

POLSKIE TOWARZYSTWO SEMIOTYCZNE

STUDIA SEMIOTYCZNE

Tom XXXI • Numer 1 • 2017



studiasemiotyczne.pts.edu.pl

„Studia Semiotyczne” powstały w 1970 r. z inicjatywy prof. Jerzego Pelca i były przez niego wydawane do 2015 r. jako czasopismo nieperiodyczne. Od 2016 r. czasopismo ukazuje się równolegle w druku i w Internecie jako półrocznik. W „Studiach Semiotycznych” są publikowane artykuły z pogranicza filozofii i semiotyki, w szczególności z zakresu: analitycznej filozofii języka, ogólnej teorii znaku, zastosowań metod semantycznych w filozofii, filozoficznych aspektów lingwistyki, psycholingwistyki i kognitywistyki, semiotycznych aspektów filozofii umysłu, filozoficznych konsekwencji metalogiki i teorii języków formalnych, analizy języka filozofii i argumentacji filozoficznej oraz historii idei semiotycznych i logicznych. Wydawcą „Studiów Semiotycznych” jest Polskie Towarzystwo Semiotyczne. Zasady składania tekstów do Redakcji, prawne i etyczne aspekty publikacji oraz procedura recenzyjna stosowana przez Redakcję są szczegółowo opisane w zakładce „Do Autorów” na stronie:

<http://studiasemiotyczne.pts.edu.pl>

Studia Semiotyczne (Semiotic Studies) is a journal founded in 1970 by Jerzy Pelc, who was its Editor-in-Chief up until 2015. Between 1970 and 2015 *Studia Semiotyczne* was published non-periodically. In December 2015 *Studia Semiotyczne* was formally transformed into a six-monthly published simultaneously in print and on the Internet. The journal publishes papers that fall on the borderline between philosophy and semiotics and in particular, in the fields of analytic philosophy of language, general theory of signs, applications of semantic methods in philosophy, philosophical aspects of linguistics, psycholinguistics and computer science, semiotic aspects of philosophy of mind, philosophical consequences of metalogic and the theory of formal languages, analysis of the language of philosophy and the philosophical argumentation, and history of ideas in semiotics and logic. *Studia Semiotyczne* is published by Polskie Towarzystwo Semiotyczne (The Polish Semiotic Society). The guidelines for submitting manuscripts to the Editors, the legal and ethical aspects of the publication, and the review procedure used by the Editors are described in detail in the “For Authors” section of the website:

<http://studiasemiotyczne.pts.edu.pl>

POLSKIE TOWARZYSTWO SEMIOTYCZNE

STUDIA SEMIOTYCZNE

Tom XXXI • nr 1

PÓŁROCZNIK

WYBRANE ZAGADNIENIA FILOZOFII JĘZYKA I UMYŚLU



WARSZAWA • 2017

Założyciel „Studiów Semiotycznych” (*Founding Editor*):

Jerzy Pelc

Zespół redakcyjny (*Editorial Board*):

Andrzej Biłat (redaktor naczelny),
Krystyna Komorowska (redaktor prowadząca),
Zbysław Muszyński, Joanna Odrowąż-Sypniewska, Bartłomiej Skowron

Redaktorzy numeru (*Guest Editors*):

Tadeusz Ciecierski, Tomasz Puczyłowski

Rada naukowa (*Advisory Board*):

Jerzy Bartmiński (Uniwersytet Marii Curie-Skłodowskiej), Paul Bouissac (University of Toronto), Andrzej Bronk (Katolicki Uniwersytet Lubelski), Idalia Kurcz (Szkoła Wyższa Psychologii Społecznej), Witold Marciszewski (Uniwersytet w Białymstoku), Genoveva Marti (ICREA oraz Universitat de Barcelona), Adam Nowaczyk (Uniwersytet Łódzki), Stefano Predelli (University of Nottingham), Mieczysław Omyła (Uniwersytet Kardynała Stefana Wyszyńskiego), Piotr Stalmaszczyk (Uniwersytet Łódzki), Anna Wierzbicka (Australian National University), Andre Włodarczyk (Université Paris-Sorbonne), Jan Woleński (Wyższa Szkoła Informatyki i Zarządzania)

Redakcja językowa:

Marcin Będkowski

Skład elektroniczny:

Pracownia Wydawnicza, Zalesie Górne

Adres redakcji:

Krakowskie Przedmieście 3, 00-047 Warszawa
e-mail: studiasemiotyczne@pts.edu.pl
<http://studiasemiotyczne.pts.edu.pl/>

ISSN 0137-6608

© Copyright by Polskie Towarzystwo Semiotyczne

Redakcja merytoryczna, skład elektroniczny, korekta językowa w języku polskim, zamieszczanie plików PDF na stronie internetowej „Studiów Semiotycznych”, koordynowanie prac i opieka techniczna zespołu redakcyjnego nad procedurą recenzyjną w systemie *peer review* w formule otwartego dostępu, realizacja prac redakcyjnych mających na celu zagwarantowanie oryginalności publikowanych artykułów naukowych są finansowane w ramach umowy 614/P-DUN/2016 ze środków Ministra Nauki i Szkolnictwa Wyższego przeznaczonych na działalność upowszechniającą naukę.



Ministerstwo Nauki
i Szkolnictwa Wyższego

SPIS TREŚCI

Tadeusz Ciecierski, Jerzy Pelc (ur. w Warszawie 30 września 1924 r., zm. w Warszawie 2 czerwca 2017 r.)	5
Od Redakcji	19
Paweł Grabarczyk, Directival Theory of Meaning Resurrected	23
Krzysztof Pośłajko, Semantic Deflationism, Public Language Meaning and Contextual Standards of Correctness	45
Piotr Konderak, The Conscious Semiotic Mind	67
Piotr Wilkin, Reprezentacje znaturalizowane – użyteczny cel czy użyteczna fikcja?	91
Jacek Wawer, Problem ustalania indeksu w semantyce temporalno-modalnej	109
Maciej Sendłak, U podstaw sporu o kontrmożliwe okresy warunkowe	131
Maria Matuszkiewicz, Problem aktualisty modalnego z czysto możliwymi indywiduami (o <i>Mere Possibilities</i> Roberta Stalnakera)	153

CONTENTS

Tadeusz Ciecierski, Jerzy Pelc (b. in Warsaw on September 30, 1924, d. in Warsaw on June 2, 2017)	5
From Editors	19
Paweł Grabarczyk, Directival Theory of Meaning Resurrected	23
Krzysztof Pośłajko, Semantic Deflationism, Public Language Meaning and Contextual Standards of Correctness	45
Piotr Konderak, The Conscious Semiotic Mind	67
Piotr Wilkin, Naturalized Representations – a Useful Goal or a Useful Fiction?	91
Jacek Wawer, The Problem of Index-Initialization in the Tempo-Modal Semantics	109
Maciej Sendłak, About the Basis for the Debate over Counterpossibles	131
Maria Matuszkiewicz, Modal Actualism and the Problem of Merely Possible Individuals (about Robert Stalnaker’s <i>Mere Possibilities</i>)	153

TADEUSZ CIECIERSKI

JERZY PELC

(ur. w Warszawie 30 września 1924 r.,
zm. w Warszawie 2 czerwca 2017 r.)

Dnia 2 czerwca zmarł w Warszawie Jerzy Pelc, założyciel oraz wieloletni redaktor „Studiów Semiotycznych”. Przez prawie pół wieku troszczył się On o to, aby na ich łamach ukazywały się wspomnienia pozgonne upamiętniające polskich i zagranicznych uczonych semiotyków. Wspomnienia te albo pisał sam, albo prosił o ich napisanie osoby najlepiej znające dorobek i działalność naukową zmarłego. Jego odejście postawiło nas wszystkich w sytuacji niemożliwej: sam nie może być autorem wspomnienia, a tylko On mógłby kompetentnie napisać kilka słów o swoim ogromnym dorobku naukowym.

Jerzy Pelc przyszedł na świat 30 września roku 1924 w Warszawie jako syn Romana i Teresy, z domu Dmochowskiej. Od roku 1936 do roku 1942 uczęszczał do Gimnazjum i Liceum im. Stefana Batorego (w latach 1939–1942 w ramach tajnych kompletów). Po maturze, jak sam wspomina: „Z jednej strony nęciła mnie humanistyka – może socjologia, może filozofia, może filologia, z drugiej medycyna, z trzeciej – matematyka”¹. Profesor Tadeusz Miłobędzki zasugerował Mu, aby poradził się osobiście Władysława Tatarkiewicza. Posłużył on Jerzemu Pelcowi radą, aby studiował jednocześnie filozofię i polonistykę. Zgodnie z radą profesora Tatarkiewicza rozpoczął w roku 1942 studia na obu kierunkach na podziemnym Uniwersytecie Warszawskim.

¹ Jerzy Pelc, *Wizerunki i wspomnienia*, Biblioteka Myśli Semiotycznej, Warszawa 1994, s. 16.

Uczestniczył jako ochotnik w Powstaniu Warszawskim, walczył w Śródmieściu i na Powiślu. Wyszedł z Warszawy z ludnością cywilną, a w grudniu roku 1944 wyjechał do Krakowa, gdzie został aresztowany pod zarzutem posiadania fałszywych dokumentów. Na przełomie roku 1944 i 1945 został wysłany na roboty ziemne do Kocmyrzowa pod Krakowem. O swoim udziale w konspiracji i Powstaniu oraz współpracy z osobami reprezentującymi najróżniejsze opcje polityczne powie po latach: „Co do mnie: robiłem to, co robiłem, i nie sądzę, bym miał się wstydzić, że byłem bezideowym dwulicowcem politycznym, ani szczyścić, że – już wtedy – realizowałem hasło pluralizmu”². Ten pluralizm światopoglądowy i umiarkowanie w poglądach mógł rozpoznać każdy, kto odbył z Jerzym Pelcem krótką rozmowę. Studia kontynuował po wojnie, najpierw na Uniwersytecie Jagiellońskim, potem ponownie na Uniwersytecie Warszawskim, gdzie uzyskał w roku 1946 stopień magistra, a w roku 1951 – doktora nauk humanistycznych. W roku 1961 uzyskał stopień doktora habilitowanego w Instytucie Filozofii i Socjologii PAN, a w latach 1971 i 1979 otrzymał odpowiednio tytuły profesora nadzwyczajnego i zwyczajnego.

Zawsze podkreślał, że uważa się za spadkobiercę i ucznia Władysława Tatarakiewicza, Kazimierza Ajdukiewicza oraz Tadeusza Kotarbińskiego. O pierwszym z tych filozofów napisał kiedyś, że uważa się za jego ucznia o tyle, o ile pozwala sobie na stosowanie narzędzi semiotyki logicznej do zagadnień estetyki i teorii sztuki³. O drugim, że najbliższe były mu poglądy Ajdukiewicza z okresu empiryzmu metodologicznego⁴. W wypadku trzeciego zawsze podkreślał wielki wpływ, jaki miał na niego reizm, zwłaszcza semantyczny. Swój stosunek do tej doktryny opisze w sposób następujący:

Początkującego studenta urzekła i uczyniła na czas długi odpornym na inne doktryny Kotarbińskiego zwana reizmem lub pansomatyzmem, lub konkretyzmem (...). Zrazu obie wersje reizmu Kotarbińskiego, ontologiczna i semiotyczna, stanowiły moje *credo* naukowe. Z biegiem czasu nastąpiło jednak przesunięcie: wierzenia metafizyczne, a do takich zaliczam hipotezy reizmu ontologicznego, jestem skłonny traktować jako należące do sfery spraw prywatnych, na poły emocjonalnych, choćby były silne i uporczywe. Myślę też, że nie jest zadaniem profesora

² Jerzy Pelc, „Żołnierze powstania”, [w:] J. Jadacki, B. Markiewicz (red.), *Próg Istnienia. Dziesięciokrotne pokolenie*, Warszawa 1995, s. 105.

³ Jerzy Pelc, *Wizerunki i wspomnienia*, dz. cyt., s. 67.

⁴ Tamże, s. 68.

logiki języka przekazywania *ex cathedra* własnych spraw prywatnych. Natomiast wskazówki reizmu semiotycznego uważam za wypróbowane i pożyteczne dyrektywy nauczycielskie, składające się na godny polecenia program, który staram się rozpowszechniać⁵.

Przez całą swoją karierę Jerzy Pelc związany był z Uniwersytetem Warszawskim, gdzie pracował najpierw w Seminarium Historii Literatury Polskiej UW (w latach 1945–1950), Katedrze i Zakładzie Logiki (w latach 1951–1972) oraz w Zakładzie Semiotyki Logicznej (od roku 1972), który założył oraz którym samodzielnie kierował do roku 1994. Pracował także w Zakładzie Logiki Polskiej Akademii Nauk oraz jako sekretarz Instytutu Filozofii i Socjologii PAN (w latach 1957–1962). W latach 1967–1985 prowadził także wykłady z metodologii nauk w Wyższej Szkole Muzycznej im. Fryderyka Chopina w Warszawie.

Od roku 1961 do roku 2010 organizował odczyty i konwersatoria semiotyczne, których kontynuacją jest organizowane przez Polskie Towarzystwo Semiotyczne seminarium Znak–Język–Rzeczywistość, najstarsze seminarium filozoficzne w Polsce. Po roku 2010 pozostawał niezmiennie opiekunem naukowym seminarium. Choć kłopoty z chodzeniem sprawiły, że nie mógł od lat kilku uczestniczyć w jego spotkaniach, zawsze czekało na Niego na seminarium honorowe miejsce przewodniczącego zebrań. Był współzałożycielem Polskiego Towarzystwa Semiotycznego (w roku 1968), a potem jego sekretarzem i wieloletnim prezesem (do roku 2011, gdy względy zdrowotne skłoniły go do rezygnacji z tej funkcji). Do roku 2015 był redaktorem „Studiów Semiotycznych”, pozostawał też do dnia śmierci redaktorem Biblioteki Myśli Semiotycznej.

Pewne pojęcie o wzbudzającej podziw skali działalności Jerzego Pelca jako organizatora nauki i redaktora dać może lista prelegentów i tytułów odczytów, które zostały wygłoszone w latach 1961–2004 na zebraniach seminarium semiotycznego (zawiera ją tom XXV „Studiów Semiotycznych”, natomiast lista odczytów, poczynając od roku 2005 – wraz z abstraktami i innymi materiałami – dostępna jest na stronie Polskiego Towarzystwa Semiotycznego: www.pts.edu.pl). Listę tę można uzupełnić suchymi, ale wiele mówiącymi wyliczeniami: Jerzy Pelc był redaktorem dwudziestu dziewięciu tomów „Studiów Semiotycznych”, pod jego redakcją ukazały się w Bibliotece Myśli Semiotycznej

⁵ Tamże, s. 67.

pięćdziesiąt cztery tomy (sam był autorem jednej książki oraz bezpośrednim redaktorem lub współredaktorem siedmiu tomów z tej serii). Nie mniej znaczące od suchych wyliczeń są waga i znaczenie Jego prac redakcyjnych oraz translatorskich: wydana przez PWN w roku 1967 pod Jego redakcją antologia przekładów *Logika i język* umożliwiła i nadal umożliwia kolejnym pokoleniom studentów zapoznanie się z fundamentalnymi pracami Carnapa, Churcha, Fregego, Geacha, Russella, Strawsona i Reichenbacha. Natomiast wydana w roku 1979 w *Synthese Library* (tom 110) antologia *Semiotics in Poland 1894–1969* (wersja polska ukazała się w roku 1971 sumptem PWN) pozostaje prawdopodobnie nadal najbardziej kompletnym dostępnym w języku angielskim wyborem prac polskich filozofów języka i logików.

Prace redaktorskie, organizacyjne i nauczycielskie łączył Jerzy Pelc z oryginalną działalnością naukową. Semiotykę traktował zawsze jako naukę o semiozie rozumianej jako proces używania znaków, a zatem ich „interpretacji, rozumienia, nadawania, przetwarzania i odbierania”⁶. Tak pojęta semiotyka jest ogólną nauką o poznaniu i komunikacji, ponieważ nie sposób pomyśleć tych ostatnich bez użycia znaków. Jego prace z zakresu tak pojętej semiotyki dotyczyły w równym stopniu zagadnień teoretycznych, jak i historii tej dyscypliny. Publikował prace poświęcone m.in. średniowiecznej nauce o supozycjach (*Średniowieczna nauka o supozycjach terminów (Przyczynek do genealogii semiotyki)*, „Studia Filozoficzne”, 1979), modystom (*O modystach: karta z dziejów semiotyki średniowiecznej*, „Studia Filozoficzne”, 1979), Peirce’owi (*Several Questions to Experts in Peirce’s Theory of Signs*, „VS-versus: Quaderni di Studi Semiotici”, 1990, *The Concept of Time in the Philosophy of C.S. Peirce*, [w:] Ernest W.B. Hess-Lüttich, Brigitte Schlieben-Lange (eds.), *Signs & Time / Zeit & Zeichen: An International Conference on the Semiotics of Time in Tübingen*, Tübingen, Gunter Narr Verlag, 1998), licznym polskim i zagranicznym autorom współczesnym. Świadom był tego, czym odróżniał się od większości współczesnych filozofów analitycznych, że bardzo wiele pozornie nowych pomysłów to faktycznie pomysły znane z historii filozofii, semiotyki i lingwistyki, a wrażenie, że są oryginalne, bierze się wyłącznie z historycznej ignorancji. Przestrzegał jednak też przed skrajnością przeciwną – jak powiedział kiedyś: „Twórcze myśle-

⁶ Jerzy Pelc, *Semiotyka jako nauka o poznawaniu świata i porozumiewaniu się za pomocą znaków*, [w:] *Znak – Znaczenie – Język*, ASP, Warszawa 1994, s. 20–28.

nie polega na zapominaniu, że ktoś już kiedyś coś napisał”. To synkretyczne podejście widać znakomicie w Jego *Wstępie do semiotyki* (Wiedza Powszechna, 1982), w którym znaleźć można zarówno elementy wykładów z historii semiotyki, jak i oryginalne analizy pojęciowe. Bardziej osobiste spojrzenia na najnowszą historię badań filozoficznych i semiotycznych zawiera zaś książka *Wizerunki i wspomnienia. Materiały do dziejów semiotyki* (Biblioteka Myśli Semiotycznej, 1994).

W filozofii języka z jednej strony pozostawał pod wpływem filozofii logicznej szkoły lwowsko-warszawskiej, z drugiej interesował się zawsze filozofią języka naturalnego, zwłaszcza jego pragmatyką i semantyką. Jak można byłoby dziś powiedzieć, nawiązując do znanego rozróżnienia Janiny Kotarbińskiej, świadomie łączył w swoich pracach tradycje analitycznej filozofii rekonstrukcjonistycznej oraz deskrypcjonistycznej. Jedną z głównych idei, której poświęcił swoje prace, było pojęcie użycia wyrażenia. Jego artykuł *Funkcjonalne podejście do semiotyki logicznej języka naturalnego* („Studia Filozoficzne”, 1967) oraz książki *O użyciu wyrażenia* (Ossolineum, 1971) oraz *Studies in Functional Logical Semiotics of Natural Language* (Mouton, 1971) należą do grona najlepszych prac w powojennej historii filozofii poświęconych temu pojęciu. W pracach tych sformułował metodologiczny postulat, który określił mianem „funkcjonalnego podejścia do semiotyki logicznej języka naturalnego”. Głosi on, że w badaniach nad wyrażeniami językowymi uwzględniać zawsze należy to, w jaki sposób zostały one użyte. „Sposób użycia” jest w tym postulatcie rozumiany jako zgodne z językowym znaczeniem danego wyrażenia użycie tego wyrażenia w celu odniesienia się do czegoś lub stwierdzenia czegoś. Tak pojęte podejście funkcjonalne kładzie zasadniczy nacisk na rolę pozajęzykowych okoliczności użycia wyrażenia. Jedną z konsekwencji tego funkcjonalizmu semiotycznego jest uznanie, że wyrażenia nie mają – poza kontekstem ich użycia – odniesienia, nie grają też żadnych związanych z odniesieniem ról semantycznych.

Prace Jerzego Pelca dotyczyły wielu zagadnień: zajmował się pojęciem znaku (*Znak*, „Studia Semiotyczne”, 1980, wersja angielska: *The Concept of Sign*, „Studia Semiotyczne – English Supplement”, 2015), badał zagadnienie znaczenia (*Meaning as an Instrument*, „Semiotica”, 1969 oraz *O użyciu wyrażenia*), analizował pojęcia tematu (*O pojęciu tematu*, Ossolineum 1961) i narracji (*On the Concept of Narration*, „Semiotica”, 1971), zajmował się zagadnieniem fikcji (*Wyrażenia imienne a fikcja literacka*, „Studia Estetyczne”, 1967, angielska wersja: *Nominal*

Expressions and Literary Fiction, [w:] *Studies in Functional Semiotics of Natural Language*, 1971; *O bytach fikcyjnych i tekstach fikcjonalnych*, „*Studia Semiotyczne*”, 1983, wersja angielska: *On Fictitious Entities and Fictional Texts*, „*Recherches Sémiotiques – Semiotic Inquiry*”, 1986), badał zagadnienie nazw własnych (*Imiona własne w języku naturalnym. Prolegomena do teorii*, [w:] *Fragmenty filozoficzne. Seria trzecia. Księga pamiątkowa ku czci Profesora Tadeusza Kotarbińskiego w osiemdziesiątą rocznicę urodzin*, 1967, wersja angielska: *Proper Names in Natural Language: Prolegomena to a Theory*, [w:] *Studies in Functional Logical Semiotics of Natural Language*, 1971), deskrypcji (*The i deskrypcje*, [w:] *Deskrypcje i prawda*, 2010), okresów warunkowych (*Jeżeli, to*, „*Studia Semiotyczne*”, 1986, wersja angielska: *If, then*, „*Studia Semiotyczne – English Supplement*”, 2015), metafory (*Semiotic Functions as Applied to the Analysis of the Concept of Metaphor*, [w:] *Studies in Functional Semiotics of Natural Language*, 1971), intensjonalności i intencjonalności (*Some Semiotic Considerations Concerning Intensional Expressions and Intentional Objects*, „*Logique et Analyse*”, 1977), kłamstwa (*O pojęciu kłamstwa – z punktu widzenia semiotyki*, „*Studia Semiotyczne*”, 1990, wersja angielska: *On the Concept of Lie – From the Point of View of Semiotics*, „*Studia Semiotyczne – English Supplement*”, 2015). W pracach tych konsekwentnie stosował postulat funkcjonalizmu semiotycznego.

Imponująca działalność naukowa i organizacyjna Jerzego Pelca spotykała się z uznaniem w Polsce i na świecie. Był honorowym przewodniczącym Międzynarodowego Stowarzyszenia Semiotycznego (obok uczonych takich jak Umberto Eco, Roland Posner i Eero Tarasti) oraz Institut International de Philosophie w Paryżu (obok Evandra Agazziego i Jaakka Hintikki). Był honorowym członkiem Polskiego Towarzystwa Filozoficznego, Semiotic Society of America, Canadian Semiotic Association, Österreichische Gesellschaft für Semiotik, Sociedad Mexicana de Filosofía oraz Magyar Szemiotikai Társaság. Był również członkiem Rady Języka Polskiego przy Prezydium PAN. Zasiadał w radach naukowych renomowanych czasopism takich jak „*Semiotica*” oraz „*Poetics Today*”. Gdy w roku 1974 odbywał się w Mediolanie pierwszy Międzynarodowy Kongres Semiotyczny, „*The Times*” (z dnia 14 czerwca) upamiętnił wzmianką Jego wystąpienie i udział w kongresie⁷.

⁷ Z obszernym fragmentem artykułu z „*Timesa*” może Czytelnik zapoznać się w słowie *Od Wydawcy* w VI tomie „*Studiów Semiotycznych*” (s. 7–8).

W roku 1998 otrzymał w Nowym Jorku Nagrodę Fundacji Alfreda Jurzykowskiego, a w roku 2008 – Srebrny Medal „Zasłużony Kulturze – Gloria Artis” przyznany przez Ministra Kultury i Dziedzictwa Narodowego. Na X Polskim Zjeździe Filozoficznym w roku 2015 u honorowany został medalem „Optime Meritis de Philosophia”, którego zarówno z powodów zdrowotnych, jak i motywowany niechęcią do celebrowania własnej osoby nigdy nie odebrał. Pamiętam, że gdy na prośbę dziekana Wydziału Filozofii i Socjologii UW zapytałem Go, czy można Mu przekazać medal na posiedzeniu Rady Wydziału, odparł, że nie widzi takiej możliwości, a poza tym uważa, że medal taki szybko trzeba będzie odsyłać.

Życzeniem Jerzego Pelca było, aby po Jego śmierci kontynuowane były seminaria semiotyczne, ukazywały się nadal „Studia Semiotyczne” oraz „Biblioteka Myśli Semiotycznej”. Pozostawił po sobie manuskrypty prac poświęconych reizmowi oraz współczesnej polszczyźnie. Planowane jest ich wydanie w ramach Biblioteki Myśli Semiotycznej.

Polskie Towarzystwo Semiotyczne żegnało go słowami nekrologu:

Z głębokim smutkiem i w poczuciu niepowetowanej straty żegnamy wybitnego uczonego, znakomitego nauczyciela akademickiego, oddanego wychowawcę wielu pokoleń studentów oraz młodych i dojrzałych badaczy, wytrwałego i konsekwentnego organizatora nauki, kontynuatora tradycji filozoficznej szkoły lwowsko-warszawskiej, założyciela i wieloletniego Kierownika Zakładu Semiotyki Logicznej Instytutu Filozofii Uniwersytetu Warszawskiego, założyciela i wieloletniego Prezesa Polskiego Towarzystwa Semiotycznego, redaktora założyciela „Studiów Semiotycznych” oraz Biblioteki Myśli Semiotycznej, członka rzeczywistego Towarzystwa Naukowego Warszawskiego, członka czynnego Polskiej Akademii Umiejętności, honorowego Przewodniczącego Institut International de Philosophie oraz International Association for Semiotic Studies, byłego Przewodniczącego Centralnej Komisji ds. Tytułu i Stopni Naukowych, Komitetu Nauk Filozoficznych PAN oraz Komitetu Etyki w Nauce PAN, żołnierza Armii Krajowej i uczestnika Powstania Warszawskiego.

Na nagrobku Jerzy Pelc poprosił o umieszczenie skromnego, ale – przez tę skromność – jakże dobrze pasującego do Niego, epitafium: „PROFESOR FILOZOFII, ŻOŁNIERZ POWSTANIA WARSZAWSKIEGO”.

BIBLIOGRAFIA JERZEGO PELCA OD ROKU 1996

Pełną bibliografię Jerzego Pelca do roku 1995 może Czytelnik znaleźć w książce *W świecie znaków* (pod red. J. Jadackiego i W. Strawińskiego), PTF, Warszawa 1996, s. 23–40. Bibliografię do roku 2000 podzieloną tematycznie można znaleźć w artykule biograficznym pióra Agnieszki Salamuchy opublikowanym w drugim tomie *Polskiej filozofii powojennej* (red. W. Mackiewicz, Agencja Wydawnicza Witmark, Warszawa 2001, s. 162–184). Artykuł ten zawiera także syntetyczne omówienie poglądów filozoficznych i semiotycznych Jerzego Pelca.

1996

- „Kryteria oceny naukowej”, *Nauka*, nr 3, s. 104–115
- „Odpowiedzialność i ryzyko: uwagi o pojęciach – na marginesie dyskusji na temat odpowiedzialności i ryzyka w nauce”, *Nauka*, nr 4, s. 143–146
- „Symptom and Symbol in Language”, [w:] Marcelo Dascal, Dietfried Gerhardus, Kuno Lorenz, Georg Meggle (eds.), *Sprachphilosophie: Ein internationales Handbuch zeitgenössischer Forschung / Philosophy of language: An International Handbook of Contemporary Research / La philosophie du langage: Manuel international des recherches contemporaines*, Handbücher zur Sprach- und Kommunikationswissenschaft, Bd. 2, Berlin–New York, Walter de Gruyter, s. 1292–1313

1997

- „Dyscypliny humanistyczne a humanistyka literacka”, *Kultura i Społeczeństwo*, t. XLI, nr 2, s. 21–28
- „Humanidades científicas vs. humanidades literarias”, [w:] Rebeca Barriga Villanueva, Pedro Martín Butrangueno (eds.), *Varia Lingüística y Literaria. 50 Años del CELL*, vol. I: *Lingüística*, Publicaciones de la Nueva Revista de Filología Hispánica, vol. VII, Mexico, El Colegio de Mexico, s. 265–276
- „List”, *Roczniki Filozoficzne*, t. 45, nr 1, s. 149
- „Semantics and Semiotics”, [w:] M.L. Dalla Chiara *et al.* (eds.), *Structures and Norms in Science*, Dordrecht–Boston–London, Kluwer Academic Publishers, s. 429–430
- „Sign, Thought and World. Semiotics: The Science of Cognition and Communication”, [w:] Jeff Bernard, Josef Wallmannsberger, Gloria Withalm (eds.), *Welt der Zeichen, Welt der Dinge – World of Signs, World of Things. Akten des 8. Symposiums der Österreichischen Gesellschaft für Semiotik (Innsbruck 1993)*, *Angewandte Semiotik*, vol. 15, Wien, ÖGS, s. 123–150

- Tadeusz Kotarbiński (1886–1981), [w:] *Wielcy filozofowie polscy – sześć studiów*, Warszawa, Wydawnictwo Instytutu Filozofii i Socjologii PAN, s. 31–54
- „Theory Formation in Semiotics”, [w:] Roland Posner, Klaus Robering, Thomas A. Sebeok (eds.), *Semiotik-Semiotics. Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur – A Handbook on the Sign-Theoretic Foundations of Nature and Culture*, vol. I, Berlin – New York, Walter de Gruyter, s. 617–643
- „Understanding, Explanation and Action as Problems of Semiotics”, [w:] *ibidem*, s. 644–667

1998

- „Alfred Tarski (1902–1983) o języku przedmiotowym, metajęzyku i pojęciu prawdy”, *Studia Semiotyczne*, t. XXI–XXII, s. 301–303
- „Do jakiego stopnia semiotyka jest pomostem między naturą i kulturą”, [w:] *ibidem*, s. 253–262
- „Idee semiotyczne Izydory Dąbskiej (1904–1983)”, [w:] *ibidem*, s. 305–309
- „Janina Kotarbińska” [wspomnienie pozgonne], [w:] *ibidem*, s. 17–21
- „Janusz Feliks Chmielewski” [wspomnienie pozgonne], [w:] *ibidem*, s. 23–25
- „Kazimierz Ajdukiewicz (1890–1963) o języku i wyrażeniach mowy”, [w:] *ibidem*, s. 271–291
- „Kazimierz Twardowski (1866–1938) o znaku i języku ze stanowiska psychologii i teorii poznania”, [w:] *ibidem*, s. 271–291
- „Logika języka i filozofia języka”, [w:] *ibidem*, s. 263–264
- „Olgiard Adrian Wojtasiewicz” [wspomnienie pozgonne], [w:] *ibidem*, s. 11–13
- „Stanisław Ossowski (1897–1963) o pojęciu znaku oraz języku nauki, języku poezji i języku mistyki”, [w:] *ibidem*, s. 293–295
- „Poglądy semantyczne Janiny Kotarbińskiej (1901–1997)”, [w:] *ibidem*, s. 297–299
- „Tadeusza Pawłowskiego (1924–1996) myśli o znaczeniu wyrażen”, [w:] *ibidem*, s. 311–312
- „Dobrze przeżyć życie – Zdzisław Libera”, *Rzeczpospolita*, 4–5 kwietnia, nr 80(4940), s. 14
- „Krytyka i krytycyzm w nauce”, [w:] *Krytyka i krytycyzm w nauce*, Warszawa, Fundacja na Rzecz Nauki Polskiej, s. 7–25
- „Logic of Language and Philosophy of Language in 20th-century Poland”, *Historiographia Linguistica*, vol. XXV, no 1/2, s. 163–220
- „Nestor filozofii polskiej – o dziele Stefana Swieżawskiego”, *Tygodnik Powszechny*, 31 maja, nr 22(2551), s. 14
- „Opening Speech on the Second East Asian International Semiotics Seminar”, *Supplement to Philosophical Translation Quarterly: Proceedings of the 1st and 2nd East Asian International Semiotics Seminar*, Pekin–Shanghai, s. 3
- „Przedmowa”, [w:] Tadeusz Kowzan, *Znak i teatr*, Biblioteka Myśli Semiotycznej, Warszawa, Znak–Język–Rzeczywistość, s. 11–15

- „The Concept of Time in the Philosophy of C.S. Peirce”, [w:] Ernest W.B. Hess-Lüttich, Brigitte Schlieben-Lange (eds.), *Signs & Time / Zeit & Zeichen: An International Conference on the Semiotics of Time in Tübingen*, Tübingen, Gunter Narr Verlag, s. 228–235
- „Uwagi o semiotycznych własnościach ocen”, [w:] Janusz Krassowski (red.), *Dzieło muzyczne między inspiracją a refleksją – Księga pamiątkowa na 60 lecie urodzin Marka Podhajskiego*, Prace Specjalne, t. 55, Gdańsk, Akademia Muzyczna, s. 14–18

1999

- „Filozofia i eseistyka filozoficzna a naukowe badanie i nauczanie filozofii”, *Przegląd Filozoficzny – Nowa Seria*, r. VIII, nr 1(29), s. 93–120
- „Language of the Present-Day Scientific Humanities: Clear or Vague”, [w:] Adrian Gimiate-Welsh (ed.), *La Semiótica. Intersección entre la Naturaleza y la Cultura*, VI Congreso Internacional IASS / AIS / AMES, Mexico, CD-ROM
- „Przedmowa”, [w:] Jerzy Kroh, *Moje życie i chemia radiacyjna – od II do III Rzeczpospolitej*, Łódź, Wyd. Fundacja Badań Radiacyjnych, s. 7–9
- „Spojrzenie na minione sześćdziesięciolecie: tajne studia”, *Przegląd Filozoficzny – Nowa Seria*, r. VIII, nr 4(32), s. 211–233

2000

- Język współczesnej humanistyki* (red.), Biblioteka Myśli Semiotycznej, Polskie Towarzystwo Semiotyczne, Warszawa
- „Myśli o języku humanistyki”, [w:] *ibidem*, s. 171–204
- „Posłowie: O jasnym i niejasnym stylu w humanistyce. Felieton dedykowany «jasnościowcom» (fragment)”, [w:] *ibidem*, s. 337–344
- „Przedmowa”, [w:] *ibidem*, s. 7–15
- „Nad trumną Janusza Krajewskiego”, *Przegląd Filozoficzny – Nowa Seria*, r. IX, nr 3(35), s. 279–282
- „Semiosis and Semiosics Vs. Semiotics”, *Semiotica*, vol. 128, issue 3–4, s. 425–434

2001

- „Od wydawcy”, *Studia Semiotyczne*, t. XXIII, s. 5–9
- „Irmiona Judycka” [wspomnienie pozgonne], [w:] *ibidem*, s. 15–17
- „Willard Van Orman Quine” [wspomnienie pozgonne], [w:] *ibidem*, s. 25–26
- „Od wydawcy”, *Studia Semiotyczne*, t. XXIV, s. 5–7
- „Marcin Czerwiński” [wspomnienie pozgonne], [w:] *ibidem*, s. 9–14
- „Antonina Kłowska” [wspomnienie pozgonne], [w:] *ibidem*, s. 15–16
- „Antoni Wincenty Moniuszko” [wspomnienie pozgonne], [w:] *ibidem*, s. 17–18

- „Analiza semantyczna w pismach Władysława Tatarkiewicza”, *Przegląd Filozoficzny – Nowa Seria*, r. X, nr 1(37), s. 33–41. Toż po angielsku [„Semantic Analysis in the Works of Władysław Tatarkiewicz”] w: *Organon* 28–30 (1999–2001), s. 193–202
- „Henryk Hiż («jeszcze nie cały»)", [w:] Witold Mackiewicz (red.), *Polska filozofia powojenna*, t. I, Warszawa, Agencja Wydawnicza Witmark, s. 316–359
- „Głos laika w sprawie klonowania ludzi”, *Przegląd Filozoficzny – Nowa Seria*, r. X, nr 3(39), s. 5–24
- „Odkąd zaczyna się człowiek”, *Przegląd Filozoficzny – Nowa Seria*, r. X, nr 3(39), s. 25–29
- „Słowo wstępne”, *Nauka*, nr 4, s. 93–94
- „Uwagi o wieloletowości”, *Nauka*, nr 4, s. 125–131

2002

- „Dwa pojęcia normy a poprawność”, [w:] Włodzimierz Gruszczyński (red.), *Język narzędziem myślenia i działania: materiały z konferencji zorganizowanej z okazji 100-lecia „Poradnika Językowego”*, (Warszawa, 10–11 maja 2001 r.), Warszawa, Elipsa, s. 96–102
- „En que medida la semiótica es puente de unión entre la naturaleza y la cultura?”, [w:] Jesús G. Maestro (ed.), *Nuevas perspectivas en semiología literaria*, Madrid, Arco/Libros, s. 249–266
- „Jasność a głębia”, [w:] Jerzy Brzeziński *et al.* (red.), *Odwaga filozofowania – Leszkowi Nowakowi w darze*, Poznań, Wydawnictwo Fundacji Humaniora, s. 549–568
- „Jerzego A. Wojciechowskiego *Ekologia wiedzy*. Myśl polska na obczyźnie” [recenzja], *Ruch Filozoficzny*, t. LIX, nr 4, s. 699–704

2003

- „Czy językoznawców nie obchodzi prawda?”, [w:] Ireneusz Bobrowski (red.), *Analisis – prace ofiarowane Profesor Krystynie Pisarkowej*, Kraków Lexis, s. 195–199
- „O filozofii i nie tylko: na pytania Marka Dobrzenieckiego i Pawła Sierpińskiego odpowiada Profesor Jerzy Pelc”, *Prezentacje: Magazyn Filozoficzny*, nr 4, s. 4–6

2004

- „Od wydawcy”, *Studia Semiotyczne*, t. XXV, s. 5–13
- „Thomas A. Sebeok” [wspomnienie pozgonne], [w:] *ibidem*, s. 15–20
- „Gerard Deledalle” [wspomnienie pozgonne], [w:] *ibidem*, s. 21–26
- „Sprawa pomnika Tadeusza Kotarbińskiego w Bibliotece Uniwersytetu Warszawskiego”, *Ruch Filozoficzny*, t. LXI, nr 4, s. 535–547
- „The Soldiers of the Uprising”, *Dialogue and Universalism*, no 5–6, s. 183–188

2005

- „Antyirracjonalizm – postawa filozofów szkoły lwowsko-warszawskiej”, [w:] Kazimierz Trzęsicki (red.), *Ratione et studio – Profesorowi Witoldowi Marciszewskiemu w darze*, Białystok, Wydawnictwo Uniwersytetu w Białymstoku, s. 309–318
- „Dziedzictwo myśli etycznej Tadeusza Kotarbińskiego wobec moralnych problemów naszych czasów” [głos w dyskusji], *Res Humana*, wyd. XV, nr 6(79), s. 27–29
- „Julian Krzyżanowski – nauczyciel i pryncypał z perspektywy lat sześćdziesięciu”, *Pamiętnik Literacki*, r. XCVI, z. 1, s. 213–219
- „Kazimierz Twardowski – twórca PTF”, *Ruch Filozoficzny*, t. LXII, nr 1, s. 23–27 [wypowiedź w dyskusji], *Etyka*, nr 38, s. 147–148

2006

- „Ateizm Tadeusza Kotarbińskiego”, *Przegląd Filozoficzny – Nowa Seria*, r. XV, nr 1(57), s. 99–120
- Sens, prawda wartość: filozofia języka i nauki w dziełach Kazimierza Ajdukiewicza, Witolda Doroszewskiego, Tadeusza i Janiny Kotarbińskich, Romana Suszki, Władysława Tatarkiewicza* (red.), Biblioteka Myśli Semiotycznej, Polskie Towarzystwo Semiotyczne, Warszawa
- „Filozof filozofii – o narzędziach i warsztacie umysłowym Władysława Tatarkiewicza”, [w:] *ibidem*, s. 49–54
- „Janina Kotarbińska – sojusznica interpretacja reizmu”, [w:] *ibidem*, s. 119–120
- „Kazimierza Ajdukiewicza semiotyka logiczna”, [w:] *ibidem*, s. 175–182
- „Na fundamentach reizmu i realizmu praktycznego”, [w:] *ibidem*, s. 109–118
- „Rozważania semantyczne na marginesach *Elementów leksykologii i semiotyki* Witolda Doroszewskiego”, [w:] *ibidem*, s. 265–304
- „Słowo wstępne”, [w:] *ibidem*, s. 7–10
- „Korabielnikowcy w szkolnictwie wyższym”, *Przegląd Humanistyczny*, r. L, nr 5–6 (398–399), s. 171–183
- „Humanizm i technika” [głos w dyskusji zorganizowanej przez Instytut Problemów Współczesnej Cywilizacji], *Humanizm i Technika*, z. XXXIV, s. 11–12 i 28–32
- „Wizja polskich uczelni w społeczeństwie globalnym” [głos w dyskusji zorganizowanej przez Instytut Problemów Współczesnej Cywilizacji], *Wizja polskich uczelni w społeczeństwie globalnym*, z. XXXVI, s. 21–23 i 35–36

2007

- „Dlaczego estetyka jest dla mnie za trudna – kilka pytań do znawców meta-estetyki”, *Kwartalnik Filozoficzny*, t. XXXV, z. 4, s. 5–21
- „Human Cloning and Organ Transplants vs. Definition of Human Being”, [w:] Ionna Kuçuradi, Stephen Voss, Cemal Güzel (eds.), *The Proceedings of the Twenty-*

-*First World Congress of Philosophy*, vol. 1: *Ethics*, Ankara, Philosophical Society of Turkey, s. 235–244

- „Od wydawcy”, *Studia Semiotyczne*, t. XXVI, s. 5–12
- „Henryk Hiż” [wspomnienie pozgonne], [w:] *ibidem*, s. 13–32
- „Leon Koj” [wspomnienie pozgonne], [w:] *ibidem*, s. 33–42
- „Thomas G. Winner” [wspomnienie pozgonne], [w:] *ibidem*, s. 47–54
- „Użycie wyrażen a ich znaczenie – dylematy i problemy”, [w:] *ibidem*, s. 331–356
- „The Limits of Metaphor”, [w:] Fabio Minazzi (a cura di), *Filosofia, Scienza e Bioetica nel dibattito contemporaneo. Studi internazionali in onore di Evandro Agazzi*, Roma, Istituto Poligrafico e Zecca dello Stato s. 915–926

2008

- „Nauczyć krytycznego myślenia i jasnej mowy”, [w:] *Po drogach uczonych. Z członkami Polskiej Akademii Umiejętności rozmawia Andrzej M. Kobos*, t. III, Kraków, Polska Akademia Umiejętności, s. 563–604
- „Wacław Borowy w ostatnich latach życia”, *Blok-Notes Muzeum Literatury im. Adama Mickiewicza*, nr 14, s. 425–440

2009

- „Why Is Aesthetics Too Difficult for Me: A Few Questions to Experts in Meta-Aesthetics”, [w:] Tomonobu Imamichi, Hans Lenk (eds.), *Philosophy in International Context*, Bd. 4, Berlin, Institut International de Philosophie, LIT Verlag, s. 207–222
- „Wieloletowość, pieniądze, etos nauczyciela”, [w:] *Patriotyzm wczoraj i dziś – Seminarium PAU*, Kraków, Polska Akademia Umiejętności, s. 28–62

2010

- Deskrypcje i prawda* (red.), Biblioteka Myśli Semiotycznej, Polskie Towarzystwo Semiotyczne, Warszawa
- „Przedmowa redaktora BMS: O semiotyce i Bibliotece Myśli Semiotycznej”, [w:] *ibidem*, s. 7–28
- „The i deskrypcje”, [w:] *ibidem*, s. 157–185
- Nauka* (red.), nr 3: *Etyka – Nauka – Moralność. Rozprawy Komitetu Etyki w Nauce przy Prezydium PAN*
- „Od wydawcy: W czterdziestolecie Polskiego Towarzystwa Semiotycznego”, *Studia Semiotyczne*, t. XXVII, s. 5–26
- „Poczucie niezrozumienia”, [w:] *ibidem*, s. 37–44

2012

- Semiotica. Special Issue: Semiotics and Logic* (red.), vol. 188, no 1/4, , s. VIII + 356
 „Semiotics and Logic: Pragmatization of the Common Ground”, *Semiotica. Special Issue: Semiotics and Logic*, vol. 188, no 1/4, s. 1–27

2015

- „Some Observations on the Connections Between Semiotics, Logic and Linguistics”, *Studia Semiotyczne – English Supplement*, vol. VII, s. 159–172
 „The Concept of Sign”, *Studia Semiotyczne – English Supplement*, vol. VIII–XII, s. 103–137
 „A Posthumous Tribute to Tadeusz Kotarbinski”, *Studia Semiotyczne – English Supplement*, vol. XIII, s. 4–20
 „If, then”, *Studia Semiotyczne – English Supplement*, vol. XIV–XV, s. 217–240
 „On the Concept of Lie – From the Point of View of Semiotics”, *Studia Semiotyczne – English Supplement*, vol. XVI–XVII, s. 158–168
 „Afterword to the Discussion Concerning Linguistic and Logical Description of Language”, *Studia Semiotyczne – English Supplement*, vol. XIX–XX, s. 74–86
 „Izydora Dąmbska (1904–1983): Semiotic Concepts”, *Studia Semiotyczne – English Supplement*, vol. XXI–XXII, s. 141–146
 „Janina Kotarbińska’s (1901–1997) Views on Semantics”, [w:] *ibidem*, s. 137–140
 „Kazimierz Twardowski (1866–1938) on the Sign and Language from the Point of View of Psychology and the Theory of Cognition”, [w:] *ibidem*, s. 100–107
 „Kazimierz Ajdukiewicz (1890–1963) on Language and the Expressions of Speech”, [w:] *ibidem*, s. 108–132
 „Stanisław Ossowski (1897–1963) on the Concept of Sign, and on the Language of Science, of Poetry and of Mysticism”, [w:] *ibidem*, s. 133–136
 „A Sense of Non-Understanding”, *Studia Semiotyczne – English Supplement*, vol. XXVII, s. 15–22
 „Od Wydawcy: Pożegnanie ze «Studiami Semiotycznymi»”, *Studia Semiotyczne*, vol. XXVIII–XXIX, s. 5–30
 „Barbara Stanosz” [wspomnienie pozgonne], [w:] *ibidem*, s. 33–40
 „Marian Przełęcki” [wspomnienie pozgonne], [w:] *ibidem*, s. 31
 „Edward Stankiewicz” [wspomnienie pozgonne], [w:] *ibidem*, s. 89–98

Semiotyka, jak zawsze podkreślał Jerzy Pelc, jest nauką o semiozie, a zatem używaniu i rozumieniu znaków. Ponieważ trudno sobie wyobrazić poznanie i komunikację bez użycia i rozumienia znaków, jest ona także nauką o poznaniu i komunikacji. Wreszcie, ponieważ wszelkie poznanie wymaga umysłów poznających, a wśród różnych narzędzi komunikacji na plan pierwszy wysuwa się język, można też uznać semiotykę za najogólniejszą naukę o umyśle i języku. Do takiego właśnie rozumienia semiotyki nawiązuje niniejszy numer „Studiów Semiotycznych”, który poświęcony jest wybranym zagadnieniom współczesnej filozofii języka i umysłu.

Tom otwiera praca Pawła Grabarczyka (Instytut Filozofii, Uniwersytet Łódzki oraz IT University, Kopenhaga) poświęcony możliwości zbudowania nowej wersji dyrektywnej teorii znaczenia. Tekst zawiera zarówno zwięzłą ekspozycję oryginalnej teorii Ajdukiewicza, jak i mierzy się największymi wyzwaniami, które przed nią stoją, oraz zestawia dyrektywną teorię znaczenia z późniejszymi od niej ideami semantycznymi (inferencjalizmem oraz semantyką ról pojęciowych). Praca stanowi więc nie tylko świadectwo współczesności myśli Ajdukiewicza, zwięzłe i klarownie przypominając jego niezmiennie filozoficzne ważne i pomysłowe koncepcje dotyczące znaczenia i języków zamkniętych, lecz także zapowiedź własnego fascynującego projektu filozoficznego. Warto przypomnieć, że Autor jest laureatem pierwszej nagrody im. Henryka i Danuty Hiżów, przyznanej w drodze konkursu przez Polskie Towarzystwo Semiotyczne za najlepszą pracę z zakresu semiotyki teoretycznej rozumianej jako nauki o poznaniu i komunikacji (patrz: „Studia Semiotyczne”, t. XXX, nr 1, 2016).

Problematykę znaczenia podejmuje także Krzysztof Pośłajko (filozofia, Uniwersytet Jagielloński, autor ten debiutuje na łamach „Studiów Semiotycznych”), który w swojej pracy przedstawia argument na rzecz tezy, że pojęcie znaczenia wyrażenia określonego języka, używanego przez jakąś społeczność, zakłada, że w społeczności przyjmuje się szereg sądów dotyczących standardów poprawnego użycia wyrażeń.

Te ostatnie są, zdaniem autora, zależne kontekstowo, co z kolei nakłada istotne ograniczenia na wyjaśniającą rolę samego pojęcia znaczenia. Przedstawiony w artykule argument za stanowiskiem deflacionistycznym w teorii znaczenia języka publicznego ma uzasadniać tezę o nieistotności pojęcia znaczenia w przyczynowym wyjaśnianiu ludzkich zachowań językowych. Należy tu wspomnieć, że książka Krzysztofa Posłajki poświęcona irrealizmowi semantycznemu *To wszystko nic nie znaczy* (Biblioteka Ośrodka Badań Filozoficznych, 2012, książka jest dostępna w wersji elektronicznej pod adresem: <http://obf.edu.pl/docs/poslajko.pdf>) uczyniła z niego czołowego eksperta od problematyki znaczenia. Fakt, że „Studia” przyciągają tak dobrych specjalistów może tylko cieszyć.

Ogólniejsze zagadnienie rozumienia i użycia znaków rozważa w swojej rozprawie kolejny autor debiutujący na łamach „Studiów”, Piotr Konderak (filozofia i kognitywistyka, Uniwersytet Marii Curie-Skłodowskiej w Lublinie). Jego praca stanowi próbę przedstawiania teorii użycia znaków, która odwołuje się do pewnych pomysłów autorów takich jak David Chalmers i Antonio Damasio. Autor broni m.in. poglądu, że akt przypisywania znakowi znaczenia wymaga zaangażowania tzw. świadomości fenomenalnej oraz że użytkownik znaku musi być zdolny do myślenia o znaku jako znaku – świadomość tego rodzaju jest konieczna m.in. do wstępnego wyodrębnienia spośród ogółu przedmiotów znaków jako obiektów znaczących. Czytelnik zaznajomiony z klasyczną literaturą z zakresu semiotyki łatwo dostrzeże związki rozważań Autora z tezą o przezroczystości znaku (analizowaną m.in. w pracach Husserla, Ossowskiego czy Koja).

Filozofii umysłu w węższym sensie dotyczy tekst Piotra Wilkina (filozofia i logika, Uniwersytet Warszawski, autor ten także debiutuje na łamach „Studiów”), który podejmuje zagadnienie naturalizacji reprezentacji umysłowych. Artykuł stanowi krytyczną analizę prób naturalistycznego ujęcia błędu reprezentacyjnego. Autor przedstawia argumenty mające podważyć przekonanie o możliwości całkowitej naturalizacji reprezentacji poznawczych u człowieka. Zawiera także szkic hybrydowej, społeczno-funkcjonalnej koncepcji reprezentacji. Jest to artykuł, wobec którego nie może przejść obojętnie żaden filozof lub kognitywista zainteresowany zagadnieniem naturalizacji.

Jacek Wawer (filozofia i logika, Uniwersytet Jagielloński) to z kolei autor, który publikuje w „Studiach” swoją drugą pracę (zob. *Jak przy-*

godne i jak a priori są przygodne prawdy a priori?, „Studia Semiotyczne”, t. XXVIII–XXIX). Tym razem rozważa on w swoim tekście zagadnienie tzw. „ustalania indeksu”, które wyłoniło się ze współczesnych dyskusji wokół logik intensjonalnych oraz problemu determinizmu. Jak wskazuje za autorami takimi jak Belnap, rozważanie języków zawierających operatory czasowe oraz modeli, które wprowadzają poza czasem nadania wypowiedzi możliwe historie (i zakładają tym samym jakąś wersję indeterminizmu) sprawia, że zagadnienie ustalania indeksu (albo kontekstu) przestaje mieć oczywiste rozwiązanie. Autor krytycznie omawia w swoim artykule alternatywne rozwiązania tytułowego problemu oraz optuje za konserwatywnym jego rozwiązaniem, zgodnie z którym „czasem kontekstu jest czas terazniejszy względem wypowiedzi, a historią kontekstu jest jedyna historia, która zostanie zrealizowana przez bieg zdarzeń obejmujący rozważaną wypowiedź”.

Z problematyką intensjonalności blisko związany jest kolejny tekst pióra ostatniego autora, który debiutuje na łamach „Studiów”. Maciej Sendłak (filozofia, Uniwersytet Warszawski) podejmuje zagadnienia poprawnej semantycznej analizy nierzeczywistych okresów warunkowych, których poprzedniki stwierdzają niemożliwe stany rzeczy (takich jak: „Gdyby ludzie byli z drewna, to unosiliby się na wodzie”). Zgodnie ze standardową semantyką dla nierzeczywistych okresów warunkowych wszystkie takie zdania są banalnie prawdziwe. Z poglądem tym nie zgadza się wielu nieortodoksyjnie nastawionych filozofów, w tym także autor artykułu. Podejmuje on w tekście próbę odparcia krytycznych argumentów przeciwko stanowiskom nieortodoksyjnym, które sformułował ostatnio Timothy Williamson. Chociaż przyjęcie nieortodoksyjnej interpretacji nierzeczywistych okresów warunkowych wiąże się m.in. z postulowaniem światów niemożliwych, to właśnie nieortodoksyjne stanowiska umożliwiają dużo wnikliwszą i – jak się może wydawać – trafniejszą analizę semantyczną zdań tego rodzaju.

Numer zamyka wnikliwa i obszerna recenzja pióra Marii Matuszkiewicz poświęcona książce *Mere Possibilities* Roberta Stalnackera (Princeton University Press, 2012). Może się w niej Czytelnik zapoznać z interesującą (i kontrowersyjną) teorią, która dopuszcza przygodne istnienie sądów logicznych.

* * *

Chcielibyśmy bardzo podziękować osobom, które pomogły w przygotowaniu niniejszego tomu. Profesor Andrzej Biłat wspierał od samego początku ideę powstania tomu oraz pomagał w kontakcie z autorami. Marcin Będkowski oraz Krystyna Komorowska wykonali doskonałą pracę redakcyjną. Uwagi naszych recenzentów (niestety ze względu na swoją rolę muszą oni pozostać anonimowi) pozwoliły nie tylko wybrać do tomu najlepsze teksty, lecz także znacząco ulepszyć teksty, które pozytywnie przeszły procedurę recenzyjną. Mieliliśmy również wielkie szczęście do autorów, którzy dostarczyli nam bardzo dobre prace oraz wykazali się cierpliwością przy przedłużających się pracach redakcyjnych. Wszystkim tym osobom dziękujemy.

Tadeusz Ciecierski, Tomasz Puczyłowski

Warszawa, czerwiec 2017

PAWEŁ GRABARCZYK*

DIRECTIVAL THEORY OF MEANING RESURRECTED¹

SUMMARY: The first aim of this paper is to remind the reader of a very original theory of meaning which in many aspects has not been surpassed by subsequent theories. The theory in question is Kazimierz Ajdukiewicz's Directival Theory of Meaning. In the first section I present a version of this theory which, I trust, retains the gist of the original but loses its outdated language. In the second section I analyze some problematic consequences of the directival theory (specifically Tarski's counterexample) and show how they can be addressed.

The second aim of this paper is exploiting some of the similarities between the directival theory and later theories of meaning. In the third section I argue that using the directival theory as an interpretative tool enables us to create explications of some of the notoriously vague notions which contemporary theories of meaning employ.

KEYWORDS: Kazimierz Ajdukiewicz, directival theory of meaning, semantics, indeterminacy of translation, Wilfrid Sellars

There are two aims of this paper. The first aim is to remind the reader of a very original theory of meaning which in many aspects has not been surpassed by subsequent theories. The theory in

* University of Lodz, Faculty of Philosophy and History. E-mail: pagrab@gmail.com

¹ This paper is part of the project financed by the Polish National Science Centre, Grant number 2014/15/B/HS1/01928.

question is Kazimierz Ajdukiewicz's Directival Theory of Meaning (henceforth DTM). It was the world's first foray into functional role semantics, predating Wittgensteinian intuitions of "meaning as use" (Wittgenstein 1967) by almost 20 years. Despite this it has never been widely recognized or analyzed outside of Poland (apart from passing remarks by Carnap (Carnap 1959) and Quine (Quine 2013, p. 59)). There are two reasons for this. The first reason is that its original presentation leaves a lot to be desired when it comes to accessibility. In section 1 I present a version of the DTM which, I trust, will retain the gist of the original version but lose its outdated language and will simplify it as much as possible. The second reason is that it was quickly abandoned by Ajdukiewicz because of its counterintuitive consequences. In section 2 I will analyze these consequences (specifically Tarski's counterexample) and show how they can be addressed.

The second aim of this paper is to exploit some of the similarities between the DTM and later theories of meaning, specifically the theories of Wilfrid Sellars, Ned Block, Jerry Fodor and Willard Van Orman Quine. In section 3 I will show that apart from being a theory of meaning DTM can also be used as a pretty robust interpretative tool. I argue that using DTM in this manner not only helps us to understand these theories better but also enables us to create explications of some of the notoriously vague notions these theories employ.

1. DIRECTIVAL THEORY OF MEANING EXPLAINED

The directival theory of meaning was developed by Kazimierz Ajdukiewicz over two papers: *O znaczeniu wyrażen* (*On The Meaning Of Expressions*)² and *Sprache und Sinn* (*Language and Meaning*)³. Although only the latter paper presents the full-blown version of the theory, it is important to remember about the former as it contains some

² The original Polish version has been published in (Ajdukiewicz 1985b) and can also be found in (Ajdukiewicz 1985b) the English translation can be found in (Ajdukiewicz 1978b).

³ The original paper can be found in (Ajdukiewicz 1934), the Polish translation can be found in (Ajdukiewicz 1985a), the English translation can be found in (Ajdukiewicz 1978a).

preliminary considerations that have shed much needed light on assumptions which are crucial for understanding the theory⁴.

It is worth starting with the central intuition that motivated DTM. It is so ubiquitous and common that it could be summed up in a popular slogan: “People do not argue over semantics”. What this means is that sometimes the argument between two sides reaches a point where the sides start to suspect that the disagreement is merely verbal.

What happens next is interesting because of two reasons. The first interesting thing is that (for most of the time) people know how to test their suspicion. Contrary to what might seem to be the obvious solution they do not expect their interlocutors to provide a full definition of the problematic expression. Instead, they try to detect the suspected verbal difference by asking a few key questions about the expression. So, for example, if I was to discover that my interlocutor uses the term “idea” the same way I do, I may start by asking if “ideas” are mental entities. If the answer indicates a difference in usage, it might be enough to decide that the dispute was only verbal, that she meant something different – e.g. platonic ideas.

The second interesting thing is that the moment the two sides discover that the difference was only verbal the disagreement disappears⁵. Most of the time people do not have the motivation to fight with conventions because there is no right or wrong there and some of the conventions are mandatory: either you accept the convention and stay with the community that supports it, or you do not and you are automatically excluded from that community. Starting with these common sense observations Ajdukiewicz presumed that for every noncompound expression there are mandatory conventions and that they are adhered to in the act of confirming certain sentences. When someone knows the meaning of a given expression, and are then asked about it, they have to confirm certain sentences that this expression figures in. And if they refuse to do so, they are excluded from the community of users of this particular expression. Naturally,

⁴ Some researchers consider both papers to be two different versions of the theory (Hanusek 2013).

⁵ Or is vastly diminished. The point here is that it is significantly easier to achieve agreement, even if we have different views on which of the available dictionaries is to be treated as obligatory.

the model examples of these mandatory conventions are analytic sentences. For example, if you refuse to confirm a sentence “A circle is a figure” then you will be denied the knowledge of the meaning of the term “a circle”⁶ and once it is revealed that there is a (admittedly unspecified) number of expressions you do not know the meaning of, you will not be treated as an English speaking person.

The novel idea Ajdukiewicz adds to these observations is his insistence that it should work both ways – if you accept a certain set of sentences which contain a given expression, you can be said to know its meaning. There is nothing more to it – to know the meaning of a word is to have a disposition to confirm its meaning directives (as the specified set of obligatory sentences are to be called). So, what do these meaning directives look like? In general a directive can be presented as a sentence in the form:

If u is a user of a language L and u is in a situation S then u confirms a sentence p .

It is easy to see that the normativity of meaning is built into the directives from the start. Using a simple rule of contraposition we can derive the following consequence: if someone does not confirm sentence p than either they are not in the situation S or they are not a speaker of the language L . It means that if the user is allowed to disregard language directives they are automatically excluded from a given speaking community⁷. One thing to keep in mind is that what we talk about is the act of confirmation of a sentence and not the act of utterance. It is worth pointing this difference out because ignoring it may easily lead to a significant misinterpretation. The theory does not require the user to produce utterances automatically whenever they are in a given situation but only to react accordingly whenever they are asked to confirm the sentence p in a proper situation. Again, analytic sentences are a good example here. We are not expected to walk around and whisper them to ourselves all the time. What is

⁶ Of course you might as well be denied the knowledge of the meaning of the word “a figure” but it will be tested the same way – you will be asked to accept some other sentences the term “a figure” figures in.

⁷ Needless to say it is an idealization. The forbidden behavior would have to be somewhat systematic for her to be really excluded. The important part is that the behavior would be treated as an error and not as an expression of their (even very peculiar) point of view.

expected of us instead is a constant, enduring disposition to confirm them when asked to.

We can now group the meaning directives into three sets depending on the type of situation S .

Directive type	Axiomatic	Inferential	Empirical
Character of S	Anything	Set of sentences	Physical stimulus
Example of S		" p implies q " " p obtains"	The touching of a nerve
Sentence to confirm	" A is identical to A "	" q obtains"	"It hurts!"

Table 1: Types of meaning directives

Now let me characterize the types of directives indicated above.

In the case of axiomatic directives there are no requirements specified as to what situation S has to be. It can be any possible stimulus – verbal, physical or a combination of both. There can be no stimulus at all. The point here is that in each and every situation the user is expected to confirm some of the sentences of her language (such as the identity statement used in the table).

The inferential directives seem to be another intuitive example of the idea of obligatory rules: after all, this is how most of us learn logic – we are told that whenever we confirm a given sentence we have to confirm another, subsequent sentence. If we do not follow those instructions we will not master logic because it is exactly what mastering logic boils down to. This normative aspect of logic works exactly the same way as it is supposed to work in the DTM.

Last but not least, we have empirical directives. It is important to note that the way I explain them here presents the most significant departure from Ajdukiewicz's version. As can be seen in Table 1, I have described the situation S which precedes the confirmation of the sentence p as a physical stimuli. Contrary to this, Ajdukiewicz referred to mental states rather than to their physical causes. But despite the psychological language that he was using most of his examples of empirical directives adhere to physical stimuli and not their mental correlates. Case in point: in the example I have used in the table above Ajdukiewicz talks about the expected confirmation of

the sentence “It hurts!” when a dentist touches the nerve of a patient’s tooth and not about the feeling of pain⁸.

There is an additional difficulty that most of the examples of empirical stimuli lead to. If I am presented with an object and asked to confirm the sentence “This object is red”, I may refrain from doing so because I believe that the lighting in the room is so different from normal lighting that I am no longer sure of the object’s color. It complicates matters because we have to expand the directive by a requirement that the user has a belief that the situation (understood as a state of the environment and the perception apparatus) is typical or normal. The addition of beliefs introduces a hybrid category of directives, a mix between the empirical and the inferential ones, one part of the situation S being a sentence expressing the belief and the other being a stimulus. Ajdukiewicz mentions this complication but does not elaborate on it (Ajdukiewicz 1934). I too am going to skip it in the present exposition of the DTM.

So, how is the notion of meaning to be derived from these three types of directives? Let us assume that we created a list of directives for every noncompound expression of the language. Once we have it, the next step would be to get the notion of synonymy. The intuitive formulation of the relation between the meaning directives and synonymy is this: expressions are synonymous when the meaning directives describe them identically. To present the notion of synonymy in a less metaphoric fashion we have to use an example of a very simple language. Let us say that it contains only the following axiomatic directives:⁹

$$P(a), S(c), R(c), R(d), Q(b), P(b), P(c), Q(a)$$

Now, focus on terms a and b . The interesting thing about them is that if you switch their places – replace every instance of a with b and *vice versa* you will end up with the same list of directives – the only difference being the order of the directives:

⁸The other important reason for preferring physical stimuli over mental states is that it will make our task in section 3 much easier.

⁹That these are axiomatic directives can be easily deduced from their syntactic structure. Only axiomatic directives can be presented as a single sentence.

$$P(b), S(c), R(c), R(d), Q(a), P(a), P(c), Q(b)$$

Using this observation Ajdukiewicz proposed to use this operation of systematic simultaneous replacement of terms to define the notion of synonymy:

Expressions a and b are *synonymous* iff they can be simultaneously replaced in all respective meaning directives without changing the sum of all the meaning directives of the language.

The obvious next step is to use abstraction to obtain the definition of meaning:

The meaning of an expression is the set of all the expressions which are synonymous with it.

It is easy to see that in most cases this definition yields rather disappointing results: in the case of expressions which are not synonymous with any other expression their meaning turns out to be a singleton consisting only of themselves. To counter this, Ajdukiewicz introduces a new (and at the time rather novel) idea: he proposed to define meaning by appealing to the notion of translation. To present it, we will use another example of a simple language, let us call it L . Let L contain the following terms: two one-place predicates: $-P(x)$, $Q(x)$; three constants $-a, b, c$; one zero-place predicate (a sentential constant) Z . Additionally we introduce three symbols which signify physical stimuli: α, β, γ . It is important to stress that these symbols are not parts of L . They symbolize the extra-linguistic element in empirical directives. Now assume that L contains the following directives:

Axiomatic directives:

1. $P(a)$ ¹⁰
2. $P(a) \ \& \ Q(b)$

Inferential directives:

1. $P(a) \ \vDash \ Q(b)$ ¹¹

¹⁰ Understood as: “in every situation confirm the sentence $P(a)$ ” and so on.

¹¹ Understood as: “If you confirm the sentence $P(a)$ you have to confirm the sentence $Q(b)$ ” and so on.

2. $P(a) \ \& \ Q(b) \ \vdash \ Q(c)$ ¹²
3. $Q(b) \ \vdash \ Z$

Empirical directives:

1. $\alpha; Z$ ¹³
2. $\beta; Q(b)$
3. $\gamma; Z$

Having all this we are ready to build something Ajdukiewicz called a language matrix¹⁴. A language matrix is divided into three sections corresponding to three types of directives. In our example they are designated by numerals 1, 2 and 3 in the first column of the table. Horizontally the table is divided into two parts indicating two parts of a directive: the first part (designated by the Roman numeral I) contains the situation specified by the directive (or the lack of a specified situation in the case of axiomatic directives), the second part (designated by the Roman numeral II) contains the sentence which the directive requires to be confirmed. Every sentence put into a language matrix is divided into its constituent parts using the following procedure: the first cell contains the sentence itself, the next cell contains its main connective or a predicate (in the case of an atomic sentence), the next cell contains the first argument of the connective (or an argument of the predicate). Then the same procedure applies to the first argument – we put its main connective first, then its first argument and so on. When we achieve the level of atomic parts we move on to the second argument of the main connective of the sentence we started with. The pattern is repeated for as long as there is nothing more to decompose. If we applied this procedure to our simple language we would end up

¹² You might be surprised that, given the existence of the axiomatic directive 2 and the inferential directive 2 the sentence $Q(c)$ is not an axiomatic directive as well. After all, it is a consequence of these directives taken together. The point of this example is to show that some of the consequences of the language rules are not by itself language rules and can be overlooked by the language user. This characteristic of the DTM will be used later in section 3.

¹³ Understood as “In this situation (when the situation is α) confirm the sentence Z ” and so on (I use an indexical term to stress the extra-linguistic aspect of α).

¹⁴ This part is a substantially modified version of the original example. First of all, I use a modern predicate logic notation and secondly, I present the matrices in a more visual way which I believe makes the whole idea much easier to grasp.

with the following table (note the extra-linguistic part in the left bottom corner).

	I								II							
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
1.1									$P(a)$	P	a					
1.2									$P(a) \& Q(b)$	$\&$	$P(a)$	P	a	$Q(b)$	Q	b
2.1	$P(a)$	P	a						$Q(b)$	Q	b					
2.2	$P(a) \& Q(b)$	$\&$	$P(a)$	P	a	$Q(b)$	Q	b	$Q(c)$	Q	c					
2.3	$Q(b)$	Q	b						Z							
3.1	A								Z							
3.2	B								$Q(b)$	Q	b					
3.3	Γ								Z							

Table 2: The language matrix of L

The main point about a language matrix is that it enables us to extract the structure of the language and abstract away from the actual expressions it uses. We could do that in a variety of ways but I find it the easiest to simply use some sort of visual indication. To extract the structure we are interested in we simply replace the symbols with graphical patterns; let us call it an *expressionless language matrix*.

	I								II							
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
1.1																
1.2																
2.1																
2.2																
2.3																
3.1	A															
3.2	B															
3.3	Γ															

Table 3: The semantic structure of L
(expressionless language matrix)

Now you could fill this table anew using the following rules:

1. You do not change the α, β, γ records as they are extra-linguistic elements of the table.

2. You do not fill the white records.

3. Whenever you put something in the record you have to repeat the same symbol in every record with the same pattern.

Every table obtained this way represents a language, which is translatable to the language we started with. Finally, the idea of a language matrix gives us the possibility to define meaning:

The meaning of a noncompound expression t in the language L is an ordered pair $\langle SL, P \rangle$ consisting of the structure of L (SL) and the set of places t occupies in this structure (P)¹⁵.

As you have seen, the structure can be presented in the form of a language matrix and the set of places a given expression occupies can as well be shown visually. So, for example, the meaning of the expression $Q(b)$ from our table can be presented via the following diagram¹⁶.

											■								
												■							
													■						
														■					
															■				
																■			
																	■		
																		■	
																			■
																			■

It shows that the DTM realizes the noble goal of a reductive, syntactic definition of meaning – the meaning can be literally represented as a shape, which makes it easy to handle mechanically. The fact that what we started with are the acts of confirmation of sentences just adds

¹⁵ The relation of synonymy can still be defined using the notion of mutual exchangeability in meaning directives, just like we did on page 6.

¹⁶ To stress the possibility of representing the meaning of the expression visually I omitted the extra-linguistic parts of the table. It is possible whenever a language matrix is fixed.

a dash of pragmatics to the definition. Because of this, the DTM could not be called a purely syntactic theory. The fact remains, though, that it is a theory in which no part uses any semantic notion. It is an idea entertained by many, but I guess that it is summed up most eloquently by Chomsky:

It is possible that natural language has only syntax and pragmatics; it has a “semantics” only in the sense of “the study of how this instrument, whose formal structure and potentialities of expression are the subject of syntactic investigation is actually put to use in a speech community” (Chomsky 1995).

It is important to realize that even though the language matrix contains an extra-linguistic part, the theory does not stipulate that any of the expressions present in the matrix refer to these extra-linguistic elements. Moreover, even if the theory deals with the confirmation of sentences, in no part does it assume the sentences to be true. You might assume that they are held to be true by the users but it would be an additional assumption the theory does not depend on.

2. DIRECTIVAL THEORY OF MEANING CHALLENGED

Ironically, this attractive feature of the DTM (its independence from reference) is exactly what killed it. It happened because of a very simple example that Alfred Tarski confronted Ajdukiewicz with (he did so in a conversation and it was reported many years later in (Ajdukiewicz 1978c)). Consider a very simple language of predicate logic (with identity) and add to it two new axiomatic directives:

$$A \neq B$$

$$B \neq A$$

A and B are extra-logical constants which appear only in these very directives. The problem is that the two expressions are mutually interchangeable in all the meaning directives of the language (because there are only two such directives and you can mutually replace them). On the other hand, we have to assume that both expressions do not refer to the same object, because it is precisely how we normally interpret the negation of the identity sign. It means that the DTM allows two expressions to have the same meaning but a different reference and it seems that we do not have any means within the

theory to block this unintuitive result because the theory does not say anything about the reference of the expressions¹⁷.

It turns out that in spite of deliberately ignoring all the semantic notions Ajdukiewicz still wanted his theory to be Fregean – the meaning of the expression was supposed to determine its reference. It was so obvious to him that he did not even try to argue for it and remarked only that such a consequence was unacceptable (Ajdukiewicz 1978c). Fortunately, it is a sentiment we do not have to share today as there are at least three ways out of the trouble Tarski's example puts us in – ways which do not force us to abandon the reductive, non-semantic aspect of the DTM.

First of all, we can say that the objection works only because the example language does not contain any empirical directives. If it did, they would have differentiated the terms *A* and *B*. And in the case of uninterpreted languages there is no problem of reference anyway. This is the solution suggested by Ajdukiewicz himself (Ajdukiewicz 1978c). The question of whether this solution is effective is highly debatable though (see section 3).

The second thing we can modify is the simultaneous interchangeability requirement of the synonymy relation. It has been shown that we can modify this requirement and demand only when the terms *A* and *B* can be considered synonymous if and only if it is possible to replace *A* with *B* and then *B* with *A* (but not simultaneously) without changing the character of the directive we applied this procedure to. This means that if something has been an axiomatic directive, it remains an axiomatic directive after the replacement of the term (similarly for the other two types of directives). This solution has some disadvantages, but they will not be discussed here¹⁸.

The third, and perhaps most interesting option, is that we could simply accept and embrace this surprising consequence of the theory

¹⁷ It is worth noting that Tarski's example is very similar in spirit to Fodor and Lepore's objection against functional role semantics. As Fodor and Lepore rightly argue (Fodor, Lepore 1992, p. 170) the price hybrid theories pay for their flexibility is that there is nothing that prevents a given sentence having the inferential role of "4 is a prime number" but the truth conditions of "water is greenish" (as there is no necessary connection between inferential role and truth conditions).

¹⁸ The results in question has been published only recently by (Nowaczyk 2006) and (Buszkowski 2010). Unfortunately both articles are only available in Polish.

– especially that it is not so surprising anymore. After all, this is what Putnam’s Twin Earth thought experiment was set to do – it showed us that we do not have to hold to Fregean intuitions about the relation between meaning and reference (Putnam 1975). Could not we simply decide that a sensible strategy for a theory of meaning is to contain two parallel theories – a theory of reference and a separate theory of meaning which answers the questions about synonymy, translatability and meaningfulness of expressions?

Unfortunately the DTM has more issues than that. Specifically, there are two problematic theses it holds (one of them being an assumption, the other a consequence) which we have to analyze if the theory is to be useful for contemporary philosophers. We will refer to them later, so it might be convenient to label them:

(T1) The meaning of every word in the language changes whenever a new word is added to the vocabulary.

(T2) Syntaxes of all translatable languages have to be perfectly compatible.

(T1) is a direct consequence of definition (D2) presented above. If the meaning of a particular expression is the ordered pair of a language matrix and a set of places the expression figures in, then the meaning changes whenever the matrix changes, and the matrix changes whenever a new expression is added. It is so because the new expression has to have a set of new directives which regulate its usage and these directives have to be added to the language matrix.

(T2) follows on from the way the matrices are built and from the introduced notion of translatability. Whenever a given expression *A* is to be a translation of some term *B*, both expressions have to figure in the same places in identical language matrices. Such a strict notion of translatability does not allow the translatable expression to differ syntactically. To see why it is so let us consider the opposite situation – let us say that we found two expressions which figure in exactly the same places of their respective language matrices, but one of them is atomic and the other is not. There would have to be a place in the second matrix where the second expression was decomposed into its atomic constituents but there would be no such place in the first matrix (because there was nothing to decompose there). But if the matrices are different then the expressions are by definition not translatable.

It is important to stress that neither of these claims present a serious challenge to the theory – they are simply counterintuitive. Nonetheless, I believe that it is worth pointing them out and analyzing ways of dealing with them because, as I hope to show, even small modifications to these claims produce interesting and useful variants of the theory.

In order to understand how we could deal with the thesis T1 we have to introduce an important requirement that Ajdukiewicz added to the theory. As he points out the directival theory can only be formulated for languages which are *coherent and closed*.

A given language is *coherent* if every expression it contains is connected to every other expression (directly or indirectly) *via* meaning directives¹⁹.

In other words – if the language in question is coherent, we should be able to pick any expression and “reach” any other expression by “jumping” from a meaning directive to a meaning directive.

A language is *closed* if for every new expression, which is to be introduced to it, it already contains an expression synonymous with it.

In other words – a closed language is a language that already contains all meanings which can be added to this specific language (as further enrichment would have produced either synonyms or an incoherent language).

The bad news is that Ajdukiewicz’s requirement creates bigger problems than the problem we wanted to solve with its help (thesis T1). The second requirement is simply much too strong – there are no existing closed languages and, what is worse, we could not create a closed language even if we wanted to (see Buszkowski 2010).

3. DIRECTIVAL THEORY OF MEANING RESURRECTED

In the remaining part of this paper I am going to show how we can utilize DTM as an interpretative tool for other theories of meaning – theories which often lack the precision of Ajdukiewicz’s account and

¹⁹ Two expressions are directly connected if they figure together in a single meaning directive. Expressions *A* and *B* are indirectly connected if they are not directly connected but there exists an expression *C* such that *A* and *B* are directly connected to *C*.

which can be seen as sketches DTM fleshes out. What Ajdukiewicz's theory can provide here is showing something which other theories only hint at.

Let us start with a suggestion, which, I hope, will be rather obvious for the reader – the possibility of treating DTM as a theory of narrow content. Let us use the example of Sellars-Block's account²⁰ because the similarity between it and the DTM is striking. Sellars introduced four types of language rules, depending on whether the character of the stimulus provided for the user and her response is linguistic or not (Sellars 1963). There are three obvious possibilities:

1. Extra-linguistic stimulus – linguistic response.
2. Linguistic stimulus – linguistic response.
3. Linguistic stimulus – extra-linguistic response.

There is also a fourth, less obvious option:

4. Any stimulus – linguistic response²¹.

It is not hard to see that 1. can be understood as empirical directives, 2. as inferential directives and 4. as axiomatic directives. There is nothing similar to 3. in the DTM but what prevents us from adding a new type of directive to the theory²²? This new category of directives could be called *imperative directives* – they instruct the speaker to perform a certain action whenever she acknowledges a certain sentence by confirming it.

Now, the idea Block adds to the mix is that language described this way can be understood as a network of inputs and outputs which in turn enables us to define the narrow content of an expression (or its “conceptual role”, as Block prefers to call it) as a role the expression plays in this computational structure (Block 1986). The problem with this account is that, while attractive, it does not show us how exactly a set of user actions (sentence confirmations) translates into a network of interrelated expressions of the language. Is the network just a set of beliefs connected by their inferential roles? If so, which ones – all of them? Maybe they should be decomposed somehow or perhaps even translated into language of thought? It is precisely what language matrices can help us with. They start with a set of pragmatic

²⁰ This account is neatly summarized in (Putnam 1991).

²¹ Sellar calls this type of rule a “free rule”.

²² In fact adding new directive types is a very natural way of extending the theory and deserves further inquiry.

phenomena and then break it down into syntactic constituents of expressions enabling us to see the mechanism that underlies the phenomenon of narrow content.

Speaking of the language of thought – arguably the biggest flaw of this hypothesis is the elusiveness of the language it postulates. What does it look like? What is the ontological status of its expressions? What exactly are its meanings and how can they determine the meanings of natural languages? To see how the DTM could help here let us modify the idea of closed languages and introduce a more liberal (and realistic) notion of *semantically predetermined languages*.

A language is *semantically predetermined* if every new expression introduced to the language is synonymous with a compound expression built from the expressions the language already contains.

What we mean by that is that even if the language does not contain a proper synonym for the new expression, its meaning can be construed out of the language's existing expressions and this is exactly what Fodor assumes (Fodor 1975). The other thing we have to change is (T2) – we have to decide which syntaxes of translatable languages do not have to be identical. Instead, we assume only that the syntaxes are compatible in a sense that the differences they demonstrate are only superficial and what is important is the identity of deep syntactic structures of both languages²³. It is possible that it is a solution Ajdukiewicz tacitly assumed anyway. Consider the way we build language matrices. What we look for in sentences are connectives, their arguments, their ordering and nothing else. Ajdukiewicz was a pioneer of categorial grammars so it is possible that he assumed that a working theory of meaning presumes a developed theory of universal grammar. In other words – categorial grammar could be thought of as a description of the surface grammar of two languages that is general enough so it abstracts away from unimportant details and enables us to represent two superficially syntactically different expressions as expressions of the same type. What we end up with, then, is a theory which fits the language of thought hypothesis quite well because it gives us the answer it lacked – it shows how the semantic structure of the language can be

²³ One notable complication is that the relation between a given language and the language we use to show its deep structure could not be explained by the same notion of translation we use in the DTM, but it is a small price to pay.

construed out of its non-semantic aspects. Moreover, it gives us the much needed model of linguistic structure which contains no actual labels or sentences but is still compatible with many different sets of such labels and predetermines the relations between them. The result is a detailed functional model for LOT. We can postpone the question of what the expressions of this language actually are. Instead we point at an expressionless language matrix (similar to the one presented in Table 3) and say only that LOT is anything that works “like that”. As a functional semantics the DTM is compatible with different answers to the question about actual expressions. They can turn out to be patterns of firing neurons or parts of the brain or whatever else.

Another theory that could benefit from the DTM is Quine’s behavioral theory of meaning. For the DTM to be useful here we would have to modify the requirement of coherency a bit. Let me digress for a second and say a few things about the notion of coherency I introduced earlier, because it proves to be even more useful than Ajdukiewicz had assumed. One disappointing aspect of the DTM I did not talk about is that although it provides the notions of translatability and synonymy, it does not give us any clue as to what it is for a given expression to simply “have a meaning” (as opposed to nonsense words). Does it suffice for an expression to simply be a part of a language matrix? This is where the notion of coherency can help: we can simply assume that an expression is meaningful if it is a part of a coherent language (which means that it is somehow connected to all the other expressions of the language). The problem with this idea is that it renders all the expressions meaningful. Consider the axiomatic directive of identity. The directive instructs the user to confirm every substitution of the formula $x = x$ regardless of the circumstances. What it means is that for every expression of the language there exists a meaning directive of the form $x = x$ where the expression is substituted for x . It follows that every expression is directly connected to the identity sign and the identity sign is directly connected to every other expression in the language. But it means that every expression is indirectly connected to any other expression.

Now, getting back to Quine’s theory – instead of assuming, as we did above, that the meaningfulness of an expression depends on the number of connections to all the other expressions, we should focus only on selected connections, namely on the connections with the

non-linguistic, empirical parts of the table (that is the part specified in empirical directives, denoted by Greek letters in matrices). This way we could easily provide an explication for Quine's stimulus meaning. Note that by doing that we do not have to give up the non-semantic aspect of the theory because Quine's behavioral account does not imply that the expressions refer to stimuli.

If we allow for this modification of the DTM what we get in return is a theory which can be very well understood as a description of the manual constructed by Quine's radical translator. Remember that what the translator was supposed to do was to collect data on sentence confirmation. He collected the sentences which were confirmed in every situation, sentences which were confirmed after certain different sentences were confirmed and sentences which were confirmed whenever the empirical situation was such-and-such (Quine 2013). It is not hard to see that these three sets of data can be treated as our axiomatic, inferential and empirical meaning directives. Once again – the point here is that this convergence of theories goes far beyond a mere analogy. The DTM can be used to explain the idea which was originally rather vague – the idea of a translation manual (as created by a radical translator). The translation manual is a mapping of two different sets of expressions into a common language matrix.

Compatibility between the DTM and Quine's account is so great that we can easily recreate the infamous consequence of the latter theory, namely the indeterminacy of translation thesis. Let us get back to Tarski's counterexample. One way of looking at the problem it poses is that DTM allows for two expressions to be synonymous contrary to the beliefs of language users. If two expressions play the same role in a language (that is: figure in the same directives in the same places) they are synonymous no matter what. They are synonymous even if no one knows about it. They are synonymous even if the language users believe they are not synonymous! The latter happens when two expressions function the same way but one of the directives specifies that they are not identical (in these very words). A very well-known example of this is Putnam's elm/oak distinction (Putnam 1975, p. 226). If a given language²⁴ does not have tools to

²⁴ In Putnam's example they are, of course, dialects. Although DTM has been originally construed as a theory of language and we retained this aspect of the

differentiate between two meanings, than an empty claim that they are different will not change anything. After all – the very claim still says the same thing about each expression. Let us call this peculiar type of synonymy *tacit synonymy*.

To see how it generates the indeterminacy thesis consider two languages: L_1 and L_2 . Let us say that they are translatable (in the specified sense) and that they both contain tacit synonyms: In L_1 , A_1 and B_1 are synonymous and correspondingly in L_2 , A_2 and B_2 are synonymous. Now the problem is that if you assume that A_1 is a translation of A_2 , then, because B_2 is a synonym of A_2 , B_2 is just as good a translation of A_1 as A_2 is. Of course we might rightfully say that it does not matter whether we translate A_1 to A_2 or to B_2 – all of these expressions are synonymous. The only thing we have to remember is that we keep the structure intact (so both languages still contain a pair of synonymous expressions). There is no “fact of the matter” as to which translation is better – both are just as good (unless you treat homonymy as a sufficient reason). But this is exactly how some of the researchers interpret the indeterminacy thesis (Field 2001, p. 282). Indeterminacy of translation does not present any genuine skeptical worry. The only thing it does is to subvert our expectations towards meaning.

Additionally, we have to realize that that the language we analyze might contain synonyms on various levels. It can have one-on-one mappings that we have just discussed but it could also contain surprising mappings where a single expression figures in the same set of places as a different, syntactically complex expression. Let us call it *syntactical tacit synonymy*. This is something Ajdukiewicz did not foresee but there are no reasons as to why such a situation could not happen. For example, we could realize that for every sentence which figures in language directives and contains the word “a rabbit” there is an analogous sentence which contains a complex expression “an organized set of rabbit parts”. The important difference between the current situation and the case of simple tacit synonymy we discussed in the preceding paragraph is that a syntactical tacit synonymy is much more bewildering to us than a normal tacit synonymy. We believe that

theory in this paper, it is worth pointing out that the theory can be easily tailored to function as a theory of idiolects.

there is a difference between rabbits and organized rabbit parts but try as we might, we cannot find directives to support this assumption. Tacit synonymy of syntactically different expressions can then be understood as the main mechanism of indeterminacy of translation.

What I wanted to show in this paper (apart from presenting a version of the DTM better suited for contemporary readers) is that the main reason why DTM had been abandoned by its creator and largely forgotten, namely because of the Tarski objection, is not very serious anymore, because we learned a few lessons along the way and no longer expect the theory of meaning to be also a theory of reference. The other problematic aspects of the DTM can be interpreted as challenges to be met by different, enhanced versions of the theory. Such enhanced versions can then be used as explications of existing ideas, like the notion of *n a r r o w c o n t e n t* (understood as a product of language rules), or as a framework for the behavioral theory of meaning or as a model for the language of thought hypothesis. The Directival Theory of Meaning is worth resurrecting because it can provide much needed details for propositions which are notoriously presented as sketches or outlines of possible future theories rather than as full-blown accounts.

REFERENCES

- Ajdukiewicz, K. (1934), "Sprache und Sinn", *Erkenntnis* 4(1), pp. 100–138.
- Ajdukiewicz, K. (1978a), *Language and Meaning*, in: J. Giedymin (ed.), *The Scientific World-Perspective and Other Essays, 1931–1963*, Dordrecht: Springer, pp. 35–66.
- Ajdukiewicz, K. (1978b), *On The Meaning of Expressions*, in: J. Giedymin (ed.), *The Scientific World-Perspective and Other Essays, 1931–1963*, Dordrecht: Springer, pp. 1–34.
- Ajdukiewicz, K. (1978c), *The Problem of Empiricism and the Concept of Meaning*, in: J. Giedymin (ed.), *The Scientific World-Perspective and Other Essays, 1931–1963*, Dordrecht: Springer, pp. 306–319.
- Ajdukiewicz, K. (1985a), *Język i znaczenie*, in: *Język i znaczenie*, vol. 1, Warszawa: PWN, pp. 145–173.
- Ajdukiewicz, K. (1985b), *O znaczeniu wyrażań*, in: *Język i znaczenie*, vol. 2, Warszawa: PWN, pp. 31–77.
- Block, N. (1986), "Advertisement for a semantics for psychology", *Midwest Studies in Philosophy* 10(1), pp. 615–678.
- Buszkowski, W. (2010), *O równoznaczności wyrażań w ujęciu Ajdukiewicza*, in: J. Grad, J. Sójka, A. Zaporowski (eds.), *Nauka – Kultura – Społeczeństwo. Księga jubile-*

- uszowa dedykowana Profesor Krystynie Zamiarce*, Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Carnap, R. (1959), *Introduction to Semantics, and Formalization of Logic*, Cambridge: Harvard University Press.
- Chomsky, N. (1995), "Language and Nature", *Mind* 104(413), pp. 1–61.
- Field, H. (2001), *Truth and the Absence of Fact*, New York: Oxford University Press.
- Fodor, J.A. (1975), *The Language of Thought*, New York: Crowell.
- Fodor, J.A., Lepore, E. (1992), *Holism: A Shopper's Guide*, Cambridge, Mass.: Blackwell.
- Hanusek, J. (2013), "On a Non-Referential Theory of Meaning for Simple Names Based on Ajdukiewicz's Theory of Meaning", *Logic and Logical Philosophy* 21(3), pp. 253–269.
- Nowaczyk, A. (2006), *Dyrektywalna teoria znaczenia, czyli dramat Filozofa*, in: J. Pelc (ed.), *Sens, prawda, wartość: filozofia języka i nauki w dziełach Kazimierza Ajdukiewicza, Witolda Doroszewskiego, Tadeusza i Janiny Kotarbińskich, Romana Suszki, Władysława Tatarkiewicza*, Warszawa: Biblioteka Myśli Semiotycznej.
- Plato (1998), *Cratylus*, Indianapolis: Hackett Pub. Co.
- Putnam, H. (1975), *The Meaning of "Meaning"*, in: *Mind, Language and Reality: Philosophical Papers*, vol. 2, Cambridge, Mass.: Cambridge University Press, pp. 215–271.
- Putnam, H. (1991), *Representation and reality*, Cambridge, Mass.: MIT Press.
- Quine, W.V.O. (1969), *Ontological Relativity & Other Essays*, New York: Columbia University Press.
- Quine, W.V.O. (2013), *Word and Object*, Cambridge, Mass.: MIT Press.
- Sellars, W. (1963), *Science, Perception and Reality*, London; New York: Routledge and Kegan Paul; Humanities Press.
- Wittgenstein, L. (1967), *Philosophical Investigations*, Oxford: Blackwell.

KRZYSZTOF POSŁAJKO*

SEMANTIC DEFLATIONISM, PUBLIC LANGUAGE MEANING AND CONTEXTUAL STANDARDS OF CORRECTNESS¹

SUMMARY: The paper aims at providing an argument for a deflationary treatment of the notion of public language meaning. The argument is based on the notion of standards of correctness; I will try to show that as correctness assessments are context-involving, the notion of public language meaning cannot be treated as an explanatory one. An elaboration of the argument, using the notion of ground is provided. Finally, I will consider some limitations of the reasoning presented.

KEYWORDS: deflationism, meaning, public language, grounding

INTRODUCTION

The aim of this paper is to provide an argument for the idea that the notion of public language meaning should be treated in a deflationary fashion. The argument is based on the notion of contextual standards of correctness. The argument is also intended as a partial response to the recent objection to deflationism raised by Stephen Schiffer.

*Jagiellonian University, Department of Philosophy. E-mail: krzysztof.poslajko@uj.edu.pl

¹The work on this paper was funded by National Centre of Science, grant no. UMO-2014/15/B/HS1/01928, led by dr Paweł Grabarczyk from University of Lodz. An earlier version of this paper was presented at Context, Cognition and Communication Conference in Warsaw. I am grateful to Paweł Banaś and two anonymous referees for helpful remarks.

First, I am going to provide a working definition of deflationism as applied to semantic notions in general, and elucidate the notion of public language meaning. Then, I am going to present Schiffer's objection to deflationism. After that I'll introduce the notion of standards of correctness, which plays a central role in the debates on normativity of meaning. This notion will be crucial in the argument for deflationism. Afterwards, I am going to present an elaboration of my argument, which is based on Kit Fine's notion of ground. The next part will be devoted to an attempt to extend the line of argumentation to a broader range of phenomena, namely actions based on understanding public language expressions. In the final part of the paper I will show some limitations to the line of reasoning presented and an objection to it.

1. DEFINITION OF DEFLATIONISM

"Deflationism" is a term that has come to stand for a great variety of philosophical views in different areas of inquiry. In the present paper I shall be interested in deflationism understood as a theory concerning the status of semantic notions. The paradigmatic case here is truth – most of the early versions of deflationary theories were created as theories of this concept. However, nowadays it is not uncommon to see deflationism applied to other semantic notions, like meaning or reference (many examples of such approaches and their criticisms are discussed in a collection of essays in Gross, Tebben, Williams 2015).

In what follows I will treat deflationism as a generic position, in principle acceptable to any semantic notion. I will also assume (although it might be controversial) that one can be semantic deflationist locally – i.e. only with respect to one of the semantic notions, while adopting a substantial theory to other such notions, or remaining neutral with respect to them.

Traditionally, deflationism about "S" was understood as the idea that there is no such property as S-hood; in the paradigmatic case of truth deflationism has been for a long time defined as a theory which simply denied that there is such property as being true (see e.g. Strawson 1950).

This definition however, has led to some serious theoretical difficulties. The critics pointed out that it is not feasible to claim that

“true” does not refer to a property. If there is no such property as being true, the argument goes, then it would seem that the predicate “is true” would have an empty extension (see eg. Wright 1992). But this would amount to the claim that there are no true sentences whatsoever. Such a preposterous claim has never been intended to be made by deflationists, who clearly have not intended to develop a version of error theory concerning truth: i.e. a position which would claim that every sentence ascribing a value of truth to another sentence is false. Such a radical theory would most likely be inconsistent. Moreover, an error theory about truth is irreconcilable with our folk intuitions about the truth predicate (as the folk clearly believe that there are indeed true sentences), and most deflationists has been keen on trying to preserve our commonsense intuitions about the truth predicate.

This argument seems to generalize other semantic notions. It would be quite bizarre to claim that there is no such property as reference, if this thesis were to be interpreted as the claim that nothing ever referred to anything. Again, this purely hypothetical position could be accused of inconsistency and it goes without saying that our everyday beliefs about reference defy it.

Considerations of these sorts have led most deflationists to the admission that “being true” indeed is a property, albeit only in minimal sense (see eg. Horwich 1998a). But once deflationists agree that there is such a property as “truth”, they owe us a clear conception of what makes their theory different from substantial theories of truth, which also claim that truth is a property.

The standard move here is to claim that what is characteristic to deflationism is the claim that although “truth” (or other semantic notion) denotes a property, the property in question is not a substantive one. This idea forces the deflationist to propose a criterion for distinguishing between substantive and non-substantive properties. This distinction is usually explicated in terms of explanatory relevance (see e.g. Horwich 1998a, Edwards 2013): substantive properties are thought to be those which are relevant in explanations of phenomena. Conversely, if a certain property is not relevant in explanatory practice then we should treat it as non-substantive.

Generalizing the conclusion from the previous paragraph, we might say that the point of controversy between a deflationist and a proponent of substantial theory concerning a given semantic term

“S” is whether one should treat the property S as explanatory relevant. The deflationist is the one who claims that although we might say that there is something like property S, and that the predicate “S” has a non-empty extension, there is no deep theoretical job for the property; we might use it as a purely logical device, but that is about it. The anti-deflationist, on the other hand, argues that the notion in question is needed for theoretical purposes and that invoking the property in question helps us to genuinely explain important phenomena.

2. PUBLIC LANGUAGE MEANING AND DEFLATIONISM

The semantic notion I wish to focus on is public language meaning. I do not intend to provide a formal definition of this notion, but only an informal elucidation, which I hope would be sufficient for the purposes of the paper. The basic idea is that when we ask about a meaning of an expression, we might distinguish between an idiolectal meaning, which is specific to a single user, and a meaning which is bestowed upon an expression by a wider community. A well-worn example of that distinction is that of malapropisms: in one’s idiolect it might be well true that “eventually” has the same meaning as “actually”, whereas in standard public English of the educated Anglo-American population these two words have clearly different meanings.

It is important to note that the distinction between idiolectal and public language meaning is different from the well-known distinction between a speaker’s meaning and semantic meaning (see Kripke 1977). The speaker’s meaning is the meaning of an expression as used by a speaker in a given context. Both idiolectal meaning and public language meaning are subspecies of semantic meaning, i.e. both are meanings which are systematically attributable to the expression, rather than being properties of individual use. The difference between them lies in the fact that while idiolectal use traces patterns of use of a single individual, the public one traces the patterns of use of a wider community.

Deflationism about public language meaning is, then, a position according to which public language meaning is a non-substantive notion, which means that public language meaning should not be treated as relevant in explanations of any interesting phenomena.

Such a position has already been presented in the literature as the minimalist interpretation of the so-called “sceptical solution” of the sceptical problem presented by Saul Kripke in his reading of Wittgenstein’s remarks of following a rule. As it is well known the “Kripkensteinian” sceptic questioned whether there is any fact that determines the meaning of any expression (Kripke 1982). In the most famous example, the sceptic claimed that nothing determines whether the symbol “+” means “plus”, as there is no way of excluding the possibility that this symbol denotes some other function, say quus, i.e. a function which yields the same results as plus when the arguments are lower than an arbitrary number, and yields 5 in other cases.

Kripke presents his own answer to this challenge and dubs it the “sceptical solution”. Its basic claim is that there indeed are no facts determining meaning of an expression, but nonetheless there is room for claiming that certain semantic attributions are correct (according to the communal standards of correctness). Moreover, Kripke (1982, p. 86) claims that redundancy theory of truth can be applied to such semantic attributions. Some authors have claimed that these remarks are best understood as putting forward a deflationary account of meaning (see e.g. Byrne 1996, Kusch 2006, Wilson 1998). According to them, what Kripkenstein really denies is the existence of robust/substantial semantic fact, but his sceptical solution allows for the existence of deflationary/minimal semantic facts.

In my paper I am not going to engage in exegesis of Kripke’s notoriously vague arguments. Instead, I am about to provide a new argument for the deflationary approach to public language meaning. Although this argument is not intended to be an interpretation of Kripke, it draws some inspiration from his work. But before I do that, I shall look at Stephen Schiffer’s recent critique of deflationism, which, I believe, provides an important dialectical setting for the discussion.

But before we proceed, it is important to note that the definition of deflationism about public language meaning presented in this paper is distinct from the perhaps more well known version of deflationism about meaning, namely the one promoted by Paul Horwich. The basic tenet of Horwich’s theory is that we should explain meaning in terms of a basic acceptance property, which in turns allows us to claim that the primary meaning of an expression is a concept expressed by it (Horwich 1998b, p. 45–46). Such a theory allows him to describe

meanings in terms of biconditionals like “dog” means DOG, where DOG is a concept.

The difference between the approaches preferred by Horwich and the one I want to pursue here, stems, in my opinion, from the fact that I am primarily interested in the metaphysical status of the putative meaning-property, i.e. in a meta-semantic problem, while Horwich wants to provide a (first-order) theory of meaning. In my opinion causal relevance is a plausible candidate for a test to distinguish between those properties which are to be treated in a deflationary manner and those properties which are to be treated in a strongly realist fashion. Thus, focusing on a question of whether public language meaning is causally relevant is the best way to answer the meta-semantic question whether we are dealing with a “substantial property” here.

Paradoxically enough, on my definition Horwich’s theory turns out to be a non-deflationary one, as he admits that there are indeed such things as meaning-properties; moreover, these properties have an underlying nature and those “underlying natures of meaning-properties are basic regularities of use, explanatorily fundamental generalizations about the circumstances in which words occur” (Horwich 1995, p. 356). The complaint that Horwich’s theory of meaning is not in fact deflationary was raised by Huw Price (1998, p. 111). Price claimed that Horwich’s use of “deflationism” is significantly different when it is applied to “meaning” and not to “truth”.

Horwich seems to be mostly interested in providing a philosophical account of meaning and less with its metaphysical implications. I have no intention of providing such an account. Hence, I will not try to engage in the debate, whether, for example, it is possible to characterize meaning of an expression using a biconditional modeled on the T-schema of Tarski (see Horwich 1998b, p. 14). On my take, deflationism is a negative meta-semantic thesis and is not inherently tied to any account of meaning. I think adopting such definition of public language deflationism, although it might differ from other accounts of that position, is theoretically fruitful, as it takes public meaning deflationism a special case of a generic position. This position deserves critical attention, especially in the light of the recent Schiffer’s critique.

3. SCHIFFER'S WORRY

In his *Deflationist Theories of Truth, Meaning, and Content* (forthcoming) Stephen Schiffer argues against the idea that semantic notions should be given deflationary treatment. He opposes “radical deflationism” – a strictly defined, globally applicable, hypothetical position (based, to an extent, on Harty Field’s views). So, my defense of a modest, local, deflationism about public language meaning is not in direct opposition to Schiffer’s work.

However, Schiffer’s arguments provide, in my opinion, a substantial challenge to all forms of deflationism – even those more locally focused. The line of argumentation provided in his paper is fairly intricate, but a quite simple, yet powerful argument can be extracted from it, and it is a one that all deflationists should take seriously.

According to Schiffer, the deflationists claim that it is possible to explain human language-related behaviour without referring to any semantic properties. But for Schiffer such a project is unrealistic. In everyday practice it is perfectly normal to explain human behaviour by resorting to semantic properties of the expressions used. And there is no principled reason to treat such explanations as defective (apart from general worries about causal exclusion, which Schiffer dismisses). The other worry is that deflationism provides us with no workable alternative to the common practice; in those cases when we normally appeal to semantic properties, we do not have any practically applicable methods of explaining human behaviour other than the ones that we actually employ, and these are laden with semantic properties.

In my paper I am going to focus on the first part of the challenge. At first glance, the Schiffer’s worry might look as a pretty weak argument, as it relies on description of *de facto* existing explanatory practices (and who can be sure that our actual, pre-scientific ways of explaining phenomena are above criticism?). But I think this is indeed quite a powerful argument. It aims to show that deflationism is an under-motivated position, as it provides no reason to think that appeals to the semantic in explanatory practice are defective. And that the alternative – namely substantial theories of the semantic – have had the advantage of already being tried in working practice.

To counter this line of reasoning, a deflationist must present an argument which would provide motivation for their position.

In what follows I am going to provide an argument which aims to counter the intuitions Schiffer's argument wanted to induce. At the heart of Schiffer's argument seems to be that we must treat meaning as substantial as it plays an important causal-explanatory role in psychological explanations. So, my argument would aim to show that, appearances to the contrary, the notion of public language meaning plays no important role in causally explaining human-language related behaviour. This argument will be based on the notion of standards of correctness, which is central to the contemporary debates about normativity of meaning.

4. STANDARDS OF CORRECTNESS

The idea that expressions of public language have conditions of correct use is central to the debates on normativity of meaning. The claim that meaning is normative, once considered obvious (see e.g. Kripke 1982, McDowell 1984), has been subject to many criticisms more recently (see e.g. Hattiangadi 2006, Glüer, Pagin 1999). At the heart of the debate lies the question whether meaning is normative in a "strong" or "philosophically interesting" sense. There are, of course, many ways one might precisify the normativity claim and different arguments have been waged for and against the normativist thesis (for a recent defense of normativism see e.g. Whiting 2007).

What is curious about the debate is that both sides of it seem to agree on a basic intuition that there is something like correct and incorrect use of language. (The only prominent philosopher who had qualms about this thesis was probably Davidson (2005)). The basic idea is quite straightforward: when a user of a public language uses a certain expression, we, as other users of the same language are entitled to judge this use as correct or incorrect, according to the semantic norms of the language in question. This shared assumption is central to the argument that I am going to present.

This fact might well be regarded as constitutive of notion of public language meaning. It is only possible to claim that the phenomenon we are dealing with is indeed a public language if there are standards of correct use associated with it. (This might, at least partially, explain why Davidson ended up claiming that there is no such thing as

language (Davidson 2005) – as he denied that there is such a thing as standards of correct use).

The observation that there are standards of correctness might seem relatively trivial and not particularly relevant to the deflationism debate. But in my opinion this is a crucial fact. I claim that standards of correctness of any public language are context-involving, in the sense that they include factors external to the current, internal state of the speaker. In order to appraise someone's use of language we must look beyond what is, at the moment of an utterance, going on in the head of the speaker.

I should try to argue for this claim by way of analogy. It is widely accepted in the literature on normativity of meaning that semantic norms can be compared to institutional ones (this is accepted by normativist and anti-normativists alike). Hattiangadi (2006, p. 63) made an analogy with a theme park where there is a rule stating that only kids of a certain height can go on the ride. This example serves Hattiangadi to criticize normativism; she focuses her attention on the observation that in this case the height of a child is a purely naturalistic characteristic of her/him.

Still, this example can be used to highlight a different aspect of the correctness condition thesis. If we look only at the purely internal characteristic of the child then we are in no position to judge whether she or he is of the "right height" – we might only be able to provide with a purely physical description of the child. In order to get to know whether we are dealing with a case that is "correct" according to the rules that are in force, we must look at other factors than the subject itself (in this case we must, obviously, look at the regulations of the theme park).

I think this observation generalizes to all cases of institutional correctness. Whenever there are some institutional rules in force (no matter whether trivial or serious) that allow us to judge certain actions as correct or not, the judgment must be based on comparing an agent's actions with the rules in question. These rules must refer to at least some factors external to the agents which are being judged as acting correctly or not. This is crucial because otherwise it would be impossible for one to act incorrectly. And this very possibility of incorrectness is something which makes the very assessment possible. For if it were impossible for one to act incorrectly, the very notion

of correctness in this context would have no sense. Wittgenstein has famously described such a situation as the one in which “whatever is going to seem right to me is right. And that only means that here we can’t talk about «right»” (Wittgenstein 1953, § 258).

So, if we agree that language rules are akin to institutional rules, then we should also admit that the linguistic norms are in a way external to the speakers. This observation seems obvious for every proponent of externalist theories of meaning. According to externalism the facts that determine correctness conditions for language use are external to the speaker, as they include either social facts (as social externalist, of the type of Burge proposed) or facts about the kinds of things that are in the physical surroundings of the language user (as natural-kind externalists of the Putnamian kind assert).

But for internalists the thesis that standards of semantic correctness are context involving might not be that obvious. For the theorist of internalist inclination wants to explain meaning purely in terms of psychological states of the speakers. However, in my opinion, even someone who believes that language meaning is determined by purely psychological factors must admit that some factors, which are relevant to the assessment of expressions, are in a way external to the agent using language in a certain situation. This is due to the fact that even an internalist wants to maintain the distinction between correct and incorrect use.

Therefore, what such a theorist needs is a distinction between the psychological state of the user while making an utterance and a psychological state, which determines the correctness conditions for the use. I think that this distinction is implicit in most internalistic theories of meaning. Usually, it is introduced by postulating a time difference between the act of use (and associated psychological states) and meaning-determining psychological states. Put simply, the internalist usually claims that meaning is determined by meaning-intentions, which are made previous to the acts of use. What serves as the standards of correctness for my current use are the meaning intentions which I have made in the past. And this time difference allows for that correct/incorrect distinction. For I might presently act in a different manner than I intended in the past.

This time distinction lies at the heart of many of the examples Kripke offers in his discussion on rule-following. Even the most

famous “plus” – “quus” example plays on the fact that my present use of the symbol “+” might deviate from my previous intention: what makes the putative subject, and who uses the symbol in a quus-like way err in the fact that she is unfaithful to her previous intention to use the symbol in a “standard” way.

The general picture of meaning which I am going to presuppose in the next sections of the paper might be then described as broadly externalist. This broad conception of externalism includes many conceptions of what might have been traditionally described as internalist. The position I am putting forward is to a great extent a schematic one – it insists only on the claim that whenever we want to ascribe a public meaning to a certain expression we must implicitly accept that there is something external to the occurrent psychofunctional state of the speaker which is to be taken as a standard of correctness. But this schematic theory remains neutral to the question of what these standards of correctness are in particular cases. Even on a more general level the conception presupposed in this paper remains neutral to the question whether say, Kripke’s conception of natural kind terms is the correct theory of reference for terms like “gold” or “water”. So, the phrase “contextual standards of correctness” should be treated as a sort of theoretical place-holder, whereby various externalist theories of meaning might fill in different ways.

It might also be useful to distinguish between two general kinds of broadly externalist approaches²: according to the first it is the standards of correctness that are external to the speaker. In the other what is external to the speaker are the norms stating what contextual elements are to be taken into consideration when assessing a certain utterance. My position is obviously externalist in the first sense; I claim that for each language use there is something “outside the head” of the speaker with which his use is to be compared. The second sort of externalism claims that the norms of correctness are constituted externally – for example by the societal agreement. This version of externalism seems to be plausible when we theorize about public language, but, as I want to stress, this is not an assumption which is needed in order for the argument of the next section to be sound.

² I am grateful to an anonymous referee for drawing my attention to this point.

5. THE ARGUMENT

In this section I am going to provide an argument to the effect that public language meaning should be given a deflationary treatment, which is based on the premise that meaning involves contextual standards of correctness.

First, let me introduce some definitions. I will use “*E*” to denote an expression fact, i.e. the fact that a particular language user used a given expression at a particular occasion. “*M*” will be used to denote the fact (or the totality of facts) that determines the public language meaning of the expression used in *E*.

This description of “*M*” is deliberately vague, as I want to be as non-committal as possible with regard to the different theories of meaning. I shall not argue that either of the numerous theories of meaning is correct or not. Rather, I should use “*M*” as a sort of place-holder, which denotes states postulated by whichever theory of meaning comes out right.

I should understand “*P*” as a psychological-functional state that is causally responsible for the agent’s utterance in *E*. This again is a vague description, as the exact description of what is the character of states that are causally responsible for linguistic utterances is still largely unknown. Still, even if *P* and *M* are only vaguely characterized, I think it is possible to try to establish certain truths about relations between them.

My hypotheses concerning the relations between these two kinds of facts are the following:

First and foremost, *M* cannot be identified with *P*. This is because, as the second claim goes, *P* is doing all the causal-explanatory work, and *M* does none.

The transition from the second thesis to the first one is fairly straightforward. If two putative facts differ when it comes to their causal-explanatory role, then we might safely assume that we are indeed dealing with two different facts and any attempt to identify them would be mistaken.

So, the crucial task is to justify the second thesis that it is *P* that is relevant in providing causal explanation to the *Es*, while *M* is not. Again the first part of the task seems relatively easy: *Ps* are causally relevant to the linguistic production by definition. So, what needs to be justified is the claim that *Ms* are not.

This can be supported by an observation that in a given situation we can keep *P* fixed, while *M* changes – a subject can be in the same current, internal state (and thus produce the same expressions) and mean different things, depending on context.

Let us consider the plus vs. quus example. Kripke (1982, p. 8) invites us to consider a counterfactual situation in which the “+” sign really means quus not plus. In such a situation a subject might have used the sign “+” in a way as we actually use it: namely, as if the symbol denoted addition. Such a person would commit an error according to the standards which are in force in her public language, but it is quite possible that her internal psycho-functional state at the time of making the utterance would be identical with someone’s from our linguistic community which would perform a standard addition.

So, there is a possibility of there being two persons who are identical with respect to their linguistic behaviour and psycho-functional causes of it, but whose expressions in the relevant situations have different meanings. This is obviously a direct consequence of the claim that the standards of correctness are contextually determined. In this case these contextual factors might include community agreement, previous intentions or objective mathematical facts.

Examples might be multiplied. The famed Burge’s thought experiment of arthritis can be used to prove a similar point – depending on the contextual factors, the patient who classifies any pain in the thigh as “arthritis” (Burge 1979) might be treated as using the word correctly or not. In our actual community this is of course an incorrect use, but it is not hard to imagine a different community, in which “arthritis” is used in a way the discussed subject uses it. In all such cases it is not the psycho-functional state of the speaker that influences the meaning, but rather external, contextual factors.

Now, the converse situation is also possible. We might easily imagine two subjects whose utterances have the same meanings (so we have identical *M*s), while their *P*s are different. This is because, once we allow for the possibility of error, we must admit that the psycho-functional which lead to correct and erroneous linguistic use are indeed quite different (the psycho-functional state which leads one to use “+” as a quus-denoting symbol is obviously rather different from one which leads the “normal” user who uses “+” to simply add). But we must admit that when we have two uses of the same expressions made

in the context of the same public language then they have the same meaning, even though one of the uses is incorrect. So, in the example discussed, the fact that someone uses the “+” symbol incorrectly does not (in a normal situation) change the public language meaning of the symbol. It still means plus, even if an erratic user uses it in a quus-like pattern.

This is important, because the whole idea of meaning involving correctness conditions leads inevitably to the conclusion that even wildly erring usage does not change the meaning of the expression used. If this was not so, we would lose the possibility of error: if deviation from the standards of correctness led to the alteration of meaning of expressions then it would be impossible to use linguistic expressions incorrectly.

These considerations prove that there are two kinds of possible situations. In the first type, there are two possible subjects, who share the same P-state, but their expressions have different meanings. In the situation of the second type, there are possible subjects who use the same expression with the same meaning but their psycho-functional states differ. So, it is possible to have the same *Ps* with different *Ms* and *vice versa*. But, what is crucial, in both of the situations it is the change in *Ps* which causes the change in behaviour. Change of meaning, which is not accompanied by the change in the psycho-functional state of the user, has, in itself, no causal impact on linguistic behaviour. Additionally, in order to cause the change in the use, the change in psycho-functional state does not need to be accompanied by a change in meaning.

This shows clearly, in my opinion, that we should take psycho-functional states rather than public language meanings of the expressions to be the causes of linguistic behaviour. But, this conclusion seems enough to justify a deflationary approach to public language meaning (as defined in the section II).

It is important to note that the argument presented is not a straightforward variant of the causal-exclusion argument, which has been extensively discussed in the philosophy of mind. I do not intend to claim that only physical or “basic” properties are causally relevant. On the contrary, I am open to the possibility that psycho-functional characteristics might not be reducible to the physical ones. The contrast between the psycho-functional properties and the semantic

ones is not the contrast between “ontological levels”. It is rather a matter of granularity of descriptions. When we describe expressions as correct and incorrect, we describe them taking a broader context into account, while descriptions of psycho-functional states abstract from the contextual elements.

6. ELABORATION

In this section, I am going to present a simple metaphysical model which is an elaboration of the argument presented above and which would aim at explaining two things. Firstly, why we should treat public language meaning in a deflationary way, and secondly, why we treat public language meaning as explanatory in our everyday practice.

The model will use Kit Fine’s notion of ground (Fine 2001). According to Fine, the relation of grounding is a basic metaphysical one: if *A* grounds *B*, then *B* obtains in virtue of *A*. This relation eschews a straightforward definition as it is metaphysically basic. Still, it can illuminate the question of realism. According to Fine, we should treat certain propositions in a realist fashion when they are either metaphysically basic and factual (Fine 2001, p. 17) or are grounded in some basic and factual propositions. But if there are no real grounds for certain propositions then we might claim that these are not factual propositions. It is an important feature of Fine’s proposal that it makes room for grounding relation between non-factual elements as well (Fine 2001, p. 17). So, when we are dealing with a non-factual proposition we might make hypotheses about which constitutive elements of a given proposition make it non-factual.

Let us try to apply the notion of ground to the phenomena discussed in this paper. My hypothesis is that *P* (a psycho-functional state of the speaker) is a partial ground for *M* (the meaning of the expression used). But it is important to bear in mind that it is only partial ground. The other fact that partially grounds *M*-facts concerns the contextual factors which serve as correctness conditions.

Both *P*-facts and contextual factors might be treated as factual. Yet I claim that *M*s are non-factual, even though they are grounded in *P*s and contextual factors, which are both factual. So, in order to support the claim that *M*s are non-factual, it is necessary to postulate a non-factual element which also grounds them.

In my opinion such a non-factual element is the relation between *E* and, by extension, *P* and the contextual standards of correctness. This is a relation of “being a standard of correctness for”. Even if we take that the terms of said relation to be perfectly factual in Fine’s sense, there seems to be little motivation to take the relation itself to be factual. That a certain element of the context is taken to provide the benchmark of correct use of a given expression seems to be an utterly conventional matter. Moreover, this relation seems to have no causal-explanatory import.

Such relations might be taken to ground the putative meaning-facts. The fact that a certain expression means something is rooted in the relation which binds the expression to the standard of correctness; that my utterance of the symbol “+” in a given context, means plus, is grounded by the relation of this utterance to the standard of correctness (say, my previous meaning-intention). Should this relation be different, the meaning of my utterance would be different as well.

This model, in my opinion, allows us to elevate the worry presented by Schiffer. The question was: Why do we treat public language meaning as explanatory, when it is not, at least according to the deflationists? And the answer is: Because public language meaning is partially grounded by something that really plays the causal explanatory work – namely the psycho-functional states of the speakers. And it seems to be quite normal that in everyday explanations we treat “broader” facts as explanations, especially in the situations when we lack access to the “fine-grained” facts. We do not normally know anything about the psycho-functional states of ourselves and fellow language users, so we resort to explanations in terms of public language meaning. In doing so, we tacitly assume that these meaning facts are somehow rooted in “something in the head” of the speaker, which is the genuine cause of their behaviour.

Schiffer might be perfectly right that we have no realistic alternative to meaning-based explanations, which could be used in everyday practice. Explanations resorting to the psycho-functional states might be practically unattainable. However, I do not think this is a fatal objection to deflationism, as it is meant to be a metaphysical position regarding the nature of semantic predicates, not a practically applicable theory.

7. UNDERSTANDING AND ACTION: REMAINING PROBLEMS

The arguments presented in this paper might be easily attacked for not being general enough. They might be said to show that public language meaning is not relevant in providing causal explanations of linguistic production. But this is not the only possible use of meaning in explaining human behaviour. To my knowledge, none of the existing inflationist theorists of meaning have treated the role of meaning in explaining linguistic production as the main reason for treating the public language meaning in substantial fashion. But, I believe, the argument I presented against treating public language meaning as explanatory in the context of language production can be applied to other cases, where one might want to treat this concept as relevant in causal psychological explanation.

For Schiffer the central observation speaking in favour of treating the semantic properties as substantial was that we explain action by reference to the fact that a person understands a certain expression in a certain way. When applied to the problem of status of public language meaning, Schiffer's insight might be understood as follows: the fact of understanding, which explains some action of some objects, stems from the fact that the expressions mean something in a given public language. Thus, public language meaning plays an important role in explaining behaviour.

This might sound terribly complicated, but the phenomenon is in fact quite easy and commonplace. For example, when we want to know why the children in the classroom sat down it is perfectly legitimate to say that they did so because the teacher said to them "asseyez vous" and this phrase means "please sit down" in French.

Meaning can thus enter the explanation of action differently than by explaining linguistic production. But if this kind of explanation is a legitimate one, then deflationism about public language meaning is in serious trouble, because it turns out that meaning is actually needed in explanation of some language-related phenomena.

The question then arises whether the argument presented in the previous sections of this paper can also be generalized as to cover the cases in which meaning is used to explain actions which stem from understanding expressions of public language. In what follows I will try to formulate such an argument.

The reasoning will be similar in spirit to the one presented in the section 5, and it will also be based on the notion of standards of correctness. This is because understanding public language expressions is subject to the assessment in terms of correctness, in a similar way linguistic production is. A subject might understand a certain public language expression correctly or not, and this observation seems to be central to the notion of a public language meaning.

The possibility of error is clearly visible when we focus on understanding expressions made in a foreign public language – it is quite common for people who are not native speakers to misunderstand expressions of a given language. But even within our own native language there is always the possibility of understanding an expression differently than in a way prescribed by standards of correctness operating in said public language.

In such situations we should distinguish between the meaning of the public expression used and the act of understanding, which is a psycho-functional state of the speaker. Again, I should argue that these two facts must be considered as distinct. Moreover, it is the psycho-functional state of understanding that is causally responsible for the actions of the users. I shall try to prove this using an example.

Take Tom, a native English user who is quite ignorant of the vernacular used to denote different kinds of seafood. He sees “crayfish” on a restaurant menu and understands this expression as a name of a kind of fish dish. As he strongly dislikes fish, he decides not to order. However, he is a great fan of seafood, and if he were to believe that the dish is a kind of seafood he would most likely order it. But misinterpreting the expression of his own public language prevented him from acting on his preferences.

Again, the examples might be multiplied, but I guess it is not necessary. It is quite easy to note that the fact that misunderstandings are possible leads to the conclusion that the subjective act of understanding must be taken to be distinct from the public language meaning and whatever constitutes it. And when it comes to its relevance in causal explanations it is the subjective act of understanding which might reasonably claim priority; in the situations when one misunderstands the expressions the subject will act on her subjective psycho-functional state.

A public language meaning that is not mediated by the act of subjective understanding seems to have no direct influence on the

actions of the subject. Therefore, the deflationary position concerning this notion seems justified, even if we take into account the phenomenon of actions based on understanding expressions of public language.

8. THE OVER-GENERALIZATION CHALLENGE

The argument presented in this paper can be also attacked for being, as it were, too general. The problem is that the line of argumentation presented in section 5 can be quite easily extended to other domains, in which the conclusion might seem implausible³.

The basic idea of the argument, to put it briefly, was that meaning is, at least partly, determined by contextual factors. And given this fact, we might observe that meaning cannot be thought to be causally responsible for actions of language users. This is because the mere change in contextual factors does not, by itself, change the behaviour. The change in linguistic action is brought upon by the change in the functional-internal state of the speaker. And this state cannot be identified with meaning.

The worry is that a similar argument can be produced in all contexts of institutional norms. Every fact that an institutional norm is in force is analogous to linguistic meaning in that respect that it involves a relation of the subject whose behaviour is governed by a given norm to some contextual standard of correctness. Let us take a standard example of institutional norms – road traffic rules. When we say that someone acted incorrectly according to the traffic rules, we compare the subject's behaviour with some contextual standard. The same behaviour can be described as correct or not depending on the context in which the assessment is made (driving on the left is correct in the UK but wrong in continental Europe and so on).

Again, as in the case of meaning, the change in context does not, by itself, change behaviour. This leads to the conclusion that institutional rules are not explanatorily relevant (when contrasted with psycho-functional states of the subjects). Consequently, we should claim that all institutional rules are to be treated in a deflationary fashion. But this might seem counterintuitive: it would mean that institutional rules do not have any impact on the actions of people.

³ I am very grateful to an anonymous referee for raising this point.

There are two possible ways to answer this challenge. The first would be a direct rebuke to show that the analogy does not hold – that there is a deep theoretical difference between the way language operates and the way other institutional discourses and facts do. This would be a strategy of containment of deflationism to the linguistic realm.

Unfortunately, I do not see how this could be done. The argument presented above in no way is based on peculiar characteristics of language. The argument relied only on the fact that the norms of language are institutional ones. So, if the argument is correct it should indeed be generalized to all forms of institutional rules. Thus, I must bite the bullet and say that my argument leads to global deflationism about the institutional.

This is certainly an implausible conclusion for many theorists, but it might be noted that a deflationary approach to institutional and legal facts might not be an entirely groundless position and there are philosophers who seem to endorse it. One recent example of such an approach might be found in Thomasson (2013). For her, deflationism about the institutional and the legal is a welcome consequence of her globally deflationary approach to metaphysics. In the context of the philosophy of law, James Coleman (1995) argued that Dworkin's views could be viewed as a form of deflationism.

Of course, the question whether the deflationary account of the legal and the institutional in general is an acceptable one is extremely puzzling, and answering it would require a separate paper. But I want to stress that even though deflationism in these areas might seem intuitively implausible, it seems to be a real option on the theoretical level.

9. CONCLUSIONS

The fact that the notion of public language meaning essentially involves standards of correctness allowed me to present an argument to the effect that public language meaning is not causally explanatory, even though in everyday use we might treat it as it actually were. This reasoning applies both to explanations of linguistic production and actions based on understanding, which we would normally make using the notion of public language meaning. This conclusion gives us

a strong reason to accept the deflationary account of public language meaning, as the point of contention between deflationists, as defined in this paper, and proponents of a substantial theory regarding a given semantic notion is whether the notion in question is relevant in causal explanations of the phenomena.

However, the line of argumentation presented in this paper has some serious limitations. Firstly, it might well be the case that there are some other ways in which the notion of public language meaning comes into explanatory practice, and the kind of argument which has been developed above has no application to them. So, the argument might be, at best, treated as a shift of burden of proof. The adherent of substantial theory of public language meaning must, in response to it, show which phenomena need explanation in terms of this notion.

The other limitation of the argument is that it is, in a way, a local one. It does not extend to semantic notions other than public language meaning, which leaves open the question of whether, for example, idiolectal meaning or semantic properties of propositional attitudes should not be treated in a substantive fashion. Moreover, the line of reasoning presented here relies on the notion of an internal psychofunctional state which is assumed to explain the language-related behaviour. But nothing in what has been argued for suggests that this kind of state cannot have semantic properties. But if this is so, the only upshot of this paper would be that substantial semantic properties must be located on the psychological level and not on the level of public language meaning.

This might seem too modest a conclusion for a deflationist, as it leaves room for a substantial account of at least some semantic properties. So, we are left with the question; is it possible to mount a more general argument which would show that semantic properties are not substantial ones? This is an extremely complicated issue and I will not try to resolve it in the present paper. Still, I believe that even a partially applicable argument can shed some light on the immensely intricate problem of deflationism.

REFERENCES

- Burge, T. (1979), *Individualism and the Mental*, "Midwest Studies in Philosophy" 4, pp. 73–121.
- Byrne, A. (1996), *On Misinterpreting Kripke's Wittgenstein*, "Philosophy and Phenomenological Research" 56, pp. 339–343.
- Coleman, J. (1995), *Truth and Objectivity in Law*, "Legal Theory" 1, pp. 33–68.
- Davidson, D. (2005), *Nice Derangement of Epitaphs*, in: *Truth, Language, and History*, Oxford: Clarendon Press, pp. 89–108.
- Edwards, D. (2013), *Truth as a Substantive Property*, "Australasian Journal of Philosophy" 91(2), pp. 279–294.
- Fine, K. (2001), *The Question of Realism*, "Philosophers' Imprint" 1(1), pp. 1–30.
- Glüer, K., Pagin, P. (1999), *Rules of Meaning and Practical Reasoning*, "Synthese" 117, pp. 207–227.
- Gross, S., Tebbenm N., Williams, M. (eds.) (2015), *Meaning without Representation*, Oxford: Oxford UP.
- Hattiangadi, A. (2006), *Is Meaning Normative?*, "Mind and Language" 21, pp. 220–240.
- Horwich, P. (1995), *Meaning, Use and Truth*, "Mind" 104, pp. 355–68.
- Horwich, P. (1998a), *Truth*, Oxford: Oxford UP.
- Horwich, P. (1998b), *Meaning*, Oxford: Oxford UP.
- Kripke, S. (1982), *Wittgenstein on Rules and Private Language*, Cambridge, Mass: Harvard UP.
- Kripke, S. (1977), *Speaker's Reference and Semantic Reference*, in: P.A. French, T.E. Uehling Jr, H.K. Wettstein (eds.), *Studies in the Philosophy of Language*, Minnesota: University of Minnesota Press, pp. 255–296.
- Kusch, M. (2006), *A Sceptical Guide to Meaning and Rules*, Chesham: Acumen.
- McDowell, J. (1984), *Wittgenstein on Following a Rule*, "Synthese" 58, pp. 326–363.
- Price, H. (1998), *What Should a Deflationist about Truth Say about Meaning?*, "Philosophical Issues" 8, pp. 107–115.
- Schiffer, S. (forthcoming), *Deflationist Theories of Truth, Meaning, and Content*, in: B. Hale, C. Wright, A. Miller (eds.), *A Companion to the Philosophy of Language*, 2nd ed., Oxford: Willey-Blackwell.
- Strawson, P. (1950), *Truth*, "Proceedings of the Aristotelian Society" 24, pp. 129–156.
- Thomasson, A. (2013), *Fictionalism versus Deflationism*, "Mind" 122, pp. 1023–1051.
- Whiting, D. (2007), *The Normativity of Meaning Defended*, "Analysis" 67, pp. 133–140.
- Wilson, G. (1998), *Semantic Realism and Kripke's Wittgenstein*, "Philosophy and Phenomenological Research" 58, pp. 99–122.
- Wittgenstein, L. (1953), *Philosophical Investigations*, Oxford: Blackwell.
- Wright, C. (1992), *Truth and Objectivity*, Oxford: Oxford UP.

PIOTR KONDERAK*

THE CONSCIOUS SEMIOTIC MIND

SUMMARY: The paper discusses possible roles of consciousness in a semiotic (meaning-making) activity of a cognitive agent. The discussion, we claim, is based on two related approaches to consciousness: on Chalmers' theory of phenomenal and psychological consciousness and on Damasio's neural theory, which draws a distinction between core and extended consciousness.

Two stages of cognitive-semiotic processing are discussed: the moment of perception of a sign as a meaningful entity and the metasemiotic processes understood as the human capacity to reflect on signs and their usage, analyse and control processes of recognition, interpretation of signs and to detect and correct errors in semiotic activity.

In the case of the first stage, it is argued that signs as meaningful entities have a distinctly experiential character. The feeling of meaningfulness is a result of phenomenal consciousness, in particular a result of the so-called valuation features of phenomenal experience. I claim that this aspect of cognitive-semiotic activity is possible owing to a special neural mechanism called a semiotic marker.

It is argued that semiotic systems have to be able to use signs as signs, i.e. they should display some metacognitive capacities, in particular an ability to analyse semiosis at a metalevel. It is argued that such metasemiosis is dependent on psychological consciousness (in Chalmers' terms: awareness) and is realized at the neural level in the form of extended consciousness.

The paper is based on a particular understanding of cognitive semiotics as a discipline involving analyses of cognitive processes as semiotic processes, i.e. processes requiring usage of signs.

* Maria Curie-Skłodowska University in Lublin, Faculty of Philosophy and Sociology. E-mail: kondorp@bacon.umcs.lublin.pl

KEYWORDS: cognitive-semiotic system, phenomenal consciousness, awareness, core consciousness, extended consciousness, valuation features, somatic marker, metasemiosis

1. A VIEW ON COGNITIVE SEMIOTICS

As Sonesson (2012, p. 208) notes, “cognitive semiotics has been invented many times over during the past few decades”. In the context of the statement it is not surprising that different researchers take different perspectives towards such a marriage of cognitive, linguistic and semiotic studies. Sonesson himself characterizes cognitive semiotics as an approach which aims to wed cognitive science and semiotics. In a similar vein, Zlatev (2012, p. 2) defines cognitive semiotics as a discipline whose “ultimate goal is to provide new insights into the nature and culture of human beings, as well as other meaning-making creatures [...]”. Cognitive Semiotics (henceforth CS) can be defined as an interdisciplinary matrix of (subparts of) disciplines and methods, focused on the multifaceted phenomenon of meaning. This article presents some aspects of cognitive semiotics as seen from the perspective of standard cognitive science¹ – from the Chalmersian (1996) conception of mind. The approach presented in this paper highlights the role of signs, in particular linguistic signs, in the explanations of cognitive functioning of a cognitive agent. Consequently, my understanding of cognitive semiotics is that it encourages us to study cognitive systems, either natural ones like animals or human beings, or, possibly, artificial ones such as sign-using and meaning-making systems. In other words, I assume that at least some cognitive processes involve, in a nontrivial way², the use of signs. This means that there are sign-using (semiotic) processes which are

¹ The qualification “standard” or “cognitivist” seems to be necessary here, as cognitive semiotics highlights the role of non-standard: enactive and embodied cognitive science. My approach is grounded in the “old-fashioned” or Cartesian cognitive science based on the notion of representation and cognitive modeling (either symbolic or connectionist) as a primary method. See e.g. Harnish (2002) for a systematic presentation of standard cognitive science and Rowlands (2010) for discussion on the relationship between standard cognitive science and “4e” approaches (enacted, embodied, embedded, extended).

² “Nontrivial” here means that neither one can describe, nor explain, such cognitive systems without mentioning the notions of ‘sign’ and ‘meaning’.

in fact cognitive processes. For now I leave open the question as to whether all semiotic processes or only some of them are cognitive. Even if just some of them happen to be cognitive, we can still gain some knowledge about the nature of semiosis by studying selected cognitive activities.

My cognitive reading of semiosis (and semiotic reading of cognition) is motivated by the Peircean theory of signs, semiosis as well as his epistemology. The Peircean notion of a sign states that, “a sign, or representamen is something that stands to somebody for something in some respect or capacity. [...] The sign stands for something, its object. It stands not in all respects, but in reference to a sort of idea which I have sometimes called the ground of the representamen” (CP 2.228). In consequence, the notions of a representamen, its interpretation, and a respect are of special importance. The meaning of a sign, in turn, arises in the (dynamic) process of interpretation of the sign. Instead of a dyadic relationship between a sign and its object, we have here a triadic relationship involving interpretation as the third element. I interpret the notion of a sign in terms of the Peircean definition. In addition, the Peircean theory of signs is understood here not in a narrow sense, as a description of actual and possible signs and sign systems, but is interpreted in a broader sense, as a theory addressing fundamental questions of cognition, its relation to reality as well as a logical analysis of knowledge³.

Any human being, or – more generally – any meaning-making creature (to use Zlatev’s formulation⁴) uses signs in his or her everyday cognitive activity. Sign-using agents recognize smoke as

³ It is worth mentioning that Peirce also proposed a classification of consciousness in connection with his triadic definition of a sign. According to Peircean classification, (pure) *feeling* is consciousness of the Firstness (CP 7.551); *experience* (CP 8.266) or *Alter sense* (CP 7.551) is consciousness of “otherness or secondness”, and – finally – *Medi sense* (CP 7.544) is awareness of the Thirdness. The latter may be divided further into *abstraction*, *suggestion* and *association* (CP 7.544–548). In addition, consciousness has a bodily (neural) and social dimension (CP 575). See also (Houser 1983).

⁴ It is necessary to notice that Zlatev uses the formulation – in line with Thompson’s (2007) approach – only in reference to natural, *autonomous* systems (see also: Zlatev 2009, Thompson, Stapleton 2009). In other words, Zlatev excludes possible artificial systems from the scope of meaning-making creatures; in his view artificial systems cannot be truly “cognitive”.

a sign of fire, photos as signs of real people and a red traffic light as a sign of an obligation to stop. In addition, natural-language using creatures recognize certain sounds as sounds of language and (at least sometimes) they are able to understand and interpret these sounds as signs, i.e sounds standing for something else in some respect. One can interpret in a similar way more complex (or: high-level) cognitive abilities. One of the basic cognitive activities discussed within cognitive science, namely the problem-solving activity, calls for the use of signs (in the broad sense: indexes, icons or symbols) as clues or premises and usually requires an interpretation of such signs' indications⁵. (The reader can imagine here Sherlock Holmes solving one of his cases.) In a similar vein, one can emphasize the role of signs (and semiosis) in other activities like decision making, planning, etc. I would like to stress that analyses in terms of signs and meanings are not only a fancy way of describing these activities, but I am convinced that sign using and meaning making are unavoidable elements of these activities⁶.

The main point of the paper is that all these (and similar) cognitive-semiotic activities require some form of consciousness. This is in line with the general phenomenological orientation of cognitive semiotics. Cognitive semiotics highlights the importance of the first-person perspective by stressing the role of consciousness understood as a subjective, qualitative experience. Phenomenology, in turn, is considered to be an approach which provides the right kind of method for studying the structure and content of consciousness (Zlatev 2012, p. 2).

In contradistinction to the above assumption, however, I am convinced that phenomenological experience goes well beyond phenomena involved in semiotic activity which are merely labelled as

⁵ As one of the reviewers of the paper noted, this statement concerning a problem-solving activity may be interpreted as a statement about the dependence of problem-solving on its capacity to use symbols. The remark is justified in light of enactive approaches to cognition (e.g. Noë 2004), supported by some robotic experiments (Brooks 1991, Beer 1995), where not only symbols, but representations in general are rejected. To clarify my viewpoint, I take problem-solving to be one of the higher-level, "representation-hungry" activities (cf. Clark, Toribio 1994) and – as such – involving (at least partially) usage of signs. It does not imply the necessity to use symbols.

⁶ In other words, I am taking here a realist stance towards semiotic phenomena rather than an instrumentalist one.

“conscious”; indeed, I suggest that we should broaden our perspective by including analyses of instances of so-called psychological consciousness (or awareness, cf. Chalmers 2004, pp. 618–619).

In what follows I present the conceptual background of the paper – the notion of a cognitive-semiotic system and a notion of consciousness (section 2). Section 3 presents two distinctions concerning the notion of consciousness: a philosophical one, based on the Chalmersian approach, and a neuroscientific stance based on Damasio’s theory. Sections 4 and 5 are devoted to the two stages in the cognitive processing of signs. Section 4 presents the initial stage, that of sign perception. I highlight here the role of phenomenal experience of meaningfulness and its role in the cognitive activity of a sign-using agent. In section 5 I propose a metasemiotic level of analysis of semiosis.

2. TWO BASIC NOTIONS

Before taking a look at the role of consciousness in a semiotic and cognitive activity, let me elaborate the two key notions of the paper, namely the notion of a cognitive-semiotic system and the notion of consciousness.

2.1 A COGNITIVE-SEMIOTIC SYSTEM

The notion of a mind as a cognitive system is one of the basic notions used within standard, cognitivist (Thompson 2007) cognitive science. Keeping in mind the multidimensional character of a mind (phenomenal, emotional, subjective, cognitive, even computational), I will discuss – in line with the cognitivist approach – only cognitive aspects of mind. The focus on cognitive aspects is motivated by the initial assumption of the paper, i.e. the claim about a relationship that holds between cognition and semiosis.

The term „cognitive system”, as I understand and use it, describes a complex, structured entity which is a subject of processes such as perception, action, reasoning, planning, problem solving, and natural language understanding. It is understood as a dynamic structure which receives environmental and bodily information, processes it according to its internal organization, stores the information and finally acts on the basis of this information (cf. Neřka *et al.* 2006).

As mentioned above, I am interested in a subclass of cognitive systems, namely cognitive-semiotic systems. Such systems are understood here as systems which use signs in their cognitive activity, i.e. they are able to create, distinguish, interpret signs as well as use them in directing their behavior. However, to avoid the temptation of behavioristic (cf. Fetzer 1997) interpretations of the statement⁷, one must take into account one more condition: any cognitive system should use signs as signs, i.e. the system should treat signs as something that stands for something else in some respect or other. In other words, “the behavior of the system is causally affected by the presence of a sign because that sign stands for something else iconically, indexically, or symbolically, for that system” (Fetzer 1997, p. 358). As a consequence, the system is, or at least should be, aware that its mental activity and physical behavior is influenced by semiotic processes.

2.2 CONSCIOUSNESS (AND AWARENESS)

As David Chalmers notes, “consciousness is an ambiguous term, referring to many different phenomena” (Chalmers 2004, p. 617). Contemporary literature on consciousness abounds with differing approaches to the phenomenon and various attempts to define it (cf. Jackendoff 2007, pp. 77–80 for an overview). The spectrum embraces, among others, eliminativist approaches, which treat consciousness as a useful fiction at best (Churchland 1981), reductionist theories (Place 1956, Smart 1959), functionalist approaches (Armstrong 1980, Putnam 1975) as well as theories highlighting the subjective character of conscious experience (Nagel 1974, Searle 1992). One can hardly disagree with Damasio, who claims that “the conflation of so many meanings around the word consciousness renders it almost unusable without qualification, and this conflation is probably responsible for the supreme status to which consciousness has been elevated” (Damasio 1999, p. 309). To avoid the danger of conflation of this kind, I would like to put my philosophical cards on the table: I understand

⁷ I am not going to justify such avoidance – behaviorist approaches to mind have been severely criticized by philosophers representing different stances and repetition of all the arguments seems to be pointless. Consult e.g. (Kim 2011) for an overview.

the phenomenon of consciousness in the sense of Chalmers' (1996) naturalistic and nonreductive theory of consciousness. Chalmers distinguishes between phenomenal and psychological consciousness, stressing both an experiential character of consciousness and a role of consciousness in a mental activity (functional aspect). Phenomenal consciousness is – in the context of this paper – an answer to the question: How is it like to experience signs or meaningful entities? Awareness (or psychological consciousness), in turn, answers the question of what the role of conscious states (processes) in recognition, comprehension and usage of signs is?

Even if one rejects materialistic approaches, which reduce consciousness to a brain activity (*pace* Place), one nowadays can hardly deny that it is impossible to discuss consciousness independently from the achievements of neuroscience⁸. This is the reason why I wish to suggest a kind of interpretation of Chalmers' distinction in terms of a neuroscientific approach to consciousness – from the point of view of Damasio's distinction between core and extended consciousness (Damasio 1999).

3. WHAT IS CONSCIOUSNESS?

To sum up the above terminological considerations, I treat consciousness as a heterogenous phenomenon, which involve two distinctions: a philosophical distinction between phenomenal and psychological notions of consciousness and a neuroscientific distinction between core and extended consciousness. The two approaches are presented below.

3.1 CHALMERS' APPROACH TO CONSCIOUSNESS

In his nonreductive theory of consciousness, Chalmers attempts to explain a wide spectrum of phenomena called in commonsense language “conscious phenomena”. The phenomena include, *inter*

⁸ The need of neuroscientific grounding is appreciated also within phenomenological tradition. Neurophenomenology (Varela 1996) is seen as an important project integrating phenomenological research on consciousness and results of neurosciences. In the case of standard, functionalist cognitive science the connection between consciousness studies and neuroscience is evident.

alia, perceptual experiences (experience of redness, auditory experience of loud sound or tactile experience of a sheer surface), experience of pain, reportability of mental states (“I see red”), belief formation and revision (“I believe I should stop”), decision making (“I deliberately choose not to obey the rules and proceed despite the red light”), problem solving (“How to explain it to the policeman?”), planning, etc. All these phenomena may be treated as conscious ones. Analysing such and similar examples of mental activities, commonly acknowledged as “conscious”, Chalmers claims that these phenomena should be grouped into two classes: phenomenal and psychological, reflecting in this way the two ways of thinking or talking about consciousness⁹.

Certain cognitive subjects, particularly human beings¹⁰, sense the world and have feelings or experiences connected with sensory data. They experience – subjectively and privately – their world and their bodies. In that sense, cognitive agents are *sentient*. On the other hand, in the context of standard cognitive science and studies on cognitive systems, cognitivists highlight the sensitivity of an agent to information and they stress the role of information in controlling agents’ actions. In this sense cognitive agents are conscious as to whether they are able to adjust their mental or physical activity to incoming stimuli, state of knowledge, data in memory, etc. In other words, cognitive agents are *sapient*.

Both aspects of conscious phenomena: sentience and sapience have been stressed in the philosophy of mind: the first one has been elaborated and discussed in the phenomenological approaches as well as in the “subjectivist” theories (Nagel 1974, Searle 1992); the second one can be traced back to the behaviorist descriptions of cognitive systems¹¹ and is present in contemporary materialist and functionalist theories (e.g. Kim 2000). Accordingly, one can associate, as Chalmers

⁹ The distinction is somehow grounded in (and motivated by) Ned Block’s (1995) distinction between access consciousness and phenomenal consciousness.

¹⁰ This formulation raises the question about a class of sentient creatures. Zlatev (2009, p. 1981) notes that a subject should be a “minimal self” in the sense of Gallagher (2005) and enumerates – on the basis of first-, second- and third-person arguments – monkeys, dogs, cats, rats as possible sentient creatures. My argumentation concerns primarily human beings.

¹¹ I do not mean here eliminativist behaviorism.

does, the above-mentioned two kinds of mental phenomena with the following types of consciousness.

Psychological consciousness (awareness) is a state in which a cognitive system has access to information which he or she uses in controlling and directing their cognition. Most typical examples include reportability of mental states, belief formation and revision, discrimination and categorization as well as decision-making, problem solving, planning, etc. One is conscious psychologically when one is aware of the environment and its particular state (“There is red light”) as well as of his/her own bodily state (“I am cold”) or mental state (“I am too stressed”). In addition, the agent is able to report these states, draw conclusions and use the knowledge in directing his/her behavior. To apply the above characteristics to semiotic activity such as the detection or recognition of a red light (as distinguished from a green light) may result in awareness of an obligation to stop, to stop at an intersection (i.e. the sign here influences one’s behavior) or break the law (with an awareness of the consequences of such behavior). All these mental activities: distinguishing, reacting, reasoning about consequences are examples of awareness. As Chalmers notices (1996, p. 28), in everyday settings we use the word “consciousness” in reference to such a situation¹².

Phenomenal consciousness is, in turn, a state in which a cognitive agent experiences subjectively the perceptual stimuli. In other words, there is something it is like to be a cognitive agent; in particular, there is something it is like to be a conscious creature. (Nagel 1974, Chalmers 2004, p. 619) When an agent is suffering pain, if he or she is enjoying experienced sounds of someone’s speech, if a cognitive system is experiencing redness (or roundness) of a signal on a traffic light or coldness of the day, all this is a manifestation of phenomenal consciousness¹³. The reader may have noticed some correlations between both psychological and phenomenal examples. On the one hand, I am aware of a red light: I can report it, I can react

¹² From the functional point of view, it is the only aspect of consciousness that is explainable. As Putnam notices (1981) even if there is something more, it cannot be explained in a functionalist framework.

¹³ The reference to Nagel is somehow misleading: subjectivity – according to his approach – consists of two aspects: phenomenal content and particular individual point of view (perspective).

in the presence of it; on the other hand, I experience subjectively redness. This is no coincidence: it may be the case that the two types of consciousness are closely related¹⁴.

I would like to stress that the difference between phenomenal and psychological consciousness presented above is not only terminological (“two senses of the word”) or conceptual (“two concepts of consciousness”). The distinction may reflect the ontology of the world containing cognitive agents.

3.2 DAMASIO'S APPROACH

The second distinction has been suggested by Antonio Damasio, who has drawn a distinction between core and extended consciousness, based on the results reported in neuroscience. According to Damasio, core consciousness is a very basic process which enables a sense of self: a sense of the here and now. That is, an agent is aware of feelings occurring at the moment when his or her internal state changes. Core consciousness is a result of the interaction between a mind and an external entity (Damasio uses the term “object”). To quote Damasio: “the brain of the organism creates an image of its internal state, an image of the object, and an image of the internal state as it is modified by its interaction with the object. In addition, it creates a second-order image that includes all of these and may result in the feeling of the core self experiencing the resulting qualia” (Damasio 1999; my emphasis).

To highlight the basic features of core consciousness: it is a simple, biological phenomenon and it is stable across an agent's lifetime. Damasio claims that we, human beings, share this type of consciousness with some other species.

Such a characterization suggests that core consciousness may be closely correlated with phenomenal consciousness (in Chalmers' sense). I would say that the philosophical notion of phenomenal consciousness is implemented at the neural level in the form of core consciousness¹⁵.

¹⁴Chalmers claims that it is a fact about our world (Chalmers 1997, p. 18) that psychological processes of awareness are accompanied by experiences.

¹⁵However, I do not suggest that core consciousness is limited to phenomenal consciousness; the experiential aspect of consciousness goes beyond core consciousness.

I would like to stress that, according to Damasio, core consciousness does not depend on higher cognitive processes like planning, reasoning or language (Damasio 1999, p. 16).

On the other hand, Damasio distinguishes extended consciousness, which goes beyond an agent's "here and now" and beyond his or her basic feelings. It enables "an elaborate sense of self" (Damasio 1999, p. 16), i.e. an agent's awareness of location in space and time (including memories and predictions of the future), an explicit distinction between „me and other" and between a subject and his or her environment. It enables deliberations on possible causes and results of actions as well as on failures and successes. Finally, it provides explicit metaknowledge in that it allows one to access consciousness. According to Damasio (1999, p. 16), extended consciousness is a complex biological phenomenon; it requires both long-term memory and working memory and it evolves during the lifetime of an agent. Because extended consciousness in its highest form is partially a result of language, it is supposedly present only in human beings. Characterized in such a way, extended consciousness may be treated as a neural realization of psychological consciousness. As stated by Damasio (1999, p. 201), "Extended consciousness is a bigger subject than core consciousness, and yet it is easier to address scientifically. We understand fairly well what it consists of cognitively and we also understand the corresponding behavioral features". The quotation matches Chalmers' characteristics of awareness.

4. PERCEPTION OF A SIGN

With the above distinctions and clarifications made, I can now present putative roles of consciousness in the semiotic activity of a cognitive agent.

One of the basic methods used within cognitive science is cognitive modeling. This method focuses on computational – either symbolic, connectionist or hybrid – simulation of cognition. Cognitive models are based on the initial set of facts (initial knowledge) and certain control structures specifying how to cope with the data. Cognitive models are supposed not only to produce the same or similar behavior as human beings; they should also predict behavior as well as learn task-specific knowledge (cf. Taatgen, Anderson 2008). As I have argued

elsewhere (Konderak 2015), it is possible and fruitful to model in this way a process of semiosis and, in particular, language comprehension, interpretation and production (cf. Konderak 2007). To create a model of cognitive ability one usually analyses the processes modeled into a number of stages or steps. In the present chapter I will follow the procedure, indicating steps in cognitive processing important from the point of view of semiosis. In my opinion, there are at least three areas of activity involving a semiotic (sign-using, meaning-making) mind in which the role of consciousness is indispensable, namely:

- an initial step: perception (proprioception) of an object as possibly meaningful (e.g. I see someone waving her or his hands; it may be the case that I do not understand what that waving is supposed to mean; I may even wrongly treat it as meaningful);
- establishment or recognition of the relationship between that perceived phenomenon (*Representamen* in Peircean terms) and its Object (once again, understood in the framework of Peircean theory of signs) (e.g. I interpret such waving as a sign of a windy area);
- metasemiotic processes (explicit analysis of a sign as a sign), including recognition of a ground of the relationship (indexical, iconic or symbolic), discovery of an error and ability to re-interpret a sign (e.g. I try to justify my interpretation on the basis of iconicity; I may also realize that I misinterpreted the gesture – the person observed just wanted to get rid of a fly).

All these three areas require some kind of consciousness¹⁶. In the sections to follow I analyse two stages of semiotic processing: the initial stage involves the perception of a sign as a meaningful entity and the metasemiotic processes, i.e. awareness of the semiotic activity of an agent.

4.1 EXPERIENCING A SIGN AS A MEANINGFUL ENTITY

We are surrounded by signs. At first blush, the statement seems to be false: we are surrounded by objects with certain properties, processes or events. If Peirce is right, anything from our environment

¹⁶ As Zlatev noticed, the argument may be related to the one from phenomenology (Zlatev 2010): consciousness is needed to have a world (of reference); the differentiation between the expression and the referent is based on consciousness, as well as the asymmetrical relation between the two.

may be considered as a sign (once again: a red light at an intersection as a sign to stop, someone's statement: „It's red", a person stopping before a pedestrian crossing as signs of the same; pain in my stomach as a sign of e.g. stomach ulcer; doctor's words being a sign of the same disease, etc.). Everything may be a sign, but, certainly, it is not the case that everything is a sign. What is important here is the subject's perspective: stomach pain can be taken as a sign (an index) of – let us say – an ulcer, but one need not take this viewpoint on the pain in question. We, as cognitive agents, have the ability to pick up some elements of our environment and treat them as signs (meaning-inducing entities). In other words, it is a fact about our cognitive activity, that we perceive only selected subsets of surrounding objects, situations and processes, first as meaningful entities, then as signs¹⁷.

The process is often very fast and does not require much reflection. It is often the case that we perceive signs without explicit consideration or awareness of them as entities standing for something else. To illustrate, a perception of a preceding car slowing down suddenly may evoke a relevant behavior of a person driving behind (applying the brakes) or induce some emotions (fear) or beliefs (“Something happened”). That is, slowing down may function as a sign of a danger or as an unexpected event for another driver. In a similar way, the utterance „Stop!" may catch the driver's attention and cause him or her to stop the car or at least make them pay more attention to the surroundings. In both cases the reaction is immediate and neither reflection nor consideration is involved; if an analysis takes place, it follows the initial phase of the perception process.

I suggest that the initial moment of the semiotic activity of a cognitive agent is an experience of meaningfulness, a feeling that some perceived entity or event is meaningful. In other words, to paraphrase Jackendoff (2007, p. 81), a meaningful entity has a distinguished experiential character. Such a feeling may cause an immediate decision as to whether the entity mentioned is meaningful or not (cf. the discussion of a somatic marker below) and then may trigger further analysis of the experienced phenomenon.

¹⁷ A clarification is necessary here: experience of a phenomenon as meaningful does not make it a sign. It is the initial step in the process of semiosis. In other words, it might be necessary for a sign to be experienced as meaningful, but not sufficient.

4.2 AN EXPERIENTIAL BASIS: QUALIA

The above stipulation about the experience of meaningfulness requires special features of our perceptual experience – features allowing for distinguishing potentially meaningful entities from meaningless ones. “Traditional” approaches to qualia (e.g. Lewis 1929, Jackson 1982) are not sufficient to explain a special mechanism detecting “meaningfulness”. It means that – contrary to some of the researchers treating qualia as basic, unanalysable elements of our, human conscious experience – one should take a closer look at subjective experience and its features. It is possible (and probably necessary) to study the structure of qualia.

Jackendoff (2007, chapter 3) analyses phenomenal experience (in reference to perception of natural-language utterances). According to his approach, consciousness has at least two dimensions: phenomenal content (or in Jackendoffian terms, *content features* which are traditionally discussed in philosophy of mind) and *valuation features* (Jackendoff 2007, p. 87). Jackendoff enumerates a number of candidate pairs of the latter, e.g.: external (or not), self-initiated (or not), familiar (or not), affective (or not) meaningful (or meaningless), among others. Two valuations are of particular importance in the context of sign perception: the feeling that the perceived object, state or situation is meaningful and the feeling of familiarity. The idea of the two distinctions is that familiar objects (in particular utterances) – in contradistinction to unfamiliar ones – do indeed have a different experiential character (Jackendoff 2007, p. 81). Similarly, we experience differently meaningful entities (in particular utterances) and meaningless ones. According to Jackendoff, if it is true that we experience language in the form of phonological images or, in the case of signed languages, in the form of visual or proprioceptive experiences (Jackendoff 2007, p. 83), then these images (or other experiences) have an additional “felt” character: we experience them as meaningful.

I would like to push the hypothesis one step further: just as in the case of language, we tend to experience signs (including non-linguistic ones) as meaningful entities as well. Indeed, sometimes one “feels” that their experience (of object, sound or reminiscence) is meaningful despite one’s inability to grasp the meaning itself. I suppose that such

a feeling of meaningfulness is based on one's past experiences. In such cases the feeling of familiarity has some priority: if we are familiar with some stimuli (e.g. a special pattern of sounds, typical for a given language), my previous experiences (e.g. phonological images) would be responsible for the feeling of meaningfulness and later would trigger mechanisms of interpretation.

The following two properties of valuation features seem to be relevant in the context of semiotic processing. First, these features may be subject to error: it may be the case that one has a feeling of familiarity when perceiving completely new objects or situations (*déjà vu*); it may also be the case that one has a feeling of meaningfulness of an utterance while the utterance is meaningless. Second, valuation features are, in a sense, independent of perceptual modality. As Jackendoff says (2007, p. 88) they cut across the „vertical” domains of language, vision and so on.

To sum up, when one is experiencing a sign, phenomenal consciousness appears to play the first and main role, i.e. we start with qualia and their features, and to be precise – with valuation features. Such aspects of phenomenal experience bias (or guide) an agent's behavior or direct further cognitive processing, the process of interpretation included. The phenomenal valuation (e.g. the feeling of being external and meaningful) should form the basis for the initial distinction to be drawn between signs and non-signs. The above suggestion can be supported by the Damasio idea of a somatic marker, elaborated in the next section.

4.3 A HYPOTHETICAL MECHANISM: (SEMIOTIC) MARKERS

Damasio (1994) suggests the existence of a neural mechanism known as “somatic marker” which, according to him, provides an explanation for the fast (in fact immediate) decision-making process, a mechanism allowing for quick choices between available alternatives. The mechanism is based on core consciousness and in particular – on emotions. A somatic marker works as an automatic alarm, it warns against possible negative consequences of the choice made. The warning is based on our previous experiences, encoding associations between objects or events and some states of a body. One can think also about somatic markers as directing mechanisms where some

alternatives are immediately rejected, leaving a much smaller number of alternatives to be considered.

According to Damasio, “somatic markers [...] assist the deliberation by highlighting some options (either dangerous or favorable), and eliminating them rapidly from subsequent consideration. You may think of it as a system for automated qualification of predictions, which acts, whether you want it or not, to evaluate the extremely diverse scenarios of the anticipated future before you. Think of it as a biasing device” (Damasio 1994, p. 174).

It is worth noting that a somatic marker is a mechanism shaped by experience; sometimes it is created during processes of socialization.

The idea of a somatic marker was a motivation to stipulate an analogical mechanism responsible for a detection of (at least some) meaningful signals. An initial observation is that human beings, in their everyday functioning, decide quite quickly whether certain objects or events are meaningful entities. Without a mechanism allowing for a quick choice we would be “drowned” in the multiplicity of potential signs (not to mention a number of possible interpretations of each of them). I am convinced that, in some cases an efficient semiotic activity requires some “fast-track” decision mechanism. Such a mechanism, called the „semiotic marker” would be enabled by consciousness. According to this suggestion it is phenomenal consciousness (and its valuation features in particular) that could be responsible for detecting meaningful entities in an environment. The above proposition should be treated, at least at the moment, as a kind of speculation which calls for detailed empirical examination¹⁸. To reiterate, this is just an initial step in semiotic activity: experience of meaningfulness is not a sufficient reason for a sign to be recognized.

5. USING A SIGN AS A SIGN

Although the role of experience and phenomenal consciousness is unquestionable, I am convinced that we cannot explain semiotic activity

¹⁸ The results of the so-called P300 experiment (Chapman, Bragdon 1964) may be interesting in this context. The researchers presented subjects two kinds of (visual) stimuli: numbers and flashes of light. Chapman and Bragdon concluded that ERP responses to visual stimuli differed depending on whether the stimuli had meaning or not for subjects.

of a cognitive agent independently of psychological consciousness (awareness). The essence of the claim is particularly clear in the case of metasemiosis (and metacognition in general).

5.1 METACOGNITION AND METASEMIOSIS

Moses and Baird (1999) define metacognition as “any cognitive process that controls or monitors any aspect of cognition”. Metaknowledge, in turn, can be defined as “knowledge about knowledge”, which embraces, among other things, beliefs about beliefs (metabeliefs). I treat metasemiosis as a metacognitive process that utilizes metaknowledge. Consequently, metasemiosis is understood as a human capacity to reflect on signs and their usage, to analyse and control processes of recognition and interpretation of signs, to detect and correct errors in semiotic activity, etc. There are at least three reasons to discuss metasemiosis in the context of cognitive-semiotic systems.

First, as Petrilli (2014, p. xviii) points out, “human being is [...] an animal capable not only of semiosis, but also of semiotics, that is, of using signs to reflect on signs”. In other words, a cognitive semiotic system is able not only to use signs but also to discuss them: define, classify them, reflect on their properties. In general, a semiotic system is able to theorize about signs¹⁹.

Second, metasemiotic activity as characterized above is a semiotic activity *per se*. A theory of signs can be analyzed as an example of (meta-) sign-usage. For instance, a review of this chapter may be analysed as the right (or wrong) interpretation of natural language signs.

The third and main motivation for considering metaknowledge in context is the danger of a behavioristic interpretation of semiotic activity. According to such an interpretation, a cognitive agent is a semiotic system if it reacts in some way to special kinds of stimuli (called signs). I wish to claim that mere reaction is not enough as the systems displaying such ability are just „as-if” semiotic systems, i.e. systems that behave as if they use signs. The danger of such an interpretation emerges from discussions on the possibility of artificial semiotic systems: “For a causal system to be a semiotic system, of course, it has to be a system for which something can stand for something

¹⁹ It could be a kind of implicit, commonsense (or folk) semiotics.

(else) in some respect or other, where such a something (sign) can affect the (actual or potential) behavior of that system” (Fetzer 1988, p. 139)²⁰.

As stated above in line with Peircean approach, a cognitive semiotic system must be aware that it uses signs as entities standing for something else (in some respect), i.e. the system needs to have some metaknowledge embracing the usage of signs as well as be able to specify some metaprocesses that control the interpretation and usage of signs. The role of such a metalevel is implied in Fetzer’s discussion on the possibility of artificial semiotic systems. Fetzer suggests a test checking whether a cognitive system is a semiotic system as well. The criterion is the capacity to make a mistake. As he indicates, to be a real sign-user, a cognitive agent “has to take something to stand for something other than that for which it stands” (Fetzer 1988, p. 141, my emphasis)²¹. I would modify the statement: the agent has to be able to make a mistake and to realize the mistake as well. In consequence, a real cognitive and semiotic system should be able to realize (among others) that:

- there are possible alternative interpretations of a sign;
- he or she made a mistake in interpreting a sign;
- the sign used is an inappropriate one (taking into account norms of a community);
- a sign is unexpected in particular contexts etc.

5.2 METASEMIOSIS AND AWARENESS

Following Chalmers’ distinction, metasemiotic processes are examples of conscious processes in psychological sense. The processes seem to be specific in that they require explicit deliberation on semiotic activities and they distinguish a special class of semiotic systems. As mentioned above (section 3.1), of particular interest are: the ability to access and report own interpretations of signs; ability to discriminate

²⁰ As Johan Blomberg noticed (personal communication), some semioticians would reject here the applicability of the notion of sign in the context – they would treat such behavior-evoking phenomena as mere signals.

²¹ Taking something to stand for something other than what it stands for (for instance, taking the green light to stand for the obligation to stop at the intersection) implies – according to Fetzer – the capacity to take something to stand for something else (in some respect) in general.

and categorize kinds of signs, ability to revise interpretations and modify behavior, ability to make deliberate choices, plan usage of signs, etc. To justify the claim I would like to consider the typical mental capacities treated usually as instances of awareness (conscious in psychological sense) in connection with metasemiotic activities. One is aware when one is able to:

- access own mental states – one is not only stopping at a red light, but one knows the reasons for stopping;
- report mental states (the ability assumes introspection and a language faculty) – one justifies crossing the junction despite the red light – „I noticed the red light, but I am in a hurry so I decided...“;
- discriminate kinds of signs – when one wonders whether a road sign „dangerous bend to left“ is an icon or a symbol, and why;
- integrate information (and is able to solve inconsistencies) – when one sees a red light at an intersection and simultaneously one observes a police officer signaling „go“.

All the above examples of metasemiotic activity are clearly dependent on psychological consciousness.

As argued above, Chalmers' awareness seems to be neurally realized in the form of the Damasian extended consciousness. Damasio (1999, p. 195) states: „Extended consciousness goes beyond the here and now of core consciousness, both backward and forward“. What happens when a cognitive-semiotic mind perceives a sign, say, a red light? Rather than just access the fact that one experiences redness of a red light at a intersection, one can also survey the facts concerning the situation: where it is located (in front of you), what caused it (an electronic system for managing traffic), when has one experienced it before (ten minutes ago), who has also experienced it recently (one's mother), who discussed it (one's boss), the fact that one should perceive a green light soon. As can be seen from these examples, the functioning of extended consciousness requires several abilities, including the ability to learn and memorize numerous and various past experiences, the ability to reactivate those memories in connection with „a sense of self knowing“, the ability to direct attention to the content of mental states and the ability to predict and plan behaviors. Consequently, it seems that, seen from the context of semiotic activity, the characteristics of extended consciousness and the description of psychological consciousness converge. Both are connected with so-

called higher-level cognitive processes; both assume a kind of self-awareness and existence of a self-model.

Simultaneously the two approaches are formulated at different levels²². Psychological consciousness is characterized by a role of conscious mental states in the functioning of a cognitive system. To explain a cognitive function like the interpretation of an ambiguous sign, we need only to specify a mechanism that performs the function (Chalmers 2004, p. 620). It seems that extended consciousness is perfect for this task.

6. CONCLUSIONS: TOWARDS PHENOMENAL CONSCIOUSNESS

Cognitive semiotics is by stipulation closely connected with research on consciousness and priority is given to first-person methods (Zlatev 2012). In this paper, I have argued that the first step in the course of semiotic activity has such character: it is phenomenal consciousness in general and valuation features in particular that allow us to pre-select meaningful entities. As a result, a cognitive system treats certain entities as meaningful without grasping the meanings of such entities. The putative mechanism explaining the phenomenon (a semiotic marker) is based on past experiences of a system. As a consequence, the mechanism works only in reference to a subset of all possible signs. The “feeling of meaningfulness” becomes now an impulse to recall information e.g. from long-term memory and to further processing. In situations involving phenomenal consciousness, subjective experience could be (and usually is) followed by psychological consciousness (awareness): the feeling that something is meaningful may be followed by an analysis of ground of meaningfulness (similarity, convention) or attempts to elicit a meaning (cf. also Chalmers 1996, pp. 218–222). The higher-level metasemiotic processes require, it seems to me, psychological consciousness.

To appreciate fully the role of consciousness in cognitive and semiotic activities one has to include in the theory one more aspect of a sign relation. Consciousness and awareness have their role in apprehending the relation between a sign and its object (CP 2.247–

²² In the sense of the Oppenheim-Putnam hierarchy (Oppenheim, Putnam 1958, p. 9).

249). I thus stipulate that different kinds of signs (indexes, icons, symbols) require different types of consciousness. Icons appear to be more closely connected with phenomenal consciousness, whereas the use of symbols seems to primarily depend on psychological consciousness. This initial suggestion, however, can only be confirmed (or not) by further analysis.

REFERENCES

- Armstrong, D. (1980), *The Nature of Mind and Other Essays*, Ithaca: Cornell University Press.
- Beer, R.D. (1995), *A Dynamical Systems Perspective on Agent-Environment Interaction*, "Artificial Intelligence" 72, pp. 173–215.
- Block, N. (1995), *On a Confusion about the Function of Consciousness*, "Behavioral and Brain Sciences" 18, pp. 227–247.
- Brooks, R. (1991), *Intelligence without representation*, "Artificial Intelligence" 47, pp. 139–159.
- Chalmers, D. (1996), *The Conscious Mind*, New York: Oxford University Press.
- Chalmers, D. (2004), *Facing Up to the Problem of Consciousness*, in: J. Heil (ed.), *Philosophy of Mind. Guide and Anthology*, New York: Oxford University Press, pp. 617–640.
- Churchland, P. (1981), *Eliminative Materialism and the Propositional Attitudes*, "The Journal of Philosophy" 78, pp. 67–90.
- Chapman, R.M., Bragdon, H.R. (1964), *Evoked Responses to Numerical and Non-Numerical Visual Stimuli While Problem Solving*, "Nature" 203, pp. 1155–1157.
- Damasio, A. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain*, London: Papermac.
- Damasio, A. (1999), *The Feeling of What Happens*, Boston, Mass.: Houghton Mifflin Harcourt.
- Fetzer, J. (1988), *Signs and Minds: An Introduction to the Theory of Semiotic Systems*, in J. Fetzer (ed.), *Aspects of Artificial Intelligence*. Dordrecht: Kluwer Academic Publishers, pp. 133–161.
- Fetzer, J. (2016), *Peirce and the Philosophy of Artificial Intelligence*, in: M. Bergman, João Queiroz (ed.), *The Commens Encyclopedia – The Digital Encyclopedia of Peirce Studies New Edition*, URL: <<http://www.commens.org/encyclopedia/article/fetzer-james-peirce-and-philosophy-artificial-intelligence>>, [accessed: 10.01.2016].
- Harnish, R.M. (2002), *Minds, Brains, Computers. An Historical Introduction to the Foundations of Cognitive Science*, Malden: Wiley-Blackwell.
- Jackendoff, R. (2007), *Language, Consciousness, Culture. Essays on Mental Structure*, Cambridge, Mass.: MIT Press.
- Houser, N. (1983), *Peirce's General Taxonomy of Consciousness*, "Transactions of the Charles S. Peirce Society", vol. 19, no. 4, pp. 331–359.

- Kim, J. (2000), *Mind in a Physical World*, Cambridge, Mass.: The MIT Press.
- Kim, J. (2011), *Philosophy of Mind*, 3rd edition, Boulder: Westview Press.
- Konderak, P. (2007), *Model kognitywny zdolności językowych*, PhD Thesis.
- Konderak, P. (2015), *On a Cognitive Model of Semiosis*, "Studies in Logic, Grammar and Rhetoric" 40(53), pp. 129–144.
- Konderak, P. (2017), *Between Language and Consciousness: Linguistic Qualia, Awareness, and Cognitive Models*, "Studies in Logic, Grammar and Rhetoric", 48(1), pp. 285–302.
- Konderak, P. (in press), *On Evolution of Thinking about Semiosis: Semiotics Meets Cognitive Science*, "Avant. Trends in Interdisciplinary Studies".
- Nagel, T. (1974), *What Is It Like to Be a Bat?*, "The Philosophical Review" 83(4), pp. 435–450.
- Necka, E., Orzechowski J., Szymura B. (2006), *Psychologia poznawcza*, Warszawa: Wydawnictwo Naukowe PWN.
- Noë, A. (2004), *Action in Perception*, Cambridge, Mass.: MIT Press.
- Oppenheim, P., Putnam, H. (1958), *Unity of Science as a Working Hypothesis*, in: H. Feigl et al. (eds.), *Concepts, Theories, and the Mind-Body Problem*, Minnesota Studies in the Philosophy of Science, vol. II, Minneapolis: University of Minnesota Press, pp. 3–36.
- Peirce, C.S. (1931–1958), *Collected Papers of Charles Sanders Peirce*, vols. 1–6, C. Hartshorne, P. Weiss (eds.), vols. 7–8, A.W. Burks (ed), Cambridge, Mass.: Harvard University Press.
- Petrilli, S. (2014), *Sign Studies and Semioethics: Communication, Translation and Values*, Boston: Walter de Gruyter.
- Place, U.T. (1956), *Is Consciousness a Brain Process?*, "British Journal of Psychology" 47, pp. 44–50.
- Putnam, H. (1975), *Philosophy and Our Mental Life*, in: H. Putnam, *Mind, Language and Reality: Philosophical Papers*, vol. 2, Cambridge: Cambridge University Press.
- Putnam, H. (1981), *Mind and Body*, in: H. Putnam *Reason, Truth and History*, Cambridge: Cambridge University Press, pp. 75–102.
- Rowlands, M. (2010), *A New Science of the Mind. From Extended Mind to Embodied Phenomenology*. Cambridge, Mass.: MIT Press.
- Schraw, Gregory (1998), *Promoting General Metacognitive Awareness*, "Instructional Science" 26, pp. 113–125.
- Smart, J.J.C. (1959), *Sensations and Brain Processes*, "The Philosophical Review" 68, pp. 141–156.
- Sonesson, G. (2012), *The Foundation of Cognitive Semiotics in the Phenomenology of Signs and Meanings*, "Intellectica" 58, pp. 207–239.
- Taatgen N.A., Anderson, J.R. (2008), *Constraints in Cognitive Architectures*, in: R. Sun (ed.), *The Cambridge Handbook of Computational Psychology*, Cambridge, UK: Cambridge University Press, pp. 170–185.
- Thompson, E. (2007), *Mind in Life: Biology, Phenomenology and the Sciences of Mind*, London: Belkarp Press.

- Thompson, E., Stapleton, M. (2009), *Making Sense of Sense-Making: Reflections on Enactive and Extended Mind Theories*, "Topoi" 28(1), pp. 23–30.
- Varela, F.J (1996), *Neurophenomenology: A methodological remedy for the hard problem*, "Journal of Consciousness Studies" 3(4), pp. 330–350.
- Zlatev, J. (2009), *The Semiotic Hierarchy: Life, Consciousness, Signs and Language*, "Cognitive Semiotics" 4, pp. 169–200.
- Zlatev, J. (2012), *Cognitive Semiotics: An Emerging Field for the Transdisciplinary Study of Meaning*, "Public Journal of Semiotics" IV, pp. 2–24.

PIOTR WILKIN*

REPREZENTACJE ZNATURALIZOWANE – UŻYTECZNY CEL CZY UŻYTECZNA FIKCJA?

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

SŁOWA KLUCZOWE: reprezentacje poznawcze, błąd reprezentacyjny, naturalizacja

1. WSTĘP

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

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

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

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

2. REPREZENTACJE ZNATURALIZOWANE I BŁĘDNA REPREZENTACJA

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

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

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

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

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

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

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

3. BŁĘDY LUDZKIE I FUNKCJA NATURALNA U CZŁOWIEKA

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

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

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

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

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

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

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

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

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

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

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

4. CZŁOWIEK A BŁĄD NATURALNY

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

5. REPREZENTACJE A SPOŁECZEŃSTWO

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

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

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

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

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

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

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

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

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

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

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

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

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

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

6. PODSUMOWANIE

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

BIBLIOGRAFIA

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

NATURALIZED REPRESENTATIONS – A USEFUL GOAL OR
A USEFUL FICTION?

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

KEYWORDS: mental representations, representation error, naturalization

JACEK WAWER*

PROBLEM USTALANIA INDEKSU W SEMANTYCE TEMPORALNO-MODALNEJ**

STRESZCZENIE: W ramach semantyk Kripkego dla języków modalnych wartość logiczna zdania może zmieniać się wraz ze zmianą parametru indeksu semantycznego (np. świata, czasu czy miejsca). Oznacza to, że stosując tego typu semantykę do analizy zdań języka naturalnego, musimy wskazać indeks istotny dla analizy semantycznej. Zwykło się przyjmować, że odpowiedni indeks jest ustalany przez kontekst wypowiedzi. Pomysł ten został odrzucony przez semantyków analizujących języki temporalno-modalne w kontekście indeterminizmu, co wygenerowało problem ustalania indeksu. W artykule przedstawię argument owych semantyków, a następnie zarysuję kilka odpowiedzi na problem ustalania indeksu. Ostatecznie będę argumentował, że przy pewnych założeniach metafizycznych i semantycznych można udzielić klasycznej odpowiedzi na problem ustalania indeksu również w kontekstach indeterministycznych.

SŁOWA KLUCZOWE: semantyka zdań o przyszłości, semantyka dla języków modalnych, zależność kontekstowa, metafizyka modalna

Prawdziwość zdania „W Krakowie pada śnieg” zależy od czasu. Prawdziwość zdania „Wokół są góry” zależy od miejsca. Prawdziwość zdania „Świnie latają” zależy od tego, jaki jest świat. Gdy modelujemy ten fenomen przy pomocy semantyki Kripkego, postulujemy, że wartość semantyczna wyrażen może zmieniać się wraz ze zmianą indek-

* Uniwersytet Jagielloński, Wydział Filozoficzny. E-mail: jacek.wawer@uj.edu.pl

** Badania zostały sfinansowane ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/N/HS1/04805.

su semantycznego. W semantyce czasów gramatycznych zdania mogą przyjmować różne wartości w różnych chwilach. W semantyce operatorów przestrzennych zdania zmieniają wartość logiczną zależnie od wyboru koordynat przestrzennych. W semantyce możliwości i konieczności zdania mogą przyjmować różne wartości w różnych światach możliwych. Gdy pracujemy z językiem multimodalnym, indeks semantyczny musi być odpowiednio bogaty i zawierać parametr dla interpretacji każdej z modalności: parametr świata dla konieczności, parametr czasu dla modalności temporalnych, parametr miejsca dla modalności przestrzennych itp.¹

Zmienność wartości semantycznej wyrażenia wraz ze zmianą parametru indeksu jest nieodzowną cechą semantyk dla języków modalnych, gdyż funkcją operatorów modalnych jest właśnie modyfikacja odpowiedniego parametru indeksu semantycznego. Przykładowo, operator możliwości zmienia parametr świata możliwego: w świecie *w* prawdziwe jest zdanie „Świnie mogłyby latać” wtedy i tylko wtedy, gdy zdanie „Świnie latają” jest prawdziwe w pewnym świecie *w'*, dostępnym ze świata *w*. Podobnie operatory czasowe zmieniają parametr czasu: zdanie „Padał śnieg” jest prawdziwe w chwili *t* wtedy i tylko wtedy, gdy zdanie „Pada śnieg” jest prawdziwe w pewnej chwili *t'*, wcześniejszej niż chwila *t*. Można powiedzieć, że w ramach semantyk dla języków modalnych odpowiednie parametry indeksu semantycznego muszą być „ruchome”.

Podobną własność ma klasyczna semantyka logiki kwantyfikatorów, gdzie zmiennym parametrem jest funkcja wartościowania. Tak, jak w semantyce modalnej wartość semantyczna zdania może zmieniać się wraz ze zmianą świata, tak w semantyce dla logiki kwantyfikatorów wartość semantyczna formuły $P(x)$ może zmieniać się wraz ze zmianą funkcji wartościowania (formuła może być spełniona przy jednym wartościowaniu i nie być spełniona przy innym). Analogia sięga głębiej, wszak tak, jak podstawową funkcją operatorów modalnych jest „uruchamianie” modalnego parametru indeksu semantycznego, tak

¹ W całym tekście będę interpretował modyfikatory modalne raczej jako operatory zdaniowe niż jako kwantyfikatory, mimo że jest to założenie dyskusyjne (zob. np. King 2003). Przyjmuję je raczej dla prostoty ekspozycji niż z głębokiego przekonania. Zaznaczę jedynie, że problem ustalania indeksu, który dyskutuję w tekście, pojawia się niezależnie od wyboru sposobu formalnej reprezentacji modalności.

podstawową funkcją kwantyfikatorów jest „uruchomienie” (tj. odpowiednia zmiana) funkcji wartościowania.

Analogia nie jest – rzecz jasna – zupełna. Na poziomie formalnym łatwo zauważyć, że funkcja wartościowania jest dużo subtelniejszym narzędziem niż świat możliwy. Pozwala ona przykładowo na niezależne kwantyfikowanie po różnych zmiennych, podczas gdy operator modalny ma „do dyspozycji” tylko jeden świat możliwy. Analogia między semantyką operatorów modalnych i kwantyfikatorów załamuje się również na poziomie aplikacji do analizy języka naturalnego. W ramach języka logiki kwantyfikatorów zwykło się odróżniać od siebie formuły otwarte oraz zdania (formuły zamknięte). Różnica polega na tym, że w formule otwartej występuje przynajmniej jedna zmienna wolna (poza zasięgiem jakiegokolwiek kwantyfikatora). Większość, jeśli nie wszystkie, zdania języka naturalnego, które można przełożyć na język logiki kwantyfikatorów, po przekładzie przyjmują postać formuł zamkniętych (wyjątkami są być może zdania typu „To jest białe” gdzie kontekst nie dookreśla, o które konkretnie „to” chodzi).

Fakt, że typowe zdania języka naturalnego przekładalne są na formuły zamknięte, jest brzemienne w skutki, gdy aplikujemy logikę do analizy zdań języka naturalnego. Istotną cechą semantyki języka kwantyfikatorów jest to, że wartość logiczna formuł zamkniętych, w przeciwieństwie do formuł otwartych, jest niezależna od funkcji wartościowania. Oznacza to, że o ile formuła otwarta może zmieniać swoją wartość semantyczną w zależności od funkcji wartościowania (może być spełniona lub nie), o tyle formuła zamknięta jest spełniona przy każdym wartościowaniu, jeśli jest spełniona przy pewnym wartościowaniu (a jeśli nie jest spełniona przy jednym wartościowaniu, to nie jest spełniona przy żadnym innym). Parametr wartościowania jest kluczowym narzędziem pomocniczym, umożliwiającym rekursywną definicję funkcji spełniania, ale na poziomie przypisywania wartości semantycznej formułom zamkniętym, czyli zdaniom, jego wartość przestaje być istotna. Dzięki tej własności formuł zamkniętych semantyczna analiza zdań języka naturalnego przy pomocy narzędzi logiki kwantyfikatorów jest niekontrowersyjna. Mimo że formalizm domaga się wskazania pewnej funkcji wartościowania dla oceny wartości semantycznej zdania danego języka, to nie jesteśmy zmuszeni wyszczególnić, która konkretnie funkcja wartościowania jest właściwa, gdyż wybór tej czy innej funkcji nie odgrywa żadnej roli.

Sytuacja zmienia się diametralnie w przypadku próby użycia logiki modalnej do analizy zdań języka naturalnego. Elementarne formuły języka logiki modalnej są reprezentacjami zdań typu „Pada deszcz” lub „Świnie latają”, a ich wartość logiczna jest zależna od wyboru odpowiedniego parametru indeksu semantycznego (świata, czasu, miejsca itp.). O ile zatem wybór funkcji wartościowania nie ma znaczenia dla wartości semantycznej zdań logiki kwantyfikatorów, o tyle wybór możliwego świata, miejsca czy czasu ma kluczowy wpływ na wartość semantyczną zdań typu „Pada deszcz”.

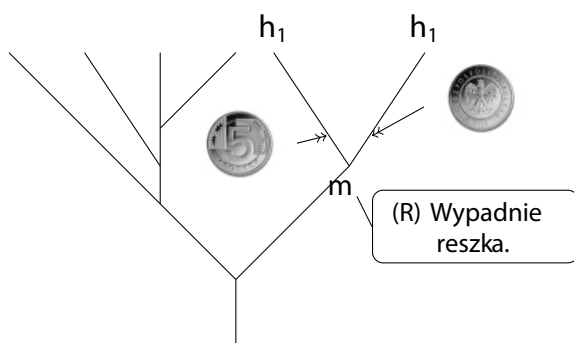
Każdy, kto chciałby używać aparatu formalnego logiki modalnej w celu semantycznej analizy zdań języka naturalnego, staje zatem przed podstawowym pytaniem: którą z wartości parametru modalnego należy wybrać, aby przypisać wartość semantyczną zdaniom języka potocznego? Pytanie to będę nazywał „problemem ustalania indeksu”.

Szkic odpowiedzi na tak postawiony problem znajdujemy już u Kazimierza Twardowskiego, który w artykule *O tzw. prawdach względnych* odpowiada na podobną trudność, pisząc: „Okoliczności, towarzyszące słowom mówiącego, uzupełniają, czego one nie wyrażają” (Twardowski 1900, s. 68). We współczesnej literaturze „okoliczności towarzyszące słowom mówiącego” zwykło się nazywać „kontekstem”, natomiast „uzupełnianie”, o którym pisze Twardowski, będę nazywał „ustalaniem indeksu”. Przedstawione jest tutaj w zarysie bardzo naturalne rozwiązanie zaprezentowanego powyżej problemu: jeśli wartość logiczna zdania zależy od wyboru parametru indeksu semantycznego, to parametr ten ustalany jest przez kontekst, w którym wypowiedziane jest owo zdanie. Tak więc przypisując wartość logiczną zdaniu „Pada deszcz” wypowiedzianemu na górze zamkowej we Lwowie 1 marca 1900 roku (w naszym świecie), należy jako odpowiednie parametry wybrać: górę zamkową we Lwowie – jako miejsce, 1 marca 1900 – jako czas, oraz nasz świat – jako świat możliwy. Podejście to zostało spopularyzowane przez Davida Kapłana, który w komentarzu do swojej przełomowej pracy *Demonstratives* dobitnie podkreślił podwójną rolę kontekstu: jako narzędzia interpretacji wyrażen okazjonalnych oraz jako narzędzia ustalania indeksu semantycznego stosownego dla interpretacji zdań języka modalnego (zob. Kaplan 1989, s. 595). Wygląda na to, że dzięki pomocy kontekstu wypowiedzi problem ustalania indeksu zniknął tak szybko, jak się

pojawił. Nie wszyscy filozofowie są jednak w pełni usatysfakcjonowani tą odpowiedzią.

PROBLEM ZE ŚWIATEM KONTEKSTU

Odpowiedź Twardowskiego–Kaplana na problem ustalania indeksu była podważana w kontekście semantyk stworzonych do analizy możliwości zależnych od czasu (możliwości, które znikają wraz z upływem czasu). Dobrym narzędziem do badania tych możliwości okazał się model rozgałęziających się historii (lub światów). W ramach tego modelu zakładamy, że historie mogą pokrywać się na początkowym interwale, a następnie się rozchodzić. Historia, która „odgałęziła” się w przeszłości danego punktu, reprezentuje możliwość, która była dostępna w przeszłości, ale zniknęła wraz z upływem czasu. Taką możliwością może być przykładowo historia, w której obywatele Wielkiej Brytanii głosują za pozostaniem w Unii Europejskiej. Była ona dostępna przed referendum, które odbyło się 23 czerwca 2016, ale faktyczny przebieg referendum unicestwił tę możliwość. Wzajemne relacje między historiami można obrazowo przedstawić w postaci drzewa, jak na rys. 1 (pierwszą wersję modelu rozgałęziających się historii zaproponował Arthur Prior, zainspirowany sugestiami Saula Kripkego, zob. (Prior 1966, 1967), (Ploug, Øhrstrøm 2012)).



Rysunek 1

Język, w którym chcemy mówić o modalnościach czasowych, zawiera zarówno operatory temporalne: „w przyszłości”, „w przeszłości”, jak i operator historycznej konieczności: „jest przesądzone”. Żeby interpretować te modalności, potrzebujemy dwóch parametrów w indeksie semantycznym: parametru czasu oraz parametru historii (lub świata). O ile ustalanie parametru czasu przez kontekst nie budzi poważniejszych wątpliwości, to ustalanie parametru historii okazało się dużo bardziej kontrowersyjne. Rozważmy zdanie (R) „Wypadnie reszka” wypowiedziane w momencie m wskazanym na rys. 1. O ile jest jasne, który czas ustalany jest przez kontekst – czas, w którym wypowiedziane jest zdanie (R) – o tyle nie ma jasności, którą ustalić historię (świat). Nie możemy udzielić analogicznej odpowiedzi – historię, w której wypowiedziane jest zdanie (R) – gdyż zdanie (R) pada zarówno w historii h_1 , jak i w historii h_2 . Kilku autorów podkreślało ten fakt: „W odróżnieniu od światów historii zachodzą na siebie, stąd pojedynczy akt mowy zazwyczaj należy do wielu możliwych historii” (Belnap i in. 2001, s. 152, wszystkie tłumaczenia autora), „Jedna wypowiedź ma miejsce w wielu światach” (MacFarlane 2008, s. 85), „Rozważ konkretny przypadek, w którym użyto zdania. (...) W ogólności będzie wiele światów, które reprezentują te same przeszłe i teraźniejsze wydarzenia. (...) Konkretny epizod użycia zdania, ma miejsce w nich wszystkich” (MacFarlane 2014, s. 208).

Jeśli jednak zdanie użyte jest jednocześnie w wielu różnych historiach, to nie sposób wskazać na „jedną historię, w której użyto zdania”. Co za tym idzie, nie ma prostej metody, by wskazać „jedną historię kontekstu”, co przywraca problem ustalania indeksu. Semantyka wymaga od nas wskazania jakiegoś indeksu – czasu i historii – dla rozpoczęcia analizy zdania „Wypadnie reszka”, podczas gdy metafizyka nie pozwala wyróżnić żadnego indeksu.

Nie możemy posłużyć się argumentem, który pomógł nam w przypadku semantyki kwantyfikatorów. Semantyka dla logiki kwantyfikatorów również wymaga od nas wskazania pewnej konkretnej funkcji wartościowania w celu oceny semantycznej zdania. Żadna taka funkcja nie jest, rzecz jasna, wyznaczana przez kontekst wypowiedzi. W przypadku logiki kwantyfikatorów okazało się jednak szybko, że nie ma znaczenia, którą funkcję wskażemy, gdyż wartość semantyczna zdania (formuły zamkniętej) jest niezależna od wyboru funkcji wartościowania. Sytuacja wygląda inaczej w naszym przypadku. Wartość

semantyczna zdania „Wypadnie reszka” jest zależna od wyboru parametru historii. Zdanie to jest prawdziwe w historii h_1 , lecz fałszywe w historii h_2 . Wartość ta jednak nie jest ustalana przez kontekst. Okazuje się zatem, że aplikacja semantyki logiki modalnej dla analizy zdań języka naturalnego nastrocza zasadniczej trudności, zwłaszcza jeśli skupimy się – jak w naszym przykładzie – na zdaniach dotyczących przygodnej przyszłości.

(POST)SEMANTYKA PRZYSZŁOŚCI

Aby uporać się z tym problemem, filozofowie i logicy zaproponowali całą gamę rozwiązań. Pierwszą próbę podjął Arthur Prior, definiując semantykę, którą nazwał peirce’owską (Prior 1967)². Teoria peirce’owska rezygnuje z operatorów możliwości i konieczności, wzbogacając jednocześnie operatory temporalne o komponent konieczności. W zwyczajnej logice temporalnej powiemy, że zdanie „Wypadnie reszka” jest prawdziwe w momencie m wtedy i tylko wtedy, gdy zdanie „Wypada reszka” jest prawdziwe w pewnym momencie m' późniejszym niż m . Semantyka peirce’owska modyfikuje ten warunek stanowiąc, że:

Zdanie „Wypadnie reszka” jest prawdziwe w momencie m wtw w każdej historii, do której należy moment m zdanie „Wypada reszka” jest prawdziwe w pewnym momencie m' późniejszym niż m . W przeciwnym razie jest fałszywe.

Trudność ze wskazaniem właściwej historii rozwiązujemy zatem, kwantyfikując po wszystkich historiach, co sprawia, że operator „w przyszłości będzie tak, że” ma komponent konieczności: „jest przesądzone, że w przyszłości będzie tak, że”. Taka zmiana znaczenia sprawia jednak, że semantyka peirce’owska staje się gorszym narzędziem do analizy czasów gramatycznych. Przykładowo, w ramach semantyki peirce’owskiej przed rzutem monetą prawdziwe jest zdanie „Wypadnie orzeł lub reszka, ale nie wypadnie orzeł ani nie wypadnie reszka” ($F(p \vee q) \wedge \neg Fp \wedge \neg Fq$), podczas gdy brzmi ono jak sprzeczność. Aby się o tym przekonać, wystarczy spojrzeć na model przedstawiony na rys. 1. Zdanie $F(p \vee q)$ jest prawdziwe w momencie m , gdyż w każdej

² Nazwa ma odsyłać do myśli Charlesa Sandersa Peirce’a, w którego pismach Prior znalazł inspirację dla swego rozwiązania.

historii biegnącej przez m w pewnym późniejszym momencie prawdą jest, że wypada orzeł lub reszka ($p \vee q$). Jednocześnie zarówno zdanie Fp , jak i Fq jest fałszywe w momencie m , gdyż nie we wszystkich historiach wypada reszka i nie we wszystkich wypada orzeł.

Inna propozycja przewycięzenia problemu ustalania indeksu polega na adaptacji do modeli rozgałęziających się historii trójwartościowej logiki Łukasiewicza³. W ramach tej adaptacji powiemy, że:

Zdanie „Wypadnie reszka” jest prawdziwe w momencie m wtw w każdej historii, do której należy moment m , w pewnym momencie m' późniejszym niż m zdanie „Wypada reszka” jest prawdziwe.

Zdanie „Wypadnie reszka” jest fałszywe w momencie m wtw w każdej historii, do której należy moment m , w każdym momencie m' późniejszym niż m zdanie „Wypada reszka” jest fałszywe.

W przeciwnym przypadku zdanie „Wypadnie reszka” przyjmuje trzecią wartość logiczną⁴.

Zasadniczą trudnością semantyki trójwartościowej jest jednak fakt, że zdanie „Wypadnie reszka lub nie wypadnie reszka” ma przypisaną trzecią wartość logiczną, podczas gdy intuicyjnie oceniamy je jako prawdziwe⁵.

Innowacją, która pozwala rozwiązać ten problem, jest semantyka nadwartościowań Richmonda Thomasona (1970, 1984). W ramach tego rozwiązania Thomason operuje równoległe dwoma rodzajami wartościowań. Dwuwartościowe wartościowania bazowe przypisują zdaniom klasyczne wartości logiczne względem par moment/historia, natomiast nadwartościowania (*supervaluations*) przypisują wartości logiczne zdaniom wyłącznie względem momentów, według schematu wyłożonego poniżej. Technika nadwartościowań pozwala na wprowadzenie luk (nad)prawdziwościowych, przy jednoczesnym zachowaniu tautologii logiki klasycznej oraz standardowej logiki modalnej i tem-

³ Co ciekawe, Prior wprowadził swoje logiki temporalne – peirce’owską i ockhamowską – jako odpowiedź na trójwartościową logikę Łukasiewicza, której wcześniej bronił. Logiki Priora byłyby jednak trudne do zaakceptowania dla Łukasiewicza, gdyż występujące w nich spójniki logiczne nie są ekstensjonalne.

⁴ Taką definicję proponuje John MacFarlane (2014, rozdz. 9.6), alternatywną definicję, pozwalającą zachować ekstensjonalność wszystkich spójników logicznych, można znaleźć w (Wawer 2016, rozdz. 4.4).

⁵ Notabene, sam Łukasiewicz zgadza się w pracy „O determinizmie”, że powinniśmy uznać to zdanie za prawdziwe (zob. Łukasiewicz 1961, s. 124).

poralnej. Thomason czerpał inspiracje dla swojego rozwiązania z prac Bena van Fraasena (m.in. van Fraassen 1966).

Analizując rozwiązanie Thomasona, John MacFarlane (2003) identyfikuje (nad)prawdziwość w momencie z prawdziwością w kontekście, a prawdziwość w parze moment/historia z prawdziwością w indeksie. Utrzymuje on, że drugie z tych pojęć pełni funkcję jedynie pomocniczą. Służy ono eksplikacji pierwszego, które ma modelować nasze potoczne intuicje dotyczące prawdziwości zdań wypowiedzianych w konkretnych okolicznościach. Teorię, która łączy pojęcie prawdziwości w kontekście z pojęciem prawdziwości w indeksie, nazywa MacFarlane „postsemantyką”. W tej terminologii postsemantyka nadwartościowań odpowiada na problem ustalania indeksu w następujący sposób.

Zdanie „Wypadnie reszka” jest prawdziwe w kontekście m wtw jest prawdziwe w każdym indeksie (m, h) , gdzie m jest elementem h .

Zdanie „Wypadnie reszka” jest fałszywe w kontekście m wtw zdanie „Nie wypadnie reszka” jest prawdziwe w kontekście m .

W innym przypadku zdanie „Wypadnie reszka” jest pozbawione wartości logicznej w kontekście m .

Jeśli zastosujemy tę definicję do przykładu przedstawionego na rys. 1, to okaże się, że w kontekście m zdanie „Wypadnie reszka” nie jest prawdziwe ani fałszywe. Jednocześnie zdanie „Wypadnie reszka lub nie wypadnie reszka” jest prawdziwe w kontekście m (gdyż jest prawdziwe w każdej historii przebiegającej przez m).

Problematiczną konsekwencją postsemantyki nadwartościowań jest fakt, że klasyczne spójniki logiczne (jak np. alternatywa) nie są ekstensjonalne w kontekście. W powyższym przykładzie alternatywa dwóch zdań pozbawionych wartości logicznej jest prawdziwa, ale łatwo znajdziemy przykłady, w którym alternatywa takich zdań pozbawiona jest wartości logicznej. Na przykład, jeśli przyjmę zakład, obstawiając, że wypadnie reszka, to alternatywa „Wypadnie reszka lub wygram zakład” pozbawiona jest wartości logicznej.

Innym problemem postsemantyki nadwartościowań, na który szczególnie silny nacisk kładzie MacFarlane, jest to, że mimo iż zdanie „Wypadnie reszka” jest pozbawione wartości logicznej w kontekście m , to w późniejszym kontekście m' , należącym do historii h_1 , zdanie „W przeszłości było prawdą, że w przyszłości wypadnie reszka” jest prawdziwe. MacFarlane twierdzi, że ta własność prowadzi do niein-

tuicyjnych konsekwencji. Zarzut MacFarlane'a jest bardzo subtelny, a ponadto zmieniał on swoją formę wraz z upływem czasu (zob. MacFarlane 2003, 2008, 2014), więc nie będę tutaj przedstawiał wszystkich detali. Podsumowanie dyskusji można znaleźć w (Wawer 2016, rozdz. 4.6).

Odpowiedzią MacFarlane'a na problemy postsemantyki nadwartościowań jest jego własny relatywizm oceny. W myśl tej teorii wartość semantyczną zdania należy ustalać, biorąc pod uwagę nie tylko kontekst wypowiedzi, lecz także kontekst oceny. Powracając do naszego przykładu, zdanie „Wypadnie reszka”, wypowiedziane w momencie m , jest pozbawione wartości logicznej, gdy oceniamy je w kontekście m ; oceniane w późniejszym kontekście w historii h_1 jest prawdziwe; natomiast oceniane w późniejszym kontekście w historii h_2 jest fałszywe. Efekt ten osiąga MacFarlane dzięki następującej definicji:

Zdanie „Wypadnie reszka” jest prawdziwe w kontekście wypowiedzi m i kontekście oceny m' wtw jest prawdziwe w każdym indeksie $\langle m, h \rangle$ gdzie zarówno m jak i m' są elementami h .

Zdanie „Wypadnie reszka” jest fałszywe w kontekście wypowiedzi m i kontekście oceny m' wtw zdanie „Nie wypadnie reszka” jest prawdziwe w tej parze kontekstów. W innym przypadku zdanie „Wypadnie reszka” jest pozbawione wartości logicznej w kontekście wypowiedzi m i kontekście oceny m' ⁶.

MacFarlane podpira swoją semantykę, powołując się na nasze intuicje w zakresie trafności wypowiedzi. Według niego akt wypowiedzi „Wypadnie reszka” nie jest trafny (*accurate*) przed rzutem monetą, podczas gdy po rzucie, w którym wypadła reszka, ten sam dokonany przed rzutem akt wypowiedzi jest już trafny (czy raczej był trafny). Wyjaśnić to można, wskazując, że wypowiedziane przed rzutem zdanie nie jest prawdziwe we wcześniejszym kontekście oceny, a jest prawdziwe w późniejszym (zakładając dodatkowo, że prawdziwość jest koniecznym warunkiem trafności wypowiedzi, tzn. że prawda jest normą asercji). Mam szereg wątpliwości w kwestii zarówno diagnozy, jak i kuracji, którą proponuje MacFarlane, ale omówienie moich wątpliwości odwiedłoby nas zbyt daleko od głównego wątku, więc odłożę je na inną okazję, a teraz omówię jeszcze jedną reakcję na problem ustalania indeksu.

⁶ Jeśli nie ma historii zawierających zarówno m jak i m' , to prawdziwość w parze kontekstów redukuje się do nadprawdziwości w kontekście m .

Reakcję tę prezentują Belnap, Perloff i Xu (2001, rozdz. 6C). Według nich pytanie o wartość semantyczną wyrażenia „Wypadnie orzeł” w kontekście m jest po prostu niedorzeczne. Porównują oni wyrażenie „Wypadnie orzeł” do formuły „ x jest biały”. Tak, jak w drugim przypadku nie ma sensu pytać, czy formuła ta jest spełniona, nie wskazawszy konkretnego wartościowania, tak w pierwszym wypadku nie ma sensu pytać o prawdziwość tego wyrażenia, nie wskazawszy odpowiedniego parametru historii. Gdy zaś wskażemy odpowiedni parametr, to odpowiedź jest prosta: „ x jest biały” jest spełniona przy wartościowaniu, które przypisuje śnieg x -owi, a „Wypadnie orzeł” przy wyborze historii, w której wypada orzeł, jako parametru ewaluacji. Zatem o wyrażeniu „Wypadnie orzeł” można myśleć, jak o formule zawierającej zmienną wolną przebiegającą zbiór historii. Można przyjąć, że głęboka struktura tego wyrażenia ma w istocie postać „W historii h wypadnie orzeł”, gdzie h jest zmienną.

Moją niepewność w stosunku do takiej analizy budzi fakt, że o wyrażeniu „Wypadnie orzeł” nie myślimy zazwyczaj jak o funkcji zdaniowej, która zmienia swoją wartość zależnie od arbitralnego wyboru wartości parametru h . Myślimy o tym wyrażeniu raczej jak o pełnoprawnym zdaniu, które, po dookreśleniu momentu wypowiedzi, pretenduje do prawdziwości. W codziennej praktyce nie przychodzi nam nawet do głowy, że wypowiedzianym przez nas zdaniom o przyszłości nie można przypisać wartości logicznej, o ile wcześniej nie wyznaczy się jednej z możliwych przyszłych historii (nie mówiąc już o tym, że nie do końca wiadomo, na czym miałyby polegać wskazywanie możliwej historii).

Co więcej, o ile nikt rozsądny nie użyje wyrażenia „ x jest białe”, by przekazać jakąś myśl, to nie mamy problemu z użyciem w tym celu zdań takich jak „Wypadnie orzeł” czy „W przyszłym tygodniu będę w Lublinie”. Jedno z wyjaśnień braku analogii mówi, że jest tak dlatego, że (wbrew Belnapowi) w pierwszym przypadku nie można, podczas gdy w dwóch pozostałych można, rozsądnie pytać o wartość logiczną tych wyrażen. Belnap i in. (2001) proponują nam alternatywne wyjaśnienie tej rozbieżności (odpowiedź ta jest szerzej dyskutowana w (Belnap 2002)). Uważają oni, że różnica na poziomie pragmatyki – dokonujemy asercji zdań o przyszłości, nie dokonujemy asercji formuł otwartych – wynika z innego profilu modalnego tych wyrażen. Mimo że formuła „Wypadnie reszka” nie jest prawdziwa ani fałszywa,

to w przyszłości będzie przesądzone, że to wyrażenie było teraz prawdziwe, lub będzie przesądzone, że było ono teraz fałszywe⁷. Można powiedzieć, że wraz z upływem czasu wypowiedziane dziś zdanie o przyszłości „niezależnia” się od wyboru parametru historii, co sprawia, że można go używać w praktyce językowej. Zamiast jednak omawiać szczegółowo idee Belnapa, zaproponuję alternatywną odpowiedź na problem ustalania indeksu.

MOŻLIWOŚĆ FUTURYZMU

Wbrew opinii utrwalonej wśród badaczy rozgałęziających się historiami będę argumentował, że nie trzeba odrzucać naturalnej interpretacji operatorów temporalnych lub zmieniać logiki, by odpowiedzieć na problem ustalania indeksu. Uważam, że nic nie stoi na przeszkodzie, by powołać się na kontekst jako na źródło, które ustala zarówno czas, jak i historię, nawet w obliczu niezdeterminowania przyszłości. Będę optował za konserwatywną odpowiedzią Twardowskiego–Kapłana na problem ustalania indeksu.

Myślę, że wrażenie, iż model rozgałęziających się możliwości wyklucza tradycyjne rozwiązanie problemu ustalania indeksu, wypływa z bardzo szczególnej interpretacji tego modelu, którą nazywam „realizmem rozgałęzień”. W myśl tej interpretacji alternatywne historie przypominają pod pewnymi względami światy możliwe Davida Lewisa (zob. Lewis 1986). Tak jak światy Davida Lewisa, wszystkie historie są równie realne i przypominają pod względem metafizycznym historię (a raczej historie), której (których) jesteśmy częścią. Wszystkie historie złożone są z konkretnych zdarzeń i żadna z nich nie jest metafizycznie wyróżniona. Co więcej, tak, jak mieszkańcy każdego ze światów Lewisa mogą słusznie powiedzieć o swoim świecie, że jest światem aktualnym, tak mieszkańcy każdego z punktów drzewa mogą słusznie powiedzieć, że ich sytuacja jest aktualna.

Mimo że teoretycy modelu rozgałęzień starają się unikać jednoznacznych deklaracji metafizycznych, wiele z ich tez sugeruje, że u podstaw ich refleksji leży tego typu realizm. Jeden z klasyków teorii rozgałęzień, Richmond Thomason, pisze w następujący sposób:

⁷ Ta obserwacja Belnapa zainspirowała MacFarlane’a do stworzenia relatywizmu oceny.

Rozważ dwie gałęzie b_1 i b_2 biegnące przez [moment] t , gdzie $t < t_1 \in b_1$ i $t < t_2 \in b_2$. Z punktu widzenia t_1 , b_1 jest aktualna (przynajmniej do momentu t_1). Z punktu widzenia t_2 , b_2 jest aktualna (przynajmniej do momentu t_2). Żaden z punktów widzenia nie jest trafny w żadnym absolutnym sensie (Thomason, 1984, s. 215, podkreślenie moje).

A następnie dodaje:

Zobacz (Lewis 1970) i podstaw „aktualną przyszłość” za „aktualny świat” w jego tezach. O to jest pogląd dogłębnego indeterministy (Thomason 1984, s. 215, przyp. 14, podkreślenie oryginalne).

W podobnym duchu piszą Belnap, Perloff, i Xu:

Zakładać, że jedna spośród historii Naszego Świata [tak autorzy nazywają model rozgałęzień – J.W.] jest absolutnie faktyczną historią, to tak jak utrzymywać, że stojąc na zewnątrz dziedziny konkretnych możliwości Lewisa, można wskazać tę, która jest faktyczna. W obu przypadkach jest to błędem (Belnap i in. 2001, s. 163)⁸.

Niektóre tezy Johna MacFarlane’a również sugerują modalny realizm:

Nie ma niczego w modelu rozgałęzień, co odpowiadałoby samochodowi poruszającemu się po rozgałęzionych drogach, ani niczego, co odpowiadałoby decyzjom, by wybrać jedną czy drugą gałąź, które będzie musiał podjąć samochód. Jeśli świat się rozgałęzia, my też się rozgałęziamy (MacFarlane 2014, s. 212, podkreślenie oryginalne).

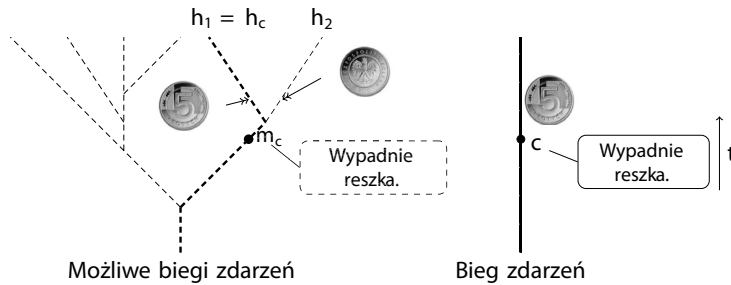
Podobna metafizyczna wizja przebija również przez semantyczne zarzuty Belnapa i MacFarlane’a cytowane powyżej. Autorzy ci zgodnie przyznają, że konkretna wypowiedź jest częścią wielu różnych historii/światów. Taką wizję sugeruje również rys. 1. Rysunek ten przedstawia wypowiadającego jako mieszkańca drzewa, którego wszystkie części są równie realne jak mówiący i jego wypowiedź.

Warto w tym miejscu zaznaczyć, że realizm rozgałęzień różni się pod istotnymi względami od realizmu światów Davida Lewisa. Przede wszystkim każde dwie historie (linie na drzewie) częściowo się pokrywają, podczas gdy światy Lewisa są rozłączne. Warto jednak zauważyć, że stosunek Lewisa do zachodzących na siebie światów nie jest jednoznacznie krytyczny. Uważa on, że tak pojęte światy przeczą pewnym zdroworozsądkowym tezom (Lewis 1986, s. 207–208, skądinąd uwa-

⁸ Realistyczny wydzźwięk ma również ich definicja „Naszego Świata”, którą można znaleźć na s. 139–140.

zam, że Lewis nie ma w tej kwestii racji). Jednak zauważa również, że tak pojęty realizm jest w stanie stosunkowo łatwo rozwiązać problemy, z którymi on sam musiał się zmagać (np. problem trans-światowej tożsamości, zob. (Lewis 1986, s. 209)). Dodaje również, że „częściowe pokrywanie się (*overlap*) psuje najprostszy sposób unifikacji światów przez relacje: mianowicie, przez mereologiczny odpowiednik definicji klasy równoważności. Jednak dostępne są odmienne sposoby, więc sądzę, że realista modalny, który chciałby zaakceptować częściowe pokrywanie się, nie miałby poważnych trudności w tym zakresie” (Lewis 1986, s. 209). Realiści spod znaku rozgałęzień również dystansują się od pewnych poglądów Lewisa (zob. przede wszystkim (Belnap i in. 2001, rozdz. 7A.6)), ale w ogólnym rozrachunku, więcej jest między nimi podobieństw niż różnic. W szczególności panuje wśród nich zgoda, że absolutne odróżnienie tego, co faktyczne, od tego, co możliwe, jest błędem. Uważają oni, że każda możliwość jest aktualna, ze swojej własnej perspektywy, i żadna z modalnych perspektyw nie jest wyróżniona.

Nie jest to jednak jedyna możliwa interpretacja struktury rozgałęziających się możliwości. Zamiast przyjmować Lewisowską wizję możliwości, można przystać do Adamsa (1974), Plantingi (1974), Stalnakera (1976) czy Kripkego (1980) i zaakceptować pewną formę modalnego aktualizmu. Z punktu widzenia naszego problemu kluczowym aspektem modalnego aktualizmu jest postulat **a b s o l u t n e g o**, a nie tylko względnego rozróżnienia między tym, co faktyczne (aktualne), a tym, co możliwe. Wbrew temu, co twierdzi Lewis, możliwe światy nie przypominają metafizycznie świata, w którym mieszkamy. Świat, którego jesteśmy częścią, jest bytem innej natury – bytem, który realizuje jedną z możliwości. Przy takim podejściu, model rozgałęzień jest obrazowym przedstawieniem **m o ż l i w y c h** temporalnych ewolucji naszego świata, jednak owe możliwe ewolucje są zasadniczo różnej natury niż świat, którego jesteśmy częścią. Co ważne, o ile możliwości rozgałęziają się w czasie, o tyle świat nie ma takiej rozgałęzionej struktury. Rozwija się on w sposób liniowy i wraz z upływem czasu realizuje jedną z dostępnych możliwości.



Rysunek 2

Gdy przyjmiemy taką perspektywę, to problem ustalania indeksu przedstawia się zupełnie inaczej. Nie możemy powiedzieć, jak modalni realiści, że konkretny akt wypowiedzi jest częścią wielu możliwych historii. Wypowiedzi nie zdarzają się w możliwościach, ale w konkretnej rzeczywistości. Każda taka wypowiedź jest częścią jednego i tylko jednego świata („naszego” świata). Sytuację tę obrazowo przedstawia rys. 2. Linia po prawej stronie przedstawia wszystkie zdarzenia, które zdarzyły, zdarzają i zdarzą się w naszym świecie, podczas gdy drzewo po lewej przedstawia wszystkie możliwe sposoby, na jakie mogły być potoczyć się dzieje. Jeden z możliwych sposobów, na jaki mogły być (i mogą) potoczyć się dzieje, jest rzecz jasna sposobem, na jaki dzieje faktycznie się potoczyły i potoczą (świat rozwija się w sposób „spójny”, realizując tylko jedną z możliwości). Oznaczyłem tę możliwość pogrubieniem.

Takie ujęcie relacji między tym, co faktyczne, a tym, co możliwe, pozwala na zupełnie inną odpowiedź na problem ustalania indeksu. Wbrew temu, co twierdzą Belnap i MacFarlane, wypowiedź zdania „Wypadnie reszka” nie ma miejsca w wielu różnych historiach/światach. Wypowiedź ta pada w dokładnie jednym świecie, co pozwala powrócić do standardowej odpowiedzi na problem ustalania indeksu – świat kontekstu jest światem, w którym pada wypowiedź. Ścisłej rzecz biorąc, świat/historia wyznaczana przez kontekst jest jedyną możliwą historią, która trafnie reprezentuje to, jaki świat był, jest, i będzie. Jako że świat rozwija się na dokładnie jeden z możliwych sposobów, mamy gwarancję, że istnieje dokładnie jedna historia trafnie reprezentująca ten rozwój. Właśnie tę historię należy wybrać jako historię kontekstu. Na poziomie intuicyjnym sprowadza się to do trywialnej obserwacji,

że zdanie „Wypadnie reszka” jest prawdziwe wtedy i tylko wtedy, gdy w przyszłości faktycznie wypadnie reszka, co można ująć bardziej formalnie w postaci tezy, którą nazywam „futuryzmem”:

Zdanie „Wypadnie reszka” jest prawdziwe w kontekście c wtw jest prawdziwe w momencie kontekstu m_c i w historii kontekstu h_c .

Momentem kontekstu jest domyślnie moment teraźniejszy, a historią kontekstu jest domyślnie historia faktyczna. Na problem ustalania indeksu można zatem udzielić konserwatywnej odpowiedzi w stylu Twardowskiego, nawet gdy zdanie, które analizujemy, odnosi się do przygodnej przyszłości. Abyśmy mogli to jednak zrobić, potrzebujemy powołać się na metafizyczną tezę aktualizmu, tj. na twierdzenie, że świat, którego jesteśmy częścią (i w którym wypowiadamy zdania), jest metafizycznie innej natury niż sposoby, na jakie może rozwinąć się świat. Przyjmując takie założenie, możemy obronić się przed zarzutami modalnych realistów podnoszonymi przeciwko pojęciu świata kontekstu.

Połączenie pomiędzy metafizyką aktualizmu a semantyką uzyskałem dzięki obserwacji, że akty wypowiedzi są częścią jednego, konkretnego świata, który różni się co do natury od możliwych ewolucji. Wobec takiego twierdzenia można jednak postawić pewien zarzut. Wszak nawet aktualiści, którzy metafizycznie odróżniają akt od potencji, często akceptują parafrazę zdań modalnych w kategoriach światów możliwych. Co więcej, nie będą oni mieli problemu z akceptacją tezy, że pewne wypowiedzi, które nigdy nie padły i nie padną, mogłyby paść. Przykładowo, senator Elizabeth Warren mogła być poprzec Berniego Sandersa w prawyborach Partii Demokratycznej w 2016 roku, ale jednak tego nie uczyniła. W związku z tym nawet aktualiści chętnie przyznają, że istnieje możliwy świat, w którym Elizabeth Warren wypowiada zdanie „Zrobię wszystko, by Bernie Sanders został kolejnym prezydentem USA”. Zatem, wbrew temu, co twierdziłem powyżej, nawet w ramach aktualizmu wypowiedzi występują nie tylko w naszym świecie, ale również w światach możliwych. Jeśli zaś tak jest, to nasz świat nie ma wyłączności na wypowiedzi, a co za tym idzie, nie można jej użyć, by rozwiązać problem ustalania indeksu semantycznego⁹.

⁹ Dziękuję anonimowemu recenzentowi za podniesienie tego zarzutu.

Na tak postawioną trudność można odpowiedzieć na dwa sposoby: w duchu elitarnym lub egalitarnym. W ramach pierwszej strategii skupiamy się na wyjątkowym statusie naszego świata i negujemy twierdzenie, że w innych możliwych światach padają jakiegokolwiek wypowiedzi (strategię w tym duchu proponuję w (Wawer 2014)). Wszak w ustach aktualisty twierdzenie, że istnieje możliwy świat, w którym senator Warren coś mówi, jest jedynie użyteczną parafrazą, czy metaforą. To, co jest parafrazowane, zależy od konkretnej wersji modalnego aktualizmu. Twierdzenie, że istnieje możliwy świat, w którym Elizabeth Warren mówi „A”, można rozumieć na przykład jako:

- Elizabeth Warren mogłaby powiedzieć „A”.
- E.W. miała dyspozycję, by powiedzieć „A”.
- Istnieje (abstrakcyjny) niesprzeczny zbiór sądów, który reprezentuje E.W. jako mówiącą „A”.
- Istnieje (abstrakcyjny) maksymalny stan rzeczy, którego częścią jest E.W. Mówiąca „A”.
- Istnieje (abstrakcyjny) sposób, na jaki mógłby potoczyć się świat, w ramach którego E.W. wypowiada „A”.

Istotne jest dla nas to, że żadna z tych parafraz nie sugeruje, że oprócz konkretnych aktów wypowiedzi, które zdarzają się w naszym świecie, istnieją podobne do nich akty zachodzące w innych światach. Przykładowo, sąd głoszący, że E.W. wypowiada zdanie „A” jest bytem radykalnie innej natury niż realna wypowiedź realnej E.W. W związku z tym nie musimy kłopotać się aktami wypowiedzi zachodzącymi w innych światach, gdyż, dosłownie rzecz biorąc, takie akty nie istnieją (istnieją jedynie stany rzeczy czy sądy reprezentujące takie akty, dyspozycje do takich aktów, możliwości zajścia takich aktów itp.). Naszym zadaniem było wskazać mechanizm łączący konkretną wypowiedź z odpowiednim indeksem semantycznym, istotnym dla interpretacji semantycznej tejże wypowiedzi. Skoro zaś wypowiedzi padają wyłącznie w jednym świecie, to mamy pełną gwarancję, że kontekst wypowiedzi ustali odpowiedni indeks semantyczny (aktualną historię i czas teraźniejszy). Tego typu elitarny aktualista musi rzecz jasna zmierzyć się w jakiś sposób z oczywistą obserwacją, że E.W. mogłaby powiedzieć „Bernie Sanders będzie następnym prezydentem”, czy wręcz E.W. mogłaby prawdziwie powiedzieć „Bernie Sanders będzie następnym prezydentem”. Jednakże analiza zdania „E.W. mogłaby prawdziwie powiedzieć «Bernie Sanders będzie następnym pre-

zydentem»” nie wymaga tego, by zakładać, że w jakimś innym miejscu E.W. rzeczywiście wypowiada zdanie „Bernie Sanders będzie następnym prezydentem” (propozycję analizy raportów wypowiedzi zagnieżdżonych w zasięgu operatorów modalnych można znaleźć w (Wawer 2016, rozdz. 6.3.6)).

Można również zaproponować bardziej egalitarną, polubowną linię odpowiedzi na naszkicowaną wcześniej trudność. W ramach tej strategii życzliwiej podchodzimy do możliwych wypowiedzi i zgadzamy się, że każdą taką wypowiedź można potraktować tak, jakby była faktyczna. To znaczy, że ściśle rzecz biorąc, jedynie faktyczne wypowiedzi mają miejsce, ale można założyć, czy wyobrazić sobie, że dana możliwa wypowiedź jest faktyczna i postawić problem analogiczny do naszego problemu ustalania indeksu: Założywszy, że Elizabeth Warren faktycznie mówi „Barnie Sanders będzie kolejnym prezydentem”, którą z historii biegnących przez tę możliwą wypowiedź należy użyć do analizy semantycznej jej wypowiedzi? Problem może wydawać się dojmujący, gdyż wcześniej argumentowałem, że to konkretny, faktyczny świat i jego bieg ustala możliwą historię istotną dla procesu analizy semantycznej, a w naszym przykładzie *explicite* zakładam, że wypowiedź E.W. nie jest częścią tego biegu zdarzeń (E.W. nigdy faktycznie nie wypowiedziała tych słów). Zatem w możliwej sytuacji wypowiedzi „brakuje świata”, który mógłby pomóc ustalić odpowiedni indeks semantyczny.

Uważam, że tego typu trudność wynika z nieporozumienia, które polega na swoistym dwójmyśleniu. Polega ono na tym, że z jednej strony traktujemy wypowiedź E.W., jakby była faktyczna, a z drugiej strony podkreślamy, że jest ona jedynie możliwa. Tego typu wizja istotnie jest problematyczna i prowadzi do kontrowersyjnych wniosków¹⁰. Niemniej aktualista nie musi, a nawet nie powinien jej przyjmować. Jeśli preferujemy egalitarne podejście do modelu rozgałęzień, to decydujemy się założyć, że każda możliwa sytuacja może być kontekstem wypowiedzi. Powinniśmy jednak wtedy pamiętać, że traktując daną możliwą sytuację jako kontekst, musimy jednocześnie założyć, że owa sytuacja jest sytuacją faktyczną i jako taka jest częścią faktycznego bie-

¹⁰ Nota bene takie właśnie dwójmyślenie jest powszechne wśród krytyków aktualizmu w kontekście modelu rozgałęzień, takich jak Nuel Belnap czy John MacFarlane.

gu rzeczy, który realizuje dokładnie jedną z czasowych możliwości dostępnych w momencie wypowiedzi. Jeśli tak jest, to indeks semantyczny możemy ustalić w dokładnie taki sam sposób, w jaki ustalamy go w przypadku rzeczywistych wypowiedzi. Czasem kontekstu jest czas teraźniejszy względem wypowiedzi, a historią kontekstu jest jedyna historia, która zostanie zrealizowana przez bieg zdarzeń obejmujący rozważaną wypowiedź. Zatem niezależnie, czy przyjmiemy elitarny, czy egalitarny stosunek do możliwych wypowiedzi, to dojdziemy do wniosku, że analizując semantycznie wypowiedź użytą w danym kontekście, musimy traktować ją jako będącą częścią świata faktycznego, a co za tym idzie, możemy powołać się na tenże świat, by ustalić odpowiedni indeks semantyczny.

Warto zauważyć, że akceptując tradycyjne, Kaplanowskie rozwiązanie problemu ustalania indeksu, opowiadamy się po jednej ze stron ciągnącego się od starożytności sporu o wartość logiczną zdań dotyczących przygodnej przyszłości. W przedstawionej powyżej (post)semantyce – futuryzmie – każde zdanie ma dokładnie jedną z dwóch wartości logicznych, a zdania dotyczące przygodnej przyszłości mogą być prawdziwe. Nie chcę twierdzić, że aktualista jest zmuszony przyjmując to rozstrzygnięcie, może on zdecydować się, przyjmując którąś z (post)semantyk przedstawionych w poprzedniej części i odmówić używania pojęcia świata kontekstu. Uważam jednak, że decyzja taka jest natury semantycznej, a nie metafizycznej.

Zauważmy, że wybierając jedną z historii jako historię kontekstu wskazałem na historię, która „trafnie reprezentuje to, jaki świat był, jest i będzie”. Co za tym idzie, aby ustalić wartość logiczną zdania wypowiedzianego w danym kontekście, powołałem się *implicite* na przeszłe i przyszłe stany świata. Aktualizm gwarantuje, że w każdym momencie czasu jest (był, będzie) dokładnie jeden taki stan. Aby jednak użyć tego stanu dla naszych potrzeb, musimy założyć, że możemy się na niego powołać, ustalając wartość semantyczną wyrażenia. Założenie to nazywam „semantycznym transtemporalizmem”. Wedle tej tezy wartość logiczna zdania „W czasie t , φ ” oceniana w czasie t' zależy od tego, jaki jest (był, będzie) świat w czasie t , a nie od tego, jaki jest w czasie t' .

Uważam, że przedmiotem sławnego sporu między Kotarbińskim (1913) oraz Leśniewskim (1913) jest w istocie pytanie o zasadność transtemporalizmu. Kotarbiński odrzuca tę tezę, podczas gdy Le-

śniewski jej broni. Kotarbiński, jak się zdaje, dał się przekonać argumentom Leśniewskiego, ale jego sposób myślenia o zależności między prawdą a czasem kontynuował Łukasiewicz¹¹. Niezmiennie obstawał on przy lokalizmie, argumentując, że aby było prawdą w czasie t zdanie „W czasie t' wypada reszka”, w czasie t muszą istnieć warunki, które przesądzają, że w czasie t' wypada reszka. Jeśli t' jest późniejsze niż t , to warunkami tymi mogą być np. kąt oraz siła, z jaką została podrzucona moneta, o ile łącznie gwarantują wypadnięcie reszki. Jeśli t' jest wcześniejsze niż t , warunkami tymi są ślady, jakie pozostawiło wypadnięcie reszki (np. wspomnienia). Jeśli w czasie t nie istnieją warunki gwarantujące prawdziwość lub fałszywość danego zdania, to nie może ono przyjąć jednej z dwóch klasycznych wartości logicznych. Pogląd taki wyraża Łukasiewicz już w roku 1922 (zob. Łukasiewicz 1961, s. 122) i powtarza go w nieomal niezmienionej formie niedługo przed śmiercią (zob. Łukasiewicz 1957, s. 154–155).

Nie jest to jednak jedyny sposób myślenia o zależności między prawdą a czasem. Można argumentować, w zgodzie z Leśniewskim, że wartość logiczna w momencie t zdania „W czasie t' wypada reszka” powinna zależeć od tego, jaki był lub będzie stan monety w momencie t' , a nie od tego, w jakim stanie jest moneta w momencie t . Nie chcę tutaj wdawać się w dyskusję nad zaletami i wadami obu podejść, ale zaznaczę, że nie uważam, by transtemporalizm stał w tym sporze na straconej pozycji.

Reasumując, tradycyjne rozwiązanie problemu ustalania indeksu nie jest wykluczone nawet dla semantyk temporalno-modalnych modelujących sytuacje indeterministyczne. Warunkiem aplikacji tego rozwiązania jest jednak przyjęcie metafizycznego aktualizmu oraz semantycznego transtemporalizmu. Są to realne zobowiązania, z których warto zdać sobie sprawę. Uważam jednak, że jeśli na szali leży logika klasyczna pozwalająca na naturalną analizę języka temporalno-modalnego, to przyjęcie ich nie jest ceną przesadnie wysoką.

¹¹ Nie jest w pełni jasne, czy Kotarbiński inspirował w tej kwestii Łukasiewicza, czy wprost przeciwnie (zob. Woleński, 1990).

BIBLIOGRAFIA

- Adams, R. M. (1974), *Theories of Actuality*, „Noûs” 8(3), s. 211–231.
- Belnap, N. (2002), *Double Time References: Speech-Act Reports as Modalities in an Indeterminist Setting*, w: D. Wolter, H. Wansing, M. de Rijke, M. Zakharyashev (eds.), *Advances in Modal Logic*, vol. 3, Singapore: World Scientific, s. 37–58.
- Belnap, N., Perlo, M., Xu, M. (2001), *Facing the Future: Agents and Choices in Our Indeterministic World*, New York: Oxford University Press.
- van Fraassen, B.C. (1966), *Singular Terms, Truth-Value Gaps, and Free Logic*, „Journal of Philosophy” 63(17), s. 481–495.
- Kaplan, D. (1989), *Afterthoughts*, w: J. Almong, J. Perry, H. Wettstein (ed.), *Themes from Kaplan*, New York: Oxford University Press, s. 565–614.
- King, J.C. (2003), *Tense, Modality, and Semantic Values*, „Philosophical Perspectives” 17(1), s. 195–246.
- Kotarbiński, T. (1913), *Zagadnienie istnienia przyszłości*, „Przegląd Filozoficzny” 16(1), s. 74–92.
- Kripke, S. (1980), *Naming and Necessity*, Cambridge, Mass.: Harvard University Press.
- Leśniewski, S. (1913), *Czy prawda jest tylko wieczna czy też odwieczna*, „Nowe Tory” 10, s. 493–528.
- Lewis, D. (1970), *Anselm and Actuality*, „Noûs”, 4(2), s. 175–188.
- Lewis, D. (1986), *On the Plurality of Worlds*, Oxford: Blackwell Publishers.
- Łukasiewicz, J. (1957), *Aristotle’s Syllogistic from the Standpoint of Modern Formal Logic*, Oxford: Oxford University Press, 2nd edition.
- Łukasiewicz, J. (1961), *O determinizmie*, w: tenże, *Z zagadnień logiki i filozofii*, red. J. Ślupecki, Warszawa: PWN, s. 114–126.
- MacFarlane, J. (2003), *Future Contingents and Relative Truth*, „The Philosophical Quarterly” 53(212), s. 321–336.
- MacFarlane, J. (2008), *Truth in the Garden of Forking Paths*, w: M. García-Carpintero, M. Kölbel (ed.), *Relative Truth*, Oxford: Oxford University Press, s. 81–102.
- MacFarlane, J. (2014), *Assessment Sensitivity: Relative Truth and Its Applications*, Oxford: Clarendon Press.
- Plantinga, A. (1974), *The Nature of Necessity*, New York: Oxford University Press.
- Ploug, T. i Øhrstrøm, P. (2012), *Branching Time, Indeterminism and Tense Logic*, „Synthese” 188(3), s. 367–379.
- Prior, A. (1966), *Postulates for Tense-Logic*, „American Philosophical Quarterly” 3(2), s. 153–161.
- Prior, A. (1967), *Past, Present and Future*, Oxford: Clarendon Press.
- Stalnaker, R. (1976), *Possible Worlds*, „Noûs” 10, s. 65–75.
- Thomason, R. H. (1970), *Indeterminist Time and Truth-Value Gaps*, „Theoria” 36, s. 264–281.
- Thomason, R. H. (1984), *Combinations of Tense and Modality*, w: D. Gabbay i F. Guenther (ed.), *Handbook of Philosophical Logic*, vol. 2. Dordrecht: D. Reidel Publishing Company.

- Twardowski, K. (1900), *O tzw. prawdach względnych*, w: *Księga Pamiątkowa Uniwersytetu Lwowskiego ku uczczeniu pięćsetnej rocznicy Fundacji Jagiellońskiej*, Uniwersytet Lwowski, s. 64–93.
- Wawer, J. (2014), *The Truth about the Future*, „Erkenntnis” 79, s. 365–401.
- Wawer, J. (2016), *Branching Time and the Semantics of Future Contingents*, dysertacja doktorska, Uniwersytet Jagielloński, Kraków.
- Woleński, J. (1990), *Kotarbiński, Many-Valued Logic, and Truth*, w: J. Woleński (ed.), *Kotarbiński: Logic, Semantics and Ontology*, Dordrecht: Kluwer Academic Publishers, s. 191–198.

THE PROBLEM OF INDEX-INITIALIZATION IN THE TEMPO-MODAL SEMANTICS

SUMMARY: In Kripke-semantics for modal logic, the truth value of a sentence depends on the choice of a semantic index (e.g. world, time, or place). It means that application of such semantics to natural language analysis requires indication of an index relevant for semantic analysis. It is commonly accepted that the relevant index is initialized by the context of an utterance. The idea has been rejected by the semanticists investigating tempo-modal languages in the framework of indeterminism what generated the problem of initialization of the semantic index. I present the main argument of those semanticists and describe several responses to the initialization problem. I finally argue that under certain metaphysical and semantic assumptions, one can respond to the initialization failure in the classical way, even in indeterministic contexts.

KEYWORDS: future contingents, semantics of modal languages, context dependence, modal metaphysics

MACIEJ SENDŁAK*

U PODSTAW SPORU O KONTRMOŻLIWE OKRESY WARUNKOWE¹

STRESZCZENIE: Przedmiotem prezentowanego artykułu jest spór o wartości logiczne kontrfaktycznych okresów warunkowych z niemożliwym poprzednikiem (tzw. kontrmożliwych okresów warunkowych). Zgodnie z rozpowszechnionymi analizami każdy taki okres warunkowy jest prawdziwy. Pogląd ten określany jest mianem ortodoksyjnego i przeciwstawiany jest mu pogląd nieortodoksyjny, który uznaje, że tylko niektóre spośród kontrmożliwych okresów warunkowych są prawdziwe, a inne są fałszywe. W jednym ze swoich ostatnich tekstów (*Counterpossibles*) Timothy Williamson argumentował na rzecz tezy, że konsekwencje poglądu nieortodoksyjnego są niespójne z motywacjami leżącymi u podstaw tego poglądu. Celem artykułu jest polemika z krytyką Williamsona oraz wskazanie na możliwość spójnego odczytania postulatów zwolenników stanowiska nieortodoksyjnego.

SŁOWA KLUCZOWE: kontrfaktyczne okresy warunkowe, kontrmożliwe okresy warunkowe, semantyka światów możliwych, metodologia, Timothy Williamson

* Uniwersytet Warszawski, Wydział Filozofii i Socjologii. E-mail: maciej.sendlak@gmail.com

¹ Artykuł powstał w ramach realizacji grantu badawczego „Rozumowania hipotetyczne i okresy warunkowe”, finansowanego przez Narodowe Centrum Nauki (nr 2016/20/S/HS1/00125). Pragnę podziękować Arkadiuszowi Chrudzimskiemu oraz dwóm anonimowym recenzentom pierwotnej wersji prezentowanego artykułu. Wierzę, że ich sugestie oraz krytyczne uwagi pozwoliły mi precyzyjniej przedstawić bronioną tezę.

Przedmiotem niniejszego tekstu jest spór o wartości logiczne kontrmożliwych okresów warunkowych, tj. okresów warunkowych z niemożliwym poprzednikiem². Jedną ze stron tego sporu stanowią zwolennicy tzw. stanowiska ortodoksyjnego, które swe korzenie ma w klasycznych pracach Roberta Stalnackera (1968) oraz Davida Lewisa (1973), a którego broni m.in. Timothy Williamson (2007, 2016b, 2016c). Zgodnie z ortodoksyjną interpretacją każdy okres warunkowy, którego poprzednik wyraża niemożliwość, jest prawdziwy. Przeciwni tej tezie są zwolennicy stanowiska nieortodoksyjnego, którzy argumentują, że tylko niektóre spośród spornej klasy okresów warunkowych są prawdziwe. Stanowiska tego bronią m.in. Takashi Yagisawa (1988), Daniel Nolan (1997), Graham Priest (2009), Berit Brogaard oraz Joe Salerno (2013)³.

Głównym celem tego artykułu jest polemika z dwoma argumentami, jakie przeciwko motywacjom stojącym za stanowiskiem nieortodoksyjnym przedstawił Timothy Williamson (2016b, 2016c). Pierwszy z nich wskazuje na to, że intuicja znajdująca się u podstaw stanowiska nieortodoksyjnego, stoi w konflikcie z powszechnie akceptowanymi zasadami dotyczącymi okresów warunkowych. Drugi z zarzutów dotyczy tego, co można uznać za pośrednią motywację stanowiska nieortodoksyjnego, a co jest pewną niepokojącą konsekwencją stanowiska ortodoksyjnego. Będąc obrońcą tego ostatniego, Williamson wskazuje na obecność nieporozumienia interpretacyjnego, skutkującego jego zdaniem przypisaniem zwolennikom stanowiska ortodoksyjnego poglądu, którego faktycznie nie głoszą. Oba argumenty mają wykazać bezpodstawność postulatów zwolenników stanowiska nieortodoksyjnego.

Naszym zadaniem jest analiza tych zarzutów oraz obrona stanowiska nieortodoksyjnego. Zamierzamy przede wszystkim wykazać, że istnieje możliwość spójnego odczytania postulatów zwolenników tego stanowiska. Ponadto rozważymy jedną z trudności sporu o kontrmoż-

² W literaturze anglosaskiej tego typu zdania określa się mianem „counterpossibles”. Z uwagi na to, że zbliżony termin „counterfactuals” zwykle się tłumaczy na języku polski jako „kontrafaktyczne okresy warunkowe” (lub dokładniej „okresy warunkowe z nierealnym poprzednikiem”), stosując pewną kalkę językową, używać będziemy określenia „kontrmożliwe okresy warunkowe” lub „okresy warunkowe z niemożliwym poprzednikiem”.

³ Określenia „ortodoksyjne” oraz „nieortodoksyjne” używamy za Timothy Williamsonem (2016b).

liwe okresy warunkowe, za którą Williamson uznaje błędną interpretację stanowiska ortodoksyjnego przez jego krytyków. Wierzymy, że pozwoli to uargumentować zasadność analiz nieortodoksyjnych.

Warto zaznaczyć na samym początku dwa aspekty tego sporu. Po pierwsze, głównym jego przedmiotem jest wartość logiczna okresów warunkowych. Niektórzy krytycy stanowiska ortodoksyjnego sugerują, jakoby konsekwencją uznania banalnej prawdziwości kontrmożliwych okresów było również uznanie, że są one pozbawione informacji lub że ich znaczenie jest niezależne od następnika (Brogaard, Salerno 2013). Kwestią sporną jest to, czy faktycznie takie są konsekwencje stanowiska ortodoksyjnego. Problemu tego nie będziemy jednak podejmować w tej pracy i skupimy się jedynie na wyrażonej *explicite* tezie stanowiska ortodoksyjnego, zgodnie z którą każdy kontrmożliwy okres warunkowy jest prawdziwy.

Po drugie, celem naszym nie jest argumentacja na rzecz braku spójności stanowiska ortodoksyjnego lub braku podstaw do jego obrony. Należy bowiem mieć na uwadze, że również niektórzy zwolennicy tego stanowiska traktują problem kontrmożliwych okresów warunkowych jako zasługujący na szczególną uwagę. Część z nich stara się przenieść ciężar tego problemu z semantycznego pytania o wartość logiczną spornej klasy okresów warunkowych na pragmatyczne pytanie o warunki ich asercji (Emery, Hill 2016). W opinii zwolenników tego podejścia każdy kontrmożliwy okres warunkowy jest trywialnie prawdziwy, a potoczne intuicje, kryjące się za stanowiskiem przeciwnym, mają swoje źródło w tym, że ze względów pragmatycznych nie każdy z tych okresów warunkowych spotyka się z podobną asercją (Emery, Hill 2016, s. 5).

Choć propozycja ta jest interesującą alternatywą, a akcentowanie pragmatycznych aspektów badań nad okresami warunkowymi ma długą tradycję⁴, to jej analizy wykraczają poza rozważania, na których skupimy się w prezentowanym tekście. Wynika to z tego, że w większym stopniu interesuje nas wskazanie na zasadność rozwijania stanowiska nieortodoksyjnego, jako spójnej oraz umotywowanej alternatywy dla stanowiska ortodoksyjnego. Wszak m.in. brak spójności oraz brak zasadności został temu pierwszemu zarzucony przez Williamsona.

⁴ Por. m.in. klasyczne prace Paula Herberta Grice'a (1975/1977) oraz Franka Jacksona (1988).

Chcąc osiągnąć zarysowany powyżej cel, zaczniemy od krótkiej charakterystyki kontrfaktycznych okresów warunkowych oraz wskazania na motywacje stojące u podstaw obu stanowisk. W następnej części przyjrzymy się argumentowi, który z założenia miał wskazać na nie-spójność stanowiska nieortodoksyjnego. Kolejno przybliżymy argumentację Williamsona, która wykazać ma nieporozumienie interpretacyjne stanowiska ortodoksyjnego, oraz przedstawimy, w jaki sposób można polemizować z krytyką Williamsona. Część ostatnia poświęcona jest metodologicznemu aspektowi tego sporu.

KONTRFAKTYCZNE OKRESY WARUNKOWE

Kontrfaktyczne okresy warunkowe są sędami, które często wyrażamy za pomocą zdań takich jak „Gdyby było tak, że A , to byłoby tak, że C ” (symbolicznie: „ $A > C$ ”), gdzie A (poprzednik okresu warunkowego) i C (następnik okresu warunkowego) są zdaniami w sensie logicznym, np.:

- (1) „Gdyby na terenie obecnej Irlandii znaleziono ślady kultu Świętowita, to byłoby to interesującym świadectwem ekspansji Słowian”.
- (2) „Gdyby Shakespeare był Hiszpanem, to pisałby w języku francuskim”.
- (3) „Gdyby Szczecin był stolicą Polski, to stolica Polski znajdowałaby się nad Odrą”.

W zdaniach takich wskazujemy na występowanie związku pomiędzy zajściem jednego nierzeczywistego stanu rzeczy oraz zajściem innego. Tego typu konstrukcje wpisane są na stałe zarówno do praktyki życia codziennego, jak i w sferę dociekań naukowych oraz filozoficznych. Powszechność ta wynika z tego, że stanowią one językowy wyraz myślenia hipotetycznego, które jest nieodzownym elementem zdobywania oraz przekazywania wiedzy, a także wyrazu poglądów, postaw lub wyobrażeń (Ingarden 1935), (Edgington 1995), (Bennett 2003), (Williamson 2016a).

Znaczenie, jakie przypisujemy użyciu okresów warunkowych, w dużej mierze opiera się na tym, że niektóre z nich uznajemy za prawdziwe, a inne – za fałszywe. Choć możemy zgodzić się ze zdaniem „Gdyby na terenie obecnej Irlandii znaleziono ślady kultu Świętowita, to byłoby to interesującym świadectwem ekspansji Słowian”, to raczej nie zgodzimy się ze zdaniem „Gdyby Shakespeare był Hiszpanem, to pisałby w języku francuskim”. O ile uwaga o występowaniu różnych wartości

logicznych okresów warunkowych ociera się o banał, o tyle kwestia ustalenia odpowiednich kryteriów prawdziwości dla takich złożonych zdań jest daleka od oczywistości. Należy zaznaczyć, że problem ten nie dotyczy tego, czy Świątowit należał do panteonu bóstw słowiańskich, lub tego, w jakim języku faktycznie piszą Hiszpanie. Nie pytamy bowiem o wartości logiczne zdań takich jak „Świątowit był bóstwem Słowian” lub „Hiszpanie piszą w języku francuskim”. Ustalenie odpowiedzi na te pytania jest zadaniem dla specjalistów poszczególnych dziedzin. Choć odpowiedzi te nie są bez znaczenia, to podejmowany problem dotyczy tego, co z punktu widzenia analiz językowych stanowi o prawdziwości zdań, w których wskazujemy na określony związek poprzednika oraz następnika okresu warunkowego.

ANALIZA MOŻLIWOŚWIATOWA

Dominującym obecnie modelem analiz okresów warunkowych jest semantyka światów możliwych. Zgodnie z jej podstawowymi założeniami, zdania zawierające operatory modalne możliwości – „jest możliwe, że p ” (lub „mogłoby być tak, że p ”) – należy rozumieć jako zdania stwierdzające, że w pewnych możliwych okolicznościach p jest prawdziwe. Owe możliwe okoliczności nazywane są światami możliwymi i uznaje się, że zdanie „jest możliwe, że p ” jest zdaniem prawdziwym wtedy i tylko wtedy, gdy w pewnym świecie (aktualnym lub nieaktualnym) jest tak, że p . Wobec tego zdanie „Mogłoby być tak, że Szczecin byłby stolicą Polski” należy interpretować jako zdanie mówiące o tym, że w pewnym świecie możliwym jest tak, że Szczecin jest stolicą Polski. Analogicznie zdania, które poprzedzone są operatorem konieczności, np. „Jest konieczne, że $2 + 2 = 4$ ” (lub „Musi być tak, że $2 + 2 = 4$ ”), należy rozumieć jako zdania, które są prawdziwe we wszystkich światach możliwych. Gdyby było inaczej, tj. gdyby w pewnym świecie było tak, że $2 + 2$ nie równa się 4, to musielibyśmy uznać prawdziwość zdania „Mogłoby być tak, że $2 + 2$ nie równa się 4”.

Semantyka światów możliwych, dzięki interpretacji terminów modalnych, stała się atrakcyjnym narzędziem wyjaśnienia funkcjonowania okresów warunkowych. Na jej podstawie powstały dwie niemal bliźniacze analizy, które zostały przedstawione przez Roberta Stalnaker (1968) oraz Davida Lewisa (1973). Zgodnie z nimi okres warunkowy o postaci „gdyby było tak, że A , to byłoby tak, że C ” należy uznać

za prawdziwy w świecie aktualnym wtedy i tylko wtedy, gdy spełniony jest jeden z dwóch warunków:

(i) nie istnieje świat możliwy, w którym jest tak, że A

lub

(ii) istnieje świat możliwy w_1 , w którym jest tak, że A i C , oraz świat ten jest bardziej podobny do świata aktualnego aniżeli jakikolwiek świat możliwy w_2 , w którym choć jest tak, że A , to nie jest tak, że C .

W świetle powyższych warunków zdanie „Gdyby Szczecin był stolicą Polski, to stolica Polski leżałaby nad Odrą” jest uznawane za zdanie prawdziwe, ponieważ istnieje świat możliwy, w którym jest tak jak stwierdzają poprzednik oraz następnik tego okresu warunkowego oraz świat ten jest bardziej podobny do świata aktualnego aniżeli jakikolwiek świat, w którym pomimo tego, że Szczecin jest stolicą Polski, stolica Polski nie leży nad Odrą.

Choć tzw. „możliwoświatowa” analiza od blisko pół wieku cieszy się dużą popularnością, to nie jest ona wolna od trudności. Jedną z nich jest pewien analogon paradoksów materialnej implikacji. Jak wskazuje warunek (i), każdy okres warunkowy, który zawiera niemożliwy poprzednik, jest zdaniem prawdziwym. W konsekwencji prawdziwe jest każde z poniższych zdań:

(4) „Gdyby istniał okrągły kwadrat, to prawa geometrii byłyby różne od obecnie obowiązujących”.

(5) „Gdyby istniał okrągły kwadrat, to prawa geometrii nie byłyby różne od obecnie obowiązujących”.

(6) „Gdyby deszcz padał i nie padał jednocześnie, to istniałyby prawdziwe sprzeczności”.

(7) „(Nawet) gdyby deszcz padał i nie padał jednocześnie, to (i tak) nie istniałyby prawdziwe sprzeczności”.

(8) „Gdyby wieloryby były rybami, to oddychałyby one za pomocą skrzel”.

(9) „Gdyby wieloryby były rybami, to oddychałyby one za pomocą płuc”.

Z uwagi na to, że zdania (4)–(9) zawierają poprzedniki wyrażające niemożliwości (odpowiednio: matematyczną, logiczną, metafizyczną), każde z nich uznane jest za prawdziwe⁵. Wszak każde z nich spełnia

⁵ Przykłady te pokazują, że niemożliwości nie ograniczają się jedynie do niemożliwości logicznych, które najczęściej przyjmują postać koniunkcji dwóch wy-

warunek (i). Zdania te są prawdziwe niezależnie od następnika, co czyni je zdaniami trywialnie (lub banalnie) prawdziwymi (*vacuously true*). Oznacza to, że niezależnie od tego, co wyrażałyby następniki zdań (4)–(9), zdania te byłyby prawdziwe.

Uznanie każdego kontrmożliwego okresu warunkowego za prawdziwy zdaje się stać w konflikcie z potoczną intuicją. Wprawdzie zdania (4), (6) oraz (8) uznajemy za prawdziwe, lecz nie są one prawdziwe w sposób trywialny. Wynika to po części z tego, że za fałszywe uznajemy zdania (5), (7), (9). Wszak fakt, że żaden kwadrat nie jest okrągły ma swoje podstawy w prawach geometrii, podręczniki logiki uczą, że jednoczesna prawdziwość zdań o postaci p oraz $\neg p$ jest sprzecznością logiczną. Dzięki osiągnięciom biologii wiemy, że jedną z cech konstytutywnych ryb jest to, że oddychają one za pomocą skrzelii. Tym samym uzasadnione wydaje się oczekiwanie, ażeby właściwa teoria okresów warunkowych uwzględniała ten fenomen.

Filozofowie, którzy podzielają zasygnalizowaną powyżej obawę, wskazują na niedostatki klasycznej analizy możliwościowej oraz argumentują na rzecz jej poszerzenia o postulowanie światów niemożliwych, tj. światów, w których faktem jest to, co z punktu świata aktualnego uznajemy za niemożliwe. I tak postulowane są przez nich światy, w których istnieją okrągłe kwadraty, prawdziwe sprzeczności oraz wieloryby, które są rybami. Zmiana ta skutkuje dostarczeniem alternatywnych kryteriów prawdziwości, zgodnie z którymi zdanie „ $A > C$ ” należy uznać za prawdziwe w świecie aktualnym wtedy i tylko wtedy, gdy istnieje świat (możliwy lub niemożliwy), w którym jest tak, że A i C , i świat ten jest bardziej podobny do świata aktualnego niż jakikolwiek świat (możliwy lub niemożliwy), w którym jest tak że A , ale w którym nie jest tak, że C .

Zmiana ta wprawdzie pozwala do pewnego stopnia uczynić zadość intuicji o nietrywialnej prawdziwości kontrmożliwych okresów warunkowych, lecz rodzi przy tym również istotne pytania, dotyczące m.in. logicznej oraz metafizycznej natury światów niemożliwych⁶. Choć dys-

kluczających się zdań, p oraz p (np. poprzedniki (6) oraz (7)). Za niemożliwy uznawany jest każdy taki stan rzeczy, który nie znajduje swojej realizacji w żadnym ze światów możliwych. Jeśli wobec tego uznamy, że obok prawd logicznych prawdami koniecznymi są również prawdy matematyki oraz metafizyki, to również poprzedniki zdań (4), (5), (8) oraz (9) wyrażają pewne niemożliwości.

⁶ W sprawie dyskusji nad metafizyczną naturą światów niemożliwych zob. (Berto 2013) oraz (Sendlak 2013).

kusje te są wysoce interesujące, to sensowność ich prowadzenia opiera się na założeniu o zasadności postulowanej zmiany. Założenie to spotyka się często z krytyką zarówno filozofów, jak i logików (Lewis 1986, s. 7), (Stalnaker 1996). Wśród wielu pomniejszych uwag dotyczących komplikacji filozoficznych wynikających z konieczności postulowania światów niemożliwych znajduje się również argument wskazujący na brak spójności poglądu głoszącego konieczność odróżnienia wartości logicznych okresów warunkowych z niemożliwym poprzednikiem (Williamson 2016b). Analizę tego argumentu poprzedzimy eksplikacją stanowiska ortodoksyjnego.

STANOWISKO ORTODOKSYJNE

Punktem wyjścia dla zwolenników ortodoksyjnego poglądu – argumentuje Williamson – jest wskazanie na to, że na gruncie semantyki intensjonalnej każdy okres warunkowy z niemożliwym poprzednikiem posiada tę samą intensję, a co za tym idzie – tę samą wartość logiczną⁷. Nie przesądza to jeszcze o tym, czy wszystkie okresy warunkowe z niemożliwymi poprzednikami są prawdziwe, czy fałszywe. Dodatkowym założeniem jest to, że dowolny okres warunkowy, którego następnik jest powtórzeniem poprzednika (np. „ $A > A$ ”) powinniśmy uznać za prawdziwy. Założenie to nie powinno wzbudzać większych kontrowersji, wszak jeśli możemy być pewni prawdziwości jakiegoś stwierdzenia, to $A > A$ wydaje się jednym z nich. Z uwagi na to, że $A > A$ jest prawdziwe niezależnie od tego, co wyraża A , uznaje się że założenie to w równym stopniu dotyczy okresów warunkowych z możliwym poprzednikiem (np. „Gdyby Szczecin był stolicą Polski, to Szczecin byłby stolicą Polski”), jak również tych z niemożliwym poprzednikiem (np. „Gdyby istniał okrągły kwadrat, to istniałby okrągły kwadrat”). Jeśli zatem zgodzimy się z tym, że każdy kontrmożliwy okres warunkowy ma tę samą intensję oraz z tym, że $A > A$ jest prawdziwy dla dowolnego A , powinniśmy przystać również na to, że każdy okres warunkowy z niemożliwym poprzednikiem jest prawdziwy (Williamson 2016b, s. 1).

⁷ Źródłem tego jest naturalnie fakt, że obrońcy tego stanowiska odmawiają istnienia światom niemożliwym, które mogłyby reprezentować różne niemożliwości.

Zwolennik stanowiska nieortodoksyjnego mógłby starać się wskazać na to, że jednym z powodów, dla których uznajemy $A > A$ za zdanie zawsze prawdziwe, jest to, że przeciwne mu $A > \neg A$ uznajemy za zawsze fałszywe. Wszak nawet jeśli nie dysponujemy jakąkolwiek wiedzą na temat A , możemy z powodzeniem założyć że $\neg A$ jest z nim niespójne, wobec czego nie może być tak, że zarazem A oraz $\neg A$. Tym samym, podstawy, dla których uznaliśmy, że okres warunkowy $A > A$ („Gdyby Szczecin był stolicą Polski, to Szczecin byłby stolicą Polski”, „Gdyby istniał okrągły kwadrat, to istniałby okrągły kwadrat”) jest zawsze prawdziwy, są pośrednio podstawami do uznania $A > \neg A$ („Gdyby Szczecin był stolicą Polski, to Szczecin nie byłby stolicą Polski”, „Gdyby istniał okrągły kwadrat, to nie istniałby okrągły kwadrat”) za zawsze fałszywy. Mogłoby to sugerować, że przy akceptacji prawdziwości $A > A$ uzasadnione jest uznanie fałszywości $A > \neg A$.

Wbrew powyższemu zwolennicy ortodoksyjnej analizy argumentują na rzecz tezy, zgodnie z którą jeśli poprzednik A wyraża niemożliwość, to zarówno $A > A$ jak i $A > \neg A$ są prawdziwymi okresami warunkowymi. Jak wskazuje Williamson, pogląd ten znajduje swoje pośrednie uzasadnienie w powszechnie akceptowanej zasadzie rozdziału względem koniunkcji następnika, zgodnie z którą koniunkcja dwóch okresów warunkowych o tych samych poprzednikach i różnych następnikach jest prawdziwa wtedy i tylko wtedy, gdy prawdziwy jest okres warunkowy o wspomnianym poprzedniku i następniku będącym koniunkcją następników: $((A > C) \wedge (A > D)) \equiv (A > (C \wedge D))$. W świetle tej zasady, prawdziwość $A > A$ oraz $A > \neg A$ skutkuje koniecznością uznania prawdziwości $A > (A \wedge \neg A)$. Choć akceptacja tego rezultatu może budzić wątpliwości, to wskazuje on jedynie na to, że jeśli konsekwencją poprzednika A jest sprzeczność $A \wedge \neg A$, a przy tym żadna sprzeczność nie jest możliwa, to możliwy nie jest również poprzednik (Williamson 2016b, s. 3). Tym samym uznanie prawdziwości $A > A$ oraz $A > \neg A$ oparte jest na tym, że A jest niemożliwe. W innych przypadkach, tj. w tych, w których A jest możliwe, prawdziwość $A > A$ pociąga za sobą fałszywość $A > \neg A$ (Stalnaker 1968, s. 106).

Przytoczona powyżej argumentacja Williamsona stanowi raczej uzasadnienie dla uznania trywialnej prawdziwości kontrmożliwych okresów warunkowych niż krytykę stanowisk nieortodoksyjnych. Opiera się ona w dużej mierze na założeniu o nieistnieniu światów niemożliwych. Może to sugerować, że odpowiednie rozszerzenie analiz moż-

liwoświatowych o światy niemożliwe doprowadziłyby nas do punktu, w którym stanęlibyśmy przed wyborem pomiędzy dwiema konkurencyjnymi teoriami – ortodoksyjną oraz nieortodoksyjną. Sytuacja ta byłaby interesująca wówczas jedynie, gdyby każda z nich była spójna, a właśnie brak spójności zarzuca teoriom nieortodoksyjnymi Timothy Williamson (2016b, s. 9). Niespójność ta dotyczy samej motywacji stojącej u podstaw przekonania o różnej wartości logicznej kontrmożliwych okresów warunkowych. W następnej części przyjrzymy się bliżej tej krytyce.

ZWODNICZA HEURYSTYKA

W świetle popularności stanowiska ortodoksyjnego rodzi się pytanie o to, w jaki sposób wyjaśnić występowanie intuicji, zgodnie z którą przypisujemy kontrmożliwym okresom warunkowym różne wartości logiczne. Williamson upatruje źródeł tego fenomenu w tym, co on sam nazywa heurystyką (*heuristics*), a co zdaje się oddawać potoczną intuicję, do której odwołujemy się przy wartościowaniu okresów warunkowych. Znajduje ona swój wyraz w jednym z dwóch bliźniaczych sformułowań:

(HCC) „Zakładając, że C jest niespójne z D , należy uznać że $A > C$ jest niespójne z $A > D$ ”

lub

(HCC*) „Uznając prawdziwość jednego z okresów warunkowych $A > C$ oraz $A > \neg C$, powinniśmy uznać fałszywość drugiego” (Williamson 2016b, s. 8).

Jak wskazuje Williamson, założenie o słuszności powyższych zasad leży u podstaw nieortodoksyjnego stanowiska w kwestii kontrmożliwych okresów warunkowych. Akceptując bowiem (HCC) lub (HCC*) oraz uznając prawdziwość $A > A$, powinniśmy zgodzić się z fałszywością $A > \neg A$. Konsekwencja ta – zdają się twierdzić zwolennicy nieortodoksyjnego poglądu – trafnie oddaje to, w jaki sposób posługujemy się okresami warunkowymi. Zarówno tymi z możliwymi poprzednikami, jak i tymi, których poprzedniki wyrażają niemożliwość.

Wbrew temu Williamson wskazuje na to, że choć w wielu przypadkach odwołanie do wspomnianych heurystyk ma swoje uzasadnienie, to nieograniczone ich stosowanie jest nieuprawnione. Świadczy

o tym kontrprzykład, w którym A wyraża niemożliwość. Jak wskazywaliśmy powyżej, zwolennik stanowiska ortodoksyjnego argumentuje, że w takich sytuacjach, pomimo niespójności $A > C$ oraz $A > \neg C$, oba zdania powinniśmy uznać za prawdziwe. Stąd też (HCC) oraz (HCC*) mogą być stosowane jedynie wówczas, gdy poprzednik danego okresu warunkowego wyraża możliwość (Williamson 2016b, s. 9). W pozostałych przypadkach opieranie się na tych heurystykach prowadzi do konsekwencji sprzecznych z założeniami poglądu ortodoksyjnego, tj. do uznania, że niektóre kontrmożliwe okresy warunkowe są fałszywe.

Powyższą obserwację, dotyczącą relacji pomiędzy heurystyką oraz założeniami ortodoksyjnego stanowiska, możemy jednak zinterpretować również w inny sposób. Może ona bowiem wskazywać po prostu na to, że ortodoksyjny pogląd stoi w konflikcie z tym, w jaki sposób oceniamy wartości logiczne okresów warunkowych, co może stanowić motywację, by faworyzować podejście nieortodoksyjne. Taka interpretacja byłaby uprawniona, gdyby zwolennicy tego ostatniego mogli w nieograniczony sposób zaakceptować wskazane powyżej heurystyki. Jak wskazuje jednak Williamson ani (HCC), ani (HCC*) nie mogą być przez nich w pełni zaakceptowane. Zobrazowane jest to przykładami następujących okresów warunkowych:

- (a) $(A \wedge \neg A) > A$,
- (b) $(A \wedge \neg A) > \neg A$.

Na podstawie (HCC*) należałoby uznać, że akceptując prawdziwość jednego z powyższych okresów warunkowych, drugi powinniśmy uznać za fałszywy. Jest to jednak wysoce problematyczne z trzech powodów. Po pierwsze wymagałoby to odrzucenia ogólnej zasady, zgodnie z którą konsekwencją poprzednika o postaci koniunkcji jest m.in. każdy z jej elementów, tj. $(A \wedge B) > A$, $(A \wedge B) > B$. Po drugie akceptacja tylko (a) lub tylko (b) przeczy wspomnianej zasadzie rozdziału względem koniunkcji następnika. Wszak zarówno (a), jak i (b) mają dokładnie ten sam poprzednik, co powinno nas uprawniać do akceptacji okresu warunkowego (c): $(A \wedge \neg A) > (A \wedge \neg A)$. Po trzecie niemożliwość uznania prawdziwości (c) przeczy wyjściowemu założeniu o prawdziwości okresów warunkowych o postaci $A > A$, którego jedną z egzemplifikacji jest $(A \wedge \neg A) > (A \wedge \neg A)$. Tym samym to, co uznane zostało za wstępną motywację dla rozwinięcia stanowiska nieortodoksyjnego, okazało się z nim niespójne (Williamson 2016b, s. 8).

W świetle powyższego zwolennik podejścia nieortodoksyjnego znajduje się w trudnej pozycji. Chcąc zachować spójność swojego stanowiska, powinien albo dostarczyć argumentów przeciwko trzem przytoczonym powyżej zasadom, albo dokonać modyfikacji heurystyk. Argumentację na rzecz tego ostatniego poprzedzimy uwagą dotyczącą polemiki Williamsona z często stawianym zarzutem, jakoby zwolennicy poglądu ortodoksyjnego utrzymywali, że następnik kontrmożliwego okresu warunkowego nie ma wpływu na wartość logiczną całego wyrażenia. Zarzut ten sformułowany został m.in. przez Beritt Brogaard oraz Joe Salerno:

W świetle standardowego poglądu kontrmożliwe okresy warunkowe są trywialne. Mówiąc, że są one „trywialne” mamy na myśli to, że są one banalnie (*vacuously*) prawdziwe oraz że z semantycznego punktu widzenia pozbawione są informacji. Kontrmożliwe okresy warunkowe są banalnie prawdziwe w tym sensie, że są one zawsze prawdziwe – niemożliwość poprzednika pociąga za sobą każde wyrażenie. Są one pozbawione informacji w tym sensie, że następnik kontrmożliwego okresu warunkowego nie ma wpływu na wartość logiczną, znaczenie oraz zrozumienie całego wyrażenia (Brogaard, Salerno 2013, s. 642).

Problem wyrażony w tym zarzucie jest uznawany często za pośredni powód odejścia od stanowiska ortodoksyjnego. W opinii Williamsona zarzut ten opiera się na błędnej interpretacji tego stanowiska. O ile w przekonujący sposób argumentuje on na rzecz tego, że wcale nie jest tak, że stanowisko ortodoksyjne pociąga za sobą tezę o tym, że następnik okresu warunkowego nie ma wpływu na **znaczenie** oraz **rozumienie** całego wyrażenia (Williamson 2016b, s. 4–5), o tyle są powody, by sądzić, że odparcie zarzutu o braku wpływu następnika na wartość logiczną kontrmożliwego nie jest trafne.

NASTĘPNIK KONTRMOŻLIWEGO OKRESU WARUNKOWEGO

Kontrargument Williamsona ma postać *reductio ad absurdum*, a istotnym jego elementem jest odwołanie się do innego typu trywialnie prawdziwych okresów warunkowych, którymi są okresy warunkowe z koniecznie prawdziwym następnikiem. Williamson wskazuje na to, że gdyby zwolennicy ortodoksyjnej interpretacji chcieli faktycznie utrzymać, jakoby trywialna prawda kontrmożliwych okresów warunkowych pociągała za sobą brak wpływu następnika na wartość logiczną całości, to (analogicznie) trywialna prawdziwości okresu wa-

runkowego z koniecznie prawdziwym następnikiem byłaby niezależna od poprzednika. Pozwala to na rozważenie szczególnego typu okresu warunkowego, którego poprzednik wyraża niemożliwość, natomiast następnik prawdę konieczną. Przykładem jest wyrażenie (10) „Gdyby 6 było liczbą pierwszą, to 35 byłoby liczbą złożoną” (Williamson 2016b, s. 5).

Podążając tokiem rozumowania krytyków stanowiska ortodoksyjnego – argumentuje Williamson – powinniśmy uznać, że zarówno poprzednik (10) (z uwagi na konieczny następnik), jak i następnik (z uwagi na niemożliwy poprzednik) nie mają wpływu na wartość logiczną całości. Jest to – twierdzi Williamson – niedorzecznością, ponieważ bez wskazania odpowiedniego poprzednika oraz następnika, pozostaje czysta forma okresu warunkowego (*bare form of a counterfactual sentence*), która sama w sobie nie może stanowić o wartości logicznej wyrażenia. Jeśli taka jest konsekwencja wskazanego zarzutu, to – konkluduje Williamson – jest on błędnie postawiony. Powodem tego miałyby być fakt, że nikt ze zwolenników ortodoksyjnego poglądu nie głosi tak niedorzecznej (*ridiculous*) tezy (Williamson 2016b, s. 5).

Jeśli Williamson ma racje, to krytyk stanowisk ortodoksyjnych – chcąc utrzymać w mocy swój zarzut – powinien albo argumentować na rzecz tego, że wskazana niedorzeczność faktycznie jest konsekwencją poglądu alternatywnego, albo wykazać, że wskazany przez Williamsona problem nie jest konsekwencją zarzutu Brogaard i Salerno. Chcąc zrealizować tę drugą możliwość, postaramy się wskazać, że zarzut ten nie prowadzi do konieczności uznania, że zdanie (10) jest prawdziwe na mocy bycia okresem warunkowym.

Kluczowe dla interpretacji kontrargumentu Williamsona jest wyjście od tego, czym jest czysta forma okresu warunkowego. Jeśli uznamy, że czystą formą alternatywy jest wyrażanie o postaci „ $p \vee q$ ”, to za czystą formę okresu warunkowego możemy uznać „ $A > C$ ”. Faktycznie, samo to wyrażenie nie stanowi jeszcze o wartości logicznej okresu warunkowego, o czym świadczy konieczność dostarczenia dodatkowych warunków, które przytoczyliśmy powyżej⁸. Podobnie,

⁸ Williamson nie napisał wprost, co rozumie przez czystą formę zdania (10). Stąd można podać w wątpliwość, czy jest nią przytoczone przez nas wyrażenie „ $A > C$ ”. Choć może ono zostać użyte do reprezentowania rozpatrywanego zdania, to nie oddaje ono w pełni statusu modalnego poprzednika oraz następnika tego wyrażenia. Wszak poprzednik jest niemożliwy a następnik prawdziwy

przytoczona forma alternatywy nie rozstrzyga jeszcze wartości logicznej tego wyrażenia. Choć trudno nie zgodzić się z tym, że samo wyrażenie „ $A > C$ ” (bez wskazania dodatkowych warunków) nie pozwala na rozstrzygnięcie wartości logicznej zdania (10), to można podać w wątpliwość, czy wymóg, by było inaczej, jest konsekwencją wspomnianego zarzutu Brogaard i Salerno. Wydaje się, że istnieją dwa powody, by uznać, że pogląd, jakoby następnik kontrmożliwego okresu warunkowego nie miał wpływu na wartość logiczną całości, nie prowadzi do konieczności uznania konsekwencji wskazanej przez Williamsona.

Pierwszym powodem jest to, że gdyby tak było, to zarzut Brogaard i Salerno, zastosowany do takich zdań jak (8), przypisywałby zwolennikom stanowiska ortodoksyjnego pogląd, zgodnie z którym przekonani są oni o prawdziwości zdania (8*) „Gdyby wieloryby były rybami, to C ”. Oznaczałoby to, że zarzut ten dotyczy uznania banalnej prawdziwości niepoprawnie zbudowanych formuł. Zakładając bowiem, że okres warunkowy jest funktorem zdaniotwórczym od dwóch argumentów zdaniowych, aby podjąć się oceny jego wartości logicznej, powinien on być spełniony przez dwa zdania w sensie logicznym. Wyrażenie (8*) nie spełnia tego warunku, ponieważ zawiera ono jedno zdanie oraz jedną zmienną zdaniową⁹.

Choć uznanie prawdziwości zdania (8*) może być postrzegane za niewłaściwe, to nie jest to problemem, na który wskazują krytycy stanowisk ortodoksyjnych. Jest nim to, że zakładając konieczną fałszywość zdania „Wieloryby są rybami”, dla dowolnego C , zdanie, które powstanie po jego podstawieniu w (8*), będzie zdaniem prawdziwym. W tym sensie, następnik zdania (8*), tj. wyrażenie podstawione za C , nie ma wpływu na wartość logiczną całości. Niezależnie od tego, jakie zdanie zastąpi C , powstałe z tego podstawienia zdanie będzie prawdziwe.

w sposób konieczny. Pozwala to zastanowić się, czy właściwszą formalizacją nie jest „ $\neg \Diamond A > \Box C$ ”. W takim przypadku prawdziwość (10) byłaby oparta na czystej formie okresu warunkowego. Zdanie to byłoby prawdziwe niezależnie od tego, co kryje się za „ A ” oraz „ C ”. Podobnie, jak prawdziwe są wyrażenia „Jan śpi lub Jan nie śpi” oraz „Jeśli Jan śpi, to Jan śpi”, których formami są kolejno „ $p \vee \neg p$ ” oraz „ $A > A$ ”. Kwestia ta nie jest oczywista, chcąc jednak w życzliwy sposób odczytać argumentację Williamsona, przyjmiemy, że czystą formą (10) jest „ $A > C$ ”.

⁹ Używając analogii do czystej formy alternatywy, choć każda alternatywa, której jednym z elementów jest wyrażenie „ $2 + 2 = 4$ ” jest zdaniem prawdziwym, nie wynika z tego, że wyrażenie „ $2 + 2 = 4$ lub p ” (lub „ $2 + 2 = 4 \vee p$ ”) jest zdaniem prawdziwym. Wszak wyrażenia te nie są poprawnie zbudowanymi formułami.

Wobec powyższego możemy zapytać, czy konsekwencją zarzutu Brogaard i Salerono jest przypisanie zwolennikom poglądu ortodoksyjnego przekonania, że o prawdziwości (10) stanowi fakt, że posiada ono strukturę „ $A > C$ ”. Gdybyśmy zapytali zwolennika tego podejścia o powody, dla których uznaje on zdania (4)–(9) za trywialnie prawdziwe, odpowiedziałby, że wynika to z niemożliwości poprzednika. Gdybyśmy zapytali o to, dlaczego zdanie „Nawet gdyby Szczecin był stolicą Polski, to 36 byłoby liczbą parzystą” jest zdaniem trywialnie prawdziwym, to wskazałby na to, że zawiera ono następnik, który wyraża prawdę konieczną. Każdy z tych warunków jest – w opinii zwolennika ortodoksyjnego stanowiska – warunkiem wystarczającym do uznania prawdziwości przytoczonych okresów warunkowych. Podobnie jest w przypadku zdania (10). Oznacza to, że prawdziwość tego zdania nie jest oparta na koniunkcyjnym warunku stwierdzającym, że **z a r ó w n o** poprzednik jest koniecznie fałszywy oraz następnik koniecznie prawdziwy. Możemy zatem uznać, że tym, co zgodnie z ortodoksyjnym poglądem czyni (10) prawdziwym, jest spełnienie **a l t e r n a t y w n e g o** warunku: okres warunkowy jest prawdziwy zawsze wtedy, gdy jego poprzednik jest niemożliwy **l u b** gdy jego następnik jest prawdziwy w sposób konieczny. W tym pierwszym przypadku następnik nie ma wpływu na wartość logiczną całości, w tym ostatnim wpływu takiego nie ma poprzednik.

Pokazuje to, że wbrew sugestiom Williamsona krytyka stanowiska ortodoksyjnego nie musi pociągać za sobą przypisania jego zwolennikom głoszenia niedorzecznej tezy, jakoby zdanie (10) było prawdziwe na mocy samej struktury. Niemniej nie zmienia to faktu, że akceptując ortodoksyjny pogląd, należy zgodzić się z tym, że niemożliwość poprzednika przesądza o wartości logicznej całego okresu warunkowego. Tym samym następnik kontrmożliwego okresu warunkowego (jego znaczenie, status modalny oraz wartość logiczna) nie ma wpływu na wartość logiczną całego wyrażenia.

MODYFIKACJA HEURYSTYKI

Powyższe pozwala nam powrócić do problemu heurystyk. Jak wskazywał Williamson, ich nieograniczona akceptacja jest w równym stopniu problematyczna dla zwolennika, jak i dla przeciwnika ortodoksyjnego poglądu, wobec czego nie należy ich traktować jako wy-

starczająco silnej motywacji dla odejścia od tego poglądu. Wynika to głównie z niespójności heurystyk z zasadami, które regulują użycie okresów warunkowych. Okoliczność ta rodzi pytanie, czy możliwe jest dostarczenie dla nich pewnej alternatywy, która z jednej strony oddałaby intuicje stojące za oceną wartości logicznych kontrmożliwych okresów warunkowych (oraz okresów warunkowych w ogólności), a z drugiej strony nie stałaby w konflikcie z uznaniem jednoczesnej prawdziwości zdań (a) oraz (b).

Zdaje się, że źródłem problematyczności jednoczesnej akceptacji (HCC*) oraz prawdziwości zdań (a) i (b) jest to, że choć następniki A oraz $\neg A$ są ze sobą niespójne, każdy z nich jest spójny z poprzednikiem $A \wedge \neg A$. Wszak jeśli w poprzedniku znajduje się koniunkcja, to jego konsekwencją może być dowolny z członów tej koniunkcji. Podobnie jest w przypadku ortodoksyjnego poglądu, gdzie niespójność pomiędzy A oraz $\neg A$ nie stoi na przeszkodzie, ażeby oba wyrażania były spójną konsekwencją poprzednika A , o ile ten wyraża niemożliwość. Ta ocierająca się o banał obserwacja, poprzedzona rozważaniami poczynionymi w poprzednim paragrafie, wskazuje na rolę, jaką w ocenie prawdziwości okresu warunkowego pełni poprzednik. Rola ta jest na tyle istotna, że w niektórych przypadkach poprzednik pozwala na uznanie prawdziwości zdań zawierających w sobie sprzeczność, a w innych dopuszcza absolutną dowolność w kwestii tego, co kryje się za następnikiem.

Wobec powyższego możemy pokusić się o próbę uwzględnienia tego faktu w zmodyfikowanej heurystyce. Gdybyśmy chcieli dać wyraz założeniom kryjącym się u podstaw ortodoksyjnego stanowiska, moglibyśmy wskazać na to, że „Jeśli A jest możliwe, uznając prawdziwość jednego z okresów warunkowych $A > C$ oraz $A > \neg C$, uznajemy fałszywość drugiego”. Formuła ta zdaje się oddawać to, w jaki sposób (według zwolenników poglądu ortodoksyjnego) oceniamy wartość logiczną okresów warunkowych. Wskazuje przy tym, że w przypadku, gdy A nie jest możliwe, zasada ta nie obowiązuje. Dzieje się tak oczywiście za sprawą tego, że niemożliwość A pozwala na akceptację prawdziwości zarówno $A > C$, jak i $A > \neg C$. Tym samym ortodoksyjną intuicję możemy wyrazić przy pomocy sformułowania alternatywnego dla (HCC*), zgodnie z którym:

(HCC**) „Jeśli A nie pozwala na jednoczesną prawdziwość $A > C$ oraz $A > \neg C$, jeden z tych okresów warunkowych jest fałszywy”.

Do pewnego stopnia analogiczna sytuacja ma miejsce również w przypadku teorii nieortodoksyjnych. Różnica polega jedynie na tym, że w tym ostatnim przypadku sama niemożliwość A nie jest wystarczającym warunkiem do tego, żeby móc uznać jednoczesną prawdziwość $A > C$ oraz $A > \neg C$. Nie jest to jednak nazbyt problematyczne, ponieważ (HCC**) nie przesądza o tym, jakie dokładnie warunki musi spełniać A . Tym samym, również stanowiska nieortodoksyjne mogą zaakceptować (HCC**) oraz wskazać, że tym, co czyni jednocześnie prawdziwymi $(A \wedge \neg A) > A$ oraz $(A \wedge \neg A) > \neg A$, jest to, że oba następniki spójne są z poprzednikiem.

Zarówno w przypadku stanowisk ortodoksyjnych, jak i nieortodoksyjnych znakomita większość przypadków spełnia (HCC*) bez konieczności rozpatrywania tego, czy następniki dwóch przeciwnych sobie okresów warunkowych są ze sobą spójne. Niemniej w obu przypadkach występują przykłady uniemożliwiające nieograniczone stosowanie się do (HCC*). Tym samym (HCC*) jest zawodną heurystyką i jedynie częściowo odzwierciedla to, w jaki sposób oceniamy wartość logiczną okresów warunkowych. Pełniejszą jej wersją jest przytoczona powyżej (HCC**). Wobec tego możemy uznać, że o ile (HCC) oraz (HCC*) są niedoskonałymi heurystykami, które mogły być wykorzystywane przez poszczególne stanowiska jedynie przy pewnych istotnych ograniczeniach, o tyle (HCC**) wolna jest od takich ograniczeń. Co więcej, w równym stopniu może ona być stosowana przez zwolenników stanowiska ortodoksyjnego, jak i przez zwolenników stanowiska nieortodoksyjnego. Uznając brak ograniczeń w opieraniu się na (HCC**) przy ocenie wartości logicznej okresów warunkowych, mamy powody do uznania jej nie za zawodną heurystykę, lecz za normatywną regułę, wyrażającą relacje zachodzące pomiędzy okresami warunkowymi o tych samych poprzednikach, lecz wzajemnie sprzecznych następnikach.

Jeśli jednak przystaniemy na akceptację (HCC**), pojawia się pytanie o to, dlaczego miałyby ona przemawiać na rzecz nieortodoksyjnej analizy okresów warunkowych. Wszak reguła ta w równym stopniu wspiera ortodoksyjne, jak i nieortodoksyjne stanowiska, co może sugerować, że w żadnym stopniu nie przybliżyła nas ona do opowiedzenia się po którejkolwiek ze stron. Choć faktycznie może tak być, to korzyścią płynącą z akceptacji (HCC**) jest możliwość wskazania, co było naszym celem, tj. na spójność motywacji stojących u podstaw stanowi-

ska nieortodoksyjnego z pozostałymi regułami, dotyczącymi okresów warunkowych.

Niemniej – jak twierdzi Williamson – niezależnie od domniemanej niespójności stanowisk nieortodoksyjnych można wskazać powód, dla którego powinniśmy skłaniać się ku analizie ortodoksyjnej. Jest nim to, że zwolennicy stanowiska nieortodoksyjnego zmuszeni są do postulowania m.in. światów niemożliwych, co skutkuje „niewiarygodną semantyką hybrydową”, wobec której atrakcyjniejsze wydaje się „ujednolicone podejście” przedstawione w ramach stanowiska ortodoksyjnego (Williamson 2016c). Uwaga ta skłania do podjęcia problemu metodologii analiz kontrfaktycznych okresów warunkowych.

ASPEKTY METODOLOGICZNE

Jeśli akceptacja (HCC**) pozostaje spójna z oboma stanowiskami, to może to sugerować, że tytułowy spór sprowadza się do sporu intuicji co do tego, jakie warunki powinien spełniać model wyjaśniający funkcjonowanie okresów warunkowych. Tym samym stajemy przed wyborem jednej z dwóch propozycji. Pierwszą z nich jest prosty model, który od blisko pięciu dekad cieszy się dużą popularnością, a który nakazuje uznanie każdego kontrmożliwego okresu warunkowego za prawdziwy. Alternatywą wobec niego jest relatywnie nowa propozycja, która oprócz światów możliwych postuluje również światy niemożliwe, a która uznaje, że jedynie niektóre spośród kontrmożliwych okresów warunkowych są prawdziwe.

Rozpatrując obie propozycje z punktu widzenia metodologii można zauważyć, że zwolennicy poszczególnych podejść akcentują różne walory teoretyczne swoich stanowisk. Zwolennicy pierwszego podkreślają prostotę modelu, przejawiającą się w postulowaniu jedynie światów możliwych. Choć prostota jest istotną cechą teoretyczną, to z pewnością nie jest cechą najistotniejszą¹⁰. Powodem tego jest to, że złożone alternatywy teoretyczne mogą być uznane za atrakcyjniejsze, o ile wyjaśniają większą liczbę danych. Warunek ten jest implicite wyrażony w znanej maksymie oszczędności, którą jest „brzytwa Ockha-

¹⁰ Niektórzy wskazują nawet, że prostota teorii jest jedynie wartością estetyczną i jako takiej nie powinno się jej przypisywać zbyt dużej roli w ocenie wartości poszczególnych teorii (Barcan Marcus 1995, s. 199).

ma”, a która w popularnym sformułowaniu głosi, że „nie należy mnożyć bytów ponad potrzebę”. Choć znakomita część filozofów powołuje się na tę zasadę, to niektórzy z nich skupiają się jedynie na pierwszym jej komponencie, który nakazuje, by nie mnożyć bytów. Pomijają przy tym, że byty te nie powinny być mnożone p o n a d potrzebę. Uprawione wydaje się przyjęcie, że ową potrzebą jest wyjaśnienie danych, które określa teoria uznaje za swój przedmiot. Tym samym maksymę oszczędności należy traktować jako głoszącą, że w przypadku, gdy obie teorie posiadają tę samą moc wyjaśniającą, powinniśmy faworyzować tę, która jest oszczędniejsza/prostsza, tj. tę, która postuluje mniejszą liczbę przedmiotów, hipotez, aksjomatów itp.

Wobec powyższego możemy uznać, że teoria ortodoksyjna, choć pod wieloma względami jest prostsza od nieortodoksyjnej, to złożoność tej ostatniej ma swoje uzasadnienie. Jest nim większa moc wyjaśniająca, co przejawia się podjęciem fenomenu przypisywania różnych wartości logicznych okresom warunkowym (4)–(9). Tym samym złożoność analiz nieortodoksyjnych nie musi wcale być uznawana za mnożenie bytów ponad potrzebę. Wręcz przeciwnie, owo mnożenie ma na celu zaspokojenie wskazanej potrzeby, którą jest próba wyjaśnienia nietrywialnej wartości logicznej okresów warunkowych.

Taka linia obrony stanowisk nieortodoksyjnych może spotkać się z zarzutem, który dotyczy pewnej nadgorliwości. Problem ten pojawia się wówczas, gdy określona teoria jest nazbyt wrażliwa w kwestii rozpoznawania danych, które są jej przedmiotem (Hitchcock, Sober 2004). Jak wskazują filozofowie, którzy zdiagnozowali ten problem, z uwagi na to, że często jesteśmy w błędzie w kwestii rozpoznawania tego, co jest, a co nie jest daną dla określonej teorii, dochodzi niekiedy do traktowania za rzeczywistą daną (*real data*) tego, co jest faktycznie jedynie „szumem informacyjnym” (*noise*) (Hitchcock, Sober 2004, s. 10). W takich sytuacjach mamy do czynienia z błędnym zaklasyfikowaniem zjawiska jako przedmiotu określonej teorii, co w konsekwencji prowadzi do nieuzasadnionego oczekiwania, ażeby teoria ta dążyła do jego wyjaśnienia.

Błędna identyfikacja danych może skutkować daleko idącymi konsekwencjami. Jeśli to, co jest jedynie szumem, zostanie zakwalifikowane jako dana wymagająca wyjaśnienia, to pojawia się ryzyko wprowadzania nieuzasadnionych zmian w teorii wyjściowej. Dzieje się tak zazwyczaj wówczas, gdy teoria ta zostaje odrzucona na podstawie błędnie

skonstruowanego kontrprzykładu (por. Williamson 2016d). Jak argumentował Williamson, za takie błędnie skonstruowane kontrprzykłady dla ortodoksyjnej interpretacji okresów warunkowych powinniśmy uznać zdania typu (4)–(9). Źródłem ich niepoprawności miałyby być to, że zostały one oparte na nieograniczonej akceptacji (HCC) lub (HCC*), która jest w równym stopniu problematyczna dla zwolenników obu stron tego sporu.

Jeśli jest to główny powód, dla którego postulaty zwolennika stanowiska nieortodoksyjnego miałyby być nieuzasadnione, to może on powołać się na regułę (HCC**). Jak wskazaliśmy wyżej, zdaje się ona korespondować z tym, w jaki sposób oceniamy wartość logiczną okresów warunkowych, a przy tym jest na tyle ogólna, że współgra zarówno z analizą ortodoksyjną, jak i nieortodoksyjną. Co najważniejsze pozwala ona wskazać na spójność motywacji tej ostatniej.

BIBLIOGRAFIA

- Barcan Marcus, R. (1995), *Modalities*, Oxford: Oxford University Press.
- Bennett, J. (2003), *A Philosophical Guide to Conditionals*, Oxford: Oxford University Press.
- Berto, F. (2013), *Impossible Worlds*, w: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL: <<http://plato.stanford.edu/entries/impossibleworlds>>.
- Brogaard B., Salerno J. (2013), *Remarks on Counterpossibles*, „Synthése” 190, s. 639–660.
- Edgington, D. (1995), *On Conditionals*, „Mind” 104, s. 235–329.
- Emery, N., Hill, C. (2017), *Impossible Worlds and Metaphysical Explanation: Comments on Kment's "Modality and Explanatory Reasoning"*, „Analysis” 77(1), s. 134–148.
- Grice, P.H. (1975/1977), *Logika a konwersacja*, tłum. J. Wajszczuk, „Przegląd Humanistyczny” 6, s. 85–99.
- Hitchcock, C., Sober, E. (2004), *Prediction Versus Accommodation and the Risk of Overfitting*, „British Journal for the Philosophy of Science” 55(1), s. 1–34.
- Ingarden, R. (1935), *Analiza zdania warunkowego*, w: R. Ingarden (1972), *Z teorii języka i filozoficznych podstaw logiki*, Warszawa: PWN, s. 260–313.
- Jackson, F. (1988), *Conditionals*, Oxford: Blackwell.
- Lewis, D. (1973), *Counterfactuals*, Oxford: Blackwell.
- Lewis, D. (1986), *On the Plurality of Worlds*, Oxford: Blackwell.
- Nolan, D. (1997), *Impossible Worlds: Modest Approach*, „Notre Dame Journal of Formal Logic” 38(4), s. 535–572.
- Priest, G. (2009), *Conditionals: A Debate with Jackson*, w: I. Ravenscroft (ed.), *Minds, Worlds, and Conditionals: Themes from the Philosophy of Frank Jackson*, Oxford: Oxford University Press, s. 311–335.

- Sendlak, M. (2013), *Teorie metafizyczne światów niemożliwych*, „Studia Philosophica Wratislaviensia”, Vol. VIII, Fasc. 3, s. 17–41.
- Stalnaker, R. (1968), *A Theory of Conditionals*, w: Nicholas Rescher (ed.), *Studies in Logical Theory*, Oxford: Blackwell, s. 98–112.
- Stalnaker, R. (1996), *Impossibilities*, „Philosophical Topics” 24, s. 193–204.
- Williamson, T. (2007), *Philosophy of Philosophy*, Oxford: Oxford University Press.
- Williamson, T. (2016a), *Knowing by Imagining*, w: P. Kung, A. Kind (eds.), *Knowledge Through Imagination*, Oxford: Oxford University Press, s. 113–132.
- Williamson, T. (2016b), *Counterpossibles*, „Topoi” [online first].
- Williamson, T. (2016c), *Counterpossibles in Metaphysics*, w: F. Kroon (ed.), *Philosophical Fictionalism*, Oxford: Oxford University Press [w druku].
- Williamson, T. (2016d), *Abductive Philosophy*, „The Philosophical Forum” 47, s. 263–280.
- Yagisawa, T. (1988), *Beyond Possible Worlds*, „Philosophical Studies” 53, s. 175–204.

ABOUT THE BASIS FOR THE DEBATE OVER COUNTERPOSSIBLES

SUMMARY: According to the most popular (so-called “orthodox”) theories, counterfactuals with impossible antecedents are vacuously true. Critiques of this view argue that contrary to this, we tend to consider only some of them true and others to be false. In his recent paper (“Counterpossibles”) Timothy Williamson has ingeniously explained the motivations for the orthodox view and argued that although there are some heuristic reasons that may suggest the plausibility of the unorthodox view, they are fallible. The most important of Williamson’s arguments is that the unorthodox interpretation is inconsistent with the heuristic assumption that supposedly motivates this very view. The aim of this paper is to consider Williamson’s critique and to support the unorthodox approach towards counterpossibles. In order to do so, we argue in favour of the modified version of the heuristic assumption.

KEYWORDS: counterfactuals, counterpossibles, possible worlds semantics, methodology, Timothy Williamson

MARIA MATUSZKIEWICZ*

PROBLEM AKTUALISTY MODALNEGO Z CZYSTO
MOŻLIWYMI INDYWIDUAMI
(O *MERE POSSIBILITIES* ROBERTA STALNAKERA)¹

Sądzimy, że świat mógłby być zupełnie inny, niż jest, nie tylko ze względu na to jak się rzeczy mają, lecz także ze względu na to, co istnieje. Nasze potoczne intuicje dopuszczają, że mogłoby istnieć coś innego niż to, co faktycznie istnieje. Pewne konkretne indywiduum, które istnieje (np. pisząca ten tekst), mogłoby nie istnieć, mógłby też istnieć ktoś inny (np. siostra bliźniaczka mojej babki po mieczu). Wydaje się, że możemy wypowiadać prawdziwe sądy dotyczące czysto możliwych indywiduów, np. to, że mogłabym mieć siostrę, która mogłaby być baletnicą, ale została bokserką. Przygodne istnienie indywiduów stanowi problem dla zwolennika aktualizmu, który uważa, że istnieje tylko świat rzeczywisty oraz indywidua go zamieszkujące. W książce *Mere Possibilities* Stalnaker chce pogodzić intuicję, że indywidua istnieją w sposób przygodny z aktualizmem i pokazać, jak reprezentować warunki prawdziwości zdań o czystych możliwościach na gruncie realistycznie interpretowanej semantyki światów możliwych. Drugim bardzo ważnym tematem tej książki jest relacja między metafizyką i semantyką światów możliwych oraz pytanie, czy semantyka ta dostarcza neutralnego narzędzia do opisu problemów metafizycznych.

* Uniwersytet Warszawski, Wydział Filozofii i Socjologii. E-mail: maria.j.matuszkiewicz@gmail.com. Praca nad tym artykułem była wspierana finansowo przez Narodowe Centrum Nauki w ramach grantu 2016/23/N/HS1/02180.

¹ Dziękuję Tadeuszowi Ciecierskiemu za wiele pomocnych uwag do pierwszej wersji tekstu.

MOŻLIWE ŚWIATY JAKO MAKSYMALNE WŁASNOŚCI

Stalnaker nie zgadza się ze stwierdzeniem Davida Lewisa, że jedyną alternatywą dla realizmu modalnego jest traktowanie światów możliwych jako erzaców. Jest on umiarkowanym realistą i traktuje światy możliwe jako maksymalne własności, jakie mógłby mieć świat (inaczej mówiąc: sposoby, na jakie świat mógłby istnieć). Maksymalna własność (możliwy świat) to maksymalny niesprzeczny zbiór sądów, to znaczy taki zbiór, że dla dowolnego sądu p , p albo jego negacja wynika z tego zbioru. Różnica między światem rzeczywistym a światami możliwymi, jest różnicą między własnością, która jest egzemplifikowana, i własnościami, które nie są egzemplifikowane. Stalnaker traktuje pojęcie własności jako pojęcie modalne: własności należy rozumieć, myśląc o tym, jaki byłby świat, gdyby własności te były egzemplifikowane. Taka koncepcja nie ma zatem na celu redukcji pojęć modalnych, a jedynie ich uporządkowanie. Zaletą tej koncepcji jest to, że unika ona pewnego problemu, na który – jak zwrócił uwagę Lewis – narażone są koncepcje traktujące możliwe światy jako reprezentacje. O ile zasadne jest pytanie, za sprawą czego dana reprezentacja reprezentuje taki a nie inny stan rzeczy, o tyle nie ma sensu pytać, dlaczego własność bycia światem, w którym wszystkie koty są czarne, stanowi reprezentację własności bycia światem, w którym wszystkie koty są czarne, a nie własności bycia światem, w którym wszystkie kruki są białe.

Stalnaker rozróżnia wśród własności, jakie mógłby egzemplifikować świat, własności ogólne i szczegółowe. Własność ogólną należy rozumieć jako funkcję propozycjonalną niewymieniającą konkretnych przedmiotów. Własność szczegółowa jest funkcją propozycjonalną wymieniającą konkretne przedmioty. Istnieją też własności mieszane, ogólno-szczegółowe. Na przykład własność bycia światem, w którym X ma siedem siostr jest własnością mieszaną: szczegółową, ponieważ reprezentująca ją funkcja propozycjonalna wymienia konkretną osobę, i ogólną, ponieważ nie wymienia siedmiu konkretnych osób, siostr X -a. Natomiast własność bycia światem, w którym pewne konkretne indywidua, a i b , są siostrami X -a jest własnością szczegółową.

Problemem, z jakim musi sobie poradzić zwolennik aktualizmu, są sądy dotyczące czysto możliwych indywiduów. Sądy te wyrażają pew-

ne fakty modalne. Nasze potoczne intuicje modalne dopuszczają jako sensowne (posiadające warunki prawdziwości) stwierdzenie: „Mogłabym mieć siedem starszych siostr”. Zgodnie z semantyką światów możliwych zdanie to jest prawdziwe zawsze i tylko wtedy, gdy istnieje taki świat możliwy, w którym autorka tego tekstu ma siedem starszych siostr. Istnieje zatem pewna własność ogólna, która nie jest egzemplifikowana, ale mogłaby być egzemplifikowana. Gdyby jednak ta własność ogólna była egzemplifikowana, musiałaby być zarazem egzemplifikowana odpowiadająca jej własność szczegółowa, w tym wypadku własność bycia światem zawierającym siedem indywiduów, będących moimi starszymi siostrami (oraz inne szczegółowe własności z nimi związane). Żadna szczegółowa własność tego rodzaju jednak nie istnieje, ponieważ zgodnie z aktualizmem nie istnieje ktoś, kto mógłby być moją starszą siostrą. Zatem: ogólny sąd, że mogłabym mieć siedem starszych siostr, jest prawdziwy (w naszym świecie), mimo że nie istnieje (w naszym świecie) żaden sąd jednostkowy dotyczący kogoś będącego moją siostrą. Gdyby jednak światem rzeczywistym był świat czyniący prawdziwym to ogólne stwierdzenie (a więc świat, w którym mam siedem siostr), wówczas musiałoby istnieć siedem sądów jednostkowych dotyczących moich siedmiu siostr. Twierdzenie, że mogłoby istnieć coś, co nie istnieje w świecie rzeczywistym, ma paradoksalną konsekwencję: może istnieć ogólny sąd egzystencjalny, nawet jeśli nie istnieje żaden odpowiadający mu sąd szczegółowy.

Problemem dla aktualisty są też niektóre iterowane twierdzenia modalne, wydaje się bowiem, że wyrażają one sądy *de re* o czysto możliwych indywiduach (McMichael 1983). Mogę sensownie powiedzieć: „Mogłabym mieć siedem starszych siostr, z których ta, która mogła zostać baletnicą, została bokserką, a ta, która mogła zostać bokserką, została baletnicą”. Zgodnie z semantyką światów możliwych zdanie to jest prawdziwe wtedy i tylko wtedy, gdy istnieje świat x , w którym siostra a ma własność F , a siostra b ma własność G , oraz gdy istnieje świat y , taki że jest w nim prawdą o a , że ma ona własność G , o b natomiast prawdą jest, że ma własność F . Wydaje się, że mówimy w ten sposób *de re* o czysto możliwych indywiduach a i b . Sądy *de re* są zależne od przedmiotu, zatem taki sąd istniałby i moglibyśmy go wyrazić, tylko gdyby światem rzeczywistym był świat kontrfaktyczny zawierający te dwa indywidua. Wedle aktualisty nie możemy wypowiadać żadnych sądów *de re* dotyczących czysto możliwych indywiduów.

Problemem jest też to, że po to, żeby odróżnić wspomniane dwa światy (zakładając, że nie ma między nimi żadnych innych różnic), musimy zakładać tożsamość tych czysto możliwych indywiduów w dwóch światach. Jak możemy jednak mówić o tożsamości czegoś co nie istnieje?

Stalnaker uważa, że semantyka możliwych światów powinna pozwalać reprezentować warunki prawdziwości stwierdzeń o czysto możliwych indywiduach, także iterowanych stwierdzeń modalnych, ponieważ za ich pomocą wyrażamy pewne ogólne fakty modalne dotyczące świata rzeczywistego. Jednak po to, żeby móc wyrazić te fakty, musimy móc odnosić się do czysto możliwych własności i indywiduów. Myśl tę Stalnaker ilustruje, przywołując znany przykład Kripkego: wyobraźmy sobie dwa światy możliwe składające się tylko z dwóch kostek. Możemy sensownie powiedzieć o dwóch fikcyjnych kostkach, że na jednej kostce mogłoby wypaść 5, a na drugiej 6, ale że mogłoby też wypaść odwrotnie. Wypowiadając takie zdanie, tylko pozornie – jak przekonuje nas Stalnaker – stwierdzamy coś na temat możliwych indywiduów; faktycznie, mówiąc o *possibiliach*, opisujemy tylko pewne ogólne własności modalne. Możemy definiować takie czysto możliwe własności, mówiąc o pewnych własnościach drugiego rzędu (Stalnaker 2012, s. 18–19). Zdanie o dwóch możliwych kostkach mówi jedynie o pewnej ogólnej własności modalnej, mianowicie takiej, że jeśli mamy 2 kostki, to istnieje 36 możliwych wyników, które mogą wypaść.

PRYZGODNE ISTNIENIE SĄDÓW

Stalnaker godzi twierdzenie o przygodnym istnieniu indywiduów z aktualizmem i z realistyczną interpretacją semantyki światów możliwych:

- (i) uznaje, że nie wszystkie sądy istnieją w sposób konieczny,
- (ii) przyjmuje pewną modyfikację tradycyjnej semantyki możliwych światów pozwalającej mu na oddzielenie tych własności modelu, które mają realistyczną interpretację, od tych, które są tylko artefaktami modelu, pozbawionymi własności reprezentacyjnych.

Konsekwencją przygodnego istnienia sądów jest to, że różne sądy istniałyby z perspektywy różnych światów w zależności od tego, który z nich byłby światem rzeczywistym. Sąd jednostkowy o mojej (możli-

wej) siostrze, która mogłaby zostać baletnicą, a została bokserką, istniałby w pewnym kontrfaktycznym świecie, chociaż nie istnieje w świecie rzeczywistym. Z perspektywy takiego świata istniałby także sąd jednostkowy o mojej siostrze, że ona nie istnieje. Paradoksalną konsekwencją takiej koncepcji (o czym jeszcze będzie mowa) jest to, że chociaż sąd ten nie istnieje, a zatem nie jest prawdziwy w naszym świecie, to gdyby światem rzeczywistym był świat zawierający moją siostrę, sąd ten nie tylko by istniał, lecz także byłby prawdziwy o naszym świecie (Stalnaker 2012, s. 28).

Jeśli uznamy istnienie sądów przygodnych, to maksymalny niesprzeczny zbiór sądów może być maksymalny niesprzeczny zgodnie z podaną definicją, ale może nie być maksymalny w takim sensie, że mógłby być bardziej uszczegółowiony w zależności od tego, który możliwy świat okazałby się światem rzeczywistym. Wymusza to pewną modyfikację tradycyjnej semantyki światów możliwych. Stalnaker przedstawia swoją propozycję obrazowo w następujący sposób: proponuje, żeby myśleć o światach możliwych nie jako o punktach przestrzeni logicznej (zgodnie ze standardową semantyką Kripkego), lecz jako o komórkach podziału przestrzeni logicznej, wewnątrz których można by wyróżnić nieskończenie wiele punktów. Świat rzeczywisty reprezentuje jeden z punktów. Jest to maksymalny niesprzeczny zbiór sądów, który istnieje, jeśli ten świat jest światem rzeczywistym. Każdy z punktów mógłby reprezentować pewien świat, gdyby ów świat był światem rzeczywistym. Dla każdego punktu istnieje funkcja odwzorowująca go na klasę równoważnych punktów. To te zbiory równoważnych punktów – komórki podziału przestrzeni logicznej – są światami możliwymi. Punkty wewnątrz tej samej komórki nie różnią się własnościami reprezentacyjnymi, natomiast to, że jest ich wiele, odzwierciedla fakt, że maksymalne spójne sądy mogłyby być dalej uszczegółowione.

To, że różne możliwe światy istnieją z perspektywy różnych światów, powoduje pewne trudności. McMichael zarzucił takiej koncepcji to, że rezygnuje ona z ekstensjonalności semantyki możliwych światów (McMichael 1983, s. 5) Rozważmy zdanie: „Mogłabym mieć siostrę, która mogłaby zostać baletnicą, a została bokserką”, które reprezentuje zapis: $\Diamond \exists x [S(x) \wedge F(x) \wedge \Diamond G(x)]$. Na gruncie semantyki możliwych światów jego warunki prawdziwości reprezentuje zdanie o postaci: $\exists w \exists x [S(x) \wedge F(x) \wedge \exists w' G(x)]$. McMichael stwierdza, że kwantyfikator zewnętrzny i wewnętrzny nie przebiegają po tym samym zbiorze

możliwych światów. Na gruncie takiej koncepcji nie możemy założyć uniwersalnej dziedziny światów, po której kwantyfikujemy; z każdą iteracją musimy zakładać od nowa zbiór światów, po których kwantyfikujemy.

Stawia to także pod znakiem zapytania możliwość realistycznej interpretacji semantyki możliwych światów. Co reprezentują punkty w przestrzeni logicznej? Stalnaker stara się pogodzić realistyczną interpretację semantyki z aktualizmem i z twierdzeniem, że indywidua i sądy istnieją w sposób przygodny, oddzielając te elementy semantyki, które mają wartość reprezentacyjną od tych, które jej nie mają. Metoda ta polega na wskazywaniu – przez określenie odpowiedniej klasy równoważności – niezmienników. Podobnie postępuje relacjonista przestrzenny. Relacjonista twierdzi, że wszystkie własności przestrzenne sprowadzają się do relacji między ciałami. Aby reprezentować własności i relacje przestrzenne, posługuje się on jednak abstrakcyjną przestrzenią matematyczną: punkty tej przestrzeni pełnią rolę lokalizacji przestrzennych zajmowanych przez ciała. Relacjonista może posługiwać się taką reprezentacją, nawet jeśli nie uznaje istnienia absolutnych lokalizacji przestrzennych. Twierdzi on, że rzeczywiste własności przestrzenne są tym, co jest niezmiennicze względem wszystkich modeli. Załóżmy, że mamy trzy rysunki przedstawiające dwa trójkąty:

- (i) pewien rysunek wyjściowy,
- (ii) rysunek różniący się tym, że oba trójkąty są przemieszczone o 10 cm w prawo oraz odwrócone o 90° ,
- (iii) rysunek, na którym cała konfiguracja jest przedstawiona w skali 100:1.

Możemy powiedzieć, że rysunki te reprezentują te same własności przestrzenne figur, ponieważ istnieje odpowiednie przekształcenie matematyczne odwzorowujące je wzajemnie na siebie.

Podobną strategię stosuje Stalnaker po to, żeby oddzielić te własności modelu, które stanowią reprezentację rzeczywistych indywiduów, własności i relacji, od tych, które są tylko częścią modelu, niemającą własności reprezentacyjnych. Zgodnie z propozycją Stalnakera (Stalnaker 2010, s. 24) światy x i y należące do zbioru światów możliwych danego modelu są równoważne ze względu na świat w wtedy i tylko wtedy, gdy dla dowolnego zdania p zawierającego nazwy własne elementów dziedziny w , zdanie p ma taką samą wartość logiczną w y i w x .

Analogiczne relacje równoważności można określić dla każdego świata należącego do zbioru możliwych światów. Tak wyznaczone klasy równoważności modelują maksymalne sądy, które istniałyby, gdyby dany świat był zrealizowaną możliwością. Dwa światy x i y , które są równoważne względem rzeczywistego świata nie różnią się znaczeniem reprezentacyjnym. Innymi słowy, jeśli pewne zdanie p niezawierające nazw desygnujących indywidua należące do świata rzeczywistego jest prawdziwe w x , a fałszywe w y , to p nie wyraża właściwego faktu modalnego. (Stalnaker 2010, s. 24). Intuicja metafizyczna, którą wyraża ta propozycja, jest taka, że zdania dotyczące czysto możliwych indywiduów nie reprezentują żadnych rzeczywistych możliwości.

ARGUMENT PLANTINGI

Twierdzenie, że sądy istnieją w sposób przygodny, opiera się na dwóch przesłankach:

- (i) że indywidua istnieją w sposób przygodny,
- (ii) że sądy jednostkowe są ontologicznie zależne od indywiduów, których dotyczą (egzystencjalizm).

Jeśli (i) i (ii) są prawdziwe, to prawdziwe jest także to, że niektóre sądy istnieją tylko przygodnie. Pierwszą przesłankę uznajemy, zdaniem Stalnakera, ponieważ jest ona zgodna z naszymi intuicjami i jako taka nie wymaga dodatkowego uzasadnienia. Drugie twierdzenie wymaga natomiast wyjaśnienia. Twierdzenie o ontologicznej zależności sądów od indywiduów wynika bezpośrednio z pierwszej przesłanki, jeśli jesteśmy zwolennikami strukturalnej Russellowskiej koncepcji sądów. Zgodnie z tą koncepcją sądy mają składowe odpowiadające wyrażeniom składowym zdania i składowymi sądu są indywidua oraz własności, będące desygnatami tych wyrażen. Stalnaker chce pokazać, że twierdzenie egzystencjalizmu jest prawdziwe, niezależnie od tego, jaką koncepcję sądu przyjmujemy. Jest ono także prawdziwe na gruncie preferowanej przez Stalnakera koncepcji, zgodnie z którą sądy są funkcjami ze światów możliwych w wartości logiczne i nie mają składowych. Tak rozumiane sądy także są ontologicznie zależne od przedmiotów, ponieważ tym, co zgodnie z tą koncepcją indywiduuje dany sąd, jest funkcja wartościowania po światach możliwych. Po to, żeby dany sąd miał wartość logiczną, musi istnieć przedmiot, którego doty-

czy. W wypadku, w którym przedmiot nie istnieje, nic nie indywidualuje tego sądu.

Stalnaker krytykuje konkurencyjne podejście Alvina Plantingi (Plantinga 1983), który unika przygodnego istnienia sądów. Plantinga przyjmuje przesłankę (i): że indywidua istnieją przygodnie, natomiast odrzuca przesłankę (ii): że sądy są ontologicznie zależne od przedmiotów. Argument Plantingi przeciwko drugiej przesłance (zrekonstruowany przez Stalnakera) wygląda następująco (Stalnaker 2012, s. 44):

P1. Jest możliwe, że Sokrates nie istnieje.

P2. Jeśli P1, to sąd „Sokrates nie istnieje” jest możliwy.

P3. Jeśli sąd „Sokrates nie istnieje” jest możliwy, to sąd „Sokrates nie istnieje” może być prawdziwy.

P4. Jest konieczne, że gdyby sąd „Sokrates nie istnieje” był prawdziwy, to sąd „Sokrates nie istnieje” istniałby.

P5. Jest konieczne, że gdyby sąd „Sokrates nie istnieje” był prawdziwy, to Sokrates by nie istniał.

Wniosek: Możliwe, że Sokrates nie istnieje i istnieje sąd „Sokrates nie istnieje”.

Zdaniem Plantingi argument ten pokazuje, że sądy nie mogą być ontologicznie zależne od przedmiotów. Stalnaker odrzuca argument Plantingi, wprowadzając parę pojęć: możliwości w świecie i możliwości względem świata (oraz analogiczne pojęcia prawdziwości w świecie i prawdziwości o świecie). Zdaniem Stalnakera zwolennik ontologicznej zależności sądów od przedmiotów może odrzucić argument Plantingi, twierdząc, że opiera się on na ekwiwocacji. Predykat „możliwy” może bowiem znaczyć: (i) możliwy w pewnym świecie, czyli taki, który może być prawdziwy w pewnym świecie: $\lambda x \diamond T(x)$ (T oznacza predykat prawdziwości), albo (ii) możliwy względem danego świata, czyli taki, który wynika logicznie z pewnego świata (maksymalnego spójnego sądu): $\lambda x (\exists w) (I(w, x))$ (gdzie I oznacza wynikanie logiczne). Jednocześnie podkreśla, że te dwa znaczenia byłyby równoważne, gdyby wszystkie sądy istniały w sposób konieczny, natomiast nie jest równoważne, jeśli dopuszczamy przygodne istnienie sądów.

Jeśli rozumiemy „możliwy” na pierwszy z tych sposobów (jako taki, który może być prawdziwy w pewnym świecie), zwolennik ontologicznej zależności sądów od przedmiotów powinien odrzucić przesłankę P2, uznać zaś przesłankę P3. P2 głosi: jeśli jest możliwe, że S nie ist-

nieje, to sąd „*S* nie istnieje” mógłby być prawdziwy w pewnym świecie. Zwolennik tezy o zależności ontologicznej sądów od przedmiotów odrzuca P2, ponieważ sąd „to, że *S* nie istnieje”, nie może być prawdziwy ani w świecie, w którym *S* nie istnieje, bo prawdziwość orzeczek można tylko o czymś, co istnieje, ani w takim, w którym *S* istnieje (wówczas nie może być prawdziwy). Zatem P2 jest fałszywe. Przesłankę P3 na gruncie pierwszego odczytania predykatu „możliwy” uznamy trywialnie: poprzednik P3 jest koniecznie fałszywy, ponieważ sąd „to, że *S* nie istnieje” nie może być prawdziwy w żadnym świecie.

Jeśli rozumiemy „możliwy” na drugi z tych sposobów (jako taki, który wynika logicznie z pewnego świata), zwolennik ontologicznej zależności sądów od przedmiotów powinien uznać P2 (bo jeśli istnieje pewien świat, w którym *S* nie istnieje, to sąd jednostkowy „to, że *S* nie istnieje” wynika logicznie z tego świata), a odrzucić P3 (bo jeśli sąd „to, że *S* nie istnieje” wynika logicznie z pewnego świata, to sąd ten nie może być prawdziwy w tym świecie).

Ceną, jaką trzeba zapłacić za dopuszczenie przygodnego istnienia sądów, jest przyjęcie rozróżnienia dwóch sensów predykatu „prawdziwy”: prawdziwy w świecie i prawdziwy o świecie (analogicznych do rozróżnienia dwóch sensów predykatu „możliwy”). Sąd jest prawdziwy w pewnym świecie, jeśli ma w nim warunki prawdziwości, natomiast jest prawdziwy o świecie, jeśli wynika z niego logicznie. Konieczność tego rozróżnienia wynika z czterech przesłanek:

- (1) istnieją tylko przygodne sądy,
- (2) dla każdego sądu istnieje sąd z nim sprzeczny,
- (3) z konieczności tylko sądy istniejące mają własność bycia prawdziwymi,

(4) światy możliwe są z założenia maksymalne tzn. dla dowolnego sądu *p* i świata *w*, albo *p* albo $\neg p$ wynika z *w*: $[\forall p \forall w I(w, p) \vee I(w, \neg p)]$. Jeśli *p* jest sądem przygodnym, to istnieje taki świat *w*, w którym *p* nie istnieje. Nie istnieje też $\neg p$. Na mocy (4) albo *p* albo $\neg p$ wynika z *w*. Zatem istnieje taki sąd, które nie istnieje w pewnym świecie, ale jest prawdziwy o tym świecie (Stalnaker 2012, s. 48). Przykład: Sąd „Święty Mikołaj nie istnieje” nie istnieje w świecie rzeczywistym, natomiast gdyby światem rzeczywistym był świat, w którym istnieje Święty Mikołaj, to sąd „Święty Mikołaj nie istnieje” nie byłby prawdziwy w naszym świecie (gdyż po to, żeby być prawdziwym, trzeba istnieć), ale byłby prawdziwy o naszym świecie, tzn. wynikałby logicznie z naszego świata.

Stalnaker uważa, że stanowisko uznające przygodny charakter sądów, mimo że wymusza pewne odstępstwa od ortodoksyjnej semantyki światów możliwych, ma przewagę nad konkurencyjnymi stanowiskami Alvina Plantingi i Timothy'ego Williamsona. Alvin Plantinga (1983) godzi twierdzenie, że istnieją przygodne indywidua z twierdzeniem, że wszystkie sądy istnieją w sposób konieczny, przyjmując pewne kontrowersyjne założenie. Choć indywidua istnieją w sposób przygodny, ich *haecceitates* – rozumiane przez Plantingę jako własności bycia identycznym z indywiduum – istnieją w sposób konieczny. Własności te pozwalają indywiduować sądy jednostkowe o tylko możliwych indywiduach. Uznanie, że istnieją *haecceitates* wszystkich czysto możliwych indywiduów kłóci się jednak z duchem aktualizmu. Z kolei Timothy Williamson (2005) przyjmuje ontologiczną zależność sądów jednostkowych od indywiduów, których sądy te dotyczą, odrzuca jednak przygodne istnienie indywiduów. Jego zdaniem wszystko, co istnieje, istnieje w sposób konieczny. Obaj autorzy, zdaniem Stalnakera, zmuszają nas do przyjęcia pewnych wysoce nieintuicyjnych założeń metafizycznych.

HAECCEITYZM I ANTYHAECCEITYZM

Trzecią strategią, mającą pozwolić uniknąć przygodnych sądów, jest przyjęcie antyhaecceityzmu: poglądu, że czysto jakościowa charakterystyka świata możliwego stanowi jego pełną charakterystykę. Wedle tego poglądu bycie danym indywiduum sprowadza się do posiadania pewnych własności jakościowych. Sądy dotyczące czysto możliwych indywiduów można więc wyrazić za pomocą własności jakościowych. Jest wątpliwe – jak zauważa Stalnaker – czy faktycznie przyjęcie antyhaecceityzmu pozwala nam uniknąć przygodnych sądów: jeśli nic nie wyklucza istnienia czysto przygodnych własności (np. własność bycia czerwonym nie istnieje w świecie, w którym nie ma ani fal świetlnych, ani istot zdolnych do doznawania kolorów), to przygodnie istnieją też sądy, które ontologicznie zależą od istnienia tych własności. Abstrahując jednak od tej trudności, Stalnaker zastanawia się nad tym, w jakim stopniu haecceityzm i antyhaecceityzm (w zależności od sposobu sformułowania tych stanowisk) wyrażają nasze intuicje metafizyczne. Omawiając ten spór, Stalnaker zwraca uwagę na związki między problemami semantycznymi a problemami metafizycznymi oraz sugeruje właściwe, jego zdaniem, podejście rozdzielające te kwestie.

Stalnaker przytacza trzy różne charakterystyki sporu między stanowiskiem haecceitystycznym i antyhaecceitystycznym. David Kaplan (1975) charakteryzuje doktrynę antyhaecceityzmu jako twierdzenie, że nie możemy sensownie mówić o tożsamości indywiduów w różnych światach możliwych: tożsamość dwóch indywiduów w różnych światach nie jest ugruntowana ani w podobieństwach jakościowych, ani w nieredukowalnej własności istotowej, po prostu nie istnieje. Haecceitysta twierdzi, że możemy mówić o tożsamości indywiduów w różnych światach możliwych, abstrahując od własności oraz zachowania tych indywiduów. Tożsamość indywiduum w różnych możliwych światach opiera się nieredukowalnej własności istotowej.

Kaplan opatrzył charakterystykę haecceityzmu następującym komentarzem: „choć wydawałoby się, że to antyhaecceitysta twierdzi, że żadne indywiduum nie istnieje w więcej niż jednym możliwym świecie, faktycznie pogląd ten należy przypisać haecceityście uznającemu wyjątkowo silny metafizyczny determinizm” (Stalnaker 2012, s. 55). Uwaga ta, zdaniem Stalnakera, rzuca pewne światło na stanowisko Lewisa, mieszającego w pewnym stopniu problemy metafizyczne i semantyczne. Silny determinista metafizyczny to ktoś, kto twierdzi, że żadne indywiduum nie może należeć do dwóch różnych światów (zob. Ciecierski 2003). Konsekwencją tego poglądu jest to, że wszystko ma swoje własności w sposób konieczny. Z pewnej perspektywy można przypisywać Lewisowi silny determinizm metafizyczny: zgodnie z jego teorią odpowiedników żadne indywiduum nie istnieje w więcej niż jednym świecie, a tym, co czyni prawdziwym zdanie, że pewien przedmiot mógłby być inny, niż jest, jest odpowiednik tego przedmiotu w pewnym świecie możliwym. Na przykład tym, co czyni prawdziwym sąd, że mogłabym mieć w kieszeni dwie monety zamiast trzech, jest to, że mój odpowiednik w innym świecie możliwym (a nie ja) ma w kieszeni dwie monety. Zwolennik silnego metafizycznego determinizmu mógłby zgodzić się z Lewisem, że istnieje ktoś w innym świecie możliwym, kto przypomina pod wieloma względami mnie, a różni się tylko tym, że ma w kieszeni dwie monety, a nie trzy. Nie zgodziłby się jednak z tym, że ja mogłabym mieć w kieszeni dwie monety zamiast trzech. Stalnaker zwraca uwagę, że to, o co spierają się Lewis i silny metafizyczny determinista, dotyczy tego, czy przedmiot w innym świecie możliwym może reprezentować własności modalne pewnego rzeczywistego przedmiotu, natomiast ich metafizyka jest zasadniczo podobna.

Wedle Roberta Adamsa (1979) haecceityzm i antyhaecceityzm spierają się co do tego, czy byciem tym oto (*thisness*) – przez które Adams rozumie własność bycia identycznym z danym indywiduum – jest redukowalne do pewnej własności jakościowej²: antyhaecceitysta uważa, że tak, haecceitysta – że nie. David Lewis (1986) charakteryzuje doktrynę haecceityzmu za pomocą pojęcia superweniencji: światy mogą różnić się albo dystrybucją naturalnych własności jakościowych albo mogą różnić się tym, jak reprezentują *de re* pewne indywiduum. Antyhaecceitysta to ktoś, kto uważa, że wszystkie różnice drugiego rodzaju superwenują na różnicach pierwszego rodzaju. Stalnaker zwraca uwagę, że charakterystyka Lewisa ma pewien rys szczególny: Lewis nie twierdzi po prostu, że różnice niejakościowe superwenują na różnicach jakościowych, ale że różnice związane z tym, jak możliwe światy użyte są do reprezentowania sądów *de re* o indywiduach superwenują na różnicach jakościowych. Lewis, inaczej niż większość zwolenników antyhaecceityzmu, dopuszcza bowiem istnienie jakościowo nieodróżnialnych światów, które są różne. Spór między stanowiskiem haecceitystycznym i antyhaecceitystycznym – wedle charakterystyki Lewisa – nie dotyczy tego, czy możemy zakładać tożsamość indywiduów w różnych światach możliwych (Lewis, będąc realistą modalnym i zwolennikiem teorii odpowiedników, uważa, że żadne dwa indywidua w różnych światach nie mogą być tożsame ze sobą), ale tego, czy relacja bycia odpowiednikiem superwenuje na własnościach jakościowych. Charakteryzując w ten sposób spór między haecceityzmem i antyhaecceityzmem, Lewis, co podkreśla Stalnaker, zaciera granicę między metafizyką i semantyką (Stalnaker 2012, s. 59).

Stalnaker zwraca uwagę, że przynajmniej dwie z przedstawionych charakterystyk: Lewisa i Adamsa, opierają się na nieoczywistym rozróżnieniu na własności jakościowe i niejakościowe. Obaj autorzy przeprowadzają to rozróżnienie, odwołując się do języka. Stalnaker próbuje modelować różnicę między stanowiskami haecceitystycznym i antyhaecceitystycznym za pomocą Lewisowskiej metafory silnego języka, żeby wskazać problemy związane z takim podejściem. Wyobraźmy sobie, że mamy język o maksymalnej sile wyrażalności. Wszyst-

² Adams używa pojęcia *thisness*, przez które rozumie własność bycia identycznym z danym przedmiotem (ontologicznie zależną od jego istnienia), i odróżnia je od pojęcia *haecceitas*, przez które rozumiano niejakościowy element składowy przedmiotu, indywiduujący go i mogący istnieć samodzielnie (Adams 1981, s. 12–13).

kie jego podstawowe predykaty mają charakter jakościowy, język ten jest ponadto wystarczająco bogaty, by reprezentować wszystkie możliwe różnice jakościowe. Jeśli dodamy do tego języka nazwy własne dla wszystkich indywiduów, to możemy wyrazić *haecceitates* przedmiotów za pomocą predykatu $\lambda x(x = a)$. Możemy powiedzieć, że antyhaecceitysta to ktoś, kto twierdzi, że wszystkie sądy dotyczące *haecceitates* superwenują na sądach dotyczących własności jakościowych³.

Podejście to nie jest jednak, jak pokazuje Stalnaker, satysfakcjonujące. Po pierwsze nie mamy kryterium odróżniającego predykaty jakościowe od predykatów niejakościowych: nie ma żadnej syntaktycznej różnicy między nimi. Po drugie, nawet zakładając, że takie rozróżnienie jest oczywiste, problemem dla proponowanego podejścia są własności wyrażone za pomocą operatorów modalnych. Musimy bowiem albo uznać wszystkie własności modalne za własności jakościowe, albo traktować własności modalne jako własności niejakościowe. Przyjmując pierwsze podejście uznalibyśmy za czysto jakościowe także problematyczne własności z iterowanym operatorem modalnym (np. własność, że mogłoby się mieć siedem siostr, z których jedna mogłaby być baletnicą, a została bokserką). Takie podejście prowadzi do trudności. Stalnaker ilustruje te trudności przykładem z kostkami: wyobraźmy sobie świat składający się tylko z dwóch indywiduów: dwóch kostek *A* i *B*, oraz sześciu predykatów: 1–6. Na kostkach faktycznie wypadły następujące wyniki *A*: 5, *B*: 6, i jest to jedyna własność odróżniająca je w świecie rzeczywistym. Załóżmy też dodatkowo, że mogły one wypaść odwrotnie (*B*: 6, *A*: 5). Zdanie mówiące, że kostki mogłyby wypaść inaczej, nie może w tym modelu być prawdziwe, ponieważ te dwa światy (świat 5–6 i świat 6–5) są jakościowo nieodróżnialne. Zatem antyhaecceitysta musi powiedzieć, że na tych dwu kostkach nie mogły wypaść inne wyniki, albo uznać różnicę między tymi światami za nieredukowalną różnicę niejakościową. Załóżmy jednak, że jest prawdą o każdej z kostek, że mogło na niej wypaść 5 lub 6, ale jest prawdą tylko o jednej z nich, że mogło wypaść 1. Wówczas możemy odróżnić świat 5–6 od świata 6–5 (światy te różnią się własnościami modalnymi kostek). Jak jednak wskazuje Stalnaker, reidentyfikacja kostek w innym świecie

³ Relację superwenuencji definiuje się jako relację między klasami własności. Zatem relację superwenuencji w tym kontekście należy rozumieć tak, że prawdziwość sądów z jednej klasy superwenuje na prawdziwości sądów z drugiej.

możliwym nie jest oparta na żadnej własności jakościowej, co nie jest zgodne z duchem antyhaecceityzmu (Stalnaker 2012, s. 65).

Stalnaker rozważa także takie ujęcie stanowiska antyhaecceitystycznego, które nie wymaga odwołania do językowej metafory. Podstawowe dla tej koncepcji jest pojęcie przestrzeni własności, zaczerpnięte od Basa van Fraassena (1967). Przestrzeń własności to n -wymiarowa abstrakcyjna przestrzeń, której n wymiarów reprezentuje n własności jakościowych. Opisać możliwy świat to podać liczbę wymiarów oraz funkcję odwzorowującą indywidua na lokalizacje w przestrzeni. Dwa światy x i y są nieodróżnialne, gdy każdemu indywiduum świata x zajmującego punkt m odpowiada pewne indywiduum świata y zajmujące odpowiadający m punkt m' . Stanowisko antyhaecceitystyczne głosi, że dwa nieodróżnialne światy są tożsame. Jednak i to rozwiązanie nie jest – zdaniem Stalnakera – satysfakcjonujące (Stalnaker 2012, s. 69–70):

(i) nic nie wyklucza tego, żeby w pewnym świecie ten sam punkt zajmowały dwa jakościowo nieodróżnialne – choć różne – indywidua (co jest niezgodne z charakterystyką antyhaecceityzmu podaną przez Adamsa),

(ii) własności modalne nie są redukowalne do lokalizacji indywiduum w przestrzeni własności. Ten drugi problem ponownie ilustruje przykład z kostkami: może być prawdą, że na kostce, na której wypadło 5, mogło wypaść 3, a na tej, na której wypadło 6, mogło wypaść 4. Tożsamość kostek z wynikami 5 i 3 oraz 6 i 4 (w różnych światach) nie wynika z lokalizacji tych kostek w przestrzeni własności, lecz z ich nieredukowalnej tożsamości.

Stalnaker twierdzi, że można by uratować taką charakterystykę stanowiska antyhaecceitystycznego, przyjmując teorię redukującą indywidua do wiązek własności. Wedle takiej teorii nie możemy mówić o indywiduach zajmujących miejsca w przestrzeni własności, lecz po prostu o miejscach w przestrzeni, które są zajęte. Na gruncie takiej teorii w ogóle nie powstaje problem tożsamości indywiduów w różnych możliwych światach, ponieważ w ogóle nie ma indywiduów, a tylko wiązki własności. Takiej teorii nie mógłby jednak przyjąć Lewis, który jest nominalistą i traktuje własności jako zbiory indywiduów je egzemplifikujące.

Wnioskiem, jaki wyciąga z tego Stalnaker jest to, że różne charakterystyki stanowiska antyhaecceitystycznego nie dają się ze sobą pogodzić. Antyhaecceityzm Lewisa jest, zdaniem Stalnakera, stanowiskiem spójnym, ale wynikającym z koncepcji metafizycznej, z którą on się za-

sadniczo nie zgadza. Lewis jest nieprzejednanym zwolennikiem Hume'a: historia świata obejmująca wszystkie fakty dotyczące dystrybucji własności jakościowych wyczerpuje wszystkie fakty dotyczące świata. Twierdzenie o superweniencji hume'owskiej głosi, że prawa przyrody nie są niczym więcej niż twierdzeniami najlepszej prawdziwej teorii uporządkowującej wszystkie jednostkowe fakty w sposób najbardziej ekonomiczny. Lewis oddziela własności wewnętrzne świata: te, które wynikają z kompletnej historii, oraz własności modalne, reprezentowane przez relacje między danym światem a innymi światami możliwymi. Te drugie są rzeczywiste, ale tylko ze względu na arbitralnie wybraną – a nie metafizycznie ugruntowaną – relację bycia odpowiednikiem. Lewisowski antyhaecceityzm nie daje się oddzielić, zdaniem Stalnaker, od twierdzenia o superweniencji hume'owskiej. Stalnaker jest zwolennikiem innego filozoficznego światopoglądu, zgodnie z którym własności modalne są w takim samym sensie własnościami świata jak własności faktualne. Stalnaker uważa, że nie da się oddzielić własności wewnętrznych świata od tych jego własności, które reprezentujemy za pomocą relacji z innymi światami możliwymi (Stalnaker 2012, s. 72).

PROBLEM IDENTYCZNOŚCI NIEROZRÓZNIALNYCH

Jednym z często pojawiających się w literaturze filozoficznej argumentów przeciwko antyhaecceityzmowi jest argument przeciwko zasadzie identyczności nierozróżnialnych. Zasada ta stwierdza, że jeśli dwa przedmioty mają te same własności, są one identyczne. Antyhaecceityzm – przynajmniej w większości wersji tego stanowiska – zasadę tę za sobą pociąga. Wydaje się jednak, że zasada ta narusza niektóre nasze intuicje. Klasyczny argument Maxa Blacka (1952) przeciwko identyczności nierozróżnialnych wygląda następująco: wyobraźmy sobie świat składający się z dwóch identycznych żelaznych kul, nieróżniących się ani własnościami wewnętrznymi, ani własnościami relacyjnymi. Zwolennik zasady identyczności nierozróżnialnych zmuszony jest uznać ich tożsamość. Nie może one jednak mieć racji, twierdzi Black, kule te muszą być różne, ponieważ wiemy, że są od siebie przestrzennie oddzielone. Argument ten działa, jak zauważył Stalnaker, tylko na gruncie umiarkowanej relacyjnej koncepcji przestrzeni.

Ian Hacking (1975) odrzuca argument Blacka, podważając założenie, że dwie kule są faktycznie różne. Na gruncie radykalnej rela-

cyjnej teorii przestrzeni nie tylko lokalizacje przestrzenne, ale i sama geometria przestrzeni nie są niczym więcej niż artefaktami modelu. Zatem dwie kule w przestrzeni euklidesowej albo jedna kula z zakrzywionej przestrzeni nieeuklidesowej są dwoma równoważnymi opisami tego samego faktu. Jak zauważa z kolei Adams (1979) – a za nim Stalnaker – mamy jednak co najmniej równie dobre powody, by uznać wbrew Hackingowi, że dwie kule w przestrzeni euklidesowej albo jednak kula w zakrzywionej przestrzeni stanowią dwie odrębne rzeczywiste możliwości.

Adams formułuje dodatkowo argument przeciwko identyczności nierozróżnialnych, odwołujący się do kontrfaktycznych okresów warunkowych (1979). Argument ten, jak zauważa Stalnaker, uderza także w antyhaecceityzm Lewisa, który choć nie dopuszcza nieredukowalnej tożsamości indywiduów w różnych światach, uznaje prymitywne tożsamości indywiduów w jednym świecie. Argument ten pokazuje, że jeśli chcemy móc twierdzić, że dwa faktycznie nieodróżnialne indywidua mogłyby się różnić, to musimy założyć nieredukowalną do własności jakościowych tożsamość nie tylko w obrębie świata, ale także w różnych światach. Adams każe nam pomyśleć świat z dwoma nieodróżnialnymi (w całej historii) globami Kastorem i Polluksem, zamieszkanymi przez dwóch nieodróżnialnych mieszkańców. Mogłoby się jednak zdarzyć, że któryś z tych globów zostałby w pewnym momencie unicestwiony. Możliwość tę reprezentują dwa kontrfaktyczne światy w_2 i w_3 , identyczne z w_1 aż do chwili t . W chwili t w w_2 pierwszy z globów, Kastor, zostaje unicestwiony, a drugi, Polluks, istnieje dalej, a w w_3 Polluks zostaje unicestwiony, a Kastor istnieje dalej. Z punktu widzenia mieszkańców tych dwóch globów w $w_1 - w_2$ i w_3 reprezentują dwie różne możliwości: albo zostaną unicestwiony, albo się uratują. Nie ma jednak jakościowych różnic między w_2 i w_3 .

Stalnaker stwierdza, że wobec braku jakościowych różnic Lewis musiałby stwierdzić, że w_2 i w_3 , to ten sam świat, tylko różnie opisany (Stalnaker 2012, s. 79). Lewis jednak, jak już było powiedziane, dopuszcza to, że dwa światy mogą być nierozróżnialne, a jednak różne. Problemem dla Lewisa jest raczej to, że sytuacja ta stanowi kontrprzykład dla twierdzenia o superweniencji. Zdaniem Lewisa to jak światy możliwe są użyte do reprezentowania *de re* indywiduów w superweniencji na własnościach jakościowych. Jeśli jednak w_2 i w_3 nie różnią się jakościowo, to relacja bycia odpowiednikiem nie może su-

perweniować na własnościach jakościowych. Relacja superweniencji zakłada, że nie ma różnicy we własnościach superweniujących bez różnicy we własnościach bazowych. Własnościami bazowymi są własności jakościowe, a własnościami superweniującymi to, że pewien przedmiot użyty jest do reprezentowania *de re* innego przedmiotu. Glob, który w w_2 zostaje unicestwiony reprezentuje *de re* Kastora, a identyczny jakościowo glob w w_3 , który zostaje unicestwiony reprezentuje *de re* Polluksa. Zatem reprezentowanie *de re* Polluksa (czy Kastora) nie może superweniować na tych samych własnościach jakościowych.

Odpowiedź Lewisa na ten argumentu polega na przyjęciu „taniego haecceityzmu”, uwzględniającego haecceitystyczne intuicje na gruncie antyhaecceitystycznej metafizyki. Intuicje haecceitystyczne, które chce w ten sposób uwzględnić Lewis, dotyczą faktów związanych z naszym umiejscowieniem w świecie. Lewis ilustruje to następującym przykładem: Lewis mógłby sobie wyobrazić, że zamiast być sobą, jest kimś innym (np. nieszczęśnikiem Fredem), jakie to szczęście, że jednak nim nie jest! Zdaniem Lewisa możemy sensownie pomyśleć, co by było, gdybyśmy byli kimś innym, niż jesteśmy, np. mogę się zastanawiać, co by było, gdybym była Napoleonem. Nie chodzi tu o możliwość, że mielibyśmy własności pewnej innej osoby, albo że będąc sobą, żylibyśmy czyimś innym życiem, ale o to, że bylibyśmy tą osobą zamiast sobą⁴ (Lewis 1986, s. 231). Świat, w którym byłabym Napoleonem i świat w którym jestem sobą nie różnią się jakościowo⁵. Zdaniem Lewisa różnią się one tym, co reprezentują *de re* o mnie. Ponieważ różnica ta nie jest ugruntowana w żadnej różnicy jakościowej, musiałaby być ugruntowana w różnicy niejakościowej, na co jednak Lewis nie może przystać. Tani haecceityzm zakłada, że możliwości nie muszą być reprezen-

⁴ Przykład Lewisa może budzić wątpliwości. Oczywiście, w pewnym sensie możemy wyobrażać sobie, co by było, gdybyśmy byli kimś innym, niż jesteśmy. Nie wyobrażamy sobie wówczas jednak, co by było, gdybyśmy my byli kimś innym (co jest niemożliwe, bo nie można być sobą i zarazem kimś innym). Zastanawiamy się raczej, jak by to było widzieć świat z czyjejs innej perspektywy. Taką interpretację zaproponował Thomas Nagel, Lewis ją jednak odrzuca. „Czy nie potrafimy pomyśleć, że to, że jesteśmy tym lub innym «sobą», jest w jakimś sensie kwestią przypadku i szczęścia?” (Lewis 1986, s. 232).

⁵ Dwa jakościowo nieodróżnialne światy mogą też reprezentować sytuację dwóch wszechwiedzących bogów, ilustrujących w innym tekście Lewisa problem przekonania *de se*: bogowie mają zupełną wiedzę o świecie, ale nie wiedzą, którym z bogów są (Lewis 1979).

towane przez różne światy możliwe (Lewis 1986, s. 230). Możliwość, że mogłabym być Napoleonem, nie jest możliwością dla świata, ale jest możliwością dla mnie. Możliwość tę reprezentuje relacja bycia odpowiednikiem wewnątrz tego samego świata⁶.

Lewis udziela na argument Adamsa następującej odpowiedzi: to, co Adams reprezentuje jako dwa odrębne światy w_2 i w_3 , stanowi dwie możliwości wewnątrz jednego świata (nazwijmy go w_{23}). Świat w_{23} zawiera dwie planety: Szczęściarz i Pechowiec. Każda z planet w w_1 ma dwa odpowiedniki w w_{23} : Kastor ma odpowiedniki Szczęściarz i Pechowiec i Polluks ma odpowiedniki Szczęściarz i Pechowiec. Uporządkowana para ⟨Kastor, Polluks⟩ ma dwa odpowiedniki: ⟨Szczęściarz, Pechowiec⟩ oraz ⟨Pechowiec, Szczęściarz⟩, różniące się tylko порядkiem elementów. Jeśli odpowiednikiem pary ⟨Kastor, Polluks⟩ jest ⟨Szczęściarz, Pechowiec⟩, to Kastor się ratuje, natomiast Polluks zostaje unicestwiony, jeśli odpowiednikiem tym jest ⟨Pechowiec, Szczęściarz⟩, to Polluks się ratuje, natomiast Kastor zostaje unicestwiony. Haecceitystyczna różnica, reprezentowana przez światy w_2 i w_3 (czy to ja bym się uratowała, czy mój bliźniak z bliźniaczej planety) na gruncie koncepcji Lewisa jest różnicą z perspektywy pewnego kontrfaktycznego świata: czy w tym świecie jestem jednym lub drugim bliźniakiem.

Koncepcja taniego haecceityzmu wiąże się jednak z trudnościami wskazanymi przez Delię Graff Farę, którym Stalnaker poświęca ostatnią część rozdziału o haecceityzmie. Spór haecceitysty z antyhaecceitystą jest sporem metafizycznym i Stalnaker nie podaje żadnego rozstrzygającego argumentu na rzecz któregoś z tych stanowisk (choć sam jest zwolennikiem haecceityzmu).

⁶ Można postawić pytanie, czy fakty dotyczące tego, że pewien przedmiot x reprezentuje wewnątrz tego samego świata pewien przedmiot y , superwenują na własnościach jakościowych. Kontrprzykładem dla superwenuencji byłaby sytuacja, w której ten sam przedmiot reprezentowałby (wewnątrz tego samego świata) dwa różne przedmioty. Ponieważ jednak różnice haecceitystyczne reprezentowane przez odpowiedniki wewnątrz tego samego świata to różnice dotyczące naszej lokalizacji w świecie, trudno wyobrazić sobie kontrprzykłady. O ile uznamy za Lewisa, że można wyobrazić sobie, że mogłoby się być kimś innym, niż się jest (np. mogę rozważać możliwość, że byłabym Napoleonem), o tyle nie bardzo wiadomo, co mielibyśmy na myśli, wyobrażając sobie, że ktoś inny mógłby być kimś innym, niż jest (np. David Lewis mógłby być Napoleonem). W świecie widzianym z mojej perspektywy każdy przedmiot wewnątrz tego samego świata może (w pewnym określonym kontekście) reprezentować mnie, ale tylko mnie.

METAFIZYKA I SEMANTYKA

Ważnym problemem, któremu Stalnaker poświęca oddzielny rozdział, jest kwestia rozgraniczenia sporów metafizycznych i semantycznych oraz pytanie o to, jaka semantyka jest najlepsza do modelowania tych pierwszych. Stalnaker stara się pokazać, że semantyk możliwych światów w wersji, którą on proponuje, może przyjąć zarówno aktualista, przyjmujący przygodne istnienie sądów, jak i realista modalny, taki jak Lewis czy Plantinga, będący aktualistą, ale uznający istnienie *haecceitates* czysto możliwych indywiduów. Chociaż rozstrzygnięcia semantyczne nie są całkowicie neutralne względem problemów metafizycznych, Stalnaker postuluje, by zawsze szukać takich ram pojęciowych, które są neutralne ze względu na rozważany problem. Stalnaker pokazuje na trzech przykładach, jak nie należy mieszać problemów semantycznych i metafizycznych.

Pierwszy przykład ilustruje nieuprawnione wnioskowanie z przesłanki semantycznej do wniosku metafizycznego. Stalnaker rozważa argument Thomasa Hoffwebera (2005), który z twierdzenia o bezpośrednim odniesieniu wyrażen jednostkowych wnioskuje o ontologicznej zależności sądów od przedmiotów. Uzasadnienie Hoffwebera wygląda następująco: teza o bezpośrednim odniesieniu głosi, że wkładem, jaki wnosi dane wyrażenie w treść zdania, jest sam przedmiot, a nie sens czy sposób dania przedmiotu. Zatem jeśli przedmiot nie istnieje, nie istnieje też sąd, ponieważ zdanie nie ma wówczas kompletnej treści. Stalnaker wskazuje na to, że argument ten działa tylko, jeśli przyjmiemy strukturalną teorię sądów, natomiast nie działa, jeśli jesteśmy neutralni względem różnych konkurencyjnych teorii sądów. Jeśli chcemy pozostać neutralni względem różnych koncepcji sądów (albo jesteśmy zwolennikami funkcyjnej teorii sądów), przez „wkład”, jaki wnosi wyrażenie w treść, rozumiemy sposób, w jaki dane wyrażenie wyznacza sąd, a nie to, że desygnat jest częścią składową sądu. Twierdzenie o bezpośrednim odniesieniu rozumiemy wówczas tak, że wyrażenie jednostkowe odnosi się do przedmiotu i to ten przedmiot wraz z funkcją propozycjonalną (wartością semantyczną reszty zdania) wyznacza wyrażony przez to zdanie sąd. To, że pewna funkcja nie wyznacza sądu, nie znaczy jeszcze, że odpowiedni sąd nie istnieje. Stalnaker ilustruje to na przykładzie wyrażenia wyznaczającego pewien przedmiot jako funkcję innego przedmiotu. Rozważmy wyrażenie „matka

Napoleona Bonaparte”. Gdyby nie istniał Napoleon Bonaparte, to wyrażenie to nie wyznaczałoby żadnego przedmiotu. Nie znaczy to jednak, że nie istniałaby Letycja Bonaparte. Analogicznie, to, że wyrażenie jednostkowe wyznacza sąd jako funkcję przedmiotu, do którego się ono odnosi, nie oznacza, że sam sąd jest zależny od istnienia tego przedmiotu.

Choć Stalnaker zgadza się zarówno z przesłanką, jak i z wnioskiem (sam uważa, że sądy jednostkowe są ontologicznie zależne od przedmiotu) sądzi, że nie zachodzi tu wynikanie. Samo występowanie wyrażenia jednostkowego w zdaniu nie powoduje jeszcze, że sąd wyrażony przez to zdanie jest ontologicznie zależny od przedmiotu. Rozważmy zdanie: „Sokrates był filozofem albo ktoś inny niż Sokrates był filozofem” (Stalnaker 2012, s. 100). Każdy z sądów wyrażonych przez człony alternatywy jest zależny od istnienia Sokratesa, natomiast sąd wyrażony przez alternatywę nie zależy od istnienia Sokratesa. Ten sam sąd może być wyznaczony na różne sposoby.

Drugi przykład podany przez Stalnakara ma pokazać, że koncepcja metafizyczna (dopuszczająca przygodne istnienie sądów) da się wyrazić na gruncie ortodoksyjnej semantyki (z modyfikacjami, o których była mowa), nie wymaga natomiast – jak chciał Adams – zasadniczych zmian w semantyce, prowadzących do naruszenia zasady kompozycyjności. Trzeci przykład dotyczący pewnej zagadki Kita Fine’a (2005)⁷. Stalnaker rozważa, czy wskazany przez Fine’a problem jest natury metafizycznej, czy semantycznej, i na jakim gruncie należy szukać jego rozwiązania.

Głównym celem Stalnakera w *Mere possibilities* jest pogodzenie twierdzenia aktualizmu z intuicją, że przedmioty istnieją w sposób przygodny i że mogłoby istnieć coś innego niż to, co faktycznie istnieje. Stalnaker przedstawia koncepcję semantyki modalnej – minimalną i neutralną ze względu na spory metafizyczne – która ma pozwolić wyartykułować różne spory metafizyczne. Pogodzenie aktualizmu z przygodnym istnieniem indywiduów jest możliwe, twierdzi Stalnaker, jeśli odrzucimy powszechnie przyjmowane założenie, że sądy istnieją

⁷ Zagadka ma postać następującego, na pierwszy rzut oka poprawnego, rozumowania prowadzącego od prawdziwej przesłanki do budzącego sprzeciw wniosku: (1) Jest konieczne, że Sokrates jest człowiekiem. (2) Jest możliwe, żeby Sokrates nie istniał. (3) Zatem jest możliwe, że Sokrates jest człowiekiem i nie istnieje (Fine 2005; Stalnaker 2012, s. 102).

w sposób konieczny. Stalnaker pokazuje pewne konsekwencje tego stanowiska. Inne stanowiska aktualistyczne, Plantingi i Williamsona, unikają przygodnych sądów, ale za cenę przyjęcia kontrowersyjnych założeń metafizycznych. Choć książka nie jest polemiką z realizmem modalnym, wiele miejsca poświęca Stalnaker przeciwstawieniu swojego podejścia koncepcji Lewisa, którego poglądy na temat możliwości łączą się z jego nominalizmem oraz twierdzeniem o superweniencji hume'owskiej.

ZAKOŃCZENIE

Tym, co łączy niektórych aktualistów (m.in. Plantingę i Williamsona) z Lewisem, jest pogląd, że cała przestrzeń możliwości istnieje niezależnie od tego, która z możliwości zrealizowała się w rzeczywistości. Pogląd ten jest metafizycznym odpowiednikiem epistemologicznego stanowiska, nazwanego przez Davida Chalmersa „racjonalizmem modalnym”, mówiącego, że cała przestrzeń możliwości jest dana *a priori*: to, co możliwe, pokrywa się z tym, co można pomyśleć. Stalnaker określa wyznawany przez siebie przeciwstawny pogląd mianem „naturalizmu modalnego”. Stanowisko to głosi, że przestrzeń możliwości nie tylko nie jest dostępna *a priori*, lecz także że nie istnieje niezależnie od tego, która z możliwości się zrealizowała. Jedną z trudności, na jakie musi odpowiedzieć naturalista modalny, jest to, w jaki sposób przebiega proces poznawania rzeczywistości. Zgodnie z naturalizmem modalnym nie mamy zewnętrznego punktu widzenia, z którego moglibyśmy najpierw wyznaczyć wszystkie możliwości, a potem pytać, której z tych możliwości odpowiada świat rzeczywisty. Z naszej lokalizacji w przestrzeni możliwości pewne możliwości nie tylko nie są dostępne, lecz wręcz nie istnieją. Proces badania polega jednak także na stawianiu pytań dotyczących natury rzeczy, której wszak nie znamy *a priori*, a więc także na rozważaniu czegoś niemożliwego np. tego, że pewna rzecz miałaby inną naturę, niż ma. Nie jest możliwe, żeby rzeczy miały inną naturę, niż mają – np. woda nie może mieć innego składu chemicznego niż H_2O , możemy natomiast reprezentować epistemiczne możliwości, że istnieje coś, co wygląda jak woda, ale ma strukturę inną niż H_2O . Możemy reprezentować epistemiczne możliwości, posługując się środkami, jakich dostarcza nam świat rzeczywisty. To jednak, zdaniem Stalnakera, nie jest argumentem na rzecz racjo-

nalizmu modalnego: nie znaczy to, że możemy odnosić się za pomocą własności jakościowych do czystych możliwości.

Książka Stalnaker jest nie tylko obroną aktualizmu i naturalizmu modalnego, lecz także bardzo wnikliwą refleksją na temat związków metafizyki i semantyki. Jej ogromną wartością jest też światło, które rzuca ona na inne zagadnienia filozoficzne, na przykład na problem intencjonalności.

BIBLIOGRAFIA

- Adams, R. (1979), *Primitive Thisness and Primitive Identity*, „Journal of Philosophy” 76, s. 5–26.
- Adams, R. (1981), *Actualism and Thisness*, „Synthese” 49, s. 3–41.
- Black, M. (1952), *The Identity of Indiscernibles*, „Mind” 61, s. 153–64.
- Ciecierski, T. (2003), *Leibniz o konieczności, możliwości i wolnej woli*, „Przegląd Filozoficzny” 45, s. 135–142.
- Fine, K. (2005), *Necessity and Non-Existence*, w: *Modality and Tense: Philosophical Papers*, Oxford: Oxford University Press, s. 321–354.
- van Fraassen, B. (1967), *Meaning Relations among Predicates*, „Nous”, s. 161–179.
- Hofwebber, T. (2005), *Supervenience and Object-Dependent Properties*, „Journal of Philosophy” 102, s. 1–28.
- Kaplan, D. (1975), *How to Russell a Frege-Church*, „Journal of Philosophy” 72, s. 716–729.
- Lewis, D. (1986), *On the Plurality of Worlds*, New York: Basil Blackwell.
- Lewis, D. (1979), *Attitudes de dicto and de se*, „The Philosophical Review” 88, s. 513–543.
- McMichael, A. (1983), *A Problem for Actualism About Possible Worlds*, „The Philosophical Review” 92, s. 49–66.
- Plantinga, A. (1983), *On Existentialism*, „Philosophical Studies” 44, s. 1–20.
- Stalnaker, R. (2010), *Merely Possible Propositions*, w: B. Hale, A. Hoffman (eds.), *Modality: Metaphysics, Logic and Epistemology*, Oxford: Oxford University Press, s. 21–32.
- Stalnaker, R. (2012), *Mere Possibilities: Metaphysical Foundations of Modal Semantics*, Princeton: Princeton University Press.
- Williamson, T. (2001), *Necessary Existents*, w: A. O’Hear (ed.), *Logic, Thought and Language*, Cambridge: Cambridge University Press, s. 231–251.

NOTY O AUTORACH

TADEUSZ CIECIERSKI – dr hab., adiunkt, Instytut Filozofii, Zakład Historii Filozofii Nowożytnej, Uniwersytet Warszawski, Krakowskie Przedmieście 3, 00-927 Warszawa

PAWEŁ GRABARCZYK – dr, adiunkt, Instytut Filozofii, Katedra Filozofii Współczesnej, Uniwersytet Łódzki, ul. Lindleya 3/5, 90-131 Łódź

KRZYSZTOF POŚLAJKO – dr, adiunkt, Instytut Filozofii, Pracownia Retoryki Logicznej, Uniwersytet Jagielloński, ul. Grodzka 52, 31-044 Kraków

PIOTR KONDERAK – dr, adiunkt, Instytut Filozofii, Zakład Logiki i Kognitywistyki, Uniwersytet Marii Curie-Skłodowskiej, pl. Marii Curie-Skłodowskiej 4, 20-031 Lublin

PIOTR WILKIN – dr, Ośrodek Badań Filozoficznych

JACEK WAWER – dr, adiunkt, Instytut Filozofii, Zakład Epistemologii, Uniwersytet Jagielloński, ul. Grodzka 52, 31-044 Kraków

MACIEJ SENDŁAK – dr, adiunkt, Instytut Filozofii, Zakład Filozofii Analitycznej, Uniwersytet Warszawski, Krakowskie Przedmieście 3, 00-927 Warszawa

MARIA MATUSZKIEWICZ – mgr, doktorantka, Instytut Filozofii, Zakład Historii Filozofii Nowożytnej, Uniwersytet Warszawski, Krakowskie Przedmieście 3, 00-927 Warszawa

STUDIA SEMIOTYCZNE

Tom XXXI • Numer 1 • 2017

W numerze między innymi:

TADEUSZ CIECIERSKI Jerzy Pelc (ur. w Warszawie
30 września 1924 r., zm. w Warszawie 2 czerwca 2017 r.)

PAWEŁ GRABARCZYK Directival Theory of Meaning Resurrected

KRZYSZTOF POSŁAJKO Semantic Deflationism, Public Language
Meaning and Contextual Standards of Correctness

PIOTR KONDERAK The Conscious Semiotic Mind

PIOTR WILKIN Reprezentacje znaturalizowane – użyteczny cel
czy użyteczna fikcja?

JACEK WAWER Problem ustalania indeksu w semantyce
temporalno-modalnej

MACIEJ SENDŁAK U podstaw sporu o kontrmożliwe okresy
warunkowe

MARIA MATUSZKIEWICZ Problem aktualisty modalnego z czysto
możliwymi indywiduami (o *Mere Possibilities* Roberta Stalnakera)

ISSN 0137-6608