



Tarského definice pojmu pravdy a její kritika

Jan Štěpánek, FF MU Brno

Abstrakt

Záměrem tohoto článku je popsat a prozkoumat legendární sémantickou koncepci pravdy Alfreda Tarského spolu s námitkami proti ní vznesenými. První část tohoto článku je rozdělena do pěti oddílů: kritéria kladená na každou adekvátní teorii pravdy jsou probírána v prvním oddílu; druhý je věnován výkladu Tarského Konvence T; distinkce mezi objektovým jazykem a metajazykem spolu s Tarského postojem k hovorovým a formálním jazykům popisuje třetí oddíl; čtvrtý oddíl je věnován Tarského definici pravdy za pomoci pojmu „splňování“; poslední oddíl zhodnocuje Tarského řešení paradoxů. Ve druhé části tohoto článku jsou předloženy vybrané námitky proti Tarského koncepci s důrazem na jejich různorodost.

Abstract

This paper aims to describe and examine Alfred Tarski's famous semantic conception of truth as well as some of the critiques presented against it. The first part of this paper is divided into five segments: criteria imposed upon every adequate definition of truth are discussed in the first segment; the second is dedicated to Tarski's Convention T; distinction between object language and metalanguage, as well as Tarski's attitude toward formalized and colloquial languages, is described in the third segment; fourth segment is devoted to Tarski's definition of truth using the notion of "satisfaction"; the last segment is the evaluation of Tarski's solution of paradoxes. In the second part of this paper the selected objections against Tarski's conception are presented with emphasis on their diversity.

Klíčová slova

Alfred Tarski, definice pravdy, pravda, sémantika, paradox lháře, metajazyk, objektový jazyk, T-schéma, konvence T, splňování

Keywords

Alfred Tarski, definition of truth, truth, semantics, liar paradox, metalanguage, object language, T-schema, convention T, satisfaction

doi: [10.5817/pf11-1-15](https://doi.org/10.5817/pf11-1-15)

Úvod

Většina filosofů by souhlasila, že dílo Alfreda Tarského týkající se pravdy bylo nesmírně vlivné. Na tomto místě však shoda obecně končí, neboť ústřední téma Tarského teorie, její aplikovatelnost na filosofické záležitosti a její samotná podstata jsou všechno zevrubně diskutovaná témata.¹

Alfred Tarski patří bezesporu mezi nejvlivnější logiky dvacátého století, a ačkoliv on sám se považoval spíše za matematika a logika než za filosofa, jeho dílo se stalo velmi inspirativním právě pro filosofy. Jednalo se pak především o jeho práce na poli sémantiky a teorie pravdy, které se dočkaly velkého ohlasu.² Mým záměrem v tomto článku je prozkoumat právě tu část Tarského díla, která je zaměřena na teorii pravdy.

Při svých zkoumáních budu vycházet ze dvou nejslavnějších Tarského statí věnovaných problematice pravdy. Tou první je „The Concept of Truth in Formalized Languages“, (dále jen [CTFL]) která poprvé vyšla roku 1933 v polštině, ale postupem času se dočkala německého a posléze i anglického překladu. Tato práce se však nijak výrazněji nerozšířila v povědomí filosofů. Tarski tento neúspěch přičítal jejímu příliš technickému charakteru. Proto se rozhodl svoje výsledky prezentovat ještě jednou, tentokrát ale ve „filosoficky přívětivější“ formě. Výsledkem jeho snažení se stala stať „The Semantic Conception of Truth: and the Foundations of Semantics“, (dále jen [SCT]) která vyšla poprvé roku 1944. Tarski se v tomto článku omezil na shrnutí nejdůležitějších principů své sémantické koncepce pravdy a vynechal specializované partie věnované konstrukci definice pravdy pro jazyky konečného nebo nekonečného řádu, kterým byla věnována podstatná část [CTFL]. Kromě toho je [SCT] obsahuje polemickou část, ve které se Tarski vypořádává s kritikami proti jeho koncepci vznesenými.

Moje studie bude mít po vzoru [SCT] rovněž dvě části – ta první bude věnována výkladu Tarského sémantické koncepce pravdy, druhá pak vybraným kritikám proti této koncepci namířeným. Ve výkladové části bude mým hlavním zdrojem především [SCT], přičemž na některých místech se budu snažit doplnit výklad o informace obsažené pouze v [CTFL]. Druhá a kratší část předkládané studie je věnována některým kritikám vzneseným proti sémantické koncepci pravdy, přičemž se jedná o kritiky, které jsou většinou novějšího data, než kritiky, se kterými se vyrovnává sám Tarski v [SCT].

Sémantická koncepce pravdy a její důsledky

1 Předpoklady adekvátní teorie pravdy

1.1 Tarského teoretické východisko

Termín „pravdivý“, tak jak je používán v každodenní konverzaci v hovorovém jazyku, má nejednoznačný význam. Už od antiky se filosofové, a nejen oni, pokoušejí uspokojivým způsobem vysvětlit, v čem vlastně pravda spočívá a jaký tedy vlastně je význam termínu „pravdivý“. Ale analýza námi v každodenní konverzaci používaného termínu „pravdivý“ není to, o co Tarskému jde: „Důkladná analýza významu termínu „pravdivý“ přítomného

¹ „Most philosophers would agree, that Alfred Tarski’s work on truth has been immensely influential. But that is where agreement generally ends, for the point of Tarski’s theory, its applicability to philosophical concerns, and its very nature are all deeply disputed.“ (Lynch, 2001, s. 323)

² Srovnej (Suppes, 1988, s. 80–91) a (Heck, 1997, s. 533–534).

v každodenním životě zde není zamýšlena. [...] Budu se zabývat výhradně uchopením intencí obsažených v takzvané *klasické* koncepci pravdy.“³

Z citovaného je evidentní, že Tarski si mezi různými teoriemi pravdy vybral teorii korespondenční, jejíž tradice vychází z Aristotela. Z toho lze vyvodit závazky, kterým bude muset Tarski při vyjádření svojí koncepce dostát. Formulaci Aristotelovy koncepce pravdy, ke které Tarski odkazuje, je možné najít ve čtvrté knize *Metafyziky*:

- (1) Omylem jest, jestliže řekneme, že jsoucí není anebo že nejsoucí jest; pravdou jest, řekneme-li, že jsoucí je a nejsoucí není.⁴

Třebaže Aristotelovu definici pravdy uvádí Tarski v obou svých hlavních spisech věnovaných pravdě, tak pouze v [SCT] ji srovnává s ostatními, modernějšími formulacemi korespondenční teorie pravdy:

- (2) Pravdivost věty spočívá v jejím souhlasu (či korespondenci) se skutečností.
(3) Věta je pravdivá, designuje-li existující stav věcí.⁵

Další moderní formulaci korespondenční teorie pravdy nabízí Tarski v [CTFL]:

- (4) [P]ravdivá věta je taková věta, která říká, že stav věcí je takový a takový, a stav věcí je vskutku takový a takový.⁶

V [SCT] Tarski uvádí, že Aristotelova definice je mnohem jasnější a přesnější než novodobé formulace, ale ani ona není bezchybná. Podrobnější analýzu zde uvedených formulací korespondenční teorie nabízí Donald Davidson v (Davidson, 1996), který také pokládá za nejlepší z uvedených tu Aristotelovu. Oba se však rozcházejí ohledně hodnocení obecnosti této formulace – Davidson totiž nesdílí Tarského názor, že Aristotelova definice není dostatečně obecná. Aristotelova definice je ostatně Davidsonem vyzdvižována jako velmi zdařilá, přičemž v její prospěch jsou uvedeny tři argumenty. Zaprvé: definice (3) a (4) odkazují ke stavům věcí [state of affairs], které jsou podle Davidsona v sémantice k ničemu. Zadruhé: Aristotelova formulace neobsahuje „podivné mezery“ jako je „takový a takový“ které se vyskytují v Tarského formulaci (4).⁷ Zatřetí: z Aristotelovy formulace zřetelně plyne, že pravdivost věty závisí na struktuře této věty. Osobně bych přidal ještě čtvrtý důvod – Aristotelova definice neobsahuje žádné problematické pojmy, jako jsou např. pojmy „designuje“ z definice (3) nebo „koresponduje“ z definice (2).

1.2 Aplikace termínu „pravdivý“ na věty

V této kapitole se zmíním o tom, na jaké entity je podle Tarského aplikovatelný termín „pravdivý“. Vhodné bude začít vysvětlením pojmu pravděnec, což je to, čemu je přisuzována pravdivostní hodnota (respektive to, co může pravdivostní hodnotu nabývat). V každodenním životě a hovorovém jazyku nemá termín „pravdivý“ jasně stanovené vymezení – ostatně např. v angličtině je možné termín „true“ použít i ve slovním spojení „true friend“, které má

³ „A thorough analysis of the meaning current in everyday life of the term ‘true’ is not intended here. [...] I shall be concerned exclusively with grasping the intentions which are contained in the so-called *classical* conception of truth.“ (Tarski, 1944, s. 153)

⁴ *Met.* G 1011b26.

⁵ Jak věta (2), tak věta (3) pocházejí z (Tarski, 2006, s. 138).

⁶ „[A] true sentence is one which says that state of affairs is so and so, and the state of affairs indeed is so and so.“ (Tarski, 1944, s. 155)

⁷ Viz (Davidson, 1996, s. 267).

s problematikou pramálo společného. Logikové a filosofové nám jednoznačnou odpověď na otázku, co lze jednoznačně označit za pravděnec, také neposkytnou – „Mezi kandidáty jsou přesvědčení, propozice, soudy, tvrzení, výroky, teorie, zmínky, ideje, akty myšlení, prohlášení, větné tokeny, větné typy, věty (nespecifikované) a řečové akty.“⁸

Vzhledem ke skutečnosti, že se Tarski snaží o vybudování teoretické sémantiky, kterážto je součástí logiky, mohl by někdo navrhnout, aby byl termín „pravdivý“ aplikován na propozice. Tuto možnost sice Tarski uvažuje, ale okamžitě ji zavrhuje, jelikož shledává pojem propozice příliš mnohoznačným a jeho význam kolísavým.

Pokud jde o termín „propozice“, je jeho význam neblaze proslulým tématem rozvláčných rozepří mezi mnoha filozofy a logiky a nezdá se, že by byl kdy vysvětlen jasně a jednoznačně. Z několika důvodů se jeví nejvhodnějším *aplikovat termín „pravdivý“ na věty*, a my se této praxe přidržíme. (Tarski, 2006, s. 137)

V době, kdy Tarski psal své dvě hlavní práce zabývající se pravdou, nebyl pojem propozice jednoznačně definován a Tarského odpor vůči němu je tak do značné míry oprávněný.

Tarského postoj vůči entitám hrajícím v jeho koncepci roli pravděnců shrnuje Peter Milne následovně: „[P]ro Tarského je pravda vypovídána [...] o větách, nikoli o obsazích vět, [...] ani o větách společně s jejich obsahy.“⁹ Je však také nutné vzít v úvahu, že Tarski se nebránil tomu, aby byla jeho koncepce pravdy uzpůsobena tak, aby v ní mohly hrát roli pravděnců i jiné entity, což explicitně vyjadřuje v [SCT]: „To, že nás zde v první řadě zajímá pojem pravdy pro věty, samozřejmě nevylučuje možnost jeho následovného rozšíření i na jiné druhy předmětů.“ (Tarski, 2006, s. 137)

Problematikou aplikovatelnosti termínu „pravdivý“ na věty se budu dále zabývat ve druhé části této práce, kde představím koncepci Jonathana Harrisona napadající Tarského aplikování termínu „pravdivý“ na věty.

1.3 Závislost pravdy na jazyku

Z Tarského rozhodnutí aplikovat pojem „pravdivý“ na věty vyvstávají jisté velmi závažné konsekvence, z nichž jednou z nejdůležitějších je nutnost relativizace tohoto pojmu vzhledem k jazyku. Nyní se pokusím dokázat, že není obtížné zdůvodnit požadavek na relativizaci pravdy vzhledem k jazyku. Uvažte následující příklad. Pokud je někde napsaná věta „horká čokoláda je tekutina“, pak tuto větu můžeme považovat za pravdivou větu v češtině, nepravdivou větu ve slovenštině a za nesmyslnou větu v prakticky jakémkoliv jiném jazyku. Tento jednoduchý příklad dokládá, že „jeden a týž výraz, jenž je pravdivou větou v jednom jazyce, může být nepravdivou nebo nesmyslnou větou v jazyce jiném“. (Tarski, 2006, s. 137)

Dále je nutné si uvědomit, že důležitou roli při připisování pravdivosti libovolné větě hraje temporální faktor, respektive diachronie jazyka. To si dobře uvědomují např. lingvisté zabývající tzv. diachronní lingvistikou. Ta se zabývá studiem jazyků s ohledem k jejich postupnému vývoji v průběhu času. Do jazyka se dostávají slova nová – tzv. computerové revoluci vděčíme za obohacení slovníku češtiny o slovo „harddisk“ –, význam již existujících slov se mění – příkladem budiž slovo „myš“ –, a jiná naopak z jazyka (běžně používané slovní zásoby) mizí – např. disketa. Kromě toho je zapotřebí brát v úvahu i genezi a vývoj celých jazyků.

⁸ „Among the candidates are beliefs, propositions, judgements, assertions, statements, theories, remarks, ideas, acts of thought, utterances, sentence tokens, sentence types, sentences (unspecified), and speech acts.“ (Kirkham, 1995, s. 54)

⁹ „[F]or Tarski truth is predicated, [...] of sentences, not of the contents of sentences nor [...] of sentences jointly with their contents.“ (Milne, 1997, s. 1–2)

Zatímco některé jazyky mizí a stávají se tak mrtvými jazyky – dokladem může být mykénština –, jiné jazyky zase vznikají – jako např. esperanto.

Logikové a filosofové si často vytvářejí vlastní „jazyky“ pouze za účelem dokázání vlastních tezí. Prakticky vždy jde buď o fragmenty jazyka, nebo o jazyky málo obsáhlé. Tyto jazyky jsou občas založeny na již existujících jazycích, ale občas jsou to jazyky diametrálně odlišné ode všech známých jazyků. Z tohoto si lze dovodit, že zohlednění diachronie jazyků není dostatečné. Je totiž myslitelné, že by mohly být i jiné jazyky, jejichž výrazy a jejich významy (stejně jako pravidla jejich syntaxe) by byly odlišné od všech diachronně uvažovaných jazyků. Je tak zapotřebí brát v úvahu nejenom reálně existující jazyky, ale zároveň s nimi všechny možné jazyky.

Tímto ovšem není vyčerpáno vše, co jsem chtěl o problematice relativizace k jazyku v této práci uvést. Nadále se tomuto tématu budu věnovat v souvislosti s T-schématem v kapitole 2.2.

1.4 Kritérium formální správnosti

Tarski klade na definici pravdy dva požadavky. Nyní se budu zabývat prvním z nich – požadavkem formální správnosti. Tarski sám se tomuto požadavku nikdy nevěnoval dostatečně podrobně. Toto tvrzení podporuje např. Milne:

Tarského druhou podmínkou kladenou na definice pravdy je, že musí být *formálně správné*, tuto podmínku občas vyjadřuje jako *metodologickou správnost*. Tato podmínka je poněkud méně jasně specifikována než podmínka materiální adekvátnosti, ale (kromě omezení toho, co může patřit do metajazyka) vyžaduje, aby definice odpovídaly standardním podmínkám kladeným na definice.¹⁰

Bohužel ani v sekundární literatuře není situace o mnoho lepší. Interpretátoři a komentátoři Tarského dávají ve svých pracích přednost spíše druhému požadavku – materiální adekvátnosti. Tarski sám prohlašuje o formální správnosti následující: „Musíme tedy specifikovat slova či pojmy, jichž chceme k definování pravdy použít, a také musíme předložit formální pravidla, kterým by měla definice vyhovovat. Řečeno obecněji, musíme popsat formální strukturu jazyka, ve kterém bude definice podána.“ (Tarski, 2006, s. 136–137)

Při konstrukci své definice nehodlá Tarski zavrhnout žádné běžně akceptované logické zákony a principy. Jedním z logických principů, který hraje velmi důležitou roli v Tarského koncepci je princip bivalence – podle něho má věta vždy jednu z pravdivostních hodnot „pravda“, nebo „nepravda“. Tarského přijetí principu bivalence bývá kritizováno, jak bude ukázáno v kapitole 10, věnované Kripkeho kritice Tarského pravdivostního predikátu.

Cílem kritéria formální správnosti tedy je, aby z Tarského definice nebylo možné odvodit nepravdivá tvrzení (což úzce souvisí s požadavkem materiální adekvátnosti), popř. nějaký paradox. Tarski se zabýval především možností vytvoření paradoxu lháře.¹¹ Paradox plodící větu lze formulovat např. následovně:

(LP) Věta (LP) je nepravdivá.

¹⁰ „Tarski’s second constraint on definitions of truth is that they must be *formally correct*, a condition he sometimes renders as *methodological correctness*. This is somewhat less clearly specified than the material adequacy condition but, in addition to limits on what can belong to the metalanguage, requires that the definitions conform to standard constraints imposed on definitions.“ (Milne, 1997, s. 4)

¹¹ V částech této práce věnovaných variantám paradoxu lháře budu přebírat terminologii, kterou při zkoumání paradoxů používá dr. Raclavský.

Jak jsem uvedl dříve v této kapitole, pro větu (LP) musí platit normální logická pravidla. U jiných vět by to samozřejmě nepředstavovalo žádný problém, ale v případě věty (LP) dochází ke vzniku paradoxu, který lze schematicky znázornit takto:

- (1) (LP) je věta
- (2) Každá věta je buď pravdivá, nebo nepravdivá, ale nikdy nemůže být pravdivá a nepravdivá zároveň
- (3) (LP) je tedy buď pravdivá, nebo nepravdivá, ale nikoli obojí zároveň
- (4) Je-li (LP) pravdivá, pak je pravdivé to, co říká; (LP) je tedy nepravdivá
- (5) Je-li (LP) nepravdivá, pak není to, co říká, pravda; (LP) je tedy pravdivá
- (6) (LP) je tedy pravdivá i nepravdivá zároveň – podle (4) a (5)
- (7) (6) je ale zjevně v rozporu s (2) a (3) a tedy vzniká paradox¹²

Zde je jasně vidět, že paradoxní situace spočívá v současné pravdivosti a nepravdivosti věty (LP). Pokud by se mělo zkoumání vzniku paradoxu lháře a možností jeho vytvoření omezit pouze na předpoklady obsažené v tomto schématu, pak by mohl leckdo nabýt dojmu, že vzniku paradoxu by bylo možné zabránit odmítnutím nebo pozměněním některého z nich. S tímto by však Tarski nesouhlasil, neboť v tomto schématu nejsou zohledněny všechny možné předpoklady, které mohou mít za následek možnost vzniku paradoxu lháře. Těmto předpokladům bude věnována kapitola 3.2 a kapitoly následující. Tam bude ukázán Tarského přístup k řešení paradoxu lháře. Toto řešení však bývá vystavováno kritice založené na tzv. mstivém lháři. Stejně tak bývá Tarského práce napadána pomocí Jourdainova a Buridanova paradoxu, což jsou taktéž speciální varianty paradoxu lháře. Obě tyto kritiky budou objasněny v kapitolách 6.1 a 6.2.

1.5 Kritérium materiální adekvátnosti

Zde je třeba navázat na to, co bylo o Tarském a jeho přístupu k problému pravdivosti uvedeno již dříve. Především je třeba znovu uvést, že Tarskému nešlo o analýzu termínu hovorového jazyka „pravdivý“ používaného v každodenní konverzaci. Dále je zapotřebí vzít v úvahu, že Tarski chtěl, aby jeho definice pravdy „naplňovala intuice, jež se pojí s *klasickým aristotelským pojetím pravdy*“ (Tarski, 2006, s. 138), o kterém jsem již pojednal v první kapitole. Přesnější vymezení Tarského cíle a metody lze najít u Soamese:

[P]odle Tarského a mnoha ostatních je náš běžný pojem pravdy defektní právě proto, že jeho neomezenost dává vzniknout paradoxům. Analýza pravdy, kterou se Tarski snaží poskytnout, je taková, která by eliminovala tento údajný defekt, zatímco by zachovala hlavní a užitečné prvky tohoto pojmu. Ve skutečnosti se Tarski nesnažil specifikovat, jak termín „pravdivý“ skutečně chápeme, ale jak by měl být pochopen, pokud má fungovat v našich logických, matematických a vědeckých teoriích způsobem normálně zamýšleným.

Analýzy tohoto druhu jsou často zvány „explicace“.¹³

¹² Srovnej (Kirkham, 1995, s. 272–273).

¹³ „[A]ccording to Tarski and many others, our ordinary notion of truth is defective precisely because its unrestrictedness gives rise to paradox. The sort of analysis of truth that Tarski attempted to provide is one that would eliminate this alleged defect while preserving the central and useful features of the notion. In effect, Tarski tried to specify not how ‘true’ is actually understood but how it ought to be understood if it is to function in our logical, mathematical, and scientific theories in the ways normally intended. Analyses of this kind are often called “explications”.“ (Soames, 1999, s. 99)

Tato Tarského snaha by šla vyjádřit za pomoci článku Rudolfa Carnapa O explikaci následovně: „Řekneme-li, že daný předvědecký pojem [explikandum] je třeba přetransformovat do exaktního pojmu, znamená to ovšem, že je nutno zavést přesný pojem [explikát], který odpovídá danému pojmu.“ (Carnap, 1968, s. 170) Roli explikanda nemůže v Tarského koncepci hrát termín „pravdivý“ převzatý z hovorového jazyka. To by totiž bylo zcela nevyhovující zadání explikanda. Přílišná vágnost zadání explikanda by totiž klidně mohla mít za následek, že by se explikát minulo účinkem, jelikož by bylo explikováno něco úplně jiného, než o co Tarskému šlo.

Na počátku obou prací se Tarski přihlásil ke korespondenční teorii pravdy a intuicím s ní spojovaným. Problém pravdivosti tak dostává docela zřetelné obrysy a je jasné, co má být náplní zbytku obou Tarského zmiňovaných prací.

Cílem Tarského snažení bylo podání adekvátní explikace korespondenční teorie pravdy, přičemž jejím nejjasnějším příkladem byla klasická definice. Tato Aristotelova formulace jasně určuje dva případy, kdy je to, co říkáme, pravdivé – totiž pokud o tom, co je, prohlásíme, že to je; a pokud o tom, co není, řekneme, že to není. Zdá se, že Aristotelés by souhlasil s tvrzením, že věta „sníh je bílý“ může být pravdivá pouze tehdy, když sníh je bílý – tedy pouze tehdy, když platí následující ekvivalence:

- (5) Věta „sníh je bílý“ je pravdivá tehdy a jen tehdy, když sníh je bílý. (Tarski, 2006, s. 139)

Tato ekvivalence musí podle Tarského být důsledkem definice pravdy, protože všechny pravdivé věty musí být z definice pravdy odvoditelné.

2 Konvence T

2.1 Konvence *T* a *T*-schéma

K tomu, abychom věděli, jaké všechny věty by měly plynout z definice pravdy je zapotřebí zobecnit větu (5) z předchozí kapitoly, což lze provést poměrně jednoduše. Stačí si povšimnout jednoho velmi důležitého detailu. V ekvivalenci (5) se vyskytuje výraz „sníh je bílý“ dvakrát. Zatímco na pravé straně se jedná o výraz samotný, na levé straně věty se vyskytuje jeho jméno.

Výskyt jména výrazu ve výše uvedené ekvivalenci (5) má zcela klíčový význam. Tarski o výskytu jmen v ekvivalencích jako je (5) prohlásil: „chceme-li o nějaké větě např. říci, že je pravdivá, musíme použít jméno této věty, a nikoli tuto větu samotnou.“ (Tarski, 2006, s. 139) Je otázkou, nakolik výstižná je tato formulace, jelikož at' už se mluví o nějaké větě nebo o nějakém objektu, vždy je automaticky používáno jméno této věty nebo objektu. Např. ve výpovědi o větě, kterou kdosi vyslovil, se nemůže vyskytovat ona věta sama, ale pouze její jméno – podobně jako se v promluvě o Slunci nemůže objevit samo Slunce. Samotnému tvoření jmen vět a jejich distinkci je věnována následující kapitola.

Nyní je zapotřebí vytvořit obecné schéma, ze kterého by byla výše uvedená ekvivalence odvoditelná. Toto schéma vznikne, když za větu, o které chceme prohlásit, že je pravdivá, dosadíme písmeno „*p*“ a za jméno této věty dosadíme písmeno „*X*“. Výsledkem zobecnění je toto schéma:

- (T) *X* je pravdivá tehdy a jen tehdy, když *p*. (Tarski, 2006, s. 140)

Pro toto schéma¹⁴ se v literatuře zažilo označení T-schéma. Každou ekvivalenci vzniklou dosazením do T-schématu je zvykem označovat jako ekvivalenci tvaru (T), potažmo jako T-větu. Roli T-schématu v rámci Tarského koncepce lze lépe pochopit po seznámení se s obsahem tzv. konvence T:

KONVENCE T. *Formálně správná definice symbolu ‚Tr‘ formulovaná v metajazyku bude nazývána adekvátní definicí pravdy, pokud má následující důsledky:*

- (α) *všechny věty získané z výrazu ‚ $x \in Tr$ tehdy a jen tehdy, když p ‘ nahrazením symbolu ‚ x ‘ strukturálně-deskriptivním jménem jakékoliv věty zkoumaného jazyka a nahrazením symbolu ‚ p ‘ výrazem tvořícím překlad této věty do metajazyka;*
 (β) *věty ‚pro jakékoli x , jestliže $x \in Tr$, pak $x \in S$ ‘ (jinými slovy ‚ $Tr \subseteq S$ ‘).¹⁵*

Klíčovou roli hraje především důsledek (α), podle kterého musí mít každá adekvátní definice pravdy za následek všechny ekvivalence tvaru (T) získané dosazením vět zkoumaného jazyka do T-schématu. Zde je zapotřebí si uvědomit, že toto je požadavek, který musí splnit každá teorie pravdy, ale rozhodně se zde nejedná o teorii pravdy samu. T-schéma nemůže být teorií pravdy právě proto, že jde jen o větné schéma.

Ačkoliv T-schéma není definicí pravdy, mohl by padnout návrh, že definicí pravdy bude konjunkce všech ekvivalencí tvaru (T), které je možné vytvořit pro daný zkoumaný jazyk. Jak dodává Tarski, každou jednotlivou ekvivalenci tvaru (T) „lze považovat za dílčí definici pravdy, jenž vysvětluje, v čem spočívá pravdivost této jedné jednotlivé věty“. (Tarski, 2006, s. 140) Konjunkce všech ekvivalencí tvaru (T) by opravdu byla definicí pravdy, nicméně toto řešení naráží na jisté metodologické problémy. Pokud by zkoumaný připouštěl konstrukci pouze omezeného množství vět, nepředstavovala by tvorba definice pro tento jazyk výraznější problémy. Ty by však nastaly v případě, pokud by bylo možné ve zkoumaném jazyku konstruovat nekonečné množství vět. Definice pravdy pro takový jazyk by pak musela být konjunkcí nekonečného množství vět. Tarského přístup k tvorbě definice pravdy je proto odlišný a bude představen v kapitole 4.

2.2 Rozšíření T-schématu

V předchozí kapitole jsem představil T-schéma tak, jak ho ve svých pracích používá Tarski. Jeho verze je však poněkud „chudá“, jelikož obsahuje jako pravdivostní predikát pouze „pravdivý“, popř. „Tr“. To je poněkud zvláštní, jelikož Tarski explicitně uvádí, že termín „pravdivý“ musí být vždy relativizován vzhledem k jazyku. A nejjednodušším způsobem, jak dát tuto relativizaci explicitně najevo, se jeví její zavedení do T-schématu, což však Tarski sám neprovedl. Šťastnou se v této souvislosti jeví skutečnost, že na relativizaci pravdivostního predikátu v T-schématu nezapomněli Tarského komentátoři.

Používat v T-schématu pravdivostní predikát s explicitním zohledněním jazyka, na který je tento predikát aplikovatelný, je v literatuře běžné. Příkladem T-schématu, ve kterém je závislost na jazyku zohledněna můžeme najít např. u Soamese:

¹⁴ Výraz (T) není větou, ale větným schématem, věta se z něj stává až dosazením (metajazykového překladu) libovolné věty za p a jejího jména za X .

¹⁵ „CONVENTION T. *A formally correct definition of the symbol ‘Tr’, formulated in the metalanguage, will be called an adequate definition of truth if it has the following consequences:*

- (α) *all sentences which are obtained from the expression ‘ $x \in Tr$ if and only if p ’ by substituting for the symbol ‘ x ’ a structural-descriptive name of any sentence of the language in question and for the symbol ‘ p ’ the expression which forms the translation of this sentence into the metalanguage;*
 (β) *the sentence ‘for any x , if $x \in Tr$ then $x \in S$ ’ (in other words ‘ $Tr \subseteq S$ ’).“ (Tarski, 1944, s. 187–188)*

Schéma T. X je T v L tehdy a jen tehdy, když P.¹⁶

Prvky tohoto schématu jsou následující: X je jméno věty, zatímco P je jejím překladem do metajazyka, T je pravdivostní predikát a L je jazyk, na jehož věty je pravdivostní predikát aplikovatelný.¹⁷ V literatuře lze najít i mnohé jiné varianty notace relativizovaného pravdivostního predikátu, jako např.: „pravdivý-v-L“, „pravdivý_L“, „Tr_L“ a podobně.

Jelikož Tarského zajímala pouze možnost konstrukce definice pravdy pro formalizované jazyky, mohla by se jevit relativizace pravdivostního predikátu vzhledem k jazyku dostatečnou. Mnozí autoři však počítají i s možností aplikace Tarského koncepce na hovorové jazyky, přičemž relativizace pravdivostního predikátu pouze k jazyku se nezdá být pro jejich účely dostatečná. Proto bývá pravdivostní predikát vztahován i k jiným hlediskům.

Boisvert ve své kritice Harrisona používá rozšířenou verzi T-schématu, aby mohl Tarského koncepci pravdy aplikovat na příklady obsahující věty hovorového jazyka:

TN. S je pravdivá_{u,t} v L tehdy a jen tehdy, když p .¹⁸

V tomto schématu představuje S jméno věty, p její metajazykový překlad a L objektový jazyk, pro který je definována teorie pravdy. Oborem parametru u jsou mluvčí a oborem parametru t jsou momenty v čase. Soames se zmiňuje o koncepci, v níž je pravda relativizována vzhledem k okolnostem, které mohou být chápány jako částečné stavy světa, nebo jako možné světy. Pravdivostní predikát v ní má podobu „pravdivý v C “.¹⁹ Tímto se ale již příliš vzdalují od samotné Tarského koncepce, pročež radši přejdu k dalšímu tématu.

2.3 Dva druhy jmen vět

V předchozí kapitole jsem uvedl, že v T-schématu se vyskytuje jednak věta a pak i její jméno. Tuto kapitolu tedy věnuji problematice vytvoření jména věty. V [CTFL]²⁰ Tarski uvádí a definuje dva různé způsoby, jak vytvořit jméno věty (resp. libovolného výrazu), které lze posléze dosadit do T-schématu.

Prvním druhem jsou citační jména.²¹ Vytvoření jména věty proběhne jednoduše uzavřením dané věty do uvozovek – jméno se tedy bude skládat pouze z věty samotné a uvozovek na obou jejích stranách.

S citačními jmény může být zacházeno jako s jednotlivými slovy jazyka, tudíž jako se syntakticky jednoduchými výrazy. Jednotlivé konstituenty těchto jmen – uvozovky a výrazy stojící mezi nimi – plní stejnou funkci jako písmena a komplexy následujících písmen v jediném slově. [...] Každé citační jméno je tedy konstantním individuálním jménem určitého výrazu (výrazu uzavřeného do uvozovek) a ve skutečnosti jménem stejné povahy, jako je vlastní jméno člověka.²²

¹⁶ „Schema T. X is T in L iff s.“ (Soames, 1999, s. 68)

¹⁷ L je tedy jazyk, pro který je definice pravdy konstruována.

¹⁸ „TN. S is true_{u,t} in L if and only if s.“ (Boisvert, 1999, s. 378)

¹⁹ Srovnej (Soames, 1999, s. 137–138).

²⁰ Viz (Tarski, 1944, s. 156–158).

²¹ Tarského termín „quotation-mark names“ pro účely této práce překládám jako „citační jména“.

²² „Quotation-mark names may be treated like single words of a language, and thus like syntactically simple expressions. The single constituents of these names – the quotation marks and the expressions standing between them – fulfil the same function as the letters and complexes of successive letters in single words. [...] Every quotation-mark name is then a constant individual name of a definite expression (the expression enclosed by the quotation marks) and in fact a name of the same nature as the proper name of a man.“ (Tarski, 1944, s. 159)

Citační jména jsou jednoduchá právě v tom smyslu, že každé slouží jako individuální jméno daného výrazu. Vnitřní struktura citačního jména je sice složená, zvláště jedná-li se o jméno věty, ale vždy je s ním zacházeno jako s jediným slovem/jménem. Tarského přirovnání k vlastním jménům lidí je velmi trefné – ostatně jméno „Bertrand Arthur William, 3. hrabě Russell“ má rozhodně složenou vnitřní strukturu, ale musíme s ním zacházet jako s jediným slovem/jménem určité osoby.

V praxi by použití citačních jmen při dosazení do T-schématu vypadalo následovně:

- (5) Věta „sníh je bílý“ je pravdivá tehdy a jen tehdy, když sníh je bílý. (Tarski, 2006, s. 139)

Dalším druhem jmen vět jsou strukturálně-deskriptivní jména. Tato jména necitují výraz jimi denotovaný, ale jistým způsobem ho popisují. K vytvoření strukturálně-deskriptivního jména věty je zapotřebí mít jména pro všechna písmena jazyka, do kterého daná věta náleží, a také jména pro všechny ostatní v něm použité symboly (interpunkční znaménka, matematické a logické symboly a podobně) – tato jména ale sama nemohou být citačními jmény. Samotné vytvoření lze připodobnit ke skládání puzzle – vezmeme jména pro všechny symboly ve větě se vyskytující a poskládáme je tak, aby uspořádání jmen symbolů odpovídalo uspořádání symbolů jimi denotovaných. Při dosazení strukturálně-deskriptivních jmen do T-schématu dostaneme následující ekvivalenci tvaru (T):

- (6) Věta skládající se ze tří slov, z nichž první má čtyři písmena – Es, En, Dlouhé-I a Há –, druhé má dvě písmena – Jé a É – a třetí ze čtyř písmen – Bé, Dlouhé-I, El a Dlouhé-Y, je pravdivá tehdy a jen tehdy, když sníh je bílý.

Oba druhy pojmenování v této kapitole probírané je samozřejmě možné mezi sebou zaměňovat – citační jména je možné nahradit jmény strukturálně-deskriptivními a naopak.

3 Tarského dělení jazyků

3.1 Jazyky se specifikovanou strukturou, formalizované jazyky

V této kapitole se budu zabývat jazyky, pro které chtěl Tarski podat definici pravdivosti. Tarski explicitně uvádí: „Problém definice pravdy nabývá přesného významu a může být rigorózně vyřešen jen pro jazyky, jejichž struktura byla exaktně specifikována.“ (Tarski, 2006, s. 143) Bude tedy záhodno definovat jazyky se specifikovanou strukturou, což bude nejjednodušeji proveditelné tak, že uvedu kritéria, kterým musí tyto jazyky vyhovovat:

- i. Třída všech slov a všech výrazů považovaných v daném jazyku za smysluplné musí být jednoznačně popsána
- ii. Musí být uvedeny všechny primitivní (nedefinované) termíny a zároveň musí být dána pravidla definice umožňující zavádění definovaných termínů do jazyka
- iii. Musí být stanovena kritéria, podle kterých lze z třídy výrazů daného jazyka vyčlenit věty a následně musí být stanovena pravidla určující podmínky, za nichž je možné větu tvrdit
- iv. Aby bylo možné odvozovat nové věty, je zapotřebí určit axiomy (primitivní věty) daného jazyka a pravidla vyplývání (pravidla inference), pomocí kterých

z axiomů odvodíme nové věty, které budou nazývány teorémy (tedy věty dokazatelné v daném jazyku)²³

Jedinými jazyky se specifikovanou strukturou byly v době, kdy psal Tarski svoje články, formalizované jazyky různých systémů deduktivní logiky, jejichž účelem bylo zkoumání deduktivních věd. Tarského cílem bylo podat definici pojmu „pravdivý“ právě pro formalizované jazyky, jelikož právě ony se hodí pro účely matematiky a přírodních věd.

Formalizované jazyky jsou všechny takové jazyky, v nichž mohou být tvrzeny pouze teorémy – tedy pouze ty věty, které jsme schopni odvodit za pomoci usuzovacích pravidel z axiomů daného jazyka. V takovém jazyku je význam každého výrazu jednoznačně dán jeho formou. Podobně, jako kritéria pro jazyky se specifikovanou strukturou uvádí Tarski na počátku §2 [CTFL]²⁴ kritéria pro formalizované jazyky, avšak vzhledem k jejich veliké vzájemné podobnosti je zde již nebudu uvádět.

Je vhodné dodat, že formalizované jazyky a jazyky se specifikovanou strukturou nejsou vzájemně zaměnitelné termíny. Jak totiž Tarski uvádí, lze si „představit jazyky, jež by měly exaktně specifikovanou struktur, aniž by byly formalizované“. (Tarski, 2006, s. 143) V těchto jazycích není tvrditelnost vět závislá výlučně na jejich formální stránce, ale může být závislá i na mimojazykových faktorech. Takové jazyky, pokud by byly vytvořeny, by podle Tarského teoreticky mohly převzít roli hovorových jazyků ve vědeckém diskursu.

Na tomto místě se vnučuje otázka, jestli existuje možnost vytvořit takovou definici i pro hovorové jazyky a jestli Tarski o této možnosti uvažoval. Tarski se samozřejmě vyjádřil i k této otázce, já se jí budu zabývat v následujících kapitolách.

3.2 Sémanticky uzavřené jazyky

Než se pokusím pojednat o Tarského postoji k hovorovým jazykům, připravím si teoretické základy pojednáním o jazycích, které Tarski ve svých zkoumáních zavrhuje jakožto nekonzistentní. Ostatně i Tarski se k problematice hovorových jazyků a možnosti vytvoření definice termínu „pravdivý“ pro tyto jazyky dostává jaksi oklikou přes takzvané sémanticky uzavřené jazyky.

Sémanticky uzavřený jazyk je dle Tarského takový jazyk, který vyhovuje prvním dvěma z následujících předpokladů:

- a) Tento jazyk obsahuje kromě vlastních výrazů také jména těchto výrazů (nebo jiný způsob umožňující mluvit o větách tohoto jazyka) a sémantické termíny odkazující k větám tohoto jazyka²⁵
- b) V tomto jazyku je možné tvrdit všechny věty, vymezující adekvátní použití těchto sémantických termínů²⁶
- c) V tomto jazyku platí obvyklé logické zákony

V jazyku vyhovujícím všem těmto předpokladům je možné zformulovat paradox lháře, jak to bylo ukázáno v kapitole 1.4. Nyní se znovu zaměřím na paradox plodící větu (LP) a následně poukážu, jak výše zmíněné předpoklady vedou ke vzniku paradoxu:

(LP) Věta (LP) je nepravdivá.

²³ Volně podle (Tarski, 2006, s. 142–143).

²⁴ Viz (Tarski, 1944, s. 166).

²⁵ Srovnej (Soames, 1999, s. 67).

²⁶ Srovnej (Tarski, 2006, s. 145).

Podle předpokladu (a) mají věty ve zkoumaném jazyku možnost odkazovat samy k sobě a navíc mohou obsahovat na sebe aplikovatelné sémantické predikáty. Věta (LP) této podmínce nepochybně vyhovuje. Obsahuje totiž svoje vlastní jméno, díky kterému odkazuje sama k sobě. Dále věta (LP) obsahuje sémantický termín „nepravdivý“, který je aplikován na větu (LP) samu, čímž naplňuje i druhou polovinu předpokladu (a). Podle předpokladu (b) musí být věta (LP) příkladem adekvátního použití termínu „nepravdivý“. Tvrzení, že nějaká věta je nepravdivá, jsou v jazyku běžná a (LP) tedy vyhovuje i tomuto předpokladu. A konečně vznik paradoxu je těsně vázán na bod (c), jelikož kdyby neplatily obvyklé logické zákony, jako je zákon vyloučeného třetího, ke vzniku paradoxu by nemuselo dojít.

Podle Tarského není akceptovatelné, aby byl kvůli odstranění možnosti tvorby paradoxu odmítnut předpoklad (c). Aby mohl být zkoumaný jazyk konzistentní, pak podle Tarského nesmí obsahovat žádné prostředky, kterými by bylo možné odkazovat k jeho vlastním větám. Tarski tímto požadavkem z daného jazyka odstraňuje všechny autoreferenční věty, což mu bude také vytýkáno. Ve druhé části této práce představím kritiku tohoto postupu pocházející od Roberta L. Martina, podle které je odmítnutí všech autoreferenčních vět krokem špatným směrem. Dále je nepřijatelné, aby zkoumaný jazyk obsahoval sémantické termíny aplikovatelné na výrazy tohoto jazyka, což je důsledkem Tarského teorému. Podle něj je nemožné, aby byla definice pravdy vytvořena pouze s použitím pojmů objektového jazyka. Tomuto požadavku budu věnovat pozornost v kapitole 3.4.

3.3 *Status hovorových jazyků*

V minulé kapitole jsem se zabýval sémanticky uzavřenými jazyky, tedy jazyky, pro které je nemožné zkonstruovat definici pravdy. Nyní se budu zabývat tím, co Tarski napsal o hovorových jazycích.

V předchozí kapitole jsem uvedl, že každý sémanticky uzavřený jazyk musí vyhovovat třem podmínkám – takový jazyk pak Tarski považuje za nekonzistentní. V souladu s podmínkami (a) a (b) je zapotřebí zmínit jeden důležitý rys hovorového jazyka – jeho univerzálnost.

můžeme tvrdit, že ‚pokud o něčem můžeme smysluplně mluvit, pak o tom můžeme mluvit v hovorovém jazyce‘. Pokud má být zachována univerzálnost každodenního jazyka, [...] musíme v něm, abychom byli konzistentní, připustit kromě jeho vět a dalších výrazů také jména těchto vět a výrazů, věty obsahující tato jména, stejně tak i sémantické výrazy jako ‚pravdivá věta‘, ‚jmenovat se‘, ‚denotovat‘ atd.²⁷

Úkolem hovorového jazyka je tak podle Tarského mluvit o všem, o čem se mluvit dá. Jelikož hovorový jazyk je primárním prostředkem komunikace mezi lidmi, musí být dostatečně expresivní, aby mohl svoji roli dobře plnit. To však znamená, že takový jazyk bude muset (aby si zachoval svoji univerzálnost) mít i ty vlastnosti, které má sémanticky uzavřený jazyk. Z výše uvedeného citátu si lze odvodit, že hovorový jazyk vyhovuje podmínkám (a) a (b) charakterizujícím sémanticky uzavřený jazyk. Lze tedy očekávat, že ho Tarski označí jako nekonzistentní – on však poněkud překvapivě konstatuje:

Náš každodenní jazyk jistě není jazykem s exaktně specifikovanou strukturou. [...] Z tohoto důvodu nemá problém konzistence ve vztahu k tomuto jazyku přesný význam. [...] jazyk, jehož struktura by byla exaktně specifikována a jenž by se našemu

²⁷ ‚it could be claimed that ‘if we can speak meaningfully about anything at all, we can also speak about it in colloquial language’. If we are to maintain this universality of everyday language [...] we must, to be consistent, admit into the language, in addition to its sentences and other expressions, also the names of these sentences and expressions, and sentences containing these names, as well as such semantic expressions as ‘true sentence’, ‘name’, ‘denote’, etc.“ (Tarski, 1944, s. 164)

každodennímu jazyku podobal do takové míry, do jaké by to jen bylo možné, by byl nekonzistentní. (Tarski, 2006, s. 146)

I přes toto tvrzení je poměrně běžně uznávaným faktem, že Tarski považoval hovorové jazyky za nekonzistentní. Jak je ale vidět, toto tvrzení není úplně oprávněné.

3.4 Objektový jazyk a hierarchie metajazyků

Aby bylo možné zkonstruovat uspokojivou definici pravdy pro daný jazyk, je zapotřebí nejdříve zavést jednu nezbytně nutnou distinkci – distinkci mezi objektovým jazykem a metajazykem. Objektový jazyk je jazyk, *pro* který bude zkonstruována definice pravdy – tedy jazyk, o kterém budeme mluvit. Mimo to ale bude zapotřebí jazyk, ve kterém budeme mluvit. Tím bude metajazyk – tedy jazyk, ve kterém budeme pronášet výpovědi o objektovém jazyku a *ve* kterém bude zkonstruována definice pravdy pro objektový jazyk.

Ačkoli hraje distinkce mezi objektovým jazykem a metajazykem velmi důležitou roli v Tarského koncepci, on sám není jejím tvůrcem – za jejího autora je považován Stanisław Leśniewski.²⁸

Tarski uvádí o metajazyku následující: „[M]etajazyk obsahuje jak individuální jména, tak i překlady všech výrazů (konkrétně všech vět) studovaného jazyka.“²⁹ Jelikož účelem metajazyka je hovořit o objektovém jazyku, musí být metajazyk dostatečně bohatý, aby v něm bylo možné odkazovat k větám objektového jazyka. Toho lze docílit tak, že metajazyk bude obsahovat prvky potřebné ke zkonstruování jmen všech výrazů objektového jazyka. Toto však samo o sobě nestačí. Dále je totiž nezbytně nutné, aby metajazyk obsahoval překlady všech výrazů objektového jazyka.³⁰ Tyto překlady hrají důležitou roli při tvorbě ekvivalencí tvaru (T). Každá taková ekvivalence musí mít na levé straně jméno zkoumané věty v metajazyku, a na pravé straně adekvátní překlad věty označované tímto jménem do metajazyka. Kdyby se v objektovém jazyku vyskytovala věta, jejíž jméno a překlad by se nenacházely v metajazyku, pak by nebylo možné o této větě v metajazyku nic vypovědět a zkoumání tohoto objektového jazyka by tedy nemohlo být nikdy úplné.

Výčet prvků metajazyka se ovšem nevyčerpává překlady všech výrazů objektového jazyka spolu s jejich jmény. Co všechno ještě metajazyk může a naopak nesmí obsahovat, je možné ilustrovat si na příkladu T-schématu.

(T) *X* je pravdivá tehdy a jen tehdy, když *p*.

Již bylo uvedeno, že metajazyk obsahuje překlady výrazů (věty) objektového jazyka – těm odpovídá v T-schématu „*p*“ – a dále jejich jména – jimž odpovídá *X*. Mimo to je zapotřebí, aby byly v metajazyku obsaženy termíny obecně logické povahy, tedy především logické termíny jako je např. „tehdy a jen tehdy, když“.

Dalším prvkem metajazyka jsou sémantické termíny vztahující se k výrazům objektového jazyka, k jehož zkoumání a popisování daný metajazyk používáme. Ty však musejí být do metajazyka zavedeny pouze za pomoci definic. Metajazyk může obsahovat i nedefinované termíny, ale pouze pokud jsou tyto termíny obsaženy již v objektovém jazyku. Druhou výjimku tvoří termíny „odkazující k formě výrazů objektového jazyka, kterých je použito k tvorbě jmen těchto výrazů“. (Tarski, 2006, s. 148)³¹ Tarski zdůrazňuje, že je nezbytně nutné, aby

²⁸ Srovnej (Milne, 1997, s. 2).

²⁹ „[T]he metalanguage contains both an individual name and a translation of every expression (and in particular of every sentence) of the language studied.“ (Tarski, 1944, s. 173)

³⁰ Srovnej (Tarski, 1944, s. 188).

³¹ Tarski zde má nejspíš na mysli prostředky k vytvoření strukturálně-deskriptivních jmen.

byly do jazyka zaváděny sémantické termíny, které by odkazovaly k výrazům objektového jazyka, pouze definicemi. Kdyby tomu tak nebylo, hrozilo by obnovení paradoxu lháře.

Je důležité si uvědomit, že jak objektový jazyk, tak i metajazyk mají relativní smysl. Relativní v tom ohledu, že pokud je zkoumán nějaký objektový jazyk O , pak je toto zkoumání prováděno v metajazyku M^1 , který obsahuje (překlad) O jako svoji součást. Pokud se stane cílem zkoumání metajazyk M^1 , pak bude muset být toto zkoumání provedeno v jiném metajazyku M^2 , který bude muset obsahovat překlad M^1 jako svoji součást. Tím se z M^1 stal ve vztahu k M^2 objektový jazyk. Jestliže by měl být zkoumán podroben M^2 , pak by se stal sám objektovým jazykem pro nový metajazyk M^3 . Toto by mohlo pokračovat do nekonečna. Nemáme zde tedy jediný objektový jazyk a metajazyk, ale rovnou hierarchii metajazyků. Tarskému jde především o to, že pokud se vyjadřujeme o větech daného objektového jazyka, pak se nacházíme v metajazyku k tomuto jazyku. Kdybychom totiž naše výpovědi o zkoumaném jazyku neformulovali v metajazyku, ale bylo by možné je formulovat přímo ve zkoumaném objektovém jazyku, pak by se jednalo o jazyk sémanticky uzavřený.

Ještě než přistoupím k samotné Tarskému definici pravdy, je zapotřebí zmínit se o důležitém principu, který je známý pod jménem Tarského teorému.³² Vychází z Teorému I. v (Tarski, 1944, s. 247) a jeho důsledek se dá formulovat následovně: „Za jistých obecných předpokladů je dokázáno, že je nemožné zkonstruovat korektní definici pravdy, pokud jsou použity pouze takové kategorie, které se vyskytují ve zkoumaném jazyce.“³³ Tedy jinými slovy, podle Tarského teorému je nemožné zkonstruovat adekvátní definici pravdy pro objektový jazyk, pokud používáme pouze prostředky objektového jazyka. Právě toto je důvod, proč musí být adekvátní definice termínu „pravdivý“ provedena v metajazyku, který navíc musí obsahovat sémantické pojmy aplikovatelné na výrazy (především na věty) objektového jazyka.

Než se však dostanu k samotné formulaci definice termínu „pravdivý“, bude třeba vyložit některé doplňující požadavky, kterých je zapotřebí se přidržovat, aby byla výsledná definice adekvátní. Úspěch konstrukce definice pravdy je závislý na tom, jestli je metajazyk, ve kterém bude definice provedena, „bytostně bohatší“ než jazyk, pro který je definice konstruována. Bohužel o podmínce bytostné bohatosti se Tarski příliš podrobně nevyjadřuje. Jako kompenzace tohoto však může posloužit Tarského ukázka, co by měl takový jazyk obsahovat, nacházející se v [SCT]: „Omezíme-li se na jazyky založené na logické teorii typů, pak je jazyk „bytostně bohatší“ než objektový jazyk, jestliže obsahuje proměnné vyšších logických typů, než jsou logické typy proměnných objektového jazyka.“ (Tarski, 2006, s. 140) Metajazyk, který by nevyhovoval této podmínce, by bylo možné rekonstruovat v objektovém jazyku. Tudíž by v takovém metajazyku bylo možné vytvořit paradox plodící větu.

4 Konstrukce definice pravdy

Záměrem této kapitoly je pojednat o samotné Tarskému definici pravdy. Tarského cílem je použít ke konstrukci definice pojem splňování. Splňování je sémantický termín, což by mohlo představovat pro Tarského definici problém, jelikož metajazyk (ve kterém je konstrukce prováděna) nesmí obsahovat žádné nedefinované sémantické termíny. Pojem splňování je však podle Tarského snadné definovat v nesémantických termínech a lze tak vyhovět předchozímu požadavku. Důležitost toho, aby byla definice pravdy provedena v nesémantických pojmech, shrnuje Milne následovně:

pokud může být pravda – a ostatní sémantické pojmy – definována v nesémantických termínech, potom jsou sémantické pojmy (a) očištěny od náznaku paradoxu dlícího

³² Nebo také Tarského teorém nedefinovatelnosti, srovnej (Gómez-Torrente, 2004, 27–37).

³³ „Under certain general assumptions, it proves to be impossible to construct a correct definition of truth if only such categories are used, which appear in the language under consideration.“ (Tarski, 1944, s. 254)

nad nimi, (b) vysvětleny z hlediska pojmů, jejichž význam je jasný a (c) vypadají, že vyhovují požadavkům fyzikalismu³⁴

Nyní tedy adekvátnost definice závisí na tom, jak se Tarskému povede definovat termín „pravdivý“ za pomoci pojmu splňování. „Splňování je relace mezi libovolnými předměty a jistými výrazy, které se nazývají „*větné funkce*“.“ (Tarski, 2006, s. 150) Jelikož věty/větné funkce jsou výrazy jazyka, potom je splňování relací mezi výrazy jazyka a předměty, což odpovídá Tarského definici sémantických termínů z počátku [SCT].

Větná funkce je takový výraz, který má obdobnou formální strukturu, jako má věta. Zatímco věty mohou obsahovat proměnné pouze vázané, větné funkce mohou obsahovat i volné proměnné. Aby se větná funkce stala větou (a tudíž získala pravdivostní hodnotu), musí být všechny volné proměnné v dané větné funkci nahrazeny jmény objektů.

Pro definování větné funkce použil Tarski metodu rekurzivní definice, která spočívá v tom, že nejdříve definuje nejjednodušší větné funkce a posléze jsou určena pravidla pro jejich spojování.³⁵ Pokud máme definovány elementární větné funkce a pravidla pro jejich slučování, nic nám nebrání vytvořit za pomoci těchto pravidel libovolně dlouhou větnou funkci. Ovšem, jak poukazuje Kirkham, i rekurzivní metoda má jedno slabé místo. Rekurzivní definici je totiž možné úspěšně vytvořit jen tehdy, když existuje pouze omezené množství elementárních členů dané třídy a omezené množství pravidel, podle nichž je možné kombinovat elementární členy dané třídy.

Jakmile jsou definovány větné funkce, je možné přejít k definici pojmu splňování. Jak jsem uvedl výše, k tomu, aby se z větné funkce stala věta, musí dojít k nahrazení volných proměnných jmény objektů: „dané předměty splňují danou funkci, jestliže se z této funkce stane pravdivá věta, když v ní nahradíme volné proměnné jmény těchto předmětů.“ (Tarski, 2006, s. 150)

Tedy pokud máme větnou funkci „x je bílý“, pak je tato funkce splňována např. sněhem, jelikož dosazením jména „sníh“ do zmíněné větné funkce získáme pravdivou větu „sníh je bílý“. K definování pojmu splňování je opět využita metoda rekurzivní definice. Prvním krokem jsou určeny předměty splňující nejjednodušší větné funkce. Když jsou tyto určeny, pokračuje proces určením podmínek, kterým musejí předměty splňující elementární větné funkce vyhovovat, aby mohly splňovat složené větné funkce.³⁶ Problém nastává, když má větná funkce více proměnných, pak totiž přestane být splňování dvojčlennou relací, ale stane se relací troj- a vícečlennou (podle počtu proměnných ve větné funkci). Proto je podle Tarského vhodné upravit definici splňování tak, že splňování je definováno jako dvojčlenná relace mezi větnou funkcí a nekonečnou posloupností objektů. Jelikož věty neobsahují žádné proměnné, mohou nastat pouze dva případy – věta je splňována všemi posloupnostmi nebo není splňována žádnými. Definici pravdivé věty lze tedy formulovat následovně: Věta je pravdivá tehdy, když je splňována všemi (nekonečnými) posloupnostmi objektů.³⁷

5 Adekvátnost pro řešení paradoxů

Na závěr této části hodlám pojednat o tom, jestli Tarského koncepce vyhovuje požadavkům na adekvátní řešení paradoxů, jak je uvádí Kirkham.³⁸ Tím si vytvořím jakýsi „most“ pro přechod ke druhé části článku.

³⁴ „if truth – and other semantic notions – can be defined in non-semantic terms then semantic notions are (a) cleared of the suspicion of paradox that lingers over them, (b) explained in terms of concepts whose meaning is clear, and (c) seem to satisfy the demands of physicalism“ (Milne, 1997, s. 11)

³⁵ Srovnej (Kirkham, 1995, s. 146–150).

³⁶ Srovnej (Tarski, 1944, s. 190–191)

³⁷ Srovnej (Tarski, 1944, s. 190–198)

³⁸ Viz (Kirkham, 1995, s. 273).

1. *Specifičnost* – Řešení musí ukázat, v čem spočívá možnost tvorby paradoxu. Tarski jako příčinu paradoxů uvedl sémantickou uzavřenost jazyků, ve kterých jsou formulovány paradox plodící věty. Proti tomuto bodu je namířena kritika Roberta L. Martina.
2. *Absence ad hoc postulátů* – Nabízené řešení by nemělo obsahovat takové prvky, které byly přijaty pouze na základě toho, že řeší paradoxy. Tarskému je v tomto ohledu vytýkáno zavedení hierarchie metajazyků s pravdivostními predikáty relativizovanými k jazykům, na něž jsou aplikovány. Kritiku tohoto Tarského přístupu podal např. Saul Kripke.
3. *Zamezení přílišným restrikcím* – Řešení paradoxu by mělo být takové, aby nebyly na jeho úkor odstraněny i neproblematické věty. Důsledkem odmítnutí sémanticky uzavřených jazyků však je, že jsou z objektového jazyka odstraněny veškeré autoreferenční věty. Kritiku tohoto bodu lze nalézt u R. L. Martina.
4. *Kompletnost* – Nabízené řešení by mělo být aplikovatelné na všechny varianty paradoxu. Jestli Tarski tomuto kritériu vyhovuje, se pokusím určit za pomoci prací dr. Jiřího Raclavského věnovaných paradoxům lháře.
5. *Soulad s intuicemi* – Řešení paradoxu by se nemělo protivit našim intuicím o významu termínu „pravdivý“. Kritikové Tarskému vytýkají, že hierarchie metajazyků a s nimi i hierarchie pravdivostních predikátů neodpovídají běžnému úzu. Do této skupiny kritiků by se dal zařadit např. Jonathan Harrison, podle kterého není v souladu s našimi intuicemi aplikovat termín „pravdivý“ na věty, a Saul Kripke se svou kritikou zmnožení pravdivostních predikátů v Tarského koncepci.³⁹

Tímto je výkladová část této studie vyčerpána a nezbyvá, než přejít k její kritické části.

Kritika Tarského sémantické koncepce pravdy

6 Problémy hierarchie metajazyků

6.1 Mstivý lhář a nedostatečnost hierarchií metajazyků

V této a v následující kapitole se budu zabývat kritikou založenou na „vylepšených“ variantách paradoxu lháře.⁴⁰ Půjde mi především o takové varianty, které je možné nasměřovat proti Tarského sémantické koncepci pravdy.

Řeč nyní bude o takzvaném mstivém lháři [revenge liar], který patří spolu se zesíleným lhářem [strengthened liar] mezi tzv. silnější lháře [stronger liar]. Oproti zesílenému lháři není mstivý lhář zaměřen pouze na klíčový predikát řešení klasického lhářského paradoxu, ale je zaměřen proti principu daného řešení jako celku – v případě Tarského koncepce jím je existence (údajně nedostatečného počtu) hierarchií (meta)jazyků. Existují přinejmenším dva způsoby, jak mstivého lháře zformulovat. Jak se pokusím ukázat dále, obě tyto formulace jsou ve své podstatě zakotveny ve stejném předpokladu a ani jedna nepředstavuje žádné nebezpečí jak pro koncepci metajazyků, tak pro Tarského koncepci jako celek.

První varianta mstivého lháře má pravděpodobně nejlépe viditelné nedostatky a je snáze napadnutelná:

(RL₁) Věta (RL₁) není pravdivá v žádném z jazyků J¹–Jⁿ.⁴¹

³⁹ Volně podle (Kirkham, 1995, s. 273).

⁴⁰ Různé varianty paradoxu lháře, spolu s návrhem jejich řešení založeném na systému transparentní intenzionální logiky navržené Pavlem Tichým lze nalézt v (Raclavský, 2007), potažmo v (Raclavský, 2009).

Na první pohled je jasné, že tato varianta je zaměřena na jazyky řádů 1– n . Princip, kterým má tento paradox zpochybnit Tarského koncepci je následující:

- (1) Máme dány pouze jazyky řádů 1– n
- (2) Věta (RL₁) tedy musí být formulována v jednom z jazyků J¹–Jⁿ
- (3) Věta (RL₁) o sobě vypovídá, že není pravdivá ani v jednom z jazyků J¹–Jⁿ
- (4) Avšak k tomu, aby mohla věta (RL₁) vypovídat o jazycích řádu 1– n , by sama musela být v jazyku vyššího řádu, tedy alespoň Jⁿ⁺¹
- (5) Z výše uvedeného však dostáváme spor, jelikož (RL₁) by měla být zároveň v některém z jazyků J¹–Jⁿ a zároveň v jazyku Jⁿ⁺¹
- (6) Hierarchie metajazyků vede k paradoxu a je tudíž nedostačující

To vypadá jako platná námitka, protože věta (RL₁), vyjadřující se o větách řádu 1– n , sama vypadá jako věta řádu $n+1$ – pokud však můžeme použít pouze jazyky n řádů, vzniká nám problém.

Druhá formulace tohoto paradoxu vypadá nebezpečněji, ačkoli je její znění velmi podobné:

(RL₂) Věta (RL₂) není pravdivá na žádné úrovni (dané) hierarchie jazyků.

V této formulaci nenalezneme žádný explicitní odkaz na počet hierarchií jazyků, je tedy zapotřebí počítat s „nejhorší“ variantou, tedy s jejich nekonečným množstvím. I kdyby byla hierarchie jazyků nekonečná, stejně by se v případě věty (RL₂) zdálo, že ani nekonečná hierarchie nestačí, podobně jako tomu bylo v případě věty (RL₁). Mstivý lhář tak napadá hierarchie jazyků tvrzením, že ať je jejich množství jakékoliv, nikdy není dostatečné.

Již dříve jsem uvedl, že mstivý lhář je založen na předpokladu, že disponujeme určitým množstvím metajazyků. Jenomže s tímto předpokladem mstivý lhář coby námitka proti Tarskému padá. Ten totiž nikde netvrdí, že by mohl existovat pouze jistý počet metajazyků a také nikde neuvádí, že v tom kterém případě máme k dispozici konkrétní počet metajazyků. Z jeho koncepce je sice možné odvodit existenci hierarchie metajazyků (která může být nekonečná), ale to ještě samo o sobě neospravedlňuje napadení jeho koncepce mstivým lhářem. K pochopení, proč tomu tak je, se stačí znovu podívat na závěr kapitoly 3.4. Jak jsem tam uvedl, pro Tarského má distinkce mezi objektovým jazykem a metajazykem relativní smysl. Pokud mluvíme o libovolném jazyku, vždy o něm musíme mluvit v metajazyku, který bude dostatečně bohatý, aby v něm mohly být formulovány výpovědi o všech větách zkoumaného jazyka. K tomu, aby bylo možné zformulovat smysluplné a konzistentní výpovědi o větách nějakého jazyka, je zapotřebí nacházet se v metajazyku, který bude obsahovat jak jména a překlady všech výrazů zkoumaného jazyka, tak i sémantické termíny aplikovatelné na výrazy tohoto jazyka. Vzniká tak sice hierarchie metajazyků, ale to není základem Tarského koncepce: „Jenže podstatou hierarchických řešení není určité (omezené) číslo hierarchií, ale princip komentujících „metapozicí“.“ (Raclavský, 2009, s. 21)

Klíčovým principem Tarského koncepce je nutnost vystoupit do metajazyka a definici pravdy pro objektový jazyk zkonstruovat až v metajazyku. Tím však není řečeno, že musíme použít určité množství jazyků. Mstivý lhář se tak mívá účinkem.

⁴¹ Volně podle (Raclavský, 2009, s. 21).

6.2 Nerozhodnutelnost zařazení k metajazyku

Stejně jako předchozí kapitola, i tato je založena na napadení Tarského koncepce variantou paradoxu lháře. Zatímco mnou dosud použité varianty by se daly označit jako autoreferenční, nyní se budu zabývat variantami heteroreferenčními,⁴² a to jednak variantou nekvantifikační:

JP₁: Věta JP₂ je nepravdivá.
 JP₂: Věta JP₁ je pravdivá. (Raclavský, 2009, s. 23)

a variantou kvantifikační:

Platón: Vše, co řekl Sókratés, je pravdivé.
 Sókratés: Vše, co řekl Platón, je nepravdivé. (Raclavský, 2009, s. 24)

Tyto dvě obměny znázorňují, že paradoxní situace může vzniknout i jinak, než za pomoci autoreference. Zde máme v obou případech dvě věty, přičemž ani jedna není sama o sobě nebezpečná. Díky jejich spojení však dojde k zacyklení – jedna věta odkazuje k větě druhé a druhá věta zase odkazuje k větě první. Schéma plodící paradox je pro nekvantifikační variantu následující:

- (1) Předpokládáme, že věta (JP₁) je pravdivá
- (2) Jestliže je věta (JP₁) pravdivá, pak je pravdivé to, co říká
- (3) Věta (JP₁) říká, že věta (JP₂) je nepravdivá
- (4) Věta (JP₂) je tedy nepravdivá
- (5) Jestliže je (JP₂) nepravdivá, pak to, co říká, není pravda
- (6) Věta (JP₂) říká, že věta (JP₁) je pravdivá
- (7) Podle (5) a (6) je tedy věta (JP₁) nepravdivá
- (8) (JP₁) je tedy pravdivá a nepravdivá zároveň

Je samozřejmě možné začít předpokladem pravdivosti věty JP₂, nebo i předpokladem nepravdivosti vět JP₁ nebo JP₂ – konstrukci odpovídajících schémat ponechám na čtenáři. Prakticky stejné schéma je možné vytvořit i pro buridanovskou insolubilitu – tento krok opět ponechám na čtenáři.

Heteroreferenční varianty paradoxu lháře mají za úkol ukázat, že v některých případech je nemožné určit řád jazyka, ve kterém jsou dané věty formulovány. Pro pochopení této námítky je vhodné ilustrovat ji dalším schématem:

- I. Věta (JP₁) zmiňuje větu (JP₂)
- II. Aby se mohla nějaká věta zmiňovat o jiné větě, musí být formulována v (meta)jazyku vyššího řádu
- III. Předpokládáme, že (JP₂) je v jazyku řádu k (J^k)
- IV. Aby mohla (JP₁) zmiňovat (JP₂), musí být formulována alespoň v jazyku řádu $k+1$ (J^{k+1})
- V. (JP₁) je tedy v J^{k+1} a (JP₂) v J^k
- VI. (JP₂) zmiňuje (JP₁)
- VII. (JP₂) tedy musí být podle (II) formulována v jazyku vyššího řádu, než (JP₁), tedy alespoň v J^{k+2}

⁴² R. L. Martin používá k označení variant paradoxu lháře, které já označuji jako heteroreferenční, termín nepřímá autoreference [indirect self-reference].

VIII. Z (V) a (VII) vzniká spor, jelikož (JP₂) by měla být formulována v J^k a J^{k+2} zároveň

Na základě tohoto schématu je tedy údajně oprávněné tvrdit, že v Tarského systému není možné rozhodnout, ve kterém jazyku jsou zformulovány věty heteroreferenčního paradoxu. Ovšem to by platilo pouze, pokud by bylo výše uvedené schéma adekvátním rozbořením heteroreferenčního lháře. Nyní se pokusím za pomoci řešení dr. Raclavského ukázat, proč není Tarského koncepce heteroreferenčními variantami paradoxu lháře ohrožena.

Nejdříve je důležité si uvědomit, že výše uvedené schéma počítá pouze s jedním možným způsobem, jak lze údajně zkonstruovat heteroreferenční variantu paradoxu lháře, kde východiskem je předpoklad, že věta (JP₁) je formulována v jazyku J^{k+1} a věta (JP₂) je v jazyku J^k – tuto variantu označím jako (α). Je samozřejmě možné tyto předpoklady prohodit a začínat tak s větou (JP₂) formulovanou v jazyku J^{k+1} a naopak s větou (JP₁) v jazyku J^k (schéma by i přes tuto změnu zůstalo prakticky stejné) – tuto variantu označím jako (β). Nabízí se také možnost vyjít z předpokladu, že jak věta (JP₁), tak i věta (JP₂) jsou formulovány v jazyku stejné úrovně, např. J^k – varianta (γ).

Jelikož důkazy neadekvátnosti (α) a (β) se od sebe liší pouze prohozením jednotlivých vět, důkaz neadekvátnosti (β) ponechám na čtenáři. Začátek důkazu je stejný, jako ve výše uvedeném schématu – věta (JP₂) je formulována v jazyku J^k a věta (JP₁), která ji zmiňuje, je tudíž v jazyku J^{k+1}. Dosud zde není žádný problém – ten ale vyvstává hned v dalším kroku (VI). O větě (JP₁) jsme oprávněni tvrdit, že zmiňuje větu (JP₂), jelikož je formulována v jazyku vyššího řádu. Je totiž formulována v jazyku J^{k+1}, který obsahuje prostředky umožňující pojmenovat a tedy i zmiňovat všechny věty jazyka nižšího řádu, tedy i J^k. Věta (JP₂), která je sama formulována v J^k obsahuje odkaz k (JP₁) – to však v žádném případě nesmí. Žádný jazyk nesmí podle Tarského obsahovat kromě jmen pro vlastní věty i jména vět jazyků vyššího řádu. To je ale v případě věty (JP₂) zcela zjevně porušeno a proto je nutné ji považovat za nesmyslnou (bez ohledu na (JP₁)). Tímto mizí problém se zacyklením a žádný paradox nevzniká.

V případě (γ) je situace ještě jednodušší. Jelikož obě věty jsou v tomto případě formulovány v jazyku J^k, pak ani jedna z nich není smysluplná. Jelikož kdyby tyto věty mohly odkazovat k větám stejného jazyka, ve kterém jsou zformulovány, pak by byly samy zformulovány v sémanticky uzavřeném jazyku. A to Tarski rozhodně nepřipouští. Jelikož důkaz neškodnosti heteroreferenční kvantifikační varianty vypadá prakticky shodně, ponechám jeho studium na čtenáři.⁴³

7 Záchrana autoreferenčních vět

Jak bylo uvedeno již v první části této studie, Tarski se zabýval problémem definování pravdy pouze pro jazyky se specifikovanou strukturou a odmítl všechny jazyky sémanticky uzavřené. Důsledkem toho je, že Tarského definice jsou aplikovatelné pouze pro jazyky, které neobsahují žádné autoreferenční výrazy – jelikož prostředky pro odkazování k výrazům jazyka J^k mohou být v Tarského koncepci pouze v jazyku vyššího řádu J^{k+1}. Podle některých je toto omezení příliš restriktivní a Tarski tak odstraňuje i množství vět, které jsou neproblematické. To je podle kritiků příliš velká cena za odstranění paradoxů. V této kapitole rozeberu přístup reprezentanta „populárních“ námitek, Roberta L. Martina, jehož cílem bylo ukázat, že problémy s autoreferenčními větami lze vyřešit, aniž bychom použili hierarchie jazyků a odstranili tak z jazyka autoreferenční výrazy.

Námítka v této kapitole projednávána má své kořeny v Tarského definici sémanticky uzavřených jazyků. Kritéria, kterým sémanticky uzavřený jazyk vyhovuje, lze nalézt v kapitole 3.2. Klíčovou roli zde budou hrát kritéria, která jsem tam uvedl pod bodem (a):

⁴³ Srovnej (Raclavský, 2009, s. 24–25).

- i. Sémanticky uzavřený jazyk obsahuje kromě vlastních výrazů také jména těchto výrazů (nebo jiný způsob umožňující mluvit o větách tohoto jazyka)
- ii. sémantické termíny odkazující k větám tohoto jazyka

Bod (i) nám o sémanticky uzavřených jazycích říká to, že každý takový jazyk v sobě obsahuje prostředky umožňující v něm formulovat věty odkazující k jiným větám tohoto jazyka nebo k těmto větám samým. Jedná se zde tedy o tzv. autoreferenci – jev, jehož vzniku se Tarski brání přesunutím prostředků odkazujících k větám daného jazyka do jazyka vyššího řádu.

Příkladem nebezpečnosti autoreferenčních vět může být mnou již zmíněný paradox lháře ve znění:

(L) Věta (L) je nepravdivá.

Tarski spatřoval nebezpečí pro adekvátní teorii pravdy jak v bodu (i), tak i v bodu (ii). Zabránit vzniku paradoxů lze podle Tarského přesunutím sémantických pojmů a prvků umožňujících pojmenování výrazů daného jazyka do metajazyka k tomuto jazyku. Pokud ovšem z jazyka odstraníme veškeré prostředky umožňující jeho větám zmiňovat samy sebe, pak z tohoto jazyka zmizí veškeré autoreferenční výrazy.

Martinova námitka je založena na tom, že autoreferenční výrazy mohou být jak „dobré“, tak i „špatné“. Zatímco „špatná“ autoreference vede ke vzniku paradoxů, „dobrá“ autoreference zahrnuje výrazy, které „mohou být mluvčím jazyka snadno pochopeny a mohou být posouzeny logicky neproblematickým způsobem jako jasně pravdivé (nebo nepravdivé)“.⁴⁴ Příkladem „dobré“ autoreference může být např. věta „Tato věta je v češtině“.

Martin se snaží přijít s řešením, které nebude natolik restriktivní a nebude vyžadovat přesouvání sémantických termínů a jmen výrazů objektového jazyka do metajazyka. Svoji argumentaci Martin zahajuje poukázáním na jisté rozdíly mezi variantami paradoxu lháře. Uvažme následující možnosti:

1. „Tato věta je nepravdivá.“
2.
 - a. „Následující věta je pravdivá.“
 - b. „Předchozí věta je nepravdivá.“

Paradox plodící větu (1) Martin označuje jako „crux“ verzi paradoxu lháře. Varianta (2) není klasickou autoreferenční variantou, jelikož obsahuje dvě navzájem k sobě odkazující věty. Martin používá pro vztah mezi větami v této variantě termín nepřímá autoreference. Zatímco paradox plodící větu (1) k sobě odkazuje naprosto vždy, věty (2a) a (2b) musejí být ve správném kontextu, aby odkazovaly jedna k druhé (musejí být umístěny u sebe). Z toho lze vyvodit, že vznik paradoxu z těchto vět je závislý na kontextu, ve kterém se věty vyskytují. Kdyby byla mezi věty obsažené ve (2a) a (2b) vložena nějaká další věta, ke vzniku paradoxu by již dojít nemuselo.

Pokud má být věta považována za smysluplnou a pokud jí má být přiřazena pravdivostní hodnota, pak musí taková věta splňovat dva požadavky. Prvním z nich je, aby tato věta byla v souladu s pravidly gramatiky daného jazyka. Podle tohoto pravidla by byly jako nesmyslné vyloučeny věty jako např. „paruce v okenice bambusová zpívala měsíční“. Toto pravidlo samo o sobě však nestačí. Je totiž možné formulovat gramaticky korektní větu, která bude i

⁴⁴ „can be readily understood by a speaker of the language, and be judged, in a logically unproblematic way, to be straightforwardly true (or false).“ (Martin, 1967, s. 279)

přesto nesmyslná, jako např. „bambusová okenice zpívala v měsíční paruce“. Martin k tomuto dodává, že kromě gramatiky je třeba brát v úvahu i sémantiku.⁴⁵

Na věty, které jsou sémanticky nekorektní, není podle Martina možné vztahovat princip bivalence, protože by jeho použití vedlo k paradoxním situacím. Martinovou modelovou větou je „Jedenáctka je fialová“. Pokud bychom o ní řekli, že je nepravdivá, pak bychom se dostali do situace, že bychom museli připsat pravdivost některé alternativní větě, jako je „Jedenáctka je žlutá“ nebo „Jedenáctka je červená“. (Martin, 1967, s. 289) Proto by se takovými větám neměla přisuzovat žádná pravdivostní hodnota. Dalším Martinovým krokem je návrh procedury, jejímž cílem by bylo rozpoznat veškeré deviantní věty. Rozebírání této metody by však bylo nad rámec této práce.

Martin ve své koncepci neodstraňuje z jazyka veškeré autoreferenční věty, ale pouze ty, které jsou deviantní. Na paradox plodící věty není možné vztahovat princip bivalence a jsou tak ponechány bez pravdivostní hodnoty. Neproblematické (nejenom) autoreferenční věty však touto procedurou projdou a může jim být připsána pravdivostní hodnota. Martin se tak zbavuje problémů s deviantními větami, aniž by eliminoval všechny autoreferenční věty a musel provést vzestup do metajazyka.

8 Nekonzistence hovorových jazyků

Tato kapitola bude věnována kritice pocházející od Hanse Herzbergera, podle kterého nemůžeme o žádném jazyku vypovědět, že je nekonzistentní, aniž by tato naše výpověď byla kontradikcí. Svoji argumentaci zahajuje Herzberger konstatováním, že pokud má být nějaký jazyk nekonzistentní, pak jediné proto, že je nekonzistentní množina analytických vět v tomto jazyku obsažených. To je ale podle Herzbergera nepřipustné, což se pokouší dokázat argumentem založeným na *reductio ad absurdum*. Jeho argument má následující schéma:

- 1) Daný jazyk L není logicky konzistentní kvůli své třídě analytických vět
- 2) L obsahuje neprázdnou třídu vět A takovou, že
 - a) každá věta patřící do A je analytická a
 - b) třída A je logicky nekonzistentní třída vět
- 3) Každá věta patřící do třídy A je pravdivá
- 4) Alespoň jedna věta patřící do A musí být nepravdivá
- 5) V jazyku L existuje alespoň jedna věta, která je pravdivá a nepravdivá zároveň⁴⁶

Pokud by neplatily první dva body, nemělo by další zkoumání smysl. Bod (2b) je přímým důsledkem bodu (1) – nekonzistence třídy analytických vět lze dosáhnout pouze tak, že bude tato třída obsahovat alespoň jednu větu, která bude nepravdivá. Bod (3) je jistým rozvedením bodu (2a), jelikož z definice pojmu „analytický“ musejí být všechny analytické věty pravdivé. Aby byl jazyk nekonzistentní, pak musí jeho třída analytických (a tedy pravdivých vět) obsahovat alespoň jednu nepravdivou větu. To je ale zjevný rozpor, jelikož by musela třída A obsahovat alespoň jednu větu, která by byla pravdivá i nepravdivá zároveň.

Další argumentace je založena na zkoumání pravdivostních podmínek. Podle Herzbergera je totiž myšlenka nekonzistence jazyků založena na představě, že mohou nastat takové podmínky, díky nimž by se stala pravdivou alespoň jedna dvojice vzájemně kontradiktorických vět – daný jazyk by pak byl nekonzistentní díky svým pravdivostním podmínkám. Herzberger vrací i tuto možnost vytvoření nekonzistentního jazyka.

I. Máme daný jazyk L , který obsahuje logicky nekonzistentní třídu vět Q

⁴⁵ Srovnej (Martin, 1967, s. 286).

⁴⁶ Schéma volně podle (Herzberger, 1967, s. 29).

- II. Je dána podmínka w taková, že pokud w nastane, učiní pravdivými všechny věty Q
- III. Jelikož Q je logicky nekonzistentní třída vět, je nemožné, aby byly všechny věty v ní obsažené pravdivé
- IV. Jelikož je nemožné, aby byly všechny věty Q pravdivé, je nemožné, aby nastala podmínka w ⁴⁷

Jelikož neexistují podmínky, za kterých by mohly být pravdivé všechny věty patřící do logiky nekonzistentní třídy vět, není možné o žádném jazyku prohlásit, že by byl nekonzistentní na základě pravdivostních podmínek. Tento druhý argument je podle Herzbergera generalizací prvního argumentu. Zatímco v tomto argumentu bylo prokázáno, že neexistují podmínky, za nichž by mohly nabýt pravdivost vzájemně kontradiktorické věty, první byl postaven na pojmu analytické věty, která je pravdivá za každých podmínek (a tedy věty, která nemůže být za žádných podmínek nepravdivá).

Z výše uvedeného je zřejmé, jak chápe nekonzistentní jazyky Herzberger. Odpověď na otázku, zda stejně chápal nekonzistentní jazyky i Tarski, je nejen dle mého názoru záporná. Jak ostatně uvádí i Soames,⁴⁸ pro Tarského nebyl jazyk nekonzistentní kvůli nekonzistenci třídy svých analytických vět, neboť Tarski nepovažoval ekvivalence tvaru (T) za analytické věty daného jazyka. Jeho požadavkem bylo, aby ekvivalence (T) vzniklé dosazením vět do T-schématu byly pravdivé – požadavek, aby tyto věty byly analytické, nevznášá a ani není z jeho požadavků odvoditelný.⁴⁹ Třída analytických vět se tak i v případě dosazení paradox plodící lhářské věty do T-schématu nestává nekonzistentní. Je tak možné říci, že Herzbergerova kritika se poněkud mívá účinkem, jelikož u Tarského není hovorový jazyk nekonzistentní proto, že by byla nekonzistentní třída jeho analytických vět.

9 Aplikovatelnost predikátu „pravdivý“ na věty

V první části jsem se sice již problematice aplikace pravdivostního predikátu na věty věnoval, ale nezabýval jsem se přímo kritikou tohoto Tarského rozhodnutí. To hodlám učinit až na tomto místě. Předložím zde argumenty, které proti možnosti aplikovatelnosti predikátu „pravdivý“ na věty předložil Jonathan Harrison.

Harrison zastává stanovisko, že Tarski a všichni zastánci T-schématu se dopouštějí chyby, když považují za pravděnce věty, jelikož věty jimi být nemohou. Svůj článek proto věnuje kritice vět coby pravděnců a místo nich navrhuje přijetí propozic. O propozicích prohlašuje Harrison, že je chápe jako „věci“, které jsou, nebo mohou [...] být tvrzeny a které, kdyby byly tvrzeny, by byly pravdivé bez ohledu na slova, v nichž byly formulovány“.⁵⁰ Pročež je podle něho neadekvátní mluvit o tom, že věta je pravdivá nebo nepravdivá – úkolem vět je totiž vyjadřovat propozice. Dále je Harrisonovým cílem dokázat, že T-věty nejenže nejsou nutně pravdivé (což prý tvrdí zastánci T-schématu), ale mohou být dokonce i nepravdivé.

Pro pochopení další Harrisonovy argumentace je důležité seznámit se s klíčovými pojmy, se kterými operuje. Prvním je pragmatická kontradikce. Pragmatickou kontradikci popisuje následovně: „Tvrzení [...] jsou pragmaticky kontradiktorická, protože se pokouší udělat něco, co nemůže být provedeno, pokud jsou pravdivá.“⁵¹ Pragmatickou kontradikcí je tedy tvrzení, které se vyslovením stává kontradikcí. Jako příklad pragmatické kontradikce může sloužit

⁴⁷ Volně podle (Herzberger, 1967, s. 32–33).

⁴⁸ Srovnej (Soames, 1999, s. 62–64).

⁴⁹ Srovnej (Tarski, 1944, s. 165).

⁵⁰ „‘things’ that are or might [...] be asserted, and which, if they were asserted, would be true regardless of the words in which they were formulated.“ (Harrison, 1998, s. 13)

⁵¹ „Assertions [...] are pragmatically contradictory because they attempt to do something that could not be done if what they asserted was true.“ (Harrison, 1998, s. 6)

věta „Já jsem němý“. Pokud kdokoliv pronese tuto větu, pak ji právě jejím vyslovením učiní kontradiktorickou, jelikož kdyby byla pravda to, co věta říká, pak by ji nikdy nemohl vyslovit. Dalším důležitým pojmem Harrisonovy argumentace, stojícím v opozici k pragmatické kontradikci, je pragmatická nutnost. Pragmatickou nutností chápe Harrison takové tvrzení, jehož pravdivost zaručuje právě akt jeho tvrzení. Tedy jinými slovy pragmatická nutnost je takové tvrzení, které je pravdivé právě proto, že ho tvrdíme. Podobně jako v případě pragmatické kontradikce, kdy tvrzením věty „Já jsem němý“ tuto větu právě aktem jejího tvrzení činíme kontradikcí, pragmatickou nutnost činí akt jejího tvrzení pravdivou, jako v případě, kdy vyslovím větu „Já mluvím“. V souvislosti s tímto Harrison upozorňuje, že je zapotřebí rozlišovat dva druhy vět – věty tvrdící nutnou pravdu a věty tvrdící nutně pravdu. Věty tvrdící nutnou pravdu jsou takové věty, které tvrdí něco, co je logicky nutné. Stejně tak věty tvrdící nutně pravdu jsou vždy pravdivé, avšak díky způsobu, jakým jsou formulovány (např. Jiřího jméno je Jiří). To, co tato věta tvrdí není logicky nutná pravda, nicméně tato věta je formulována tak, že nemůže nastat případ, aby byla nepravdivá. Popřením této věty by došlo ke vzniku pragmatické kontradikce. Vzhledem k právě uvedené distinkci je podle Harrisona možné rozlišit dva způsoby, jak by se mohlo specifikovat T-schéma:

- a),s' vyjadřuje nutnou pravdu tehdy a jen tehdy, když s.
- b),s' vyjadřuje nutně pravdu tehdy a jen tehdy, když s.

Harrison je názoru, že T-schéma by mělo být interpretováno v souladu s bodem (b), jelikož význam vět jako „sníh je bílý“ se učíme aposteriorně. To by znamenalo, že T-věty nejsou nutně pravdivé věty, ale jsou to věty tvrdící nutně pravdu, a jako takové jsou pragmaticky nutné. Popřením T-věty bychom se tedy dopouštěli pragmatické kontradikce. To však znamená, že pravdivost T-vět je výsledkem toho, jak jsou formulovány a nikoliv toho, co tvrdí. „Mám dojem, že to, že ‚Sníh je bílý‘, je pravdivé tehdy a jen tehdy, když sníh je bílý, je pragmatická nutnost, vám říká velmi málo o významu ‚pravdivý‘.“⁵² Podle Harrisona však stoupenci T-schématu však zastávají názor, že T-věty jsou nutnými pravdami, což je v rozporu s Harrisonovými výsledky. Proto je třeba odmítnout T-schéma jako neadekvátní. Kromě toho, že T-věty nejsou nutné pravdy, mohou být podle Harrisona i nepravdivé. K tomu, aby byla T-věta pravdivá, musí být obě její strany pravdivé nebo obě nepravdivé. Podle Harrisona je ale možné, aby byla pravdivá jen jedna strana, zatímco druhá by byla nepravdivá. Kdybychom mluvili o větě jazyka J^X , kde by slovo „bílý“ znamenalo černý, pak by věta

(A) Věta „sníh je bílý“ je pravdivá v J^X tehdy a jen tehdy, když sníh je bílý.

byla pragmatickou kontradikcí. Jestliže však podle kritéria materiální adekvátnosti musí mít každá adekvátní definice pravdy za následek všechny T-věty, pak vzniká problém. T-věty musí být pravdivé, což ale (A) není – proto musí být Tarského koncepce chybná. Závěrem Harrisonova zkoumání tedy je, že predikát „pravdivý“ nemůže být aplikovatelný na věty. Tudíž je nepřipustné používat predikát „pravdivý v L “, ale místo něj by se měl používat predikát „vyjadřuje pravdu v L “. Je zde však nutné položit si otázku, zda je Harrisonova kritika adekvátní vzhledem k Tarského koncepci. Asi největším problémem Harrisonovy kritiky je, že nepoužívá verzi T-schématu, která by odpovídala konvenci T. Zatímco Tarski požaduje, aby na levé straně T-schématu bylo jméno věty a na pravé straně její překlad do metajazyka, Harrison ve své argumentaci adekvátní překlady do metajazyka nepoužívá. Příkladem budiž věta (A), na jejíž pravé straně se nevyskytuje adekvátní překlad věty z její levé strany. Za

⁵² „I suspect that that ‘Snow is white’ is true if and only if snow is white is pragmatically necessary tells you very little about the meaning of ‘true’.“ (Harrison, 1998, s. 10)

předpokladu, že metajazykem je soudobá čeština a objektovým jazykem je J^X , by věta (A) měla správně vypadat takto:

(B) Věta „sníh je bílý“ je pravdivá v J^X tehdy a jen tehdy, když sníh je černý.

Zatímco věta (B) obsahuje na pravé straně adekvátní metajazykový překlad věty na své levé straně, u věty (A) tomu tak není, tudíž věta (A) není ekvivalencí tvaru (T). Důsledkem toho je, že ani důsledky, které si z možnosti její formulace odvodil Harrison, nejsou platné. Je tedy diskutabilní, zdali je Harrisonova kritika opravdu kritikou Tarského koncepce, nebo se spíše jedná o kritiku Harrisonova vlastního konstruktů. Navíc, jak jsem již uvedl na konci kapitoly 1.2, Tarski netrval na tom, že pouze věty jsou jedinými možnými pravděnci, což opět devaluje Harrisonovu kritiku.

10 Saul Kripke – proti bivalenci a relativizaci

Již v kapitole věnované problematice autoreferenčních vět jsem okrajově narazil na princip bivalence. Na tomto místě se budu zabývat Kripkeho teorií, ve které hraje odmítnutí principu bivalence klíčovou úlohu.

Stejně jako Tarski, i Kripke tvrdí, že nejvhodnější je aplikovat pravdivost na věty. Velký rozdíl je však v přístupu obou k požadavkům kladeným na adekvátní teorii pravdy. Tarski byl zastáncem bivalence, což znamená, že všechny věty mají buď pravdivostní hodnotu „pravda“, nebo „nepravda“. Na rozdíl od něho ale Kripke hodlá pro svoje účely používat trojhodnotovou logiku vyvinutou Stephenem C. Kleenem, přičemž pravdivostní predikát v ní používaný je parciální. V takovéto logice je kromě hodnot „pravda“ a „nepravda“ přípustná ještě třetí hodnota – „nedefinováno“. I Kripke tak, podobně jako Martin, nevylučuje ze zkoumaného jazyka všechny autoreferenční věty a jeho teorie proto není natolik restriktivní, ale kvůli použití trojhodnotové logiky je jeho systém napadnutelný zesíleným lhářem.

Tarski ve své koncepci odmítl, aby zkoumaný jazyk L^k obsahoval jako svoji část pravdivostní predikát aplikovatelný na výrazy tohoto jazyka samého. Pravdivostní predikát pro L^k tak musel být až v jazyku vyššího řádu, tedy v L^{k+1} . Toto zmnožení pravdivostních predikátů Kripke podrobuje kritice: „Náš jazyk zajisté obsahuje pouze jedno slovo ‚pravdivý‘, nikoliv sekvenci různých frází [pravdivý_n], aplikovatelných na jazyky vyšších a vyšších úrovní.“⁵³ Pokud by takto fungoval náš jazyk, pak by každý, kdo by používal termín „pravdivý“, k němu vlastně vždy připojoval index (i když pouze implicitně), který by odkazoval k úrovni jazyka, ke které se daný predikát vztahuje. Navíc by pro mluvčího bylo v některých případech složité nebo dokonce nemožné zjistit, jaké úrovně by měl být konkrétní pravdivostní predikát. Tuto kritiku podporuje Kripke vlastní variantou heteroreferenčního paradoxu lháře. Na základě těchto námitek Kripke zavrhuje Tarského hierarchii metajazykových pravdivostních predikátů.

Kripkeho cílem je vytvořit koncepci, v níž bude jen jeden parciální pravdivostní predikát aplikovatelný i na výrazy, které ho obsahují. Navíc, zatímco Tarski má na každé úrovni jazyků nový pravdivostní predikát, Kripke používá pouze jediný predikát, jehož extenze (třída pravdivých vět) i antiextenze (třída vět, které nejsou pravdivé) rostou tak dlouho, dokud není dosažen tzv. *fixpoint*, neboli pevný bod. Tento pravdivostní predikát je stejný pro všechny jazyky, ale v různých úrovních jazyka se mění jeho interpretace, která je dána jeho extenzí a antiextenzí. Při vzestupu do vyšší úrovně jazyka dochází k rozšiřování interpretace pravdivostního predikátu – tedy na čím vyšší úrovni jazyka se nacházíme, tím větší je množství vět, které mají stanovenou pravdivostní hodnotu. Tento proces v určitém okamžiku dospěje do stádia, kdy se interpretace pravdivostního predikátu vzestupem do vyšší úrovně jazyka ne-

⁵³ „Surely our language contains just one word ‘true’, not a sequence of distinct phrases [true_n], applying to sentences of higher and higher levels.“ (Kripke, 1975, s. 695)

změní (nerozšíří) – v takovémto případě bylo dosaženo (minimálního) pevného bodu. Vzestup nad pevný bod už neobohacuje třídy vět, u nichž je určena pravdivostní hodnota, jelikož byly přiděleny pravdivostní hodnoty všem větám, kterým být přiděleny mohly. Důležitým rysem tohoto procesu je, že pokud je stanovena pravdivostní hodnota nějaké věty, pak ji má tato věta již napořád.

Věty, které mají pravdivostní hodnotu v minimálním pevném bodu Kripke označuje jako uzemněné, jestliže věta nemá pravdivostní hodnotu v pevném bodu, pak je neuzemněná. To, jestli je věta uzemněná nebo nikoliv je mnohdy záležitostí empirických faktů. Pokud je daná věta výpovědí o jiných větách, pak její pravdivost (a uzemněnost) závisí na pravdivosti právě těchto vět. Rysem paradox plodících vět je, že takové věty nemají pravdivostní hodnotu v žádném pevném bodu.

Toto byla velmi stručná charakterizace Kripkeho teorie pravdy. Její přesnější popis lze najít např. u Scotta Soamese.⁵⁴ Závěrem však zbývá dodat několik poznámek a shrnutí. Jak jsem uvedl na začátku kapitoly, Kripke napadá dva fundamenty Tarského koncepce. Prvním je princip bivalence – Kripkeho odpověď spočívá ve využití trojhodnotové logiky. I když tím uniká některým nepříjemným důsledkům Tarského teorie, zároveň se tím vystavuje napadení zesíleným lhářem. Druhou napadenou součástí Tarského koncepce je hierarchie pravdivostních predikátů. Zatímco Tarski má pravdivostní predikát pro každý jazyk v jazyku o řád vyšším, Kripkeho řešení spočívá ve snaze vytvořit jediný parciální pravdivostní predikát aplikovatelný na sebe sama. Ačkoliv Kripke zredukoval počet pravdivostních predikátů na jeden, který se nenachází v metajazyku, stejně se mu nepodařilo úplně se vyhnout všem důsledkům Tarského teorie, jelikož pojmy jako „uzemněnost“ a „paradoxní“ se nacházejí v metajazyku.⁵⁵

Závěr

Záměrem první části mé studie bylo podat stručný výklad Tarského sémantické koncepce pravdy, tak, jak je prezentována v dvou jeho nejznámějších statích [CTFL] a [SCT]. Hlavním zdrojem byla sice první jmenovaná stať, ale jelikož byla v některých ohledech Tarským možná až příliš zjednodušená, pokusil jsem se svůj výklad doplnit za pomoci informací ze statí druhé. To se týká především kapitol 2.1, 2.3 a 4. Oproti tomu kapitola 2.2 byla prakticky celá inspirována sekundární literaturou. Snažil jsem se podat přehled základních principů Tarského koncepce tak, aby byly zmíněny všechny její důležité aspekty, včetně těch, které nejsou v [SCT] explicitně uvedeny, přičemž mým cílem byla snaha, aby tento výklad byl co možná nejsrozumitelnější.

Stejně jako Tarski v [SCT], i já se rozhodl věnovat druhou část svého článku kritice Tarského sémantické koncepce pravdy. Mým záměrem nebylo podat úplný výčet kritik vůči této koncepci vznesených, jelikož takový výčet by sám vydal na velmi obsáhlou práci. Místo toho jsem se pokusil vyložit alespoň ty z nich, které napadají nosné myšlenky Tarského pojetí, a zároveň jsem se u některých pokusil poukázat na jejich nedostatky.

I když Tarského koncepce rozhodně není bezchybná, jak ukázaly i tyto kritiky, přesto je i nadále velmi inspirativní a nezbyvá mi než spolu s Kripkem dodat: „Duch Tarského hierarchie je stále s námi.“⁵⁶

Literatura

Aristotelés. (2003) *Metafyzika*. Přel. Antonín Kříž. Praha: Nakladatelství Petr Rezek.

⁵⁴ Viz (Soames, 1999 s. 181–195).

⁵⁵ Srovnej (Field, 2008, s. 222–223).

⁵⁶ „The ghost of the Tarski hierarchy is still with us.“ (Kripke, 1975, s. 714)

Boisvert, D. R. (1999) The Trouble with Harrison's 'The Trouble with Tarski'. *The Philosophical Quarterly*, vol. 49, no. 196, s. 376–383. Dostupné z WWW: <<http://www.jstor.org/stable/2660401>>.

Carnap, R. (1968) *Problémy jazyka vědy: (výbor prací Rudolfa Carnapa)*. Praha: Svoboda.

Davidson, D. (1996) The Folly of Trying to Define Truth. *The Journal of Philosophy*, vol. 93, no. 6, s. 263–278. Dostupné z WWW: <<http://www.jstor.org/stable/2941075>>.

Field, H. (2008) *Saving Truth from Paradox*. Oxford: Oxford University.

Gómez-Torrente, M. (2004) The indefinability of truth in the "Wahrheitsbegriff". *Annals of Pure and Applied Logic*, vol. 126, s. 27–37.

Harrison, J. (1998) The Trouble with Tarski. *The Philosophical Quarterly*, vol. 48, no. 190, s. 1–22. Dostupné z WWW: <<http://www.jstor.org/stable/2660369>>.

Heck, R. G. (1997) Tarski, Truth, and Semantics. *The Philosophical Review*, vol. 106, no. 4, s. 533–554. Dostupné z WWW: <<http://www.jstor.org/stable/2998511>>.

Herzberger, H. G. (1967) The Truth-Conditional Consistency of Natural Languages. *The Journal of Philosophy*, vol. 64, no. 2, s. 29–35. Dostupné z WWW: <<http://www.jstor.org/stable/2023768>>.

Kirkham, R. L. (1995) *Theories of Truth: A Critical Introduction*. Cambridge (MA): MIT Press.

Kripke, S. A. (1975) Outline of a Theory of Truth. *The Journal of Philosophy*, vol. 7, no. 19, s. 690–716. Dostupné z WWW: <<http://www.jstor.org/stable/2024634>>.

Lynch, M. B. (ed.) (2001) *The Nature of Truth: Classic and Contemporary Perspectives*. Cambridge (MA): MIT Press.

Martin, R. L. (1967) Toward a Solution to the Liar Paradox. *The Philosophical Review*, vol. 76, no. 3, s. 279–311. Dostupné z WWW: <<http://www.jstor.org/stable/2183621>>.

Milne, P. (1997) Tarski on Truth and Its Definition [online]. [cit. 2010-02-12]. Dostupné z WWW: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.712&rep=rep1&type=pdf>>.

Raclavský, J. (2009) Lhářský paradox, význam a pravdivost [online]. [cit. 2010-02-12]. Dostupné z WWW: <http://www.phil.muni.cz/~raclavsk/texty/lharske_paradoxy.pdf>.

Raclavský, J. (2007) Paradox lháře a jeho řešení [online]. [cit. 2010-02-12]. Dostupné z WWW: <http://www.phil.muni.cz/~raclavsk/texty/paradox_lhare.pdf>.

Soames, S. (1999) *Understanding Truth*. Oxford, New York: Oxford University Press.

Suppes, P. (1988) Philosophical Implications of Tarski's Work. *The Journal of Symbolic Logic*. Vol. 53, no. 1, s. 80–91. Dostupné z WWW: <<http://www.jstor.org/stable/2274428>>.

Tarski, A. (2006) Sémantická koncepce pravdy a základy sémantiky. In Peregrin J. (ed.) *Logika 20. století: Mezi filosofií a matematikou: výběr textů k moderní logice*. Praha: Filosofia. s. 135–176.

Tarski, A. (1956) The Concept of Truth in Formalized Languages. In *Logic, Semantics, Metamathematics*. Oxford; New York: Oxford University Press. s. 152–278.

Tarski, A. (1944) The Semantic Conception of Truth: and the Foundations of Semantics. *Philosophy and Phenomenological Research*, vol. 4, no. 3, s. 341–376. Dostupné z WWW: <<http://www.jstor.org/stable/2102968>>.