O równoznaczności wyrażeń w ujęciu Ajdukiewicza

Wojciech Buszkowski
Wydział Matematyki i Informatyki
Uniwersytetu im. Adama Mickiewicza w Poznaniu

1 Wprowadzenie.

Kilka spostrzeżeń o Ajdukiewiczowskiej koncepcji znaczenia wyrażeń dedykuję Pani Profesor Krystynie Zamiarze, z którą miałem okazję współpracować w Zarządzie Polskiego Towarzystwa Logiki i Filozofii Nauki w latach 1993-1996. Mam nadzieję, że ta praca koresponduje z badaniami Pani Profesor, kontynuującymi tradycje Szkoły Lwowsko-Warszawskiej, a także Jej pracami o psychologizmie, np. [PPR], [PBH]. Ajdukiewicz [OZW] przedstawia swoją koncepcję jako alternatywę względem asocjacjonistycznej teorii znaczenia, mającej charakter psychologistyczny.

Ajdukiewicz nie aprobuje definicji znaczenia, odwołującej się do treści myśli użytkowników języka, wypowiadających dane wyrażenie. Zamiast tego bada możliwość zdefiniowania relacji równoznaczności wyrażeń w systemie języka, który jest określony poprzez dyrektywy znaczeniowe. Jak się zdaje, podejście Ajdukiewicza jest bliskie praktyce naukowej. Chociaż twierdzimy, że aksjomaty logiki, arytmetyki lub innej nauki opierają się na znaczeniu podstawowych pojęć tych nauk, mamy jednak na myśli znaczenie intuicyjne, nie poddane ścisłej analizie. Dopiero gotowy system, najlepiej system sformalizowany, może stanowić podstawę precyzacji znaczenia wyrażeń tego systemu. Trzeba zaznaczyć, że pomimo licznych prób nie udało się dotychczas sformułować całkowicie przekonującej teorii znaczenia nawet dla stosunkowo ubogich języków formalnych.

Sytuacja jest jeszcze trudniejsza w przypadku języków naturalnych i dojrzałych systemów wiedzy naukowej, które zwykle nie poddają się pełnej formalizacji. O takich właśnie systemach myślał Kazimierz Ajdukiewicz, przedstawiając swoje koncepcje znaczenia w latach trzydziestych ubiegłego wieku, w pracach "O znaczeniu wyrażeń" [OZW], "Język i znaczenie" [JiZ] oraz "Obraz świata i aparatura pojęciowa" [OSAP]. Pierwsza praca została opublikowana w języku polskim, a dwie pozostałe w języku niemieckim; korzystam z tłumaczenia na język polski, opublikowanego w antologii "Język i Poznanie" [JiP1,JiP2]. Główny pomysł Ajdukiewicza polegał na określeniu stosunku równoznaczności wyrażeń jako relacji wzajemnej zastępowalności tych wyrażeń w dyrektywach aksjomatycznych, dedukcyjnych i empirycznych języka. Teoria Ajdukiewicza została później nazwana dyrektywalną teorią znaczenia. Autor uważał ją za swoje najważniejsze dokonanie filozoficzne tych lat. W okresie powojennym uznał ją jednak za niezadowalającą, powołując się na argument, przypisywany Tarskiemu, który jakoby godził w podstawowe założenia tej teorii. O ile mi wiadomo, w pracach ani listach Tarskiego nie ma żadnej wzmianki o tym argumencie, musimy więc polegać na świadectwie samego Ajdukiewicza z pracy [ZEKZ, 1960]. Argument ów był kilkakrotnie cytowany przez innych autorów, piszących o teorii Ajdukiewicza; por. Woleński [FSLW], Wójcicki [ATZ].

W niniejszej pracy podejmuję się obrony teorii Ajdukiewicza przed krytyką Tarskiego. Dokładna analiza tego argumentu prowadzi do wniosku, że dotyczy on pewnej bardzo ubogiej teorii formalnej i pewnej specjalnej definicji relacji zastępowalności wyrażeń. Dla bogatych systemów i/lub innych, bardziej naturalnych pojęć zastępowalności tego argumentu nie można odtworzyć. Zatem Ajdukiewicz przedwcześnie uznał, iż argument Tarskiego jest "nie do odparcia". Z drugiej strony koncepcja Ajdukiewicza rzeczywiście prowadzi do różnych trudności innego rodzaju, które jednak nie podważają jej sensowności.

Praca jest gruntownie przerobioną wersją referatu p.t. "Uwagi o Ajdukiewiczowskiej teorii znaczenia" [UATZ], napisanego jako materiał do dyskusji na konferencji "Filozoficzne Podstawy Nauki - Naukowe Podstawy Filozofii" w Rucianem-Nidzie w 2008 r. i zamieszczonego w sieci na stronie Studia Philosophica - Forum. W 2008 r. ukazała się obszerna książka J. Maciaszka [ZPP], w której autor proponuje oddalenie krytyki Tarskiego poprzez odmienne rozumienie relacji zastępowalności, co jest zbieżne z argumentacją mojej pracy. Warto może nadmienić, że to rozwiązanie podałem już w dyskusji na Polskim Zjeździe Filozoficznym w Toruniu w 1995 r. (sesja prowadzona przez Ryszarda Kleszcza). Niniejsza wersja pracy dyskutuje te zagadnienia obszerniej, uwzględniając uwagi Urszuli Wybraniec-Skardowskiej, Pawła Grabarczyka i innych uczestników konferencji.

Rozdział 2 omawia różne pojęcia zastępowalności w związku z teorią Ajdukiewicza i argumentem Tarskiego. Zajmuję się głównie różnicą między podstawianiem a przestawianiem oraz kwestią, czy dyrektywy aksjomatyczne i dedukcyjne należy ograniczyć do dyrektyw podstawowych, czy warto uwzględnić również dyrektywy wyprowadzalne. Rozdział 3 zajmuje się szczególnym zagadnieniem istnienia języków zamkniętych w sensie Ajdukiewicza, które grają istotną rolę w jego tezie skrajnego (radykalnego) konwencjonalizmu. Twierdzę, że kwestia istnienia takich języków nie jest oczywista, a przy pewnej eksplikacji koncepcji Ajdukiewicza i dla pewnej naturalnej klasy języków formalnych ma negatywne rozwiązanie.

Ponieważ następne rozdziały ograniczają się do tych szczegółowych problemów, więc chciałbym już w tym miejscu sformułować pewne uwagi ogólne.

Nie jest celem tej pracy całościowa prezentacja tematyki znaczenia wyrażeń języka, ani nawet zasadniczych elementów teorii Ajdukiewicza; np. pomijam interesującą koncepcję przekładu jednego języka na drugi. Odsyłam czytelników do dyskusji koncepcji Ajdukiewicza w książkach Woleńskiego [FSLW], Wójcickiego [ATZ] i Maciaszka [ZPP]. Mam nadzieję, że moja praca, poruszająca tylko niektóre zagadnienia o charakterze formalnym, przyczyni się do wyjaśnienia pewnych subtelnych aspektów tej teorii. Pokonywanie trudności, które się pojawiają w trakcie realizacji programu Ajdukiewicza, powinno owocować głebszym rozumieniem pojęcia znaczenia, pojmowanego czysto językowo (syntaktycznie), a także naświetlić subtelne własności języków formalnych.

Czy trafna jest zasadnicza myśl Ajdukiewicza, że znaczenie wyrażeń można niejako odzyskać z dyrektyw znaczeniowych języka? Można wysunąć zarzut, że odwrócono tu naturalną kolejność rzeczy: to dyrektywy znaczeniowe opierają się na znaczeniu występujących w nich wyrażeń i powinny być uzasadnione na tej podstawie. Jak zauważyłem w drugim akapicie, ta opinia jest intuicyjnie słuszna, ale w rzeczywistości postępujemy inaczej nawet w przypadku standardowych języków formalnych logiki. Podejście Ajdukiewicza wydaje się zgodne z praktyką naukową.

Czy teoria Ajdukiewicza prowadzi do użytecznej definicji znaczenia, czy znaczenia wyrażeń mogą być skutecznie określone tymi metodami? Niewątpliwie dla bogatych systemów wiedzy i języka naturalnego nie da się łatwo ustalić, które wyrażenia są zastępowalne w dyrektywach znaczeniowych, chociażby dlatego, że tych dyrektyw jest bardzo wiele i nie znamy ich pełnej listy. Wójcicki [ATZ] analizuje problemy, związane z określeniem znaczenia nazwy "Warszawa" w języku polskim, konfrontując podejście Ajdukiewicza z innymi koncepcjami. Nie znaczy to jednak, że definicja Ajdukiewicza jest bezużyteczna. Może stanowić teoretyczny punkt odniesienia dla innych metod. Z podobną sytuacją mamy często do czynienia w nauce. Definicja prawa logiki pierwszego rzędu jako formuły prawdziwej we wszystkich strukturach (interpretacjach) dla danego języka ma analogiczny charakter. Możemy stosować ją do falsyfikacji, nie weryfikacji praw. Twierdzenie o pełności gwarantuje, że weryfikacja praw jest równoważna ich wyprowadzeniu w systemie dedukcyjnym tej logiki. Niewykluczone, że pojęcie równoznaczności w sensie Ajdukiewicza dałoby się scharakteryzować za pomocą pewnych reguł formalnych, przynajmniej dla niektórych, stosunkowo prostych języków.

Czy na gruncie tej teorii otrzymujemy adekwatną definicję znaczenia? To pytanie wydaje się źle postawione. Pojęcie znaczenia jest wieloznaczne, nawet jeżeli ograniczymy się do znaczenia wyrażeń w tym sensie, jaki miał na myśli Ajdukiewicz. To samo wyrażenie ma różne znaczenia dla różnych osób, mówiących tym samym językiem, w zależności od ich wykształcenia, doświadczeń życiowych, wiedzy. Ajdukiewicz rozważał sytuację idealną, w której wszyscy użytkownicy języka stosują wyrażenia tego języka w tym samym znaczeniu; z taką sytuacją mamy do czynienia w przypadku języków formalnych (i tu z wieloma zastrzeżeniami), raczej nie w przypadku języków nauk przyrodniczych, ani języka naturalnego. Ten idealny model może być jednak właściwy dla rozważań teoretycznych.

Zasadne wydaje się przypuszczenie, że rzeczywistym powodem, dla którego Ajdukiewicz porzucił swoją teorię znaczenia, był nie tyle wspomniany argument Tarskiego, co upowszechnienie semantyki Tarskiego. Stało się jasne, że pojęcia semantyczne mogą być ściśle określone bez popadania w sprzeczności i za ich pomocą można skutecznie rozwijać rozmaite teorie znaczenia (Ajdukiewicz pisze o tym wyraźnie w [ZETZ]). Trudno nie zgodzić się z tym poglądem. Z drugiej strony, sukcesy semantyki Tarskiego nie zmuszają nas wcale do porzucenia innych semantyk, opartych na odmiennych zasadach. W literaturze logicznej znane są semantyki alternatywne względem semantyki Tarskiego, np. semantyki konstruktywistyczne, w których znaczenie zdania sprowadza się do dowodów tego zdania (na podstawie założeń), przy czym istotną rolę odgrywa relacja równoważności dowodów. Takie podejście można też zastosować do składników zdań (por. Francez i inni [PTS]). Widoczna jest analogia z teorią Ajdukiewicza; tu i tam znaczenie definiuje się za pomocą pewnej relacji równoważności między obiektami syntaktycznymi, nie korzystając z modeli pozajęzykowych.

2 Relacja równoznaczności.

W pracy [OZW] Ajdukiewicz formułuje następującą definicję równoznaczności wyrażeń.

Definicja 1. "[...] wyrażenie A jest w języku J równoznaczne z wyrażeniem B, to tyle, co: każde i tylko takie przeżycie P, z którego wedle dyrektyw języka J można bezpośrednio wywieść pewne zdanie Z języka J, zawierające wyrażenie A, jest zarazem przeżyciem, z którego wedle dyrektyw języka J bezpośrednio można wywieść zdanie, różniące się od zdania Z jedynie tym, że na miejscu A (wszędzie lub nie wszędzie) figuruje wyrażenie B, a nadto każde i tylko takie zdanie, które bezpośrednio można wedle dyrektyw języka J wywieść ze zdania Z, zawierającego wyrażenie A, daje się też wedle dyrektyw języka J bezpośrednio wywieść ze zdania Z przez zastąpienie (wszędzie lub nie wszędzie) wyrażenia A przez wyrażenie B". ([JiP1], str. 132.)

Występujący w tej definicji termin "przeżycie" nie został zadowalająco wyjaśniony w [OZW]. Z dalszych fragmentów pracy dowiadujemy się, że przeżycie może polegać na uznaniu pewnych zdań. Przystosowując definicję Ajdukiewicza do języków formalnych, można - jak się zdaje - przyjąć, że P jest po prostu pewnym zbiorem zdań uznanych lub założonych. W swoim komentarzu do [UATZ] prof. Wybraniec-Skardowska słusznie zwraca uwagę, że Ajdukiewicz podaje jeszcze drugą, zmodyfikowaną definicję, w której zwrot "bezpośrednio wywieść" zostaje zastąpiony zwrotem "bezpośrednio wywieść w sposób dla danego wyrażenia istotny" ([JiP1], str. 134.). Pomijam cytowanie tej drugiej definicji, której forma nie różni się zasadniczo od powyższej. Pomijam też precyzyjną (zawiłą) definicję "bezpośredniej wywodliwości w sposób istotny dla danego wyrażenia"; jest to ograniczenie dyrektyw bezpośredniej wywodliwości do takich, w których dane wyrażenie odgrywa istotną rolę, tzn. nie można go zastąpić na wszystkich miejscach dowolnym innym wyrażeniem tej samej kategorii. Zgodnie z przykładem Ajdukiewicza, zdanie "Każdy bicz jest bezwładny" jest bezpośrednio wywodliwe ze zdania "Każdy bicz jest ciałem i każde ciało jest bezwładne", lecz nie jest wywodliwe w sposób istotny dla wyrażenia "bicz", skoro analogiczna inferencja jest poprawna po zastąpieniu słowa "bicz" dowolną inną nazwą ogólną, np. słowem "miecz". Ajdukiewicz zapewne uznałby stwierdzenie, że to wnioskowanie jest istotne dla wyrażenia "każdy", ponieważ nie można go zastąpić np. wyrażeniem "pewien", zachowując poprawność wnioskowania. Zastrzeżenie istotności jest potrzebne Ajdukiewiczowi do wyeliminowania sytuacji, gdy zastąpienie pewnego wyrażenia, np. "bicz", wyrażeniem równoznacznym, np. "bat", na niektórych miejscach w pewnej dyrektywie zakłóca poprawność tej dyrektywy. Na przykład, zdania "Każdy bat jest bezwładny" nie można bezpośrednio wywieść ze zdania "Każdy bicz jest ciałem i każde ciało jest bezwładne"; ten fakt nie pociąga nierównoznaczności wyrażeń "bicz" i "bat", skoro dana dyrektywa (cytowana powyżej) nie jest istotna dla wyrażenia "bicz".

W świetle powyższego, można chyba przyjąć, że Definicja 1 określa równoznaczność wyrażeń jako ich wzajemną zastępowalność w dyrektywach (bezpośredniej) wywodliwości języka J, istotnych dla danego wyrażenia. Podamy symboliczną postać tej definicji (nie występującą u Ajdukiewicza).

> Definicja 1'. A jest równoznaczne z B wtedy i tylko wtedy, gdy spełnione są dwa warunki: (1) dla wszelkich $P, Z[-]$ $P \vdash_A Z[A] \Leftrightarrow P \vdash_B Z[B])$, (2) dla wszelkich $Z[-], Z'$ $(Z[A] \vdash_A Z' \Leftrightarrow Z[B] \vdash_B Z')$.

$Z[-]$ reprezentuje kontekst zdaniowy, tj. wyrażenie, powstające z pewnego zdania przez zastąpienie pewnego składnika tego zdania specjalnym symbolem '-', za który wolno podstawiać dowolne wyrażenia tej samej kategorii, co zastąpiony składnik (jest to pojęcie czysto techniczne, wygodne dla naszych rozważań). $Z[A]$ oznacza wynik prawidłowego podstawienia wyrażenia A za '-' w $Z[-]$. P reprezentuje przeżycie; w dalszym ciągu utożsamiamy przeżycia ze zbiorami zdań. Wreszcie $\vdash_A$ (odp. $\vdash_B$) oznacza relację bezpośredniej wywodliwości w sposób istotny dla wyrażenia A (odp. B). Przyjmujemy, że odpowiednia relacja wywodliwości nie zachodzi, jeżeli wyrażenie $Z[A]$ albo $Z[B]$ jest wadliwie skonstruowane, tzn. dane podstawienie jest nieprawidłowe.

Nasza formalizacja warunku (2) wydaje się poprawna, jeżeli Z' nie zawiera A, ani B. W przeciwnym wypadku, Definicja 1 nie jest całkiem jasna. ale jej wariant z zastrzeżeniem istotności usuwa tę niejasność: "każde i tylko takie zdanie, które bezpośrednio i w sposób istotny dla wyrażenia A daje się wedle dyrektyw języka J wywieść z uznania zdania Z, zawierającego wyrażenie A, daje się też wywieść wedle dyrektyw języka J bezpośrednio i w sposób istotny dla wyrażenia B z uznania zdania powstającego ze zdania Z przez zastąpienie (wszędzie lub nie wszędzie) wyrażenia A przez wyrażęnie B." ([JiP1], str. 134.) To sformułowanie jest zgodne z naszą formalizacją.

Zastrzeżenie istotności utrudnia formalną eksplikację Definicji 1. W następnych pracach Ajdukiewicz zrezygnował z tego ograniczenia, modyfikując jednak sens zastępowalności. Przyjrzyjmy się drugiej definicji równoznaczności, podanej w pracy [JiZ]. Tu Ajdukiewicz formułuje "warunek konieczny równoznaczności [...] wyrażeń A i A' tego samego języka S" (poprzednia definicja formułowała warunek konieczny i dostateczny).

Definicja 2. "[...] jeśli A i A' mają w języku S to samo znaczenie, to muszą się zachowywać jednakowo w całkowitym zakresie dyrektyw znaczeniowych języka, tj. całkowity zakres dyrektyw znaczeniowych nie może doznać zmiany przez to, że we wszystkich jego elementach postawi się A' za A i A za A'. To znaczy: 1) jeśli według jakiejś aksjomatycznej dyrektywy znaczeniowej zdanie Z

musi być bez zastrzeżeń uznane, to istnieje też aksjomatyczna dyrektywa znaczeniowa, według której należy uznać bez zastrzeżeń zdanie powstałe ze zdania Z przez wstawienie A' za A i A za A'; 2) jeśli istnieje dedukcyjna reguła znaczeniowa, według której można ze zdania (czy też klasy zdań) Z_1 wywnioskować zdanie Z_2, to istnieje także dedukcyjna reguła znaczeniowa, według której ze zdania powstałego z Z_1 przez wstawienie A' za A i A za A', można wywnioskować zdanie, powstałe z Z_2 przez wstawienie A' za A i A za A'; 3) jeśli według empirycznej reguły znaczeniowej na podstawie pewnych danych można uznać zdanie Z, to istnieje także empiryczna reguła znaczeniowa, według której na podstawie tych samych danych należy uznać zdanie powstałe ze zdania Z przez wstawienie A' za A i A za A'". Cytujemy za [JiP1], str. 163.

Powyższa definicja różni się od poprzedniej pod wieloma względami. Z naszego punktu widzenia najbardziej istotna jest różnica w rozumieniu wzajemnej zastępowalności wyrażeń. Poprzednio w warunkach (1) i (2) zamienialiśmy $Z[A]$ na $Z[B]$, czyli wyrażenia A na B w $Z[-]$ i odwrotnie, bez zmiany kontekstu dyrektywy, tj. przeżycia P i kontekstu $Z[-]$ w warunku (1) oraz zdania Z' i kontekstu $Z[-]$ w warunku (2), natomiast Definicja 2 postuluje zmianę kontekstu (patrz punkt 2) oraz wymaga *równoczesnego* wstawienia A' za A i A za A'. Owa równoczesna zamiana dwóch wyrażeń, zwana dalej przestawianiem, jest żródłem kłopotów wskazanych przez Tarskiego. Zacytujemy komentarz samego Ajdukiewicza z późnej pracy [ZEKZ].

"Przyjmowałem mianowicie, że reguły sensu są inwariantne wobec przestawiania wyrażeń równoznacznych, tzn. że reguły sensu nie ulegną zmianie, gdy w nich nazwa jednego z wyrażeń równoznacznych zostanie zastąpiona przez nazwę drugiego. Przyjmowałem jednak prócz tego odwrócenie tej tezy, byłem mianowicie skłonny twierdzić, że jeżeli reguły sensu są inwariantne wobec przestawienia jakichś wyrażeń, to wyrażenia te są równoznaczne."

"Te dwie tezy pozwalały mi podać definicję stosunku równoznaczności wyrażeń, która głosiła, iż dwa wyrażenia są w języku J równoznaczne zawsze i tylko, gdy reguły sensu tego języka są inwariantne wobec przestawienia tych wyrażeń, czyli - co na jedno wychodzi - gdy reguły sensu języka to samo mówią o jednym z tych wyrażeń, co o drugim."

"Ta definicja znaczenia może się łatwo spotkać z krytyką, gdy się reguły sensu ogranicza tylko do reguł aksjomatycznych i dedukcyjnych. Wyobraźmy sobie język, w którym reguły aksjomatyczne nakazują bezwzględne uznawanie aksjomatów rachunku funkcyjnego z identycznością, a nadto dwóch aksjomatów dodatkowych:

$$A \neq B,$$

$$B \neq A,$$

gdzie A i B są nowymi stałymi terminami pierwotnymi. Jako reguły dedukcyjne niech obowiązują w tym języku reguły dowodzenia rachunku funkcyjnego z identycznością."

"Łatwo dostrzec, że reguły sensu tego języka są inwariantne wobec przestawienia stałych A i B. Wobec tego przyjęta definicja równoznaczności kazałaby terminy A i B uważać za równoznaczne. Ale wobec aksjomatów $A \neq B$ i $B \neq A$ terminy te mają różną denotację. Widać więc, że w rozważanym języku miałoby się dwa terminy o tym samym znaczeniu, a o różnej denotacji, co wydaje się konsekwencją nie do przyjęcia."

"Zarzut ten, który w ustnej rozmowie postawił mi, wkrótce po ukazaniu się pracy *Sprache und Sinn*, Tarski, wydaje mi się nie do odparcia, przynajmniej w zastosowaniu do języków, w których nie ma innych reguł sensu, jak tylko aksjomatyczne i dedukcyjne. Zarzut ten wskazuje też bodaj na to, że pojęcia znaczenia nie można definiować za pomocą samych środków syntaktycznych bez zastosowania pojęć semantyki, w węższym tego słowa rozumieniu."

"Zarzut ten jednak godzi tylko w tezę, że dla równoznaczności dwóch wyrażeń wystarcza inwariancja reguł sensu, nie podważa jednak bynajmniej jej odwrócenia, że inwariancja reguł jest niezbędnym warunkiem równoznaczności". Cytujemy za [JiP2], str. 396–397.

Powyższy argument Tarskiego wykazuje nietrafność definicji relacji równoznaczności w stylu Definicji 2, lecz nie Definicji 1.

Niech T oznacza teorię w rachunku funkcyjnym z identycznością, której jedynymi aksjomatami pozalogicznymi są $a \neq b$ i $b \neq a$, gdzie a i b są stałymi przedmiotowymi (zgodnie ze współczesną praktyką oznaczamy je małymi literami). Niech $Z[-]$ będzie kontekstem zdaniowym $a \neq -$. Wtedy $Z[b]$ to $a \neq b$, natomiast $Z[a]$ to $a \neq a$. Zdanie $Z[b]$ jest aksjomatem teorii T, lecz zdanie $Z[a]$ nie jest aksjomatem, ani nawet twierdzeniem teorii T (jest sprzeczne z twierdzeniem $a = a$). Zatem nie jest spełniony warunek (1) Definicji 1 (przyjmując za P pusty zbiór założeń), więc wyrażenia a i b nie są zastępowalne w regułach wywodliwości teorii T. Zauważmy, że - zgodnie z koncepcją Ajdukiewicza - $Z[b]$ jest wywodliwe w T w sposób istotny dla b, skoro $Z[a]$ nie jest twierdzeniem, natomiast $Z[a]$ nie jest wywodliwe w sposób istotny dla a, skoro nie jest wywodliwe w ogóle. Nie jest też spełniony warunek (2), jeżeli pominiemy zastrzeżenie istotności; ze zdania $Z[a]$ możemy wyprowadzić każde zdanie, np. $c \neq c$ (c jest jeszcze jedną nową stałą przedmiotową), lecz $c \neq c$ nie jest wyprowadzalne z $Z[b]$. W konsekwencji, wyrażenia a i b nie są równoznaczne w sensie Definicji (1).

Definicja 1 i jej formalizacja 1' odwołuje się do *podstawiania* wyrażenia B za A i odwrotnie, na przynajmniej jednym miejscu w danej dyrektywie (podstawienie na więcej niż jednym miejscu sprowadza się do kolejnego podstawiania na jednym miejscu), podczas gdy Definicja 2 wykorzystuje *przestawianie* wyrażeń A i B. W Definicji 1 Ajdukiewicz pisze, że Z' różni się od Z tym, że na miejscu wyrażenia A występuje "wszędzie lub nie wszędzie" wyrażenie B; zwrot "każde i tylko takie" implikuje warunki symetryczne, w których role wyrażeń A i B są odwrócone. W tej definicji nie żąda się, aby podstawianiu B za A towarzyszyło równoczesne podstawianie A za B na wszystkich możliwych miejscach. Równoczesne podstawianie A za B i B za A na wszystkich możliwych miejscach nazywamy *przestawianiem* wyrażeń A i B. Zauważmy, że Definicja 2 wymaga przestawialności wyrażeń A i A' w całej dyrektywie, podczas gdy Definicja 1 wymagała niezmienniczości dyrektywy ze względu na podstawianie w jednym zdaniu, wchodzącym w skład tej dyrektywy.

Ponieważ w pracy [JiZ], tekst Definicji 2 jest podany jako warunek konieczny równoznaczności, a nie definicja tej relacji, więc argument Tarskiego nie podważa również tej koncepcji, na co Ajdukiewicz zwraca uwagę w cytowanym wyżej komentarzu. Może być tak, iż wyrażenia A i A' spełniają warunki Definicji 2, lecz nie są równoznaczne na podstawie innych kryteriów. Ajdukiewicz przyznaje jednak, że jego pierwotnym zamiarem było podanie warunku koniecznego i dostatecznego; porzucił ten zamiar pod wpływem krytyki Tarskiego.

Jak wykazałem powyżej, argument Tarskiego upada, jeżeli równoznaczność wyrażeń określimy jako ich wzajemną podstawialność, nie przestawialność, w dyrektywach znaczeniowych języka. Ajdukiewicz nie wyjaśnił powodów, dla których zamienił podstawialność z [OZW] na przestawialność w [JiZ] i następnych pracach. Zapewne skłoniły go do tego analizy konkretnych przykładów. Zauważmy, że w Definicja 2 nie korzysta z zastrzeżenia istotności. Omawiana powyżej dyrektywa z wyrażeniami "bicz" i "bezwładny" zachowuje ważność, jeżeli podstawimy słowo "bat" za słowo "bicz" na wszystkich miejscach.

Porównajmy podstawialność i przestawialność. Definiowanie pewnych relacji równoważności wyrażeń przez podstawialność spotykamy często w logice i lingwistyce formalnej. Podstawowa definicja logicznej równoważności formuł w danej logice $\mathcal{L}$ jest postaci: A jest równoważne B, jeżeli dla dowolnego kontekstu $F[-]$ formuła $F[A]$ jest tezą tej logiki wtedy i tylko wtedy, gdy $F[B]$ jest tezą tej logiki (tu "jeżeli" oznacza "wtedy i tylko wtedy, gdy"). W lingwistyce matematycznej, wyrażenia x i y (reprezentowane jako skończone ciągi symboli) są równoważne ze względu na język L, jeżeli dla wszelkich wyrażeń u, v $uxv \in L$ wtedy i tylko wtedy, gdy $uyv \in L$. Natomiast przestawialność nie jest typowa dla takich definicji. Zeby zrozumieć powody, musimy uściślić pojęcie przestawiania wyrażeń.

Posłużymy się operacją przestawiania (transpozycji) $\tau(A, B)$: dla dowolnego wyrażenia E wyrażenie $\tau(A, B)(E)$ powstaje przez równoczesne podstawienie B za wszystkie wystąpienia A w E oraz wyrażenia A za wszystkie wystąpienia B w E. Warunki 1 i 2 Definicji 2 można wyrazić następująco: (I) jeżeli Z jest uznane na mocy pewnej reguły aksjomatycznej, to $\tau(A, A')(Z)$ jest uznane na mocy pewnej reguły aksjomatycznej, (II) jeżeli Z_2 jest konsekwencją Z_1 na mocy pewnej reguły dedukcyjnej, to $\tau(A, A')(Z_2)$ jest konsekwencją $\tau(A, A')(Z_1)$ na mocy pewnej reguły dedukcyjnej.

Nietrudno zauważyć, że warunki (I) i (II) wynikają z równoznaczności A i A' w sensie Definicji 1', jeżeli pominiemy zastrzeżenie istotności. Koniunkcja tych warunków nie pociąga równoznaczności wyrażeń A i A' w sensie Definicji 1'. Definicja równoznaczności jako przestawialności wyrażeń w dyrektywach języka prowadzi do pewnych trudności technicznych i merytorycznych; trzy mankamenty omawiam poniżej.

(A) Operacja przestawiania $\tau(A, B)$ nie zawsze jest wykonalna. Nie jest wykonalna wtedy, gdy wyrażenia A i B nie są rozłączne, np. A jest zawarte w B i różne od B. Niech Z będzie formułą rachunku zdań $p \Leftrightarrow \neg\neg p$. Nie można wykonać transpozycji $\tau(p, \neg\neg p)(Z)$. Wykonalność tej operacji jest ograniczona do sytuacji, gdy zastępowalne wyrażenia są na pewno identyczne lub rozłączne we wszystkich kontekstach, np. są prostymi symbolami (stałymi przedmiotowymi, symbolami operacji, relacji itp.).

(B) Koniunkcja warunków (I) i (II) definiuje relację równoważności (przymując ograniczenia, gwarantujące wykonalność transpozycji). Nie jest to tak oczywiste jak w przypadku Definicji 1'. Zwrotność wynika z faktu, że $\tau(A,A)(Z) = Z$, a symetria z równości $\tau(A,B) = \tau(B,A)$. Nie zawsze zachodzi równość $\tau(A,C)(Z) = \tau(B,C)(\tau(A,B)(Z))$, na przykład $\tau(a,b)(R(a,b,c)) = R(b,a,c)$ oraz $\tau(b,c)(R(b,a,c)) = R(c,a,b)$, lecz $\tau(a,c)(R(a,b,c)) = R(c,b,a)$. Mamy jednak $\tau(A,C) = \tau(A,B) \circ \tau(B,C) \circ \tau(A,B)$, skąd wynika przechodniość tej relacji. Niestety nie jest to kongruencja w algebrze języka: z równoznaczności a i b (definiowanej przez (I) i (II)) nie wynika równoznaczność $a+a$ i $b+a$. Oczywiście przyjmując Definicję 1', otrzymujemy kongruencję: jeżeli A i B są równoznaczne, to $Z[A]$ i $Z[B]$ są równoznaczne dla dowolnego kontekstu $Z[-]$ (niekoniecznie zdaniowego). Możemy wówczas tworzyć interesujące algebry znaczeń, których elementami są klasy abstrakcji relacji równoznaczności, co nie jest możliwe w przypadku relacji definiowanej jako przestawialność.

(C) Wydaje się, że relacja definiowana przez (I) i (II) wyraża raczej dualność (symetrię), a nie równoznaczność wyrażeń. Rozważmy teorię krat z operacjami $\cup$ (supremum) i $\cap$ (infimum) i aksjomatami:

$$\text{(L1) } x \cup x = x, \ \text{(L1') } x \cap x = x,$$

$$\text{(L2) } x \cup y = y \cup x, \ \text{(L2') } x \cap y = y \cap x,$$

$$\text{(L3) } (x \cup y) \cup z = x \cup (y \cup z), \ \text{(L3') } (x \cap y) \cap z = x \cap (y \cap z),$$

$$\text{(L4) } x \cup (x \cap y) = x, \ \text{(L4') } x \cap (x \cup y) = x.$$

Zbiór aksjomatów tej teorii jest inwariantny względem transpozycji $\tau(\cup,\cap)$; jest to dobrze znana dualność w teorii krat. Czy bylibyśmy skłonni twierdzić, że symbole $\cup$ i $\cap$ są równoznaczne w tej teorii? Nie, ponieważ w modelach tej teorii, czyli kratach, elementy $x \cup y$ i $x \cap y$ są na ogół różne, a więc te symbole oznaczają różne operacje. Argument Tarskiego jest podobny; pokazuje inwariantność dyrektyw teorii T względem transpozycji $\tau(a,b)$. Wynika stąd, że stałe a i b są dualne w tej teorii.

Nie chcę forsować tezy, że definiowanie równoznaczności jako przestawialności jest błędne. Wprawdzie przestawianie jest mniej standardową operacją niż podstawianie, lecz nie jest wykluczone, że może w pewnych sytuacjach prowadzić do użytecznej definicji równoznaczności. Argument Tarskiego można oddalić w inny sposób, nie rezygnując z przestawiania. Zauważmy, że teoria wykorzystana w przykładzie Tarskiego jest skrajnie uboga. W bogatych teoriach formalnych, a tym bardziej w rozwiniętych systemach wiedzy nie zdarza się przypadek, że w języku występują dwie nazwy, o których wiadomo tylko tyle, że oznaczają różne obiekty, a poza tym ich rola jest całkowicie symetryczna. Powróćmy do przykładu teorii krat; jest to bardzo uboga teoria formalna. Wystarczy wzbogacić tę teorię o nowy aksjomat, definiujący relację porządku:

$$x \leq y \Leftrightarrow x \cup y = y,$$

żeby usunąć przestawialność $\cup$ i $\cap$. Zasada dualności jest nadal ważna, ale wymaga równoczesnego przestawiania $\cup$ z $\cap$ i $\leq$ z $\geq$ w twierdzeniach teorii, a to już nie podważa koncepcji Ajdukiewicza. Wobec tego, Definicja 2 (rozumiana jako równoważność) niekoniecznie prowadzi do paradoksalnych konsekwencji, jeżeli zastosujemy ją do teorii dostatecznie bogatych. Podobny status ma twierdzenie Gödla o niezupełności, które jest prawdziwe dla teorii dostatecznie bogatych (to pojęcie wyjaśniam w następnym rozdziale).

Słuszny wydaje się pogląd, że koncepcja Ajdukiewicza dopuszcza różne eksplikacje; przecież sam twórca próbował różnych dróg. Pewną wadą prac Ajdukiewicza jest skłonność do formułowania bardzo ogólnych definicji i postulatów bez przetestowania ich zastosowań do konkretnych języków, np. języków formalnych. Dopiero zastosowania mogą rozstrzygnąć, która z proponowanych definicji jest najbardziej owocna w danym kontekście.

W niniejszej pracy nie podejmuję się tego zadania. Omówię tylko niektóre problemy, które pojawiają się przy stosowaniu teorii Ajdukiewicza.

Pojęcie dyrektywy jest niewątpliwie kluczowe dla tej teorii. Ajdukiewicz wyróżnia dyrektywy aksjomatyczne, dedukcyjne i empiryczne. W zastosowaniach do języków formalnych nie występują dyrektywy empiryczne. Pozostają dyrektywy aksjomatyczne i dedukcyjne. Nie jest jednak jasne, czy przez dyrektywy dedukcyjne należy rozumieć podstawowe reguły wnioskowania, czy mogą to być też reguły wyprowadzalne, a także czy reguły aksjomatyczne to tylko aksjomaty systemu, czy też wszelkie tezy tego systemu. Pewne fragmenty [JiZ] przemawiają za pierwszą opcją. Wtedy jednak dla standardowych systemów formalnych relacja równoznaczności, definiowana w stylu Ajdukiewicza, pokrywa się z relacją identyczności dla pewnych klas wyrażeń. Rozważmy aksjomaty klasycznego rachunku zdań w języku implikacyjno-negacyjnym:

(A1) $\varphi \Rightarrow (\psi \Rightarrow \varphi)$,

(A2) $[\varphi \Rightarrow (\psi \Rightarrow \chi)] \Rightarrow [(\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \chi)]$,

(A3) $(\neg\varphi \Rightarrow \neg\psi) \Rightarrow (\psi \Rightarrow \varphi)$,

gdzie φ, ψ, χ oznaczają dowolne formuły. Jedyną podstawową regułą wnioskowania jest reguła odrywania. Jeżeli określimy równoznaczność jako przestawialność w aksjomatach i regule odrywania, to żadna formuła złożona nie jest równoznaczna z żadną inną formułą, natomiast wszystkie zmienne zdaniowe są równoznaczne. Na przykład, formuła $\varphi \Rightarrow \psi$ nie może być przestawiona z żadną inną w odpowiedniej instancji (A2). W rachunku predykatów mamy aksjomat $\forall x\varphi(x) \Rightarrow \varphi(a)$. Oczywiście $\varphi(x)$ nie może być zastąpione przez żadną inną formułę za wyjątkiem $\varphi(a)$, lecz nie można przestawić $\varphi(x)$ i $\varphi(a)$. Podobne problemy wystąpią, gdy zdefiniujemy równoznaczność przez podstawialność. Nie wiadomo, czy takie relacje równoznaczności są zgodne z intencjami Ajdukiewicza; niewykluczone, ze zaakceptowałby różnoznaczność dowolnych dwóćh odmiennie skonstruowanych formuł logicznych.

Pojawiają się dodatkowe kłopoty z wykonalnością przestawiania, o których była mowa powyżej. Powinno się chyba przyjąć, że przestawianie dotyczy głównych formuł danej dyrektywy; np. dla (A1) formuł $\varphi, \psi, \psi \Rightarrow \varphi$ i całej formuły

(A1), lecz nie właściwych podformuł formuł φ, ψ. Gdybyśmy ograniczyli przestawianie tylko do formuł φ, ψ, χ, występujących w (A1)-(A3) (również jako podformuł, np. φ w $\neg\varphi$), to nietrudno zauważyć, że otrzymalibyśmy relację totalną: każde dwie formuły byłyby przestawialne.

Jeżeli φ jest formułą różną od ψ, to podstawianie ψ za φ na jednym miejscu w danej dyrektywie na ogół nie zachowuje tej dyrektywy; może przeprowadzić jedną dyrektywę w inną, zwłaszcza jeśli dopuścimy dyrektywy wyprowadzalne (por. następny akapit). Operacją pośrednią między podstawianiem a przestawianiem jest podstawianie na wszystkich miejscach. Niech $Z[A/B]$ oznacza wynik podstawienia B za wszystkie wystąpienia A w Z. Próba definicji równoznaczności wyrażeń A i B jako niezmienniczości zbioru dyrektyw ze względu na podstawianie $[A/B]$ i $[B/A]$ byłaby jednak chybiona, ponieważ tak określona relacja nie jest na ogół relacją równoważności. Rozważmy język, którego jedynymi dyrektywami są dyrektywy aksjomatyczne:

$$R(a,a), R(b,b), R(c,c), R(a,b), R(a,c), R(b,c)\,.$$

Zbiór dyrektyw jest niezmienniczy względem podstawień $[a/b], [b/a], [b/c], [c/b]$, lecz nie względem $[a/c]$, skoro $R(c,b)$ nie jest dyrektywą; zatem tak określona relacja równoznaczności nie jest przechodnia.

Dla uniknięcia tych kłopotów rozsądne wydaje się dopuszczenie dyrektyw wyprowadzalnych, a więc tez i wyprowadzalnych reguł wnioskowania. Wtedy we wspomnianym systemie rachunku zdań dwie formuły są równoznaczne w sensie podstawialności (bez zastrzeżenia istotności) wtedy i tylko wtedy, gdy są logicznie równoważne. W [JiZ] Ajdukiewicz odrzuca to rozwiązanie; nie wyjaśnia jednak jak wybrnąć z kłopotów wskazanych powyżej. Z drugiej strony, często przywołuje dyrektywy języka polskiego, które są po prostu uznanymi zdaniami lub poprawnymi wnioskowaniami, lecz nie wchodzą w skład żadnej ustalonej aksjomatyzacji tego języka; w tym przypadku nie mógłby postąpić inaczej, ponieważ nie jest znana żadna lista podstawowych dyrektyw znaczeniowych języka polskiego. Jeżeli zgodzimy się uwzględniać wszystkie wyprowadzalne dyrektywy języka polskiego, to rozwiążemy problem z wyrazami "bicz" i "bat" bez wprowadzania zastrzeżenia istotności. Ponieważ zdanie "Każdy bat jest biczem" jest dyrektywą znaczeniową, więc zdanie "Każdy bat jest bezwładny" wynika ze zdania "Każdy bicz jest ciałem i każde ciało jest bezwładne" na mocy wyprowadzalnej dyrektywy dedukcyjnej języka polskiego. Uznane dyrektywy znaczeniowe języka polskiego (razem z ich konsekwencjami) odpowiadają w systemie formalnym zbiorowi wszystkich tez i wyprowadzalnych reguł wnioskowania. Co więcej, system formalny jednoznacznie określa swoje tezy i reguły wyprowadzalne, podczas gdy aksjomatyzacja może przybierać różne formy. Uzależnienie znaczenia wyrażeń od konkretnej aksjomatyzacji wydaje się nienaturalne.

Utożsamienie równoznaczności formuł zdaniowych z logiczną równoważnością nie trywializuje koncepcji Ajdukiewicza, ponieważ dla innych wyrażeń, np. nazw, wyrażeń funkcyjnych, funktorów itp., prowadzi do interesującej relacji równoznaczności. Większość konkretnych przykładów, przytacznych przez Ajdukiewicza w cytowanych pracach, dotyczy wyrażeń nazwowych i funktorów;

niewiele dowiadujemy się o znaczeniu zdań (to różni istotnie prace Ajdukiewicza od prac inncyh autorów, zajmujących się pojęciem znaczenia, które koncentrują się zwykle na znaczeniu zdań). Postulat równoznaczności zdań logicznie równoważnych jest często akceptowany w teoriach znaczenia; np. jest on jednym z postulatów, na których opiera się tzw. argument procy (por. Maciaszek [ZPP]). Z drugiej strony niepokoi fakt, że z tego postulatu wynika równoznaczność dowolnych tez systemu formalnego.

Można wyeliminować tę konsekwencję na różne sposoby, na przykład uznając, iż dyrektywy znaczeniowe stanowią właściwy podzbiór ogółu dyrektyw języka. W przypadku systemu formalnego S, możemy rozważać jego podsystem S', który wyznacza relację równoznaczności formuł zdaniowych, subtelniejszą od relacji logicznej równoważności w S. Wtedy dowolne tezy systemu S' będą równoznaczne, ale nie dowolne tezy systemu S. W logikach niefregowskich występuje operator $\equiv$, interpretowany jako równoznaczność formuł; przyjmuje się, że $\varphi \equiv \psi$ pociąga $\varphi \Leftrightarrow \psi$. Dla takich logik φ jest równoznaczne z ψ w sensie podstawialności wtedy i tylko wtedy, gdy $\varphi \equiv \psi$ jest tezą, co wydaje się zgodne z koncepcją Ajdukiewicza.

3 Języki zamknięte.

W tym rozdziale zajmiemy się Ajdukiewiczowskim pojęciem języka zamkniętego, występującym w tezie skrajnego konwencjonalizmu. Przytoczymy oryginalny tekst Ajdukiewicza z pracy "Obraz świata i aparatura pojęciowa" [OSAP], gdzie wyjaśnienie tego pojęcia jest prostsze niż we wcześniejszej pracy [JiZ].

"Mówimy, że dwa wyrażenia są bezpośrednio znaczeniowo związane, gdy bądź 1) oba występują w jednym i tym samym zdaniu dyktowanym przez jakąś aksjomatyczną dyrektywę znaczeniową, bądź 2) oba są zawarte w jednej i tej samej parze zdań wymienionych przez jakąś dedukcyjną dyrektywę znaczeniową, bądź 3) oba wchodzą w skład jednego i tego samego zdania przyporządkowanego pewnej danej doświadczenia przez jakąś empiryczną dyrektywę znaczeniową. Język nazywamy spójnym, jeżeli nie można jego zapasu wyrażeń rozłożyć na dwie niepuste klasy tak, żeby żadne z wyrażeń pierwszej klasy nie było bezpośrednio znaczeniowo związane z jakimś wyrażeniem drugiej klasy."

"Następnie odróżniliśmy języki otwarte i zamknięte. Nazywamy język otwartym, jeśli istnieje inny język, zawierający wszystkie wyrażenia pierwszego i nadający im to samo znaczenie, które mają w pierwszym, w którym jednakże występują także wyrażenia, nie występujące w pierwszym języku ani co do postaci swej, ani co do znaczenia, przy czym z wyrażeń tych co najmniej jedno jest bezpośrednio znaczeniowo związane z jakimś wyrażeniem występującym także w pierwszym języku. Język, który nie jest otwarty, nazywa się językiem zamkniętym."

"Do języka otwartego można dołączyć nowe wyrażenia, nie będące synonimem żadnego już znajdującego się w tym języku wyrażenia, i związać je bezpośrednio znaczeniowo z jakimś takim wyrażeniem, przy czym znaczenie wyrażeń już w języku się znajdujących nie ulegnie zmianie. Natomiast języki zamknięte stają się niespójnymi, jeśli dołączyć do nich nowe wyrażenie nie będące synonimem wyrażenia już w języku się znajdującego." Cytujemy za [JiP1], str. 177.

W świetle powyższych uwag, język zamknięty to taki język S, do którego nie można dołączyć nowego wyrażenia, nierównoznacznego z żadnym wyrażeniem języka S bez utraty spójności i bez zmiany znaczeń wyrażeń języka S. Dalej Ajdukiewicz pisze: "Klasę wszystkich znaczeń przynależnym wyrażeniom należącym do pewnego zamkniętego i spójnego języka nazywamy aparaturą pojęciową. Dwie aparatury pojęciowe są więc albo identyczne, albo też nie mają wspólnych elementów." Cytujemy za [JiP1], str. 177.

Pojęcie języka zamkniętego w sensie Ajdukiewicza różni się zasadniczo od pojęcia języka semantycznie zamkniętego, którym posługiwał się Tarski. Według Tarskiego [SKP], język semantycznie zamknięty to język, który "poza swoimi wyrażeniami zawiera również ich nazwy, a także terminy semantyczne - takie jak termin *'prawdziwy'*, odnoszące się do zdań tego języka; założyliśmy też, że wszystkie zdania określające trafny sposób użycia tego terminu mogą być uznane w tym języku." (Cytujemy za [PLF], str. 242.) Innymi słowy, język semantycznie zamknięty zawiera swój metajęzyk, w którym występuje predykat prawdy, stosowalny do wszystkich zdań języka. Dla Ajdukiewicza język zamknięty to język całkowicie nasycony znaczeniowo, którego nie można wzbogacić o nowe znaczenia bez zmiany istniejących już znaczeń.

Ajdukiewicz nie podał żadnego przykładu języka zamkniętego; nie jest to zaskakujące, ponieważ zagadnienie konstrukcji, a nawet istnienia takich języków wydaje się trudne. Języki semantycznie zamknięte w sensie Tarskiego istnieją, ale zawierają sprzeczności, np. antynomię kłamcy. Jest to problem szeroko dyskutowany przez Tarskiego i późniejszych autorów; nie będziemy go tu przedstawiać.

Istnienie języków zamkniętych jest ważne dla tezy skrajnego konwencjonalizmu; jeżeli takie języki nie istnieją, to mamy do czynienia z tezą prawdziwą "w próżni", ponieważ stwierdza pewną własność wszystkich par elementów klasy pustej. Ajdukiewicz nie podtrzymywał tej tezy w późniejszych latach. Nie jest też ona przedmiotem moich dalszych rozważań, poświęconym zagadnieniu istnienia języków zamkniętych.

Rozwiązanie tego zagadnienia jest niełatwe także z tego powodu, że kluczowe pojęcia, np. języka, znaczenia wyrażeń, zachowywania znaczenia, mają rozmaite interpretacje; niektóre były wspomniane w poprzednim rozdziale. Ograniczymy się do języków formalnych, a dokładnie teorii elementarnych, czyli teorii opartych na rachunku predykatów z identycznością (równością). Przyjmiemy, że relacja równoznaczności jest określona jako podstawialność wyrażeń w dyrektywach wyprowadzalnych, rozumianych tu jako tezy i reguły wyprowadzalne teorii. Ponieważ w teoriach elementarnych obowiązuje twierdzenie o dedukcji, więc podstawialność w regułach wyprowadzalnych sprowadza się do podstawialności w tezach teorii (szczegóły omawiam poniżej).

Przyjmując taką interpretację teorii Ajdukiewicza, można wykazać, że języki zamknięte nie istnieją poza trywialnym przypadkiem teorii sprzecznych (czyli nie mających żadnego modelu) albo mających tylko modele jednoelementowe.

Niech T będzie ustaloną teorią elementarną. Musimy uściślić pojęcie równoznaczności wyrażeń teorii T, czyli wyrażeń języka elementarnego teorii T. Aby uniknąć zbędnych komplikacji, ograniczymy się do wyrażeń sensownych: symboli, termów i formuł (zdania to formuły bez zmiennych wolnych). Termy oznaczamy literami s, t, u, a formuły literami φ, ψ, χ. Podobnie jak w rozdziale 2, przez kontekst $\varphi[-]$ rozumiemy formułę φ, w której pewne wyrażenie sensowne zostało zastąpione specjalnym symbolem '-'; wtedy $\varphi[v]$ oznacza wynik podstawienia wyrażenia v w miejsce '-' w $\varphi[-]$. Mówimy, że kontekst $\varphi[-]$ jest odpowiedni dla wyrażenia v, jeżeli $\varphi[v]$ jest formułą (poprawnie zbudowaną). Symbolem $\vdash_T$ oznaczamy (syntaktyczną) relację konsekwencji teorii T. Zatem $\Gamma \vdash_T \varphi$ znaczy: φ jest konsekwencją zbioru (założeń) Γ na gruncie teorii T.

Niech v, w będą wyrażeniami sensownymi. Mówimy, że v jest równoznaczne z w w teorii T (piszemy $v \sim_T w$), jeżeli dla każdego zbioru formuł Γ i każdego kontekstu $\varphi[-]$ prawdziwa jest równoważność:

$$\text{(D)}\ \Gamma \vdash_T \varphi[v] \Leftrightarrow \Gamma \vdash_T \varphi[w].$$

Powyższa definicja równoznaczności jest pewną eksplikacją Definicji 1' z poprzedniego rozdziału. Jako natychmiastowy wniosek otrzymujemy warunek: jeżeli $v \sim_T w$, to dla każdego kontekstu $\varphi[-]$, $\varphi[v]$ jest formułą wtedy i tylko wtedy, gdy $\varphi[w]$ jest formułą. Załóżmy, że $v \sim_T w$ i $\varphi[v]$ jest formułą. Ponieważ $\varphi[v] \vdash_T \varphi[v]$, więc $\varphi[v] \vdash_T \varphi[w]$, a stąd $\varphi[w]$ jest formułą. Podobnie dowodzimy przeciwną implikację. Ponieważ struktury formuł i innych wyrażeń sensownych języka elementarnego są syntaktycznie jednoznaczne (wyrażenie sensowne ma dokładnie jedną konstrukcję), więc na mocy powyższego warunku wyrażenia równoznaczne muszą należeć do tej samej kategorii syntaktycznej, np. oba są termami różnymi od zmiennych, albo oba są formułami, symbolami relacyjnymi o tej samej liczbie argumentów, stałymi przedmiotowymi, zmiennymi itp. Gdybyśmy z góry przyjęli w definicji równoznaczności, że wyrażenia v i w należą do tej samej kategorii syntaktycznej, to moglibyśmy opuścić zbiór założeń Γ, tzn. ograniczyć się do podstawialności tych wyrażeń w twierdzeniach teorii T: dla każdego kontekstu $\varphi[-]$, $\vdash_T \varphi[v]$ wtedy i tylko wtedy, gdy $\vdash_T \varphi[w]$. Wtedy równoważność (D) wynika ze zwartości (finitystyczności) logiki elementarnej i twierdzenia o dedukcji.

Dla wyrażeń v, w należących do tej samej kategorii syntaktycznej otrzymujemy równoważność: $v \sim_T w$ wtedy i tylko wtedy, gdy dla każdego kontekstu $\varphi[-]$, odpowiedniego dla v, w,

$$\text{(D')}\ \vdash_T \varphi[v] \leftrightarrow \varphi[w].$$

Symbol $\leftrightarrow$ to znak równoważności w języku elementarnym, symbol $\Leftrightarrow$ to znak równoważności w metajęzyku; podobne oznaczenia stosujemy dla implikacji.

Implikacja $v \sim_T w \Rightarrow$(D') wynika z faktu, że $\varphi[v] \leftrightarrow \varphi[v]$ jest twierdzeniem teorii T (zastosuj (D) do kontekstu $\varphi[v] \leftrightarrow \varphi[-]$ przy pustym zbiorze Γ). Implikacja przeciwna (tzn. z prawdziwości (D') dla każdego odpowiedniego kontekstu $\varphi[-]$ wynika $v \sim_T w$) jest oczywista. Warunek (2) Definicji 1' z poprzedniego rozdziału odpowiada równoważności $\varphi[v] \vdash_T \psi \Leftrightarrow \varphi[w] \vdash_T \psi$ dla każdego kontekstu $\varphi[-]$ i każdej formuły ψ; jest on konsekwencją naszej definicji relacji $v \sim_T w$.

Jako wnioski ottrzymujemy następujące twierdzenia. Dla formuł φ, ψ, $\varphi \sim_T \psi$ wtedy i tylko wtedy, gdy $\vdash_T \varphi \leftrightarrow \psi$ (implikacja ($\Leftarrow$) wynika z tzw. twierdzenia o równoważności dla rachunku predykatów). Dla termów s, t nie będących zmiennymi, $s \sim_T t$ wtedy i tylko wtedy, gdy $\vdash_T s = t$ (implikacja ($\Leftarrow$) wynika z tzw. twierdzenia o równości dla rachunku predykatów z równością). Równoznaczne zmienne to identyczne zmienne; zmienna nie jest równoznaczna z żadnym termem nie będącym zmienną, skoro nie można zamienić x na niebędący zmienną term t w formule $\forall x \varphi$. Symbole relacyjne R, S przyjmujące n argumentów są równoznaczne wtedy i tylko wtedy, gdy dla różnych zmiennych $x_1, \ldots, x_n \vdash_T R(x_1, \ldots, x_n) \leftrightarrow S(x_1, \ldots, x_n)$.

Zgodnie z definicją Ajdukiewicza, wyrażenia v i w są bezpośrednio związane znaczeniowo, jeżeli oba występują w pewnej dyrektywie aksjomatycznej lub dedukcyjnej (dyrektywy empiryczne nie występują w systemach formalnych), czyli w pewnym aksjomacie lub pewnej regule dowodzenia teorii T. Ponieważ niektóre aksjomaty i reguły dowodzenia logiki elementarnej są schematami, w których występują dowolne formuły języka (np. reguła odrywania $\varphi \to \psi, \varphi/\psi$), więc dowolne wyrażenia sensowne są bezpośrednio związane znaczeniowo; dowolne dwa wyrażenia sensowne występują w pewnej formule języka, a ta formuła wchodzi w skład pewnej instancji dyrektywy. Każda teoria elementarna jest językiem spójnym w sensie Ajdukiewicza.

Niech T będzie teorią elementarną, która ma model, składający się z przynajmniej dwóch elementów. Do języka tej teorii dodajemy nową stałą przedmiotową c, otrzymując teorię T_c. W tej teorii nie przyjmujemy żadnych specjalnych aksjomatów ani reguł, dotyczących stałej c; stała c jednak występuje w schematach aksjomatów i reguł teorii T_c, np. aksjomatów i reguł logicznych. Twierdzenie o nowych stałych przedmiotowych głosi: dla dowolnej formuły φ języka teorii T, jeżeli $\vdash_{T_c} \varphi[x/c]$, to $\vdash_T \forall x \varphi$. $\varphi[x/c]$ oznacza wynik podstawienia c za wszystkie wolne wystąpienia zmiennej x w φ.

W szczególności, jeżeli x nie jest zmienną wolną w φ, to $\vdash_{T_c} \varphi$ implikuje $\vdash_T \varphi$; odwrotna implikacja jest oczywista. Zatem T_c jest zachowawczym rozszerzeniem teorii T. W konsekwencji, teoria T_c zachowuje relację równoznaczności wyznaczoną przez teorię T dla wyrażeń języka teorii T. Z drugiej strony, stała c nie jest równoznaczna w T_c z żadnym wyrażeniem języka teorii T. Przypuśćmy, że istnieje term t, nie zawierający c i taki, że $\vdash_{T_c} c = t$. Z twierdzenia o nowych stałych przedmiotowych wynika $\vdash_T \forall x x = t$, gdzie x jest zmienną nie wystepującą w t, a to znaczy, że teoria T jest sprzeczna lub ma tylko modele jednoelementowe, wbrew założeniu.

Wykazaliśmy, że do każdej teorii elementarnej T, poza trywialnymi przypadkami, można dodać nową stałą przedmiotową c, która nie jest równoznaczna z żadnym wyrażeniem teorii T, lecz dodanie tej stałej nie zmienia relacji równoznaczności dla wyrażeń teorii T.

Można wysunąć zarzut, że wprawdzie relacja równoznaczności w T_c pokrywa się z relacją równoznaczności w T dla wyrażeń języka teorii T, to jednak nie zachowuje znaczeń tych wyrażeń, rozumianych jako klasy abstrakcji relacji równoznaczności. Na przykład, pewne formuły języka teorii T są równoznaczne w T_c z formułami zawierającymi c; dowolne dwie tezy teorii T_c są równoznaczne, a są wśród nich tezy zawierające c i tezy nie zawierające c. Zatem pewne klasy abstrakcji relacji równoznaczności w T powiększają się w T_c o wyrażenia zawierające c.

Uważam, że zachowywanie relacji równoznaczności dla wyrażeń początkowej teorii jest właściwą interpretacją zachowywania znaczenia tych wyrażeń. Jak się zdaje, do żadnej teorii elementarnej nie można dodać nowego wyrażenia w taki sposób, żeby nie istniała formuła początkowej teorii równoważna w nowej teorii pewnej formule zawierającej dodane wyrażenie. Wymóg całkowitego zachowywania znaczeń jako klas abstrakcji relacji równoznaczności prowadziłby do wniosku, że każda teoria elementarna jest językiem zamkniętym; taka konkluzja nie wydaje się zgodna z intencjami Ajdukiewicza.

Nasuwa się wątpliwość, czy dodanie do języka nowej nazwy bez jakichkolwiek nowych dyrektyw, dotyczących tej nazwy, dobrze koresponduje z koncepcją Ajdukiewicza. Jest to czysto formalna, sztuczna procedura, która nie występuje w trakcie ewolucji języków etnicznych i systemów wiedzy naukowej. Podałem ten przykład ze względu na jego prostotę. Poniżej podam inny argument, oparty na twierdzeniu Tarskiego o niedefiniowalności prawdy, który jest bardziej wyrafinowany, lecz nie tak sztuczny.

Niech T będzie teorią niesprzeczną i dostatecznie bogatą, aby dla każdego zdania ψ istniał term (bez zmiennych) $\nu(\psi)$, interpretowany jako nazwa formuły ψ, dla której prawdziwy jest lemat Gödla o autoreferencji:

> Dla każdej formuły $\varphi(x)$, w której x jest jedyną zmienną wolną, istnieje zdanie ψ takie, że $\vdash_T \psi \leftrightarrow \varphi(\nu(\psi))$.

Zakładamy również, że $\vdash_T \neg(\nu(\psi) = \nu(\psi'))$ dla dowolnych różnych zdań ψ, ψ'. Te warunki spełniają teorie, zawierające skromny fragment arytmetyki liczb naturalnych. W takich językach rolę $\nu(\psi)$ odgrywa term (liczebnik), oznaczający numer Gödla zdania ψ.

Twierdzenia Tarskiego o niedefiniowalności prawdy głosi, iż predykat prawdy nie jest definiowalny w teorii T, tzn. nie istnieje formuła $Tr(x)$, w której x jest jedyną zmienną wolną i taka, że dla każdego zdania ψ:

$$\text{(T)}\ \vdash_T Tr(\nu(\psi)) \leftrightarrow \psi.$$

Przypomnimy prosty dowód. Przypuśćmy, że taka formuła $Tr(x)$ istnieje. Na mocy lematu o autoreferencji istnieje zdanie ψ takie, że:

$$\vdash_T \psi \leftrightarrow \neg Tr(\nu(\psi)).$$

Na mocy (T), mamy też:

$$\vdash_T \psi \leftrightarrow Tr(\nu(\psi)),$$

a więc $\vdash_T \psi \leftrightarrow \neg\psi$, a to jest formuła sprzeczna z tautologią $\neg(\psi \leftrightarrow \neg\psi)$. Zatem teoria T jest sprzeczna.

Możemy jednak dodać do teorii T nowy jednoargumentowy symbol relacyjny Tr oraz nowe aksjomaty $Tr(\nu(\psi)) \leftrightarrow \psi$ dla wszystkich zdań ψ języka teorii T. Oznaczmy tak otrzymaną teorię symbolem T'.

Teoria T' jest zachowawczym rozszerzeniem teorii T, tzn. dla każdej formuły φ nie zawierającej symbolu Tr, $\vdash_T \varphi$ wtedy i tylko wtedy, gdy $\vdash_{T'} \varphi$. Implikacja ($\Leftarrow$) jest oczywista. Implikacja ($\Rightarrow$) wynika ze zwartości teorii T' i faktu, że dla każdego skończonego zbioru zdań języka teorii T istnieje formuła Tr' teorii T, która spełnia schemat (T) dla zdań z tego zbioru.

W konsekwencji teoria T' nie zmienia relacji równoznaczności dla wyrażeń teorii T. Formuła $Tr(x)$ nie jest równoznaczna w T' z żadną formułą języka teorii T. Przypuśćmy, że istnieje formuła φ bez symbolu Tr taka, że $Tr(x) \sim_{T'} \varphi$. Pokazaliśmy powyżej, że $\vdash_{T'} Tr(x) \leftrightarrow \varphi$.

Formuła φ musi zawierać zmienną wolną x; w przeciwnym razie, z (T) i powyższej równoważności otrzymalibyśmy $\varphi \leftrightarrow \psi$ oraz $\varphi \leftrightarrow \neg\psi$ dla dowolnego zdania ψ (podstawiając $\nu(\psi)$, a potem $\nu(\neg\psi)$ za x w formule $Tr(x) \leftrightarrow \varphi$), co prowadzi do sprzeczności w teorii T'. Ewentualne inne zmienne wolne w φ możemy wyeliminować, podstawiając termy bez zmiennych w ostatniej równoważności. Od tej pory przyjmujemy, że φ zawiera jedyną zmienną wolną x i oznaczamy tę formułę przez $\varphi(x)$.

Ponieważ $\vdash_{T'} Tr(x) \leftrightarrow \varphi(x)$, więc podstawiając $\nu(\psi)$ za x otrzymujemy $\vdash_{T'} Tr(\nu(\psi)) \leftrightarrow \varphi(\nu(\psi))$ dla dowolnego zdania ψ języka teorii T. Wykorzystując równoważności z (T), będące aksjomatami teorii T', dostajemy $\vdash_{T'} \varphi(\nu(\psi)) \leftrightarrow \psi$ dla dowolnego zdania teorii T. Ponieważ T' jest zachowawczym rozszerzeniem teorii T, te równoważności są też twierdzeniami teorii T, a to znaczy, iż $\varphi(x)$ jest definicją predykatu prawdy w teorii T, wbrew twierdzeniu Tarskiego.

Powyższe rozważania istotnie opierają się na przyjetej tu eksplikacji koncepcji Ajdukiewicza. Można je przystosować do wersji, wspomnianej w rozdziale 2, w której relacja równoznaczności jest wyznaczona przez pewną podteorię teorii T, ograniczoną do dyrektyw znaczeniowych, jeśli ta podteoria zawiera wszystkie prawa i reguły rachunku predykatów z równością - dla argumentu ze stałą c, albo ponadto odpowiedni fragment arytmetyki - dla argumentu z predykatem prawdy.

Problem istnienia nietrywialnych języków zamkniętych przy innych eksplikacjach teorii Ajdukiewicza pozostaje otwarty. Głównym celem tego rozdziału było postawienie tego problemu, a podana tu próba rozwiązania miała pokazać jego związki z innymi zagadnieniami logiki oraz zademonstrować przykłady konkretnego stosowania tej teorii do języków formalnych.

Literatura

[OZW] K. Ajdukiewicz, O znaczeniu wyrażeń, *Księga Pamiątkowa Polskiego Towarzystwa Filozoficznego we Lwowie*, Lwów, 1931, str. 31–77.

[JiZ] K. Ajdukiewicz, Sprache und Sinn, *Erkenntinis*, IV, 1934, str. 100–138. Przekład na język polski pt. "Język i znaczenie" w [JiP1], str. 145–174.

[OSAP] K. Ajdukiewicz, Das Weltbild und die Begriffsapparatur, *Erkenntnis*, IV, 1934, str. 259–287. Przekład na język polski pt. "Obraz świata i aparatura pojęciowa" w [JiP1], str. 175–195.

[ZEKZ] K. Ajdukiewicz, Zagadnienie empiryzmu a koncepcja znaczenia, *Studia Filozoficzne* 1(36), 1964, str. 3–14.

[JiP1] K. Ajdukiewicz, *Język i Poznanie*, tom I, PWN, Warszawa, 1985.

[JiP2] K. Ajdukiewicz, *Język i Poznanie*, tom II, PWN, Warszawa, 1985.

[UATZ] W. Buszkowski, Uwagi o Ajdukiewiczowskiej teorii znaczenia, 2008, <http://sp-forum.fr.pl/archive>.

[PTS] N. Francez, R. Dyckhoff i G. Ben-Avi, Proof-Theoretic Semantics for Subsentential Phrases, *Studia Logica.* W druku.

[ZPP] J. Maciaszek, *Znaczenie, Prawda , Przekonania*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź, 2008.

[SKP] A. Tarski, The semantic conception of truth and the foundations of semantics, *Philosophy and Phenomenological Research* 4.3, 1944. Przekład na język polski pt. "Semantyczna koncepcja prawdy i podstawy semantyki" w: [PLF], str. 228–282.

[PLF] A. Tarski, Pisma logiczno-filozoficzne, Tom 1, Prawda, (J. Zygmunt, red.), Wydawnictwo Naukowe PWN, Warszawa, 1995.

[FSLW] J. Woleński, *Filozoficzna Szkoła Lwowsko-Warszawska*, PWN, Warszawa, 1985.

[ATZ] R. Wójcicki, *Ajdukiewicz. Teoria znaczenia*, Filozofia Polska XX Wieku, Prószyński i S-ka, Warszawa, 1999.

[PPR] K. Zamiara, The Problem of Psychologistic Reduction, *Poznań Studies in the Philosophy of Sciences and Humanities*, 4.1-4, 1978, str. 138–172.

[PBH] K. Zamiara, Psychologizm w badaniach humanistycznych, w: *Blausteina koncepcja odbioru mediów*, Z. Rosińska (red.), Filozofia polska XX wieku, Prószyński i S-ka, Warszawa 2001, str. 195–203.