest synteza, której samoświadomienie umożliwia apercpepcję transcendentalną. Co istotne, wydaje się, że obie te interpretacje – funkcjonalna i procesualna – znajdują solidne podstawy w tekście *Krytyki*, choć trudno je ze sobą pogodzić. Niejednoznaczność pojęcia apercpepcji jest jednak tylko przejawem bardziej fundamentalnego napięcia w strukturze pierwszej *Krytyki*, która wznosząc się na pojęciu Ja transcendentalnego, przypisuje mu sprzeczne zestawy cech. Spróbuję sformułować je po kantowsku, jako trzy antynomie zawarte w architekturze *Krytyki czystego rozumu*.

Antynomia spontaniczności i receptywności Ja

Teza. Ja transcendentalne jest absolutnie wolne i nieuwarunkowane przyczynowo w syntezowaniu doświadczenia.

Antyteza. Ja transcendentalne jest receptywne ze względu na zmysłowość, tzn. pierwotne wrażenia wpływają na kierunek procesu syntezy doświadczenia.

³ To przeciwstawienie powraca w *Antropologii*, gdzie Kant tłumaczy je przez analogię do przeciwstawienia naoczności i intelektu: „Zmysł wewnętrzny to nie czysta apercpepcja – świadomość tego, co człowiek robi – ona bowiem właściwa jest dla zdolności myślenia. Jest to raczej świadomość doznań w tej mierze, w jakiej człowieka pobudza gra jego własnych myśli” (Kant 1798/2005: 59).

Antynomia samoświadomości

Teza. Ja transcendentalne jest w stanie poznać siebie takim, jakim jest samo w sobie.

Antyteza. Ja transcendentalne może poznać siebie tylko pośrednio, jako zjawisko.

Antynomia zmienności

Teza. Ja transcendentalne jest bezczasowym i niezmiennym, czysto logicznym warunkiem myślenia.

Antyteza. Ja transcendentalne istnieje w czasie i podlega zmianom.

Obszerną analizą pierwszej ze sformułowanych antynomii jest książka Piotra Hoffmana (1982). Czyniąc zadość kantowskiej maksymie, że pojęcia bez zmysłowości są puste, zobowiązujemy się do przyjęcia zmysłowych elementów świadomości, czyli czegoś pierwotnie danego, co podlega dopiero syntezie. Według „Estetyki transcendentальной” takimi przedpojęciowymi elementami świadomości są materia zmysłowa oraz jej formy: czas i przestrzeń. Hoffman pokazuje jednak, że w świetle transcendentальной „Analityki zasad” (zwłaszcza „Analogii doświadczenia”) formy naoczności nie są wolne od założeń pojęciowych (np. czas presuponuje pojęcie następstwa). Jeszcze bardziej komplikuje sprawę „Odparcie idealizmu”, które logicznie uzależnia aktywność Ja transcendentального od istnienia świata zewnętrznego. Co więcej, Ja transcendentalne jest w stanie przekonać się o istnieniu niezależnej od niego rzeczywistości, ale nie może zrobić tego pojęciowo (bo ujęta w kategorii rzeczywistość przestałaby być niezależna). Prowadzi to do konkluzji, że „Odparcie idealizmu” wymaga przypisania Ja transcendentalnemu pewnej receptywności, co popada w konflikt ze sposobem określenia Ja transcendentального w dedukcji transcendentальной. Hoffman nie waha się stwierdzić, że mamy tu do czynienia ze sprzecznością, której nie da się znieść w obrębie samej pierwszej *Krytyki*. Stanie się ona napędem całego idealizmu niemieckiego, by rozwiązać się dopiero w naturalistycznej antropologii Karola Marksa. Dodajmy, że rozwiązanie to polega na odrzuceniu idealizmu transcendentального. Proponowane przeze mnie rozwiązanie tej antynomii będzie miało bardziej optymistyczne konsekwencje: zgadzam się, że zmusza ona do silnej naturalizacji Ja transcendentального, ale myślę, że można to uczynić w zgodzie z duchem filozofii krytycznej i zachowując idealizm transcendentalny w całkiem silnym sensie.

Antynomia samoświadomości wynika z pewnej niespójności między analityką a dialektyką transcendentálną. Z jednej strony w dedukcji kategorii Kant stawia szereg pozytywnych tez na temat Ja transcendentального (np. że cechuje je syntetyczna jedność) i postuluje możliwość jego samoświadomości poprzez apercepcję transcendentálną, a z drugiej w „Dialektyce” przeprowadza niestrudzoną krytykę psychologii racjonalnej (tzn. spekulacji pojęciowej wychodzącej od sądu „Ja myślę”). Krytyka ta rodzi wątpliwości, czy Ja transcendentalne nie jest czymś więcej, niż, być może transcendentalnie koniecznym, ale nie-

realnym empirycznie metafizycznym fantazmatem, takim jak kantowskie idee regulatywne. Jeśli tak, to drugi tom *Krytyki* refutowałby tezy pierwszego.

Wreszcie, antynomia zmienności wynika z podobnego napięcia. Zdawałoby się, że Ja transcendentalne, jako warunek stosowalności kategorii czasu, musi być atemporalne. Co więcej, sama rola Ja transcendentalnego w kantowskiej architekturze umysłu jako inwariantu, cechującego się transcendentalną jednością, w której nie jest dane nic różnorodnego (B 135), wydaje się pojęciowo implikować niezmienność. Sam Kant wyraźnie dopuszczał jednak możliwość jego zmienności, w krytyce trzeciego paralogizmu powołując się na analogię z przenoszeniem energii mechanicznej (notabene tworząc ciekawy przyczynek do dwudziestowiecznych debat nad kryteriami tożsamości osobowej):

Kula sprężysta, która uderza drugą taką samą w prostym kierunku, udziela jej całego ruchu, a więc całego swego stanu (jeśli uwzględnia się miejsce w przestrzeni). Przyjmijmy w analogii do tych ciał substancje, z których jedna nadawałaby drugiej przedstawienia wraz z ich świadomością, to można pomyśleć ich cały szereg, z których pierwsza udzielałaby drugiej swego stanu wraz z jego świadomością, druga trzeciej swój własny stan wraz ze świadomością poprzedniej, a ta znów stany wszystkich poprzedzających wraz ze swym własnym i ich świadomością. Ostatnia substancja byłaby przeto świadoma wszystkich stanów substancji, które przed nią zostały wymienione, jak też i swoich własnych, gdyż tamte byłyby wraz ze świadomością w nią przeniesione, a mimo to nie byłaby ona przecież ściśle tą samą osobą we wszystkich tych stanach (A 363–364, przypis).

Zwracając uwagę na niezależność Ja transcendentalnego od empirycznej tożsamości osobowej, Kant przypisuje mu posiadanie pewnego stanu wewnętrznego, podlegającego zmianom i poddającemu się transferowi.

Antynomie te można traktować jako sformułowanie sporu między antynaturalistycznymi a naturalistycznymi wykładniami idealizmu transcendentalnego. Tradycyjna wykładnia antynaturalistyczna zastrzega, że Ja transcendentalne jest nieprzekraczalną granicą wszelkich wyjaśnień przyczynowych, z tego trywialnego powodu, że jako transcendentalny warunek kategorii przyczynowości samo Ja transcendentalne nie może dysponować mocami przyczynowymi. W pewnym sensie Kant miałby tu antycypować Ryle'owski argument z nieskończonego regresu homunkulusów, dochodząc jednak do zgoła innego wniosku – obalenia naturalizmu eksplanacyjnego, przynajmniej w odniesieniu do umysłu⁴. Antynaturalistyczny interpretator Kanta będzie więc widział Ja

⁴ Taki schemat argumentacji przeciw naturalizacji umysłu jest zaskakująco żywoty także we współczesnej filozofii analitycznej. Przykładowo, w sformułowaniu Putnama: filozoficzne przedsięwzięcie znaturalizowania intencjonalności zakłada pewną treść (pomyślenie o intencjonalności jako znaturalizowanej) pewnego umysłu, więc zakłada *a priori* pewną intencjonalność, więc intencjonalności nie można znaturalizować (Putnam 1998a).

transcendentalne jako nieuwarunkowane przyczynowo, zdolne do samowiedzy i pozaczasowe. Kantysta-naturalista zaś, o ile władza syntezy i aprioryczne składniki umysłu poddają się łatwej naturalistycznej interpretacji jako biologiczne *a priori* (jak u Uexküll'a czy Lorenza; krytycznie na ten temat zob. Pobjewska 1996), napotyka problem przy Ja transcendentalnym. Przekonywające rozwiązanie trzech sformułowanych wcześniej antynomii związanych z Ja transcendentalnym bez posądzania Kanta o jawną sprzeczność w samym sercu *Krytyki* można traktować jako dezyderat dla adekwatnej interpretacji tego pojęcia.

Z pewnym pomysłem wyszedł Andrew Brook (1994, 2004), interpretując pierwszą *Krytykę* przez pryzmat reprezentacyjnej teorii umysłu (*à la* Fodor). Pojawienie się apercepcji w dedukcji transcendentальной wydania A nazywa Brook „przewrotem o 90 stopni” i uważa za moment kluczowy, w którym Kant przenosi ciężar z diachronicznej jedności świadomości (czasowej ciągłości przedmiotów doświadczenia, którą zapewniają synteza ujmowania w naoczności i synteza odtwarzania w wyobraźni) na jedność synchroniczną różnych reprezentacji⁵, której warunkiem możliwości jest transcendentálna jedność apercepcji. Synchroniczną, apercepcyjną jedność świadomości rozumie więc Brook jako jeden akt świadomości nakierowany na wszystkie reprezentacje (Brook 1994: 130), wywodząc stąd pojęcie reprezentacji globalnej: metareprezentacji, której treścią są wszystkie reprezentacje danego umysłu. Ponieważ stosowanie kategorii relacyjnych (którymi posługuje się newtonowska fizyka) wymaga powiązania czasowo odległych reprezentacji, argumentuje Brook, każdy akt

⁵ Uwaga natury filologicznej. Kantowskie *Vorstellung* anglojęzyczni tłumacze Kanta (np. Kemp Smith) tradycyjnie przekładają na *representation*. (Być może to skłoniło Brooka do tak łatwego przypisania Kantowi fodorowskiego wariantu reprezentacyjnej teorii umysłu). Taki przekład może wydawać się uzasadniony, skoro sam Kant wprowadza pojęcie *Vorstellung* wiążąc je z łacińskim *repraesentatio* (A 320/B 376). Nie jest to jednak argument definitywny: Kant nieraz nadawał starym, scholastycznym pojęciom filozoficznym nowe znaczenie, które bardziej odpowiadało jego potrzebom. Należy więc zauważyć, że słowo „reprezentacja” nie chwyta intuicji, że chodzi o postawienie (*stellung*) czegoś przed (*vor*) Ja, a nie odtworzenie. Nie jest to wyłącznie kwestia stylistyczna, ale silne rozstrzygnięcie w kierunku pewnej – reprezentacjonistycznej – linii interpretacji idealizmu transcendentálnego, która nie jest jedyną możliwą interpretacją. O wiele lepszy jest tutaj polski przekład Romana Ingardena – „przedstawienie” – który nie rozstrzyga przedwczesnie statusu desygnatów tego pojęcia. Zdaje sobie z tego sprawę następne pokolenie anglojęzycznych badaczy Kanta: w swoim przekładzie *Krytyki czystego rozumu* Werner Pluhar tłumaczy *Vorstellung* jako *presentation*, zaś Patricia Kitcher w swojej książce (1996), zależnie od kontekstu, jako *cognitive state* lub *contents of a cognitive state*, podkreślając, że nie wszystkie *Vorstellungen* coś reprezentują (A 320/B 376; Kitcher 1996: 66). W tym tekście posługuję się terminologią Ingardena, z jednym wyjątkiem: gdy referuję poglądy Brooka, jego użycia słowa *representation* tłumaczę jako „reprezentacja” – wydaje się tego wymagać życzliwość interpretacyjna.

reprezentowania musi być zapośredniczony przez akt globalnego metareprezentowania (tamże: 133). Ujmując tę myśl w formie sloganu, kantowski umysł nie tyle ma reprezentacje, co jest (jedną, globalną) reprezentacją.

Oprócz roli apercepcji transcendentalnej w dedukcji Brook podkreśla też drugą jej funkcję: jest ona formą czystej samoświadomości, w której Ja transcendentalne ujmuje siebie jako podmiot swoich doświadczeń. Ja transcendentalne jest reprezentacją, w której nie jest dane nic różnorodnego (B 135), żadna treść poza byciem wspólnym podmiotem (*single common subject*). Czysta samoświadomość jest zapośredniczona przez poszczególne reprezentacje, z czego Brook wywodzi, że podmiotem i przedmiotem samoświadomości apercepcyjnej jest umysł jako globalna, samoreprezentująca się reprezentacja (*self-representing representation*).

Brook eksplikuje więc pojęcie apercepcji transcendentalnej w ramach teorii (samo)świadomości jako stanu wyższego rzędu: apercepcja ma się do globalnej reprezentacji tak, jak percepcja do pojedynczych reprezentacji. Oczywiście strategią krytyki Brooka jest powołanie się na obszerną filozoficzną krytykę klasycznej reprezentacyjnej teorii umysłu. Nie jest to atak trudny do przeprowadzenia, skoro Brook, mówiąc o reprezentacji globalnej, akceptuje dość otwarcie istnienie czegoś, co Dennett (1991: rozdz. 5) nazywał kartezjańskim teatrem – tym samym pada łatwym łupem dennettowskiej krytyki tego pojęcia: zarówno pojęciowej, jak i empirycznej. Nie będę jednak formułował tego rodzaju argumentu, zamiast tego zwrócę uwagę na to, że ujęcie Brooka nie ustosunkowuje się adekwatnie do trzech wymienionych antynomii. (I) Koncepcja Ja transcendentalnego jako globalnej (meta)reprezentacji redukuje spontaniczność tego Ja do relacji (meta)reprezentacji między Ja a reprezentacjami pierwszego rzędu. (II) Samoświadomość apercepcyjna Ja transcendentalnego jest według Brooka tylko szczególnym przypadkiem świadomości w ogóle, kiedy wszystkie reprezentacje ujmuje się jako reprezentację globalną. Wreszcie, (III) Ja transcendentalne samoreprezentujące się jako Ja jest pozbawione jakiegokolwiek treści. Brook łączy tę tezę z problematyką semantyki i pragmatyki zaimka „ja” w analitycznej filozofii języka: w pewnych kontekstach „ja” nie poddaje się analizie jako żadne inne wyrażenie (Brook 2004). Tezy (I)–(III) są rozstrzygnięciami w kierunku tez wymienionych antynomii, podczas gdy od dobrej interpretacji pojęcia apercepcji transcendentalnej należy wymagać uczynienia zadość także antytezom poprzez wyjaśnienie genezy samej antynomii.

Apercepcja transcendentalna w ujęciu epigenetycznym

Recepcja Kanta w głównym nurcie tradycji analitycznej jest w pewnym sensie wybiórcza: Kanta widzi się na ogół jako epistemologa i etyka, filozofię transcendentálną zaś jako filozofię Ja transcendentálnego – konstrukt czysto logicznego, niezależnego od faktów empirycznych. Obraz taki ignoruje szerokie pozafilozoficzne zainteresowania Kanta, dotyczące antropologii, pedagogiki, geografii fizycznej i historii naturalnej, które – jak łatwo się przekonać – notorycznie przenikały do systemu filozofii krytycznej. Jennifer Mensch (2013) nie waha się nazwać kantowskiego stylu myślenia „organicyzmem”: architektura czystego rozumu jest modelowana na wzór żywego organizmu, pojętego na modłę trzeciej *Krytyki* jako celowo samoorganizująca się i rozwijająca całość (Kant 1790/1986: § 65). Takie rozwojowe ujęcie prezentuje Kant też w „Architektonice czystego rozumu”:

Zdaje się, że systemy – jak robaki – zostały wytworzone przez pewnego rodzaju *generatio aequivoca* z samego tylko spływu nagromadzonych pojęć. Początkowo zniekształcone [stają się] z biegiem czasu zupełne, choć wszystkie w ogóle miały swój schemat za pierwotną zaródź w jedynie rozwijającym się rozumie (A 835/B 863).

Podobnych embriologicznych porównań używa Kant w „Dziejach czystego rozumu” i wiele wskazuje na to, że ich znaczenie zdecydowanie wykracza poza manierę stylistyczną. Jeśli spojrzeć na *Krytykę* w kontekście przedkrytycznych zainteresowań Kanta, jawi się ona jako próba podania wyjaśnienia naturalnej genezy rozumu, tak jak wcześniej Kant wyjaśniał genezę układów planetarnych, zjawisk klimatycznych, życia czy różnorodności fenotypowej wśród ludzi. Przykładowo, zarys wyjaśnienia „naturalnej historii rozumu” w cytowanym wyżej fragmencie ma podobną strukturę, co koncepcja z rozprawy *O różnorodnych rasach ludzkich* (1775/2010a): zakładamy istnienie pewnych zarodków odpowiedzialnych za obserwowalne cechy i podlegających dziedziczeniu, a następnie obserwujemy ich rozwój pod wpływem czynników środowiskowych.

Więcej jasności można chyba zyskać przywołując spór, na który wielokrotnie powoływał się Kant w pytaniach o genezę – spór preformacjonistów i epigenetyków w ówczesnej biologii. Dotyczył on procesu rozwoju zarodkowego istot żywych. Według dominującego do XVIII wieku paradygmatu preformacjonistycznego (bronionego na gruncie metafizycznym przez Leibniza i Malebranche’a, a eksperymentalnie przez Albrechta von Hallera) w momencie poczęcia embrión dziedziczy preegzystującą w gametach (określanych jako – *nomen omen* – homunkulusy) formę, która ściśle determinuje cały proces rozwoju, aż do śmierci; zmienność ontogenetyczna ma zaś tylko charakter

ilościowy. Jednym z ważniejszych krytyków tej koncepcji był Caspar Friedrich Wolff, który zarzucał jej, że raczej neguje i zastępuje boskim zrzędzeniem powstanie i rozwój organizmu żywego, niż je wyjaśnia. W zamian zaproponował ulepszoną wersję starego pomysłu Arystotelesa: spontanicznego różnicowania się komórek pod wpływem czynników środowiskowych oraz wewnętrznej siły życiowej (*vis essentialis*). Choć tak sformułowana koncepcja epigenetyki nie była wolna od witalizmu, zapoczątkowała nowoczesną embriologię i fizjologię⁶. (Więcej na temat teorii epigenetyki i jej związków z pierwszą i trzecią *Krytyką* znaleźć można w tomie: Huneman 2007).

Kiedy więc Kant w „Dedukcji” drugiego wydania nazywa swoją koncepcję „systemem epigenetyki czystego rozumu” (B 166–167), wydaje się właśnie oponować przeciwko ujęciu Ja transcendentalnego jako egzystującego przed doświadczeniem homunkulusa; stanowi ono raczej „zarodek” właśnie i warunek konieczny podmiotowości niż autonomiczny podmiot. Idąc dalej tym tropem: celem apercpcji transcendentalnej jest regulowanie rozwoju czystego rozumu pod wpływem doświadczenia, tzn. teleologicznie zorientowane rozwijanie i różnicowanie apriorycznego uposażenia; stanowi ona korelat procesu syntezy zwrotnie sprzęgający materię zmysłową z prawidłami, wedle których materia jest syntezowana⁷. Wydaje się bowiem, że tylko obustronne sprzężenie zmysłowości z Ja transcendentalnym pozwala uznać czysty rozum za organiczną całość.

Epigenetyczne ujęcie apercpcji transcendentalnej pozwala zgodnie z krytyczną metodologią znieść trzy wymienione antynomie. Ja transcendentalnemu przysługuje empiryczna realność, a więc zgodnie z antytezami, związana z nim czasowość, determinacja przyczynowa i charakter fenomenalny; Ja transcendentalne jest kolejnym przedmiotem przyrody, który postrzegamy zgodnie z pewnym apriorycznym uposażeniem. Z drugiej strony, w ramach samoświadomości apercpcyjnej, Ja z konieczności (pojęciowej) postrzega siebie jako absolutnie samorzutne, niezmiennie i epistemicznie dostępne – wynika to z faktu, że akt syntezy zakłada właśnie taki charakter Ja transcendentalnego, a samoświadomość apercpcyjna jest zapośredniczona przez (uświadomione) akty syntezy. W tym sensie Ja transcendentalnemu przysługują te atrybuty z tej samej pojęciowej konieczności, z jakiej wzorzec metra pod kloszem

⁶ W dużej mierze debata epigenetyka-preformacja pozostaje aktualna do dziś, nabrawszy rozpędu w kontekście syntezy biologii ewolucyjnej i genetyki. Co więcej, wydaje się, że po okresie genocentryzmu, przynajmniej od lat 90. epigenetyczny sposób myślenia o rozwoju znów wysuwa się na prowadzenie: świadczy o tym rosnąca popularność takich dziedzin jak ewolucyjna biologia rozwojowa czy wzrost zainteresowania neuroplastycznością i mechanizmami regulacji ekspresji genów.

⁷ Dlatego racji nie ma Fodor (2011), widząc w Kancie antycypatora swojej natywistycznej teorii pojęć.

w Sèvres mierzy dokładnie jeden metr – ponieważ każda próba zmierzenia jego długości zakłada właśnie taką długość.

Jeśli powyższa rekonstrukcja jest adekwatna, to cel naturalizacji apercepcji transcendentальной – ostatniego bastionu antynaturalistycznych interpretatorów Kanta – można uznać za zakończony sukcesem. Na marginesie warto dodać, że tak rozumiane pojęcie apercepcji wydaje się nie tylko zgodne z naturalistyczną ontologią, ale ma szczególną wartość eksplanacyjną także w kontekście współczesnych wyjaśnień procesów poznawczych. Dość powiedzieć, że pojęcie to *explicitie* wykorzystują kognitywiści w modelach poznania usytuowanego.

Potencjalności interaktywne organizmu nie pozostają stałe. Zmieniają się w sposób ciągły zgodnie z aktywnością organizmu, jak również wraz z procesami i aktywnościami zachodzącymi w środowisku. Sieć wiedzy sytuacyjnej wymaga więc ciągłej aktualizacji i utrzymywania. Procesy aktualizacji i utrzymywania wiedzy sytuacyjnej określam mianem *apercepcji* (Bickhard 2014: 143).

Wydaje się, że takie ujęcie jest w całości zgodne z ujęciem epigenetycznym, które wyżej imputuję samemu Kantowi: apercepcja transcendentálna jest kanałem umożliwiającym sprzężenie zwrotne (np. w modelu sieci neuronowej można ją interpretować jako połączenia modyfikujące wagi innych połączeń) w celu (teleologicznie rozumianej) optymalizacji poznawczej organizmu. Nasuwa to śmielszą hipotezę: że w istocie cała współczesna kognitywistyka presuponuje w dużej mierze kantowską koncepcję architektury czystego rozumu i próbuje ją wyjaśniać na modłę transcendentálną.

Wyjaśnienia transcendentálne w kognitywistyce

Kognitywistyka dzieli z transcendentálną filozofią umysłu więcej niż tylko podobny zestaw problemów i notoryczne niedookreślenie fundamentalnych pojęć. Istotniejszym od podobieństwa przedmiotu badań jest podobieństwo metodologii. Standardowym modelem wyjaśnienia w kognitywistyce według tradycji funkcjonalizmu jest postulowanie utajonej struktury (mechanizmu) realizującego procesy (funkcje), których skutki układają się w pewien obserwowalny wzorzec. Taką strukturą może być częściowo wrodzona kompetencja językowa, wyjaśniająca praktyki językowe i złożoność ich wytworów; może nią też być pewien zestaw zbieżności sensomotorycznych, które nakładając pewne ograniczenia na sposoby interakcji ze środowiskiem, wyjaśniają pewne cechy doświadczenia. Bardziej formalnie, w przeciwieństwie do dedukcyjnego charakteru przypisywanego tradycyjnie (także przez Kanta) wyjaśnieniom fizycznym, wyjaśnienie funkcjonalne to rozumowanie abdukcyjne: wychodzące

od zaobserwowanych wzorców zachowania systemu poznawczego i szukające zbioru hipotez, który stanowi warunek wystarczający do zaistnienia tego wzorca zachowania.

Choć pojęcie rozumowania abdukcyjnego wprowadził do logiki filozoficznej Peirce, jego rodowód sięga nieco szerszego kantowskiego pojęcia dedukcji transcendentalnej, którą wypada może nazwać szerzej: argumentacją transcendentálną. Z jedną istotną różnicą (której zdaje się nie zauważać Brooks 2004), dotyczącą modalności: argument transcendentálny mówi o warunkach koniecznych pewnego wzorca zachowania, podczas gdy wyjaśnienie funkcjonalne jako rozumowanie do najlepszego wyjaśnienia mówi o warunkach wystarczających (a dokładniej: o minimalnym lub w jakimś kontekstowo zależnym sensie najlepszym podzbiornie wszystkich warunków wystarczających).

W owej różnicy modalności tkwi też potencjalna metodologiczna przewaga wyjaśnienia o strukturze argumentu transcendentálnego nad pokrewnym mu wyjaśnieniem funkcjonalnym: w obrębie zainteresowania kognitywistyki leżą raczej rzeczywiste mechanizmy realizujące wzorce behawioralne, a nie opis pewnego mechanizmu wystarczającego, który przecież może się okazać denotacyjnie pusty. W odróżnieniu od (najlepszego) warunku dostatecznego, dopiero warunek konieczny jako *explanandum* spełnia postulat rzeczywistości, ponieważ konieczność jakiegoś mechanizmu trywialnie implikuje jego rzeczywistość.

Drugim problemem wyjaśnienia funkcjonalnego, postulującego strukturę będącą najlepszym warunkiem dostatecznym, jest tendencja do ignorowania niskopoziomowych, implementacyjnych szczegółów postulowanej struktury. Banałem jest stwierdzenie, że czysto formalny opis funkcji nie wystarcza do jej fizycznej realizacji, tak jak nie można się najeść przepisem na tartę. Funkcjonalista tradycyjnie broni się mówiąc, że istnienie *jakiegoś* mechanizmu implementacji jest *implicite* postulowane przez opis funkcjonalny, konkretny zaś kształt tego mechanizmu jest eksplanacyjny nieistotny – wyraża się to w tezie o wielorakiej realizacji (Bechtel, Mundale 1999). Problem w tym, że teza ta okazała się empirycznie fałszywa: zwolennicy poznania ucieleśnionego w kognitywistyce mają solidne empiryczne racje, by argumentować, że takie niskopoziomowe informacje na temat mechanizmu realizującego jakąś funkcję są istotnym składnikiem jej wyjaśnienia (zob. np. badania Thelen i in. 2001 nad wpływem rozwoju motorycznego na rozwój poznawczy niemowląt). Z drugiej strony, wymóg każdorazowego opisu mechanizmów procesów poznawczych na poziomie molekularnym jest równie eksplanacyjnie jałowy: w praktyce badawczej wyjaśnienie osiąga kres na pewnym (zależnym od dziedziny) poziomie podstawowym (Machamer i in. 2011). Nowa kognitywistyka potrzebuje bardziej elastycznej metodologii niż klasyczne wyjaśnianie funkcjonalne – wymagającej wyjaśnień dostatecznie ogólnych, a zarazem uwzględniających

niskopoziomowe szczegóły implementacyjne. Wyjaśnieniem spełniającym te dezyderaty okazuje się właśnie wyjaśnienie transcendentalne: konieczny charakter postulowanego *explanandum* nadaje mu neutralny względem substratu charakter, a zarazem pozwala na dookreślenie go szczegółami architektury bądź anatomii systemu poznawczego⁸.

Można zapytać, czy ta metodologiczna przewaga nie jest okupiona brakiem praktycznej efektywności i czy nie spychałaby kognitywistyki z domeny badań empirycznych w gąszcz metafizycznej spekulacji. Ostatecznie stwierdzenie konieczności (bez określenia jej siły modalnej) to teza niezwykle silna, wedle niektórych niepoddająca się empirycznemu uzasadnieniu. Wydaje się jednak, że argumentacja transcendentalna jest w pełni zgodna z naturalistycznym wyjaśnieniem naukowym: postulat naturalizmu eksplanacyjnego, uzasadniony niezależnie (najlepiej też transcendentalnie), nakłada po prostu pewne dodatkowe ograniczenia na przestrzeń dopuszczalnych eksplanandów i dookreśla wymaganą konieczność jako konieczność nomologiczną (tzn. bycie postulowanym przez najlepsze istniejące teorie naukowe)⁹. Co więcej, metodologia wyjaśnień transcendentálnych, wyjaśniająca X poprzez podanie warunków nomologicznie koniecznych do zaistnienia X, wydaje się dobrze przystawać do szeregu programów badawczych spod znaku matematycznego modelowania systemów złożonych (w tym systemów poznawczych), np. badań nad sztucznym życiem czy rozproszonych, wieloagentowych modeli poznania społecznego.

Przykładowo, teoria autopojezy Vareli i Maturany w badaniach nad sztucznym życiem (jawnie inspirowana kantowsko, choć bardziej trzecią *Krytyką*) postuluje pewien model minimalnego organizmu żywego – systemu autopojetycznego. System autopojetyczny jest zdefiniowany niezależnie od jego

⁸ Na marginesie warto dodać, że teza o ucieleśnieniu poznania jest nie tylko kompatybilna, ale też preferowana przez wyjaśnienia transcendentálne. W przedkrytycznym traktacie przyrodoznawczym *Powszechna historia naturalna i teoria nieba albo szkic o układzie oraz mechanicznym pochodzeniu całości świata opracowany zgodnie z prawami Newtona* (Kant 1755/2010) Kant spekulował na temat kompetencji poznawczych hipotetycznych istot zamieszkujących inne światy. W ramach teorii Kanta inne planety różnią się od Ziemi gęstością materii, co przekłada się na konstytucję umysłów i kompetencję poznawczą zamieszkujących je istot (na temat związków myśli Kanta z koncepcją poznania ucieleśnionego zob. też. Carpenter 2008). Wcześniejsze księgi tego samego traktatu stanowią też poparcie tezy o kompatybilności wyjaśnień transcendentálnych i naturalistycznych. Kant oferuje tu wyjaśnienie formowania się planet Układu Słonecznego, postulując empiryczny warunek konieczny: istnienie pewnego obłoku materii, i korzysta z praw mechaniki newtonowskiej (najlepszej dostępnej teorii naukowej jego czasów), by pokazać, jak wirująca chmura materii formuje się w pierścienie oraz zagęszcza się na skutek ochładzania. Po rozwinięciu i bardziej rygorystycznym sformułowaniu ilościowym przez Laplace'a, tzw. hipoteza mgławicowa Kanta-Laplace'a stała się teorią empiryczną i do dziś pozostaje najbardziej wiarygodnym modelem powstania i ewolucji Układu Słonecznego.

⁹ Zgodności filozofii transcendentálnej z naturalizmem na przykładzie transcendentálnej filozofii społecznej broni też Miłkowski (2007).

biochemicznych mechanizmów, wyznacza raczej konieczną organizację każdej możliwej istoty żywej: istnienie pewnej półprzepuszczalnej bariery, która wyznacza topologię systemu autopojetycznego (w ziemskim drzewie życia rolę tę pełni błona komórkowa) oraz pewnych mechanizmów metabolicznych wewnątrz niej, które podtrzymują system autopojetyczny przy życiu. Ponieważ teza, że system autopojetyczny jest minimalnym organizmem żywym, obejmuje zasięgiem także możliwe organizmy w ekosystemach innych niż ziemski, jej uzasadnienie musi mieć strukturę argumentu transcendentального: każdy organizm żywy przejawia pewne własności życiopodobne, a na mocy praw fizyki własności te może przejawiać tylko system autopojetyczny (bada się to np. implementując obliczeniowy model systemu autopojetycznego i manipulując jego zachowaniem *in silico*), a zatem każdy możliwy organizm żywy jest systemem autopojetycznym. (Uzasadnienie tezy odwrotnej, że każdy system autopojetyczny jest organizmem żywym, jest bardziej kontrowersyjne i przekracza ramy tego tekstu; zob. np. Thompson 2007: 122–127).

Bardziej poznawczym przykładem są wieloagentowe symulacje nabywania kompetencji do kategoryzacji pojęciowej i jej konwergencji w normy o zasięgu populacyjnym, opisane przez Steelsa i Belpaemego (2005) na przykładzie kolorów podstawowych. Badania te pokazują, że uwspólnienie kategorii pojęciowych pomiędzy (abstrakcyjnymi) agentami jest możliwe tylko, gdy są one częściowo wrodzone i podatne na biologiczną ewolucję na drodze doboru naturalnego lub gdy agenty mogą się ze sobą komunikować i istnieje kanał transmisji kulturowej. Znow więc: dzięki modelowi obliczeniowemu dochodzimy do zestawu (abstrakcyjnych) warunków (tym razem połączonych spójnikiem alternatywy, a nie koniunkcji), jakie musi spełniać minimalna społeczność zdolna do posiadania wspólnych kategorii pojęciowych. Czyż nie rzuca to światła na klasyczny problem Hegla: genezę intersubiektywności?

W jaki sposób uzasadnianiu podlega dedukcja warunków koniecznych pewnego zjawiska? Pojęcie transcendentálních warunków koniecznych dla funkcjonowania pewnego systemu poddaje się naturalizacji jako opis minimalnego egzemplarza tego systemu, przejawiającego charakterystyczną dla niego paletę zachowań. W trywialnym przypadku dedukcja warunków koniecznych jakiegoś systemu to po prostu enumeracja zobowiązań ontologicznych naszej najlepszej teorii tego systemu. Pojęcie systemu minimalnego nabiera jednak eksplanacyjnego znaczenia dopiero, gdy potrafimy je ograniczyć jeszcze bardziej, na przykład pokazując, jak pewne jego własności są w pewnych okolicznościach generowane przez pozostałe własności. Określenie warunku koniecznego jest tylko początkiem rozumowania, które ma na celu pokazanie, jak taki minimalny system potrafi przyjąć bardziej złożone postaci. Metody obliczeniowe, takie jak równoległe symulowanie ewolucji systemów danej

klasy, pozwalają względnie dokładnie pokryć przestrzeń ich możliwych architektur (*modulo* aktualny stan wiedzy, wyrażony w sposobie reprezentacji dziedziny problemowej). Względnie efektywne przeszukanie takiej przestrzeni za pomocą algorytmu genetycznego wydaje się pozwalać na wygłaszanie sądów istotnie mocniejszych (pod względem siły modalnej) od stwierdzenia funkcjonalistycznych warunków wystarczających. Postulowana przez te ostatnie ukryta struktura może być interpretowana jako lokalne maksimum krajobrazu dostosowania dla systemów danej klasy; warunek konieczny stanowi zaś maksimum globalne. Ponieważ w dziedzinie obliczeń ewolucyjnych istnieją skuteczne metody radzenia sobie z problemem utknięcia ewolucji systemu na lokalnym minimum, tezy o warunkach koniecznych można uznać za (transcendentalnie) możliwe. Z drugiej strony, dedukcja warunków (nomologicznie) koniecznych pozostaje tezą empiryczną popartą zawodnym wnioskowaniem, odwołującym się do aktualnej wiedzy — stanowi to trywialną konsekwencję naturalizacji kantowskiego programu.

Warto dodać, że metodologia wyjaśnień transcendentálnych wydaje się szczególnie użyteczna w konfrontacji ze zjawiskami cechującymi się istnieniem samoorganizacji. Tak można odpowiedzieć na nasuwający się zarzut, że wyjaśnienia transcendentálne są notorycznie niepełne, bo określają tylko zestaw warunków koniecznych, pomijając opis, jak *explanans* dzięki tym warunkom osiąga dojrzałą i złożoną formę, od której wyjaśnienie wychodzi. Przynajmniej w przypadku systemów samoorganizujących się ten opis jest już *implicite* zawarty w wiedzy tła lub zestawie warunków koniecznych. W obu wymienionych przykładach takim fundamentalnym warunkiem koniecznym jest ewoluowalność lub plastyczność (zdolność uczenia się), które wraz ze stosunkowo dobrze opisanymi zjawiskami ewolucji na drodze doboru naturalnego i nadzorowanego uczenia się maszynowego pozwalają przewidzieć obszerną klasę zjawisk. Nie są to tylko czcze deklaracje: spontaniczne podporządkowanie systemu pewnym wzorcom daje się wyjaśnić i opisać w języku teorii układów dynamicznych jako dążenie systemu do osiągnięcia stanu (lub orbity złożonej ze stanów), który w jego przestrzeni fazowej jest atraktorem, czyli stabilnym punktem stałym funkcji przejścia.

Wreszcie, wyjaśnienia transcendentálne są bardziej inkubatorem programów badawczych (bardzo płodnych, jeśli spojrzeć na rosnącą literaturę z zakresu nauki o systemach złożonych, zob. np. Mitchell 2009) niż wyjaśnieniami roszcującym sobie prawo do kompletności. I tak na przykład kantowskie wyjaśnienie transcendentálne (w ramach modelu epigenezy) władzy apercepcji transcendentálnej oferuje koncepcję minimalnego systemu poznawczego oraz stanowi załączek programu badawczego. Co więcej, owa koncepcja minimalnego systemu poznawczego pozostaje nadal empirycznie wiarygodna, a wychodzący zeń program badawczy w dużej mierze przypomina to, co zwykliśmy

nazywać „kognitywistyką”. Kantowski minimalny system poznawczy można scharakteryzować poprzez trzy warunki konieczne:

- (I) posiada jakiś rodzaj naoczności (wejścia sensoryczne, umożliwiające rejestrację bodźców);
- (II) posiada władzę wyobraźni, czyli jest zdolny do syntezy (percepcji, kategoryzowania i porównywania rejestrowanych bodźców);
- (III) posiada apercpcję transcendentalną (umie modyfikować swoje mechanizmy syntezy bodźców pod wpływem nowych bodźców).

Choć kognitywistyka nie ma jednej definicji normalnej poznania, i prawdopodobnie takiej nie potrzebuje, to warunki (I)–(III) wydają się zarazem nietrywialne i uznawane za konsens¹⁰. Należy więc przyznać Brookowi rację, że kognitywistyka jest przesiąknięta duchem filozofii krytycznej – ale w sposób zgoła inny od tego, jaki miał na myśli Brook: zamiast funkcjonalizmu, modularnego przeciwstawienia percepcji i myślenia oraz wrodzonych składników poznania, mamy tu hipotezę systemu poznawczego jako teleologicznie rozwijającego i samoorganizującego się systemu, który poznaje otaczające go środowisko poprzez pewne aprioryczne uposażenie, które samo podlega kształtowaniu przez środowisko.

Bibliografia

- Barandiaran X., Di Paolo E. (2014), *A genealogical map of the concept of habit*, „Frontiers in Human Neuroscience”, 8.
- Bechtel W., Mundale J. (1999), *Multiple Realizability Revisited: Linking Cognitive and Neural States*, „Philosophy of Science”, 2 (66).
- Bickhard M.H. (2014), *Model interaktywistyczny*, przeł. M. Miłkowski, „Przegląd Filozoficzno-Literacki”, 2 (39).
- Brook A. (1994), *Kant and the Mind*, Cambridge and New York: Cambridge University Press.
- Brook A. (2004), *Kant's View of the Mind and Consciousness of Self*, w: *Stanford Encyclopedia of Philosophy*, red. E. Zalta.
- Carpenter A. (2008), *Kant on the embodied cognition*, „Philosophy” 36 (1), s. 59–68.

¹⁰ Warunek (II) może być co prawda kwestionowany przez nonkonceptualistów w przypadku prostych zwierząt, wydaje mi się jednak, że synteza transcendentalna zakłada jedynie bardzo minimalistyczną koncepcję pojęć, np. taką jak u Noëgo (2006), dzięki czemu można mówić o syntezie u niemal dowolnych organizmów zdolnych do uczenia się poprzez wzmocnienie. To, czy jest tak w istocie, jest przedmiotem długiej debaty w filozofii kognitywistyki i filozofii umysłu, zapoczątkowanej przez Sellarsa (1991), której genealogia ewidentnie sięga Kanta. Niestety nie mam tu miejsca na jej przedstawienie i uzasadnienie mojej tezy.

- Chemero A. (2011), *Radical Embodied Cognitive Science*, MIT Press.
- Dennett D. (1991), *Consciousness Explained*, New York: Penguin.
- Fodor J. (2011), *Język myśli LOT 2*, przeł. W. Hensel, Warszawa: PWN.
- Hoffman P. (1982), *The Anatomy of Idealism: Passivity and Activity in Kant, Hegel and Marx*, Hague: Martinus Nijhoff.
- Huneman P. (2007) (red.), *Understanding purpose? Kant and the philosophy of biology*, Rochester: University of Rochester Press.
- Jackendoff R. (1999), *Czym jest pojęcie, że człowiek może je uchwycić?*, przeł. M. Kowalczyk, w: Z. Chlewiński (red.), *Modele umysłu*, Warszawa: PWN.
- Kant I. (1755/2010), *Powszechna historia naturalna i teoria nieba, albo szkic o układzie oraz mechanicznym pochodzeniu całości świata opracowany zgodnie z prawami Newtona*, przekład pod red. M. Żelaznego, w: I. Kant, *Dzieła zebrane*, t. 1: *Pisma przedkrytyczne*, Toruń: Wydawnictwo UMK.
- Kant I. (1775/2010a), *O różnorodnych rasach ludzkich*, przeł. T. Kupś, w: I. Kant, *Dzieła zebrane*, t. 1: *Pisma przedkrytyczne*, Toruń: Wydawnictwo UMK.
- Kant I. (1781/1986), *Krytyka czystego rozumu*, przeł. R. Ingarden, Warszawa: PWN.
- Kant I. (1790/1986), *Krytyka władzy sądenia*, przeł. J. Gałeczki, Warszawa: PWN.
- Kant I. (1798/2005), *Antropologia w ujęciu pragmatycznym*, przeł. E. Drzazgowska, P. Sosnowska, Warszawa: IFiS PAN.
- Kitcher P. (1990), *Kant's Transcendental Psychology*, New York: Oxford University Press.
- Machamer P., Darden L., Craver C.F. (2011), *Myślenie w kategoriach mechanizmów*, przeł. W. Hensel, „Przegląd Filozoficzno-Literacki”, 2–3 (31).
- Mensch J. (2013), *Kant's Organicism. Epigenesis and the Development of Critical Philosophy*, Chicago: University of Chicago Press.
- Miłkowski M. (2007), *Siemek, dialektyka, rzeczywistość. O transcendentnej filozofii społecznej*, „Przegląd Filozoficzno-Literacki” 1 (16).
- Mitchell M. (2007), *Complexity. A guided tour*, Oxford: Oxford University Press.
- Noë A. (2006), *Action in perception*, MIT Press.
- Pobojewska A. (1996), *Biologia i poznanie: biologiczne „a priori” człowieka a realizm teoriopoznawczy*, Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Putnam H. (1998), *Wiele twarzy realizmu*, w: tenże, *Wiele twarzy realizmu i inne eseje*, przeł. A. Grobler, Warszawa: PWN.
- Putnam H. (1998a), *Dlaczego rozumu nie można znaturalizować*, w: tenże, *Wiele twarzy realizmu i inne eseje*, przeł. A. Grobler, Warszawa: PWN.
- Sellars W. (1991), *Empiryzm a filozofia umysłu*, przeł. J. Gryz, w: *Empiryzm współczesny*, red. B. Stanosz, Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Steels L., Belpaeme T. (2005), *Coordinating Perceptually Grounded Categories through Language. A Case Study for Colour*, „Behavioral and Brain Sciences”, 28.

- Thelen E., Schöner G., Scheier Ch., Smith L. (2002), *The dynamics of embodiment: A field theory of infant preservative reaching*, „Behavioral and Brain Sciences”, 24.
- Thompson E. (2007), *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Harvard: Harvard University Press.
- Uexküll J. (2011), *Biologia lub fizjologia*, przeł. A. Pobjewska, M. Pótróla, „Przegląd Filozoficzno-Literacki”, 2–3 (31).
- Weber A., Varela F. (2002), *Life after Kant: Natural purposes and the autopoietic foundations of biological individuality*, „Phenomenology and the Cognitive Sciences”, 1, s. 97–125.

Streszczenie

Immanuela Kanta nazywa się czasem „pradziadkiem kognitywistyki”, co wydaje się adekwatne, o ile przyjąć naturalistyczną wykładnię transcendentalnej filozofii umysłu. Problem dla tego podejścia stanowią pewne cechy apercpepcji transcendentalnej i Ja transcendentalnego trudne do pogodzenia z naturalizmem. W istocie jednak zarówno interpretacje antynaturalistyczne, jak i pewne naturalistyczne, nie pozwalają na eksplikację pojęcia apercpepcji transcendentalnej w zgodzie z rdzeniem *Krytyki czystego rozumu*. W pracy argumentuję, że te „antynomie Ja transcendentalnego” można znieść w ramach silnie naturalistycznej *epigenetycznej* wykładni interpretacji transcendentalnej, nawiązującej do trzeciej *Krytyki* i przedkrytycznych pism Kanta. Rysujący się tak nowy obraz kontrastuje z tradycyjnymi ujęciami kantowskiej filozofii umysłu jako funkcjonalizmu i reprezentacjonizmu (lansowanym przez A. Brooka), ale jest zaskakująco spójny z nowymi i bardziej adekwatnymi empirycznie pomysłami ucieleśnionego nurtu kognitywistyki, ujmującego umysł jako ustanawiany przez złożone interakcje ze środowiskiem.