

Karolina Krzyżanowska

Sprawozdania z przekonań w perspektywie filozofii języka i kognitywistyki

Słowa kluczowe: sprawozdania z przekonań, ikonizacja, teoria umysłu, symulacja, myślenie kontrfaktyczne

Dyskusja dotycząca formy logicznej zdań o przekonaniach sięga samego początku filozofii analitycznej. Są one zjawiskiem niewygodnym i niepokojącym dla każdego, kto próbował sformułować semantyczną teorię języka naturalnego. Wymykają się eleganckim ujęciom oraz łamią intuicyjne na pierwszy rzut oka zasady semantyczne, innymi słowy – generują liczne łamigłówki, których wyjaśnianiem zajmują się filozofowie co najmniej od czasów Fregego. Obecne w tradycji filozoficznej próby rozwiązania tych łamigłówek nie tylko borykały się z różnymi problemami, ale nie uwzględniały ponadto kognitywnych mechanizmów odpowiedzialnych za przypisywanie przekonań innym. Ponieważ współczesne badania psychologiczne wskazują na wyjątkowy charakter myśli oraz wypowiedzi o stanach mentalnych innych osób, celem niniejszej pracy jest znalezienie takiej analizy zdań o przekonaniach, która będzie uwzględniać ich szczególne miejsce w architekturze umysłu.

I. Tradycyjne koncepcje semantyczne

Sprawozdania z przekonań są to zdania o formie: „A sądzi, że p”, gdzie A to nazwa podmiotu, który żywi przekonanie, natomiast p jest dowolnym zdaniem, wyrażającym treść tego przekonania. Czasownik epistemiczny „sądzi, że” można interpretować jako relację dwuargumentową, której charakter jest jednak przedmiotem dyskusji.

Charakterystyczne dla zdań o przekonaniach są generowane przez nie łamigłówki semantyczne. Wyrastają one z prób ujęcia języka naturalnego w ramy

teorii, których kluczowym elementem jest atrakcyjna z wielu względów sformułowana przez Fregego zasada kompozycyjności¹.

Zasada kompozycyjności. *Wartość semantyczna wyrażenia złożonego jest funkcją wartości semantycznych prostszych wyrażeń składowych oraz sposobu, w jaki są one składniowo połączone.*

Konsekwencją kompozycyjności jest możliwość podstawiania *salva veritate* wyrażeń, które mają tę samą wartość semantyczną, inaczej zwana zasadą podstawialności bądź ekstensjonalności.

Zasada ekstensjonalności. *Wartość logiczna zdania nie zmienia się, gdy podstawimy w miejsce wyrażenia wchodzącego w skład tego zdania inne wyrażenie o tej samej ekstensji.*

Sprawozdania z przekonań zdają się jednak wymykać powyższym zasadom. Zdanie:

(1) Tomasz sądzi, że Bolesław Prus jest autorem „Lalki”

może być prawdziwe, jednak gdy podstawimy za wyrażenie „Bolesław Prus” odnoszącą się do tej samej osoby nazwę „Aleksander Głowacki”, możemy otrzymać zdanie fałszywe.

Zasada kompozycyjności jest na tyle atrakcyjna z logicznego punktu widzenia, że filozofowie starają się znaleźć ostateczne rozwiązanie łamigłówek semantycznych związanych ze sprawozdaniami z przekonań w taki sposób, by nie trzeba było z niej rezygnować. Koncepcje semantyczne tradycyjnie wpisują się w dwa nurty: intensjonalny oraz ekstensjonalny, które różnią się ujęciem interpretacji języka.

Definicja 1. *Interpretacja języka L składa się z niepustego zbioru U zwanego uniwersum interpretacji oraz z funkcji interpretacji I , której argumentami są wyrażenia języka L .*

W zależności od koncepcji semantycznej funkcja interpretacji może przyjmować różne wartości: ekstensje bądź intensje wyrażeń (zob. Grudzińska 2003).

Intensja jest to funkcja, która określa, co jest ekstensją wyrażenia w każdym z indeksów. Zbiór indeksów definiujemy jako zbiór par uporządkowanych $\langle w, t \rangle$,

¹ Chociaż tradycyjnie zasadę kompozycyjności przypisuje się Fregemu, jest to jednak kwestia dyskusyjna (por. Janssen 1997).

takich, że $w \in W$ i $t \in T$, gdzie W jest zbiorem możliwych światów, a T – zbiorem punktów w czasie. Dla uproszczenia terminologii w dalszych rozważaniach ograniczę indeksy do możliwych światów. Intensja zdania, czyli sąd logiczny, jest funkcją, która przeprowadza zbiór indeksów w zbiór wartości logicznych $\{0, 1\}$. Sąd logiczny można więc utożsamić ze zbiorem tych światów możliwych, w których jest on prawdziwy (Dowty i in. 1981: 145–147).

W ujęciach intensjonalnych czasowniki epistemiczne takie jak „sądzić” czy „myśleć” denotują relację między podmiotami a intensjami zdań, czyli sądami logicznymi (Grudzińska 2003). Ponieważ wyrażenia muszą mieć tę samą intensję, nie tylko ekstensję, by można było je wzajemnie podstawiać *salva veritate*, zasada ekstensjonalności nie zostaje wcale naruszona.

Przedstawiony powyżej program intensjonalny nie wystrzega się jednak pewnych problemów, takich jak zjawisko wszechwiedzy logicznej, a także problem identyfikacji transświatowej. Nawet wprowadzona przez Kripke’go (2001) koncepcja sztywnych desygnatorów, czyli takich wyrażen, które w każdym możliwym świecie denotują te same przedmioty, nie rozwiązuje problemu. Sztywnymi desygnatorami są według Kripke’go nazwy własne: Bolesław Prus mógłby nie być pisarzem, „Lalkę” mógłby napisać ktoś inny i tak dalej, jednak nazwa „Bolesław Prus” w każdym możliwym świecie odnosi się do tej samej osoby. Rozwiązanie to niesie jednak ze sobą kolejne problemy. Nazwy „Aleksander Głowacki” i „Bolesław Prus” wyznaczają tę samą osobę we wszystkich możliwych światach, jednak ich wzajemne podstawianie w kontekstach przekonaniowych nie jest możliwe.

Jak zauważył Israel Scheffler (1955), synonimiczność dwóch wyrażen nie gwarantuje, że wywołają one tę samą psychologiczną reakcję podmiotu. Powinniśmy włączyć do analizy zdań o przekonaniach aspekt psychologiczny.

Niektórzy filozofowie zgodni z intensjonalistami co do tego, że w teorii semantycznej zasada kompozycyjności powinna obowiązywać, krytykowali jednak postulowanie istnienia abstrakcyjnych bytów, jakimi są znaczenia czy intensje. Proponowali oni zatem takie rozwiązania, tzw. ekstensjonalne, które nie wymagają odnoszenia się do żadnych szczególnych bytów spoza uniwersum języka. W świetle koncepcji ekstensjonalnych funkcja interpretacji przypisuje każdemu wyrażeniu prostemu z L przedmiot należący do zbioru U , podzbiór U lub zbiór uporządkowanych n -tek elementów należących do U , w zależności od kategorii gramatycznej wyrażenia (Grudzińska 2003).

Przykładem ujęcia ekstensjonalnego jest teoria cudzysłowowa, w świetle której czasowniki epistemiczne denotują relacje między podmiotami a zdaniami, a konkretnie – ich syntaktycznymi reprezentacjami. Podobnie jak w ujęciu intensjonalnym, wyrażenia w zasięgu czasowników epistemicznych nie odnoszą się do swoich „normalnych” odniesień. Zdania w mowie zależnej należy rozumieć tak, jakby zawierały bezpośrednie cytaty, w związku z czym (1) należy analizować w następujący sposób:

(2) Tomasz sądzi: „Bolesław Prus jest autorem «Lalki»”.

Poprzez odwołanie się do kształtu cytowanych wyrażen, nie zaś do ich odniesienia, zwolennikom teorii cudzysłowowej udaje się utrzymać zasadę kompozycyjności.

Rozwiązanie to nie jest jednak pozbawione wad. Nacisk na kształt wyrażenia w kontekście epistemicznym wyklucza wszelkie zmiany formy syntaktycznej tego wyrażenia, nawet tak niewinne jak poniższa parafraza zdania (1):

(3) Tomasz sądzi, że autorem „Lalki” jest Bolesław Prus.

Co więcej chcielibyśmy zdawać sprawę z przekonań podmiotu, nie przesądając o języku, w jakim on sam formułuje swoje przekonania. Problemem może być również odwrotna sytuacja – możliwe jest, że pewne formy syntaktyczne występują w różnych językach jako zdania znaczące zupełnie różne rzeczy (Davidson 2001: 98).

Koncepcja cudzysłowowa nie radzi sobie również ze sprawozdaniami z przekonania zawierającymi wyrażenia okazjonalne:

(4) Joanna twierdzi, że ta książka jest ciekawa.

(5) Joanna twierdzi, że ta książka jest ciekawa.

Jeśli w zdaniu (4) „ta książka” odnosi się do „Czarodziejskiej Góry”, a w (5) do książki telefonicznej, wartości logiczne powyższych zdań będą najprawdopodobniej różne.

Cudzysłowowa analiza zdań o przekonaniach boryka się również z innymi problemami. Zagadką pozostaje fakt, że bez trudu przeprowadzamy wnioskowania typu:

(6) Jan jest przekonany, że wakacje spędzi w Tatrach.

(7) Jan zawsze realizuje swoje plany.

(8) Wniosek: Jan spędzi wakacje w Tatrach.

Zdając sprawę z cudzych wypowiedzi czy przekonań, mówimy bądź myślimy *to samo*, co podmiot sprawozdania, dzięki czemu wnioskowania takie jak (6)–(8) nie sprawiają nam problemu. Ta obserwacja była punktem wyjścia rozważań Donalda Davidsona (1968). W świetle jego propozycji czasowniki epistemiczne denotują relację, której pierwszym członem jest podmiot, a drugim – zaimek wskazujący „to”, po którym następuje kropka. Zaimek wskazujący odnosi się do zdania następującego po kropce, które wyraża treść wypowiedzi podmiotu. Zdanie (1) analizujemy, rozszczepiając je na dwa następujące po sobie zdania:

(9) Tomasz sądzi to. Bolesław Prus jest autorem „Lalki”.

Co jest nowe w tym ujęciu, to brak jakiegokolwiek logicznego czy semantycznego związku między tymi zdaniem – zdanie wyrażające przekonanie podmiotu nie jest zawarte w sprawozdaniu. Podstawianie nazwy „Aleksander Głowacki” za „Bolesław Prus” w drugim zdaniu nie wpływa na jego wartość logiczną, zmienia natomiast odniesienie zaimka wskazującego „to”, a zatem również wartość logiczną pierwszego zdania z pary.

Niektórzy krytycy (np. Higginbotham 1986) zwracają uwagę na to, że w przypadku niektórych konstrukcji intensjonalnych można zaobserwować wyraźny związek między zdaniem głównym a wyrażeniami występującymi w zdaniu podrzędnym. Koncepcja Davidsona okazuje się nie być wówczas skuteczna. Rozpatrzmy zdania:

(10) Każdy chłopiec sądzi, że (on) jest fajnym facetem.

(11) Każdy chłopiec sądzi to. On jest fajnym facetem.

Drugie zdanie nie jest parafrazą pierwszego, w którym zaimek „on” odnosi się do podmiotu przekonania, gdy w rzekomej parafrazie ten sam zaimek wskazuje raczej na kogoś innego.

II. Ikoniczność

Sprawozdania z przekonań angażują procesy poznawcze szczególnego rodzaju, wymagają mianowicie umiejętności przypisywania sobie oraz innym ludziom stanów mentalnych, czyli posiadania tak zwanej „teorii umysłu”² (Carruthers i Smith 1996).

Jednym ze sposobów badania, czy dana osoba posiada teorię umysłu, jest test fałszywych przekonań³. Żeby pomyślnie przejść test, badany musi być w stanie przyjąć cudzą perspektywę i zrozumieć, że przekonania nie muszą odzwier-

² Użycie słowa „teoria” w tym kontekście nie było zbyt szczęśliwym posunięciem, gdyż teoria umysłu nie jest teorią w ścisłym sensie – należy ją raczej rozumieć jako swoistą kompetencję poznawczą, która powoduje, że ludzie na każdym kroku uprawiają zdroworoządkową, „ludową” psychologię (Fodor 1987, 1995), czyli wyjaśniają oraz przewidują zachowania innych, odwołując się do ich stanów umysłowych. Co więcej, robią to automatycznie, odruchowo i bez wysiłku (Baron-Cohen 1995).

³ Test przebiega w sposób następujący: osoba badana, w tym wypadku dziecko, obserwuje scenkę, w której uczestniczą dwie kukielki, Sally i Anne. Dziecko obserwuje Sally, która wkłada szklaną kulkę do koszyka, po czym wychodzi z pomieszczenia. Podczas jej nieobecności Anne wyjmuje kulkę z koszyka i chowa ją do pudełka. Następnie Sally ponownie wchodzi do pokoju, a dziecko ma odpowiedzieć na pytanie: „Gdzie według Sally znajduje się kulka?” (Baron-Cohen i in. 1985).

ciedlać rzeczywistości. Dzieci nabywają umiejętności potrzebne do prawidłowego rozwiązania testu fałszywych przekonań około czwartego roku życia. Błąd, który popełniają młodsze dzieci, a także dzieci autystyczne, można określić mianem „błędu realisty” – nie potrafią one w zdaniu typu „*A* jest przekonany, że *p*” podstawić za *p* zdania fałszywego.

Chociaż stanowisk w kwestii natury teorii umysłu jest wiele zarówno pośród filozofów, jak i psychologów⁴, badania psychologiczne jednoznacznie wskazują na szczególne miejsce sprawozdań z przekonań w architekturze języka oraz umysłu. Adekwatna psychologicznie teoria semantyczna powinna zatem uwzględniać trudności związane z uczeniem się zdań o przekonaniach, a zatem odróżniać je od innych konstrukcji językowych.

Z kolei nasza zdolność do wyjaśniania cudzych zachowań poprzez odwołanie się do stanów mentalnych oraz umiejętność przyjmowania cudzej perspektywy sugerują, że mówiąc o przekonaniach, mówimy jednocześnie o rzeczywistości, której dotyczy. Ktoś, kto wypowiada zdanie o formie „*A* sądzi, że *p*”, mówi nie tylko o *A*, że jest podmiotem jakiegoś przekonania, ale mówi również o tym, że *p*. Można powiedzieć, że sprawozdawca „udaje”, że twierdzi, że *p*. Jeśli potrafimy postawić siebie w cudzym położeniu i z tego położenia formułować udawane przekonania, to są one tak samo przezroczyste znaczeniowo jak reprezentacje pierwszego stopnia.

Właśnie te psychologiczne intuicje próbował uchwycić Donald Davidson. Krytykował on te teorie, w świetle których wyrażenia w kontekście przekonaniowym stanowią semantyczne anomalie – nie odnoszą się do swoich *normalnych* odniesień. Jego intencją było „odzyskanie przed-Fregiańskiej niewinności semantycznej” (Davidson 1968), czyli stworzenia teorii, w której również zdania w mowie zależnej cechowałaby przezroczystość znaczeniowa.

François Recanati, podążając za ideą Davidsona twierdzi, że ikoniczność dotyczy wszystkich metareprezentacji. Metareprezentacjami, czy inaczej reprezentacjami drugiego rzędu, Recanati nazywa takie zdania, które zawierają inne zdania, czyli reprezentacje przedmiotowe (oprócz tych, w których reprezentacje przedmiotowe są ujęte w cudzysłów). Metareprezentacją jest więc zdanie złożone *Op*, które zawiera zdanie podrzędne *p*. *O* jest tutaj operatorem wprowadzającym metareprezentację, takim jak np. „w filmie” czy „Jan uważa, że”. Istotną, choć lekceważoną przez obrońców kompozycyjności, cechą zdań powyższego typu jest ikoniczność.

Zasada ikoniczności. *Metareprezentacje zawierają reprezentację przedmiotową nie tylko syntaktycznie (tak jak Op „zawiera” p), lecz również semantycznie: sąd Q wyrażony przez Op „zawiera” jako swoją część sąd P wyrażony przez p – dlatego nie można pomyśleć, że Q, nie myśląc jednocześnie, że P (Recanati 2000b: 10).*

⁴ Por. np. Carruthers i Smith (1996), Goldman (2009).

Nie możemy zatem pomyśleć, że *A* jest przekonany, że *p*, nie myśląc jednocześnie, że *p*. Myślenie należy tutaj rozumieć jako wypowiedzanie w myśli pewnych zdań, w związku z czym „myślenie *p*” czy „pomyślenie, że *p*” nie jest równoważne z „twierdzeniem, że *p*” czy też ze „stwierdzeniem prawdziwości *p*”. To rozróżnienie pokazuje, że możemy bez popadania w sprzeczność myśleć czy wypowiadać w myślach zdania, o których wiemy, że są fałszywe.

W konfrontacji z fałszywymi przekonaniem sprawozdawca stawia siebie w położeniu podmiotu, który myśli, że *p*, chociaż prawdą jest $\neg p$. Sprawozdawca symuluje zatem stan umysłu podmiotu i myśli *p*, które dzięki temu zachowuje swoje normalne odniesienie. Rozpatrzmy więc poniższe zdania:

(12) Delfin to gatunek ryby.

(13) Anna sądzi, że delfin to gatunek ryby.

Wypowiadając (13), mamy na myśli to samo, co Anna wypowiadająca zdanie (12), mianowicie sympatyczne i inteligentne ssaki, pływające w oceanie, które zwykliśmy nazywać delfinami – nie zaś Anny myśli czy wyobrażenie delfina, który ssakiem nie jest.

Jak zauważył Recanati, zasada ikonizacji pozwala wyjaśnić różnicę między następującymi zdaniami:

(14) Tomasz stwierdził prawo wyłączonego środka.

(15) Tomasz stwierdził, że dla dowolnego zdania *p* zachodzi: $p \vee \neg p$.

Oba powyższe zdania mówią o tym samym, mianowicie o tym, że Tomasz stwierdził prawo wyłączonego środka. Jednakże treść tego prawa jest pokazana tylko w drugim przypadku, w związku z czym tylko (15) posiada cechę ikonizacji, a co za tym idzie – jedynie (15) jest metareprezentacją (Recanati 2000: 319). Nie możemy pomyśleć (15), nie myśląc, że dla dowolnego zdania *p* zachodzi: $p \vee \neg p$, chociaż możemy pomyśleć (14), nie wiedząc w ogóle, czego prawo wyłączonego środka dotyczy.

Obecne w literaturze koncepcje semantyczne można podzielić na te, które uwzględniają zasadę ikonizacji, nazwijmy je krótko *ikonizacyjnymi*, oraz na koncepcje nieikonizacyjne, zgodnie z którymi konteksty epistemiczne oraz inne metareprezentacje są nieprzezroczyste znaczeniowo. Do tych ostatnich należy zaliczyć ujęcie intensjonalne z jednej strony oraz teorię cudzysłowową z drugiej.

W ujęciu intensjonalnym właściwym obiektem sprawozdań z przekonań są sądy logiczne utożsamiane ze światami możliwymi, w przykładzie (13) nie ma więc w ogóle mowy o delfinach – zdanie „delfin to gatunek ryby” w tym kontekście nie mówi w ogóle nic o świecie. Intensjonalności explicite odrzucają tezę o przezroczystości wyrażen w kontekstach epistemicznych, a w konsekwencji

również ich ikoniczność. Dlatego też rozwiązanie intensjonalne nie może stanowić adekwatnego ujęcia dla zdań o przekonaniach.

Zasadę ikoniczności lekceważą również zwolennicy teorii cudzystłowej, zdaniem których mówiąc o przekonaniach, mówimy o zdaniach, wyrażających te przekonania. „Delfin to gatunek ryby” w kontekście (13) jest nazwą własną zdania wewnątrz cudzystłowy. Autor sprawozdania (13) nie wie, o czym mówi zawarta w nim reprezentacja przedmiotowa – metareprezentacje w świetle teorii cudzystłowej są więc zasadniczo nieprzezroczyste znaczeniowo i nieikoniczne.

Do grupy teorii respektujących zasadę ikoniczności należy zaliczyć koncepcję Donalda Davidsona, która pozwala na podtrzymanie zasady ekstensjonalności przy jednoczesnym uwzględnieniu *normalnego* odniesienia wyrażen występujących w kontekście przekonania. Analiza Davidsona spełnia sformułowany przez Recanatiego schemat – sprawozdanie „Anna sądzi, że Jan lubi cebulę” interpretujemy jako dwa zdania: podrzędne „Jan lubi cebulę”, które wyraża sąd P , oraz poprzedzające je „Anna sądzi to (*that*)”, które wyraża sąd Q . Nie możemy zatem wypowiedzieć w myśli pełnego sprawozdania, wyrażającego dwa sądy QP , nie myśląc o lubiącym cebulę Janie – P .

Wprawdzie spośród dotychczas omówionych teorii teoria Davidsona najlepiej spełnia psychologiczne postulaty, jednak i ona nie jest pozbawiona wad. Błędnie przewiduje, że nabywanie umiejętności posługiwania się sprawozdaniami z przekonań będzie się odbywać w sposób standardowy, tak jak ma to miejsce w przypadku wyrażen okazjonalnych (Grudzińska 2008), nie potrafi wyjaśnić wspomnianych faktów psychologicznych.

III. Rozwiązanie François Recanatiego

Wokół pojęcia ikoniczności zbudował swoją teorię metareprezentacji François Recanati (2000b). Zasada ikoniczności wymaga, by p w zdaniu „ A sądzi, że p ” było użyte, nie tylko zaś wymienione, w związku z tym Recanati proponuje, by interpretować czasowniki epistemiczne jako dwu-argumentowe predykaty wyższego rzędu, których argumenty są heterogeniczne. Predykat „sądzi, że” w połączeniu z pierwszym argumentem możemy traktować jako operator, podobnie jak „zdaniem Anny”, a także: „w filmie” czy „w książce” w przypadku innych metareprezentacji. Sprawozdanie z przekonań Op składa się więc z operatora metareprezentacji O oraz zdania wyrażającego przekonanie p . Ta interpretacja pozwala na jednolite ujęcie zarówno sprawozdań o formie: „Zdaniem Jana Antananarywa jest stolicą Kenii”, jak i „Jan sądzi, że Antananarywa jest stolicą Kenii”.

Wzorując się na semantyce wyrażen okazjonalnych Kaplana (1989), Recanati rozróżnia trzy poziomy wartości semantycznej. Oprócz *ekstensji* jest to treść wyrażenia, czyli jego *intensja* (*content*) oraz językowe znaczenie (*linguistic*

meaning) wyrażenia, czyli jego charakter (Recanati 2000b: 5). Intensja wyznacza ekstensję wyrażenia w sytuacji, w której jego wartość logiczna podlega ocenie. Charakter natomiast jest to funkcja przypisująca wyrażeniom ich intencje w danym kontekście. Uogólniona zasada kompozycyjności uwzględnia wszystkie te poziomy:

Uogólniona Zasada Kompozycyjności. *Wartość semantyczna wyrażenia na wszystkich poziomach jest funkcją wartości semantycznych prostszych wyrażeń składowych:*

1. *Charakter wyrażenia złożonego jest funkcją charakterów prostszych wyrażeń składowych.*
2. *Treść (intensja) wyrażenia złożonego jest funkcją treści (intensji) prostszych wyrażeń składowych.*
3. *Ekstensja wyrażenia złożonego jest funkcją ekstensji prostszych wyrażeń składowych.*

Ponieważ zasada ekstensjonalności zawodzi w kontekstach przekonaniowych, a jej utrzymanie grozi poświęceniem zasady ikoniczności, jest zdaniem Recanatiego fałszywa. Ekstensja wyrażenie nie jest funkcją ekstensji jego części, ponieważ zależy ona od intensji oraz od tego, jaki jest świat. Dlatego też podstawianie w złożonym wyrażeniu wyrażeń, które mają tę samą ekstensję, ale różnią się pod względem intensji, może zmienić ekstensję całego wyrażenia. Recanati zgadza się, że pierwszy i drugi człon uogólnionej zasady kompozycyjności powinny obowiązywać, odrzuca natomiast jej trzeci człon – jego zdaniem:

3 Ekstensja wyrażenia złożonego nie jest funkcją ekstensji prostszych wyrażeń składowych.*

Teorie, które uwzględniają własność ikoniczności, Recanati dzieli na dwie klasy: teorie modalne oraz teorię symulacji. W świetle teorii symulacji treść *Op* nie tylko zawiera treść *p*, ale jest z nią tożsama, natomiast operator *O* pełni funkcję pragmatyczną, wskazuje mianowicie, że osoba wypowiadająca sprawozdanie nie charakteryzuje świata aktualnego, ale coś w rodzaju świata przekonań Jana. Natomiast teorie modalne akcentują różnicę pomiędzy treścią metareprezentacji *Op* a treścią reprezentacji przedmiotowej *p* – ta ostatnia zawiera się w pierwszej, ale się z nią nie pokrywa.

Z symulacją mamy do czynienia zawsze wtedy, gdy wypowiadając pewne zdanie, nie twierdzimy tego, co ono wyraża – nawet jeśli wypowiadamy je w izolacji. Najlepszym przykładem tego typu wypowiedzi są ironia, metafory (Recanati 2000a: 332), a także wypowiedzi, które można zaobserwować u dzieci bawiących się w udawanie (Leslie 1992). Stwierdzenie czegoś w sposób ironiczny jest tylko *udawaną* asercją, podobnie jak w sprawozdaniu z przekonań *Op* asercja zdania *p* jest jedynie *udawana*, nie dotyczy świata aktualnego, ale wyobrażone-

go. Jak zauważa Recanati, w takiej sytuacji reprezentacja danego stanu rzeczy jest rozszczepiona (*de-coupled*) z rzeczywistością. Środkiem, który pozwala nam na to odejście od rzeczywistości, jest właśnie wspomniany mechanizm umysłowej symulacji.

Zgodnie z teorią symulacji, kiedy wypowiadamy zdanie (13), symulujemy stan umysłu Anny i stwierdzamy, że delfin to gatunek ryby jedynie w obrębie tej symulacji. Operator „A sądzi, że” można określić jako operator „udawania”.

Recanati zwraca uwagę, że na korzyść teorii symulacji działa fakt, że ludzie bez problemu potrafią dokonywać rozumowań takich, jak w poniższym przykładzie. Załóżmy, że:

(16) Jan sądzi, że Antonio Vivaldi mieszka w Warszawie.

(17) Jan sądzi, że Antonio Vivaldi jest autorem „Czarodziejskiej góry”.

Nie zważając na nieprawdopodobieństwo przekonań Jana, ze zdań (16) i (17) bez trudu wnioskujemy:

(18) Jan sądzi, że autor „Czarodziejskiej góry” mieszka w Warszawie.

Teoria symulacji w jasny sposób wyjaśnia, w jaki sposób rozumowania takie przebiegają. Ponieważ zdania podrzędne w sprawozdaniach z przekonań są „udawanymi” asercjami, które mają miejsce wewnątrz określonej symulacji, rozumowania typu (16)–(18) są przeprowadzane wewnątrz symulacji. Podmiot, który zdaje sprawę z przekonań Jana, symuluje stan umysłu Jana, by następnie udawać, że stwierdza następujące zdania:

(19) Antonio Vivaldi mieszka w Warszawie.

(20) Antonio Vivaldi jest autorem „Czarodziejskiej Góry”.

(21) Autor „Czarodziejskie góry” mieszka w Warszawie.

Mechanizm symulacji odgrywa ważną rolę w procesie formowania i rozumienia sprawozdań z przekonań, jednak nie jest wystarczający. Jak zauważa Recanati, teoria symulacji „miesza ze sobą dwie różne rzeczy: symulowanie z wykorzystaniem symulowania (2000b: 51).

Sprawozdanie z przekonań wymaga od ich autora nie tylko symulacji, ale również czegoś, co Recanati nazywa „trybem postsymulacyjnym” (*post-simulative mode*). Samo symulowanie gwarantuje semantyczną „niewinność”, ponieważ wyrażenia w zasięgu operatora metareprezentacji odnoszą się do tego, do czego odnosiłby się w izolacji – udawana asercja dotyczy tego samego, do czego odnosiłaby się dokonana na serio. Jednak symulacja to jedynie etap tworzenia sprawozdania z przekonań. W procesie wnioskowania typu (16)–(18) po dokonaniu

symulowanej asercji zdania „autor »Czarodziejskiej góry« mieszka w Warszawie” potrzebujemy wejść na poziom postsymulacyjny, by stwierdzić, że stan rzeczy, o którym mowa w (21), zachodzi wewnątrz tego wyobrażonego świata przekonań podmiotu. W przeciwieństwie do symulowanych asercji dokonywanych wewnątrz świata przekonań, który założyliśmy, sprawozdania z przekonań, jak i inne meta-reprezentacje, dotyczą przede wszystkim świata aktualnego. Analizę tę wspiera obserwacja, że udawanie, które wymaga jedynie symulacji, jest obserwowane u dzieci dużo wcześniej niż zdolność do rozumienia fałszywych przekonań (Leslie 1992).

Symulowane asercje takie jak „(bawimy się, że) banan jest telefonem” lub wyrażenia wypowiedziane ironicznie są to reprezentacje pierwszego rzędu nierzeczywistych stanów rzeczy, czyli stwierdzenia dokonywane *względem* wyobrażonej sytuacji. Metareprezentacje natomiast są prawdziwymi asercjami, *dotyczącymi* wyobrażonych sytuacji, które pozostają jednak w relacji do rzeczywistości i jako takie właśnie są reprezentowane. Wypowiadając (17), nie mówimy jedynie o tym, kim jest Vivaldi w świecie przekonań Jana, który symulujemy, żeby móc dokonać udawanej asercji, ale przede wszystkim charakteryzujemy rzeczywistą sytuację, w której znajduje się indywiduum należące do świata aktualnego, mianowicie Jan, polegającą na tym, że Jan żywi przekonanie, że Vivaldi jest pisarzem i autorem tej konkretnej powieści. Recanati zwraca uwagę, że to jest istotna cecha metareprezentacji, którą uchwyciły teorie modalne.

Według modalnych teorii metareprezentacji treść reprezentacji przedmiotowej *p* oraz treść zawierającego ją sprawozdania *Op* są różne. Prefiks *O* jest operatorem, który wpływa na zmianę sytuacji, w odniesieniu do której oceniana jest wartość logiczna *p*. Zdanie „Jan sądzi, że trawa jest zielona” jest prawdziwe w sytuacji *c* wtedy i tylko wtedy, gdy trawa jest zielona we wszystkich możliwych sytuacjach *c'* kompatybilnych z tym, o czym jest Jan przekonany w *c*.

Opowiadając się za modalną interpretacją sprawozdań z przekonań Recanati zbliża się w pewnym stopniu do intensjonalistów. Ponieważ tylko trzeci człon uogólnionej zasady kompozycyjności został odrzucony, wartość semantyczna wyrażenia złożonego nie powinna ulec zmianie, gdy podstawimy w miejsce wyrażenia wchodzącego w jego skład inne wyrażenie o tej samej intensji. To rozwiązanie zdaje się być zatem wrażliwe na typowe kontrprzykłady, w których występują wyrażenia synonimiczne, jak na przykład:

- (22) Jan sądzi, że Adam jest dentystą.
- (23) „Stomatolog” ma tę samą intensję, co „dentysta”.
- (24) Jan sądzi, że Adam jest stomatologiem.

Tu również prawdziwość zdania (22) nie gwarantuje prawdziwości (24). Recanati stwierdza jednak, że to zasada podstawialności jest fałszywa, ponieważ ope-

ratory metareprezentacji takie jak „Jan sądzi, że” są zależne od kontekstu, który ulega zmianie w momencie, gdy w zdaniu „Jan sądzi, że *p*” podstawimy za *p* inne zdanie *p'* o tej samej treści. Podstawianie jest możliwe tylko wtedy, gdy wartości semantyczne wyrażen są ustalone podobnie, jak w słynnym przykładzie Quine'a:

(25) Giorgione był tak nazywany ze względu na swój rozmiar.

Nazwy własne „Giorgione” oraz „Barbarelli” odnoszą się do tej samej osoby i można założyć, że mają tę samą intensję, jednak podstawianie *salva veritate* w zdaniu (25) nie jest możliwe. Wyrażenie „tak nazywany” znaczy tu jednak tyle, co „nazywany »Giorgione«” – jego znaczenie możemy więc zinterpretować jako ustalone. Jeśli więc Giorgione był nazywany „Giorgione” ze względu na swój rozmiar, a Giorgione to Barbarelli, to prawdą jest, że Barbarelli był „tak” nazywany (czyli nazywany „Giorgione”) ze względu na swój rozmiar. W sprawozdaniach z przekonai słowo „że” w „Jan sądzi, że” funkcjonuje podobnie jak „tak” w wyrażeniu „tak nazywany”. Należy je wtedy potraktować zgodnie z sugestią Davidsona jako zaimek wskazujący, na wzór angielskiego „that”⁵.

Chociaż Recanati twierdzi, że metareprezentacje są zasadniczo przezroczyste znaczeniowo, mogą one zyskać nieprzezroczystą interpretację w zależności od kontekstu. Przezroczysta, czyli tzw. minimalna interpretacja sprawozdania „Jan sądzi, że *p*” może zostać w pewnych kontekstach wzbogacona tak, że jej odczytanie stanie się nieprzezroczyste: „Jan sądzi, że *p* w taki właśnie sposób”, co można krócej zapisać jako „Jan tak-sądzi, że *p*”. Operator „Jan tak-sądzi, że” wprowadza zdanie, które jest jednocześnie użyte, jak i wymienione na wzór tak zwanych „mieszanych cytowań”.

Ponieważ interpretacja zdań o przekonaniach jako przezroczystych bądź nieprzezroczystych znaczeniowo zależy od czynników natury pragmatycznej, ich dwuznaczność jest przyczyną tego, że nie są uprawnione rozumowania takie, jak (22)–(24). Kontekst może zmusić nas, by interpretację zdań o przekonaniach Jana wzbogacić o wskazanie na to, w jaki sposób Jan myśli o dentyście Adamie, za pomocą zawartego w prefiksie zaimka wskazującego „tak”, który zmienia swoje znaczenie w zależności od zdania, na które wskazuje. Zdania (22) i (24) sparafrazowane następująco:

(26) Jan tak-sądzi, że Adam jest dentystą.

(27) Jan tak-sądzi, że Adam jest stomatologiem.

⁵ W języku angielskim interpretacja ta nasuwa się sama, ponieważ „that” użyte w izolacji znaczy tyle, co zaimek wskazujący „to” w języku polskim.

mówią o przekonaniach Jana ze wskazaniem na sposób, w jaki reprezentowane są one w jego umyśle. Podobnie jak w teorii cudzysłowowej kontekst przekonaniowy jest tu „nieprzezroczysty” znaczeniowo.

Również w zdaniach (22) i (24) zmienia się wartość semantyczna zaimka wskazującego, który w prefiksie „Jan sądzi, że” zawarty jest *implicite*, w związku z czym wkład semantyczny prefiksu w każde z tych zdań jest inny. Recanati proponuje unaocznic tę różnicę w następujący sposób:

(28) Jan sądzi, że Adam jest dentystą/stomatologiem w następujący sposób „Adam jest dentystą”.

(29) Jan sądzi, że Adam jest dentystą/stomatologiem w następujący sposób „Adam jest stomatologiem”.

Tutaj wyrażenie „w następujący sposób” modyfikuje czasownik „sądzi, że”. Recanati zaznacza, że w dwóch ostatnich zdaniach mamy jednak do czynienia z interpretacją „przezroczystą” i z tego względu „Sądzi, że” i „tak-sądzi, że” nie są równoważne. Zdaniem autora również i w przezroczystych kontekstach przekonaniowych podstawianie kodesygnujących wyrażen *salva veritate* nie jest możliwe. Przede wszystkim kontekst użycia może decydować, czy prefiksowi „A sądzi, że” nadamy wzbogaconą, czyli nieprzezroczystą interpretację, czy też minimalną, a zatem przezroczystą interpretację. Ponadto, jak pisze Recanati:

podstawienie jednego wyrażenia za drugie w zdaniu podrzędnym zmienia kontekst, w którym pojawia się prefiks. Wynika stąd, że podstawienie może zmienić interpretację prefiksu z przezroczystej na nieprzezroczystą, poprzez zwiększenie prawdopodobieństwa, że sprawozdawca używając właśnie tych słów zamierza uchwycić sposób myślenia samego podmiotu o danej sprawie. Innymi słowy, konteksty przekonaniowe są niestabilne. (Recanati 2000b: 160)

O nieprzezroczystym rozumieniu metareprezentacji, które zasadniczo miały być przezroczyste, decydują zatem czynniki natury pragmatycznej.

François Recanati uchwycił w swojej teorii wyjątkowy charakter zdań o przekonaniach. Złożoność ich struktury jako stwierdzenia dokonanego względem rzeczywistej sytuacji o innej, wyobrażonej sytuacji, względem której dokonuje się udawanej asercji, tłumaczy, dlaczego dzieci zaczynają generować i rozumieć zdania o przekonaniach później niż inne konstrukcje językowe. Semantyka Recanatiego uwzględnia psychologiczny mechanizm symulacji, jako podstawę ludzkiej zdolności przypisywania stanów mentalnych innym, pozostawiając jednak miejsce dla teoretycznych aspektów „ludowej” psychologii – efekty symulacji wykorzystywane są w tak zwanym trybie postsymulacyjnym jako dane wejściowe do wygenerowania stwierdzenia o stanie mentalnym drugiej osoby.

Rozwiązanie zaproponowane przez Recanatiego stanowi także swego rodzaju kompromis pomiędzy programem intensjonalnym a teorią Davidsona, pozostają-

ca w duchu ekstensjonalnym. Podobnie jak intensjoniści Recanati interpretuje wyrażenia o formie „*A* sądzi, że *p*” jako składające się z prefiksu „*A* sądzi, że”, na wzór operatorów modalnych powodującego zmianę okoliczności, względem których oceniana jest wartość logiczna zdania *p*. Niewinność semantyczna wyrażen przy jednoczesnej niemożliwości ich wzajemnego podstawiania w zasięgu czasowników epistemicznych wymagała wyjaśnienia, które podsunęła Recanatiemu paraktyczna koncepcja Davidsona: o niemożliwości wzajemnego podstawiania wyrażen w kontekstach przekonaniowych decyduje kontekst, który wpływa na znaczenie zawartego *implicite* w prefiksie zaimka wskazującego.

Teoria Davidsona nie radziła sobie ze zdaniami typu (10), gdzie zaimek „on” wskazuje na podmiot przekonania. Recanati również odwołuje się do obecności zaimka wskazującego ukrytego w operatorze „*A* sądzi to” w swojej analizie nieprzezroczystości kontekstów przekonaniowych. Należy zatem zbadać, czy udaje mu się uniknąć tych problemów, z którymi nie poradziła sobie teoria Davidsona. Jeśli zastosujemy formułę Recanatiego do zdania (10), na pierwszy rzut oka otrzymujemy (Kapitan 2002):

- (30) Każdy chłopiec sądzi, że (on) jest fajnym facetem w następujący sposób:
„on jest fajnym facetem”.

Skoro jednak to kontekst decyduje o interpretacji danego zdania, nie widzę przeszkód, by obecność zaimka osobowego w trzeciej osobie, odnoszącego się do podmiotu referowanego przekonania, po prostu wykluczała interpretację nieprzezroczystą. Pragmatyczne reguły posługiwania się wyrażeniami okazjonalnymi oraz mową zależną zmuszają nas, by w sprawozdaniach z przekonania dostosowywać ich formę do kontekstu. Zdania typu (10) nie sprawiają kompetentnym użytkownikom języka żadnych problemów – automatycznie przypisujemy zaimkowi „on” obecnemu w zdaniu podrzędnym odniesienie, którym jest podmiot przekonania. Gdy zamieniamy mowę zależną na bezpośrednie cytowanie lub odwrotnie, dostosowywanie formy zaimka osobowego również nie stanowi trudności. Ponieważ parafrazy Recanatiego służą temu, by pokazać, że wyrażenia w zasięgu czasownika epistemicznego są zarazem użyte, jak i wymienione tworząc parafrazę, rozbijamy analizowane zdanie na dwa komponenty, z których jeden odpowiada wyrażeniu użytemu, a drugi – wymienionemu. Formę wyrażenia użytego w sprawozdaniu („on sam”) pozostawiamy niezmienioną, natomiast formę wyrażenia, które ma odpowiadać bezpośredniemu cytowaniu, dostosowujemy zgodnie z regułami, rządzącymi przekształcaniem zdań w formie zależnej na bezpośrednie cytaty. Parafraza zdania (10) powinna zatem wyglądać następująco:

- (31) Każdy chłopiec sądzi, że (on sam) jest fajnym facetem w następujący sposób: „Ja jestem fajnym facetem”.

Powyższe rozważania pokazują, że rozwiązanie Recanatiego rozwiązuje łami-główki semantyczne generowane przez zdania o przekonaniach, nie poświęcając zarazem ich niewinności semantycznej oraz uwzględniając ich szczególne miejsce w architekturze umysłu. Tym, czego zabrakło w teorii Recanati'ego, a co zamierzam uwzględnić w moich dalszych rozważaniach, jest zwrócenie uwagi na wyjąt-kowość sprawozdań z przekonań również na tle innych metareprezentacji.

IV. Myślenie kontrfaktyczne

Tym, na co Recanati nie zwrócił uwagi, jest rola myślenia kontrfaktycznego, czyli umiejętności tworzenia reprezentacji niezgodnych z rzeczywistością, uwikłanej zarówno w udawanie, jak i w przypisywanie przekonań innym. Chociaż Recanati właściwie uchwycił różnicę między udawaniem a sprawozdaniami z przekonań, błędnie ocenił relację między tymi ostatnimi a innymi reprezentacjami drugiego rzędu. Wykształcenie się u dziecka zdolności do tworzenia metareprezentacji nie wystarcza, by przechodziło ono test fałszywych przekonań (Leslie 1987; Hobson 1984), istotna zdaje się być natomiast zdolność do manipulowania jednocześnie dwoma sprzecznymi reprezentacjami – jedną dotyczącą rzeczywistości i drugą odnoszącą się do sytuacji wyobrażonej (Byrne 2002). Riggs i in. (1998) oraz Grant i in. (2004) pokazali, że warunkiem koniecznym rozumienia błędnych przekonań jest myślenie kontrfaktyczne – ta sama, która pozwala nam interpretować kontrfaktyczne okresy warunkowe.

Moim zamiarem jest rozwinięcie i wzbogacenie teorii Recanatiego w taki sposób, by w rezultacie uzyskać analizę sprawozdań z przekonań, która:

1. uwzględni wyjątkowy charakter sprawozdań z przekonań, a więc pozwala wyjaśnić trudności związane z ich uczeniem się;
2. podtrzymuje tezę o ikonizacji i przezroczystości metareprezentacji, wyjaśniając ich nieprzezroczyste interpretacje poprzez odwołanie się do pragmatyki;
3. rozróżnia udawane twierdzenia o wyobrażonych sytuacjach i metareprezentacje, czyli twierdzenia o rzeczywistych przedmiotach, z którymi powiązane są wyobrażone sytuacje;
4. wyróżnia na tle innych metareprezentacji te, które wymagają myślenia kontrfaktycznego i uwzględnia pokrewieństwo sprawozdań z przekonań z kontrfaktycznymi okresami warunkowymi, pozwalając jednocześnie wyjaśnić przesunięcie w czasie między opanowaniem obu tych kompetencji.

Przytoczone powyżej badania stanowią motywację do posłużenia się semantyką kontrfaktycznych okresów warunkowych w analizie zdań o przekonaniach. Już Recanati zwrócił uwagę, że teoria symulacji, którą zaadaptował w swojej analizie metareprezentacji, przypomina ujęcie okresów warunkowych zaproponowane przez Oswalda Ducrota i Johna Mackiego (1973). Podkreślał on wprawdzie, że struktura metareprezentacji zasadniczo różni się od struktury okresów

warunkowych, jednak analizował jedynie okresy warunkowe w trybie oznajmującym, które nie angażują myślenia kontrfaktycznego – w przeciwieństwie do kontrfaktycznych okresów warunkowych, które składają się ze zdań w trybie przypuszczającym: „Gdyby było p , to by było q ”.

Narzędziem, które pozwoli mi przedstawić formalne rozwinięcie teorii meta-reprezentacji Recanatiego, uwzględniając jednak rolę myślenia kontrfaktycznego w reprezentowaniu przekonań innych osób, jest teoria zaproponowana przez Roberta Stalnakera (1968).

Kontrfaktyczne okresy warunkowe stanowią konstrukcje nieekstensjonalne, trzeba zatem było znaleźć taką metodę określania ich wartości logicznej, która nie odwoływałaby się do tabelki prawdziwościowych⁶. Rozwiązanie, które rozwinął Stalnaker (1968), znalazł Frank Ramsey (1990). Zasugerował on, żeby w celu określenia wartości logicznej okresu warunkowego przeprowadzić pewien eksperyment myślowy, który stanowi rodzaj symulacji. Chcąc ocenić wartość logiczną zdania o warunkowej formie „Gdyby było p , to by było q ”, symbolicznie „ $p > q$ ”, powinniśmy hipotetycznie dodać poprzednik p do zbioru naszych własnych przekonań, a więc *udawać*, że uznajemy jego prawdziwość. Jeśli już wcześniej uznaliśmy p za prawdziwe lub też nie mieliśmy żadnej opinii co do jego prawdziwości, pozostaje nam tylko rozważyć wartość logiczną q na gruncie zbioru przekonań wzbogaconego o p . Jeśli natomiast byliśmy uprzednio przekonani o fałszywości p , żeby oszacować wartość logiczną okresu warunkowego, potrzebujemy zmodyfikować zbiór przekonań tak, by pozostał niesprzeczny. Jeśli na gruncie wzbogaconego o poprzednik okresu warunkowego zbioru przekonań następnik jest prawdziwy – całe zdanie warunkowe jest prawdziwe.

Za powyższą propozycją stoi więc następująca intuicja: żeby ocenić wartość logiczną zdania:

(32) Gdyby Anna uczyła się pilnie, nie oblałaby egzaminu.

potrzebuję rozważyć *najbliższy* możliwy świat, czyli taki, który zasadniczo nie różni się od świata aktualnego poza tym, że poprzednik powyższego zdania jest w nim zdaniem prawdziwym. Innymi słowy, dodaję do zbioru moich przekonań zdanie „Anna uczyła się pilnie”, dokonuję w nim tylko niezbędnych modyfikacji, które pozwolą mi zachować niesprzeczność moich przekonań, a następnie oceniam na gruncie mojego nowego zbioru przekonań wartość logiczną zdania „(Anna) nie oblała egzaminu”. Jeśli następnik jest prawdziwy w wyznaczonym świecie, to całe zdanie warunkowe jest prawdziwe w świecie wyjściowym.

⁶ Kwestia wartości logicznej zdań warunkowych jest wciąż kontrowersyjna i żywo dyskutowana w literaturze przedmiotu, por. Bennett (2003).

Jeśli poprzednik okazał się prawdziwy w świecie aktualnym, to i wartość logiczną następnika oceniamy w świecie aktualnym.

Analiza Stalnaker'a opiera się na semantyce możliwych światów Saula Kripkego (1963). Struktura modelu składa się ze zbioru możliwych światów W , relacji względnej możliwości R między elementami W , a także wyróżnionego elementu zbioru możliwych światów γ , w którym zachodzą sprzeczności. $\langle w, w' \rangle \in R$ oznacza, że w' jest możliwy względem w . γ jest to świat niemożliwy: $\forall w \in W \langle w, \gamma \rangle \notin R$. Pozwala on interpretować okresy warunkowe, których poprzednik jest absurdalny.

Ponadto model zawiera dwuargumentową funkcję selekcji f , która reprezentuje wspomniany mechanizm symulacji – dla każdego zdania α w określonym świecie możliwym $w \in W$, funkcja f wybiera taki świat możliwy $w' \in W$, że α jest prawdziwe w w' :

$$f(\alpha, w) = w' \Leftrightarrow \langle w, w' \rangle \in R \wedge v(\alpha, w') = 1$$

Dla dowolnego zdania α , będącego poprzednikiem okresu warunkowego, a także dla dowolnego świata w funkcja selekcji spełnia następujące warunki:

1. Zdanie α w świecie wybranym przez funkcję selekcji przyjmuje wartość 1. Symbolicznie: $v(\alpha, f(\alpha, w)) = 1$
2. Wartością funkcji selekcji dla α w świecie w jest świat niemożliwy γ wtedy i tylko wtedy, gdy nie istnieje taki świat możliwy w' , że $\langle w, w' \rangle \in R$, w którym α byłoby prawdziwe.
3. Jeśli α jest prawdziwe w świecie w : $v(\alpha, w) = 1$, to wówczas funkcja selekcji dla α w świecie w przyjmuje wartość w .
4. Jeśli $v(\alpha, f(\beta, w)) = 1$ oraz $v(\beta, f(\alpha, w)) = 1$, to wówczas $f(\alpha, w) = f(\beta, w)$.

Jeśli następnik okresu warunkowego jest prawdziwy w świecie w' wyznaczonym przez funkcję selekcji, to zdanie warunkowe jest prawdziwe w świecie aktualnym w . Funkcja wartościowania v przyporządkowuje zdaniom w poszczególnych światach wartości $\{1, 0\}$. Wówczas dla dowolnych zdań α, β oraz dowolnego świata w zachodzą następujące warunki: zdanie warunkowe $\alpha > \beta$ jest prawdziwe w świecie w wtedy i tylko wtedy, gdy następnik β jest prawdziwy w świecie wybranym przez funkcję selekcji, czyli takim, w którym prawdziwe jest α . W przeciwnym wypadku zdanie warunkowe jest fałszywe. Warunki te zapisujemy formalnie:

$$\begin{aligned} v(\alpha > \beta, w) &= 1 \Leftrightarrow v(\beta, f(\alpha, w)) = 1 \\ v(\alpha > \beta, w) &= 0 \Leftrightarrow v(\beta, f(\alpha, w)) = 0 \end{aligned}$$

Stosując kontrfaktyczne okresy warunkowe, ludzie robią zasadniczo dwie rzeczy. Po pierwsze, stwierdzają, że w świecie aktualnym pewne sytuacje zależą

od innych, dzięki czemu pewne warunkowe stwierdzenia są prawdziwe. Po drugie, ludzie reprezentują czy też wyobrażają sobie cechy rozważanych stanów rzeczy, stosując mechanizm symulacji. Z symulacją mamy do czynienia również wtedy, gdy rozważamy coś z cudzego punktu widzenia. W przypadku myślenia kontrfaktycznego ludzie zakładają, że wydarzenia mogły się potoczyć inaczej lub że inna sytuacja mogła mieć miejsce, a następnie stwierdzają konsekwencje tej sytuacji niejako wewnątrz tej sytuacji – z jej perspektywy. Jest to udawana, symulowana asercja.

Przypisywanie innym ludziom przekonań, w szczególności w sytuacji, gdy te przekonania są fałszywe, wymaga tego samego mechanizmu symulacji. Tworząc sprawozdanie z przekonań typu (17), uruchamiamy te same mechanizmy, które potrzebne są do kompetentnego posługiwania się kontrfaktycznymi okresami warunkowymi. Zakładamy czy też wyobrażamy sobie, jak wygląda świat z punktu widzenia podmiotu, to jest zakładamy, że reprezentacja świata stworzona przez Jana odzwierciedla rzeczywistość. Wyobrażamy sobie, że Antonio Vivaldi zamiast kwartetów smyczkowych pisze powieści. Następnie pozostając wewnątrz tego wyobrażonego świata, stwierdzamy, że Antonio Vivaldi jest autorem „Czarodziejskiej Góry”.

W języku polskim zdania o następującej formie:

(33) O ile A się nie myli, to p .

są naturalnym sposobem referowania przekonań innych osób. Powyższe zdanie znaczy nic innego, jak:

(34) Gdyby A miał rację, p byłoby prawdziwe.

Zdanie (17) możemy zatem sparafrazować jako zdanie warunkowe, np. „Gdyby przekonania Jana były prawdziwe, to Antonio Vivaldi byłby autorem »Czarodziejskiej Góry«”. Ewentualnie „Gdyby Jan miał rację, to Vivaldi byłby autorem »Czarodziejskiej Góry«”.

Dzięki kognitywnemu pokrewieństwu sprawozdań z przekonań i okresów warunkowych możemy zastosować w ich interpretacji stworzony przez Stalnaker’a formalizm. Podobnie jak w przypadku okresów warunkowych strukturę modelu dla sprawozdań z przekonań stanowi trójka uporządkowana (W, R, γ) , gdzie W jest zbiorem możliwych światów, R jest relacją względnej możliwości, którą można rozumieć jako kompatybilność pomiędzy zbiorem przekonań podmiotu, o którego przekonaniach jest mowa, a pewnym światem możliwym. $\langle w, w' \rangle \in R$ oznacza, że w' jest możliwy względem zbioru przekonań, które podmiot posiada w świecie w . Jak poprzednio γ oznacza świat niemożliwy, który pozwoli nam interpretować sprzeczne przekonania podmiotu.

Funkcja selekcji f przyjmuje jako argumenty pary uporządkowane, w których skład wchodzi: zbiór przekonań podmiotu A , który będziemy oznaczać jako B_A , oraz możliwy świat w . Dla każdego podmiotu A , o którego przekonaniach mowa, istnieje taka funkcja selekcji f_A , która przyporządkowuje podmiotowi w danym świecie w taki świat możliwy (reprezentację świata) w' , który według niego jest światem aktualnym (czyli reprezentacją odzwierciedlającą rzeczywistość).

Mechanizm symulacji jest uruchamiany, gdy autor sprawozdania z cudzych przekonań musi postawić siebie w położeniu tej drugiej osoby. Załóżmy, że z punktu widzenia autora sprawozdania podmiot referowanych przekonań myli się co do prawdziwości pewnych zdań. Symulacja oznacza „uruchomienie” funkcji selekcji, która wyznacza najbliższy świat możliwy kompatybilny ze zbiorem przekonań podmiotu w świecie aktualnym, czyli takim, w którym autor sprawozdania ocenia przekonania podmiotu. Sprawozdanie z przekonań jest prawdziwe, jeżeli w świecie przekonań podmiotu wybranym przez funkcję selekcji f zdanie podrzędne zawarte w sprawozdaniu jest również prawdziwe. W przeciwnym wypadku sprawozdanie z przekonań jest fałszywe. Formalnie, dla dowolnego zdania p , podmiotu przekonań A oraz świata $w \in W$:

$$v(„A \text{ sądzi, że } p”, w) = 1 \Leftrightarrow v(p, f(B_A, w)) = 1$$

Jeśli przekonania podmiotu są prawdziwe w świecie aktualnym (lub względem reprezentacji świata preferowanej przez sprawozdawcę), wartość logiczna sprawozdania z przekonań jest oceniana w tym samym świecie w , w którym oceniana jest wartość logiczna zdania wyrażającego przekonanie podmiotu⁷:

$$\text{Jeśli } f(B_A, w) = w, \text{ to } v(„A \text{ sądzi, że } p”, w) = 1 \Leftrightarrow v(p, w) = 1$$

Ujęcie to pozwala również łatwo interpretować zagnieżdżone sprawozdania z przekonań, czyli zdania o formie:

(35) Jan sądzi, że Piotr jest przekonany, że p .

W takiej sytuacji autor sprawozdania musi wpięrow wyobrazić sobie, jaki jest świat z punktu widzenia Jana, by następnie wyobrazić sobie, jak Jan wyobraża sobie świat przekonań Piotra. Funkcja selekcji musi więc wpięrow wybrać z pozycji świata aktualnego w taki możliwy świat, w którym prawdziwe są przekonania Jana: $f(B_J, w) = u$, by następnie w świecie przekonań Jana u przypisać zbiorowi przekonań Piotra w u taki świat możliwy, w którym są one prawdziwe:

⁷ Mówiąc ściśle, funkcja selekcji, czy też jej psychologiczny odpowiednik – mechanizm symulacji, nie są tu potrzebne. Sprawozdanie z prawdziwych przekonań można potraktować jako zwykłą konstrukcję ekstensjonalną.

$f(B_p, f(B_j, w)) = f(B_p, u) = z$. Sprawozdanie z przekonań Jana dotyczących przekonań Piotra jest zatem prawdziwe, gdy zdanie p jest prawdziwe w możliwym świecie z .

W świetle kontrfaktycznego ujęcia zdań o przekonaniach symulacja jest mechanizmem, który pozwala nam zmienić warunki, w których oceniana jest wartość logiczna zdania podrzędnego zawartego w sprawozdaniu z przekonań. Ikoniczność sprawozdań z przekonań zostaje zachowana, ponieważ możliwe światy należy tutaj rozumieć jako reprezentacje świata – jednego, tego samego dla wszystkich podmiotów. Kontekst pozwala autorowi sprawozdania z przekonań zdecydować, które elementy swojej reprezentacji świata powinien on zmodyfikować, by uzyskać symulację (odpowiedniego fragmentu) świata mentalnego podmiotu, o którym mowa.

V. Podsumowanie

Celem niniejszego artykułu było przedstawienie takiej semantyki zdań o przekonaniach, która byłaby zgodna ze współczesnymi teoriami psychologicznymi, dotyczącymi zdawania relacji z przekonań innych osób. Opierając się na wynikach badań nad mechanizmem przypisywania przekonań innym osobom, zwróciłam uwagę na to, że kompetentne posługiwanie się sprawozdaniami z przekonań wymaga nie tylko umiejętności przyjmowania cudzej perspektywy oraz wykorzystywania mechanizmu symulacji, ale również zdolności do rozważania i rozumienia sytuacji, które nie zachodzą w rzeczywistości. W artykule pokazałam, że spośród dotychczasowych ujęć semantyki dla sprawozdań z przekonań najbliższa spełnieniu tych postulatów jest koncepcja François Recanatiego, a ponadto zaproponowałam modyfikację i udoskonalenie jego teorii z wykorzystaniem analizy formalnej kontrfaktycznych okresów warunkowych zaproponowanej przez Roberta Stalnakera.

Podziękowania. Chciałabym serdecznie podziękować Justynie Grudzińskiej za niezwykle cenne komentarze i sugestie. Za pomocne wskazówki dziękuję również Piotrowi Kaźmierczakowi i Jakubowi Szymanikowi, a także Agnieszce Plucie za wprowadzenie mnie w tematykę teorii umysłu.

Bibliografia

- Baron-Cohen, S. (1995), *Mindblindness. An Essay on Autism and Theory of Mind*, The MIT Press.
- Baron-Cohen, S., Leslie, A.M. i Frith, U. (1985), „Does the autistic child have a »theory of mind«?”, „Cognition”, vol. 21, s. 37–46.

- Bennett, J. (2003), *A Philosophical Guide to Conditionals*, Oxford University Press.
- Byrne, R.M. (2002), „Mental Models and Counterfactual Thoughts About What Might Have Been”, „Trends in Cognitive Sciences”, vol. 6 (10), s. 426–431.
- Cann, R. (1993), *Formal Semantics: An Introduction*, Cambridge University Press.
- Carruthers, P. i Smith, P.K., red. (1996), *Theories of Theories of Mind*, Cambridge University Press.
- Chierchia, G. i McConnell-Ginet, S. (2000), „Meaning and Grammar. An Introduction to Semantics”, rozdz. *Intensionality*, s. 204–261, The MIT Press.
- Davidson, D. (1968), „On Saying That”, „Synthese”, vol. 19, s. 130–146.
- Davidson, D. (2001), „Quotation”, w: *Inquiries into Truth and Interpretation*, s. 79–92, Oxford.
- Dowty, D.R., Wall, R.E. i Peters, S. (1981), „Introduction to Montague Semantics”, rozdz. *Montague's Intensional Logic*, s. 141–178, D. Reidel Publishing Company.
- Fodor, J.A. (1987), *Psychosemantics: the Problem of Meaning in the Philosophy of Mind*, The MIT Press.
- Fodor, J.A. (1995), „A Theory of the Child's Theory of Mind”, w: Davies, M. i Stone, T., red., *Mental Simulation*, s. 109–122, Blackwell Publishers.
- Frege, G. (1977), „Myśl – studium logiczne”, w: *Pisma semantyczne*, tłum. Bogusław Wolniewicz, s. 101–129, PWN.
- Goldman, A.I. (2009), „Mirroring, Mindreading, and Simulation”, w: *Mirror Neuron Systems: The Role of Mirroring Processes In Social Cognition*, Humana Press.
- Grant, C.M., Riggs, K.J. i Boucher, J. (2004), „Counterfactual and Mental State Reasoning in Children with Autism”, „Journal of Autism and Developmental Disorders”, vol. 34 (2), s. 177–188.
- Grudzińska, J. (2003), „Intensjonalizm a ekstensjonalizm. Porównanie dwóch programów semantycznych”, „Przegląd Filozoficzny”, vol. 1, s. 165–180.
- Grudzińska, J. (2008), „Quine o metareprezentacjach językowych”, „Przegląd Filozoficzny”, vol. 4, s. 267–281.
- Higginbotham, J. (1986), „Linguistic Theory and Davidson's Program in Semantics”, w: LEPORE, E., red., *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford: Blackwell.
- Hobson, R.P. (1984), „Early Childhood Autism and the Question of Egocentrism”, „Journal of Autism and Developmental Disorders”, vol. 14 (1), s. 85–104.
- Janssen, T.M.V. (1997), „Compositionality”, w: Van Benthem, J.F.A.K. i Ter Meulen, A., red., *Handbook of Logic and Language*, Amsterdam: Elsevier.

- Kapitan, T. (2002), „Review: Oratio Obliqua, Oratio Recta: An Essay on Metarepresentation by François Recanati”, „The Philosophical Review”, vol. 111 (3), s. 459–462.
- Kaplan, D. (1989), „Demonstratives”, w: *Themes from Kaplan*, Oxford University Press.
- Kripke, S. (1963), „Semantical Analysis of Modal Logics, I”, „Zeitschrift für mathematische Logik und Grundlagen der Mathematik”, vol. 9, s. 67–96.
- Kripke, S. (2001), *Nazywanie a konieczność*, Fundacja Aletheia, tłum. Bohdan Chwedeńczuk.
- Leslie, A.M. (1987), „Pretense and Representation: The Origins of »Theory of Mind«”, „Psychological Review”, vol. 94 (4), s. 412–426.
- Leslie, A.M. (1992), „Pretense, Autism, and the Theory-of-Mind Module”, „Current Directions in Psychological Science”, vol. 1 (1), s. 18–21.
- Mackie, J.L. (1973), *Truth, Probability and Paradox*, Oxford: Clarendon Press.
- Quine, W.V.O. (2000), „Oznaczanie i modalność”, w: *Z punktu widzenia logiki*, s. 171–193, Fundacja Aletheia, tłum. Barbara Stanosz.
- Ramsey, F.P. (1990), „General Propositions and Causality”, w: Mellor, D.H., red., *Philosophical Papers*, s. 145–163, Cambridge University Press, (original publication, 1929).
- Recanati, F. (2000a), „The Iconicity of Metarepresentations”, w: Sperber, D., red., *Metarepresentations: a Multidisciplinary Perspective*, Oxford University Press.
- Recanati, F. (2000b), *Oratio Obliqua, Oratio Recta. An Essay on Metarepresentation*, The MIT Press.
- Riggs, K.J., Peterson, D.M., Robinson, E.J. i Mitchell, P. (1998), „Are Errors in False Belief Tasks Symptomatic of a Broader Difficulty with Counterfactuality?”, „Cognitive Development”, vol. 13, s. 73–90.
- Scheffler, I. (1955), „On Synonymy and Indirect Discourse”, „Philosophy of Science”, vol. 22 (1), s. 39–44.
- Stalnaker, R.C. (1968), „A Theory of Conditionals”, w: Cornman, J.W., Hausman, A. i Echelbarger, C., red., *Studies in Logical Theory, American Philosophical Quarterly: Monographs 2*, s. 98–112, Oxford: Blackwell.

Belief Reports in Philosophy of Language and Cognitive Science

Keywords: *belief reports, iconicity, theory of mind, simulation, Recanati, counterfactual thinking*

Sentences about beliefs have been perplexing to philosophers from the beginning of analytic philosophy. All attempts to develop compositional semantics for natural language had to face a number of puzzles that belief reports generated. None of the existing theories are free from problems and, which is more worrying, none of them takes into account the rich psychological findings on how people ascribe beliefs to others. In this article certain conditions are postulated which a psychologically adequate theory should meet. Namely, the author discusses the principle of iconicity, which is the central notion in the theory proposed by François Recanati. Although Recanati tries to account for cognitive mechanisms that underpin production and interpretation of language, in his view belief reports are wrongly considered to be equivalent to other metarepresentations. Furthermore, Recanati does not notice a link between belief ascription and counterfactual reasoning. Hence, a theory is proposed that expands on Recanati's idea in a way that better satisfy psychological postulates.