LEIDEN UNIVERSITY

MASTER THESIS

Substitutional Quantification, Satisfaction and Denotation

Author:
Cas Bezembinder

Supervisor:
Prof . Göran Sundholm

A thesis submitted in fulfillment of the requirements
for the programme of History and Philosophy of the Sciences

Institute for Philosophy

June 7, 2018

LEIDEN UNIVERSITY

Abstract

Faculty of Humanities

Institute for Philosophy

History and Philosophy of the Sciences

Substitutional Quantification, Satisfaction and Denotation

by Cas Bezembinder

This thesis aimed to explain the differences between the substitutional and the referential quantifier. It did so firstly by presenting an analysis of the discussion between Wallace and Kripke, secondly by analysing Tarski's reasons for introducing satisfaction, and finally by looking at whether Kripke's definition manages to avoid the issues that motivated Tarski. It concluded that an essential part of the recursive truth-definition given by Kripke is the assumption of a pre-given truth-definition for an atomic language. However, if this atomic language is of infinite size, satisfaction is needed to provide this definition. Rather than a differing in whether they use of satisfaction, this thesis argues that the substitutional and the referential quantifiers differ in where in the definition of truth they make the connection between truth and the world.

Word Count: 17,693

Contents

Chapter 1

Introduction

Traditionally, quantification is seen as a function ranging over objects, called referential quantification. Informally, according to this interpretation, an existential quantifier $(\exists x_i)P(x_i)$ is true if and only if there is an object that has the quality P which satisfies the sentential function $P(x_i)$. For example, 'There is something that is white' is true if there is an object such as snow, which is white. Similarly, with the universal quantifier, a statement $(\forall x_i)P(x_i)$ is true if every object has the quality P. However, this is not the only interpretation of the quantifiers. An alternative is substitutional quantification, in which quantification is not seen as ranging over objects, but over terms of a language. According to this interpretation, an existential statement of the form $(\exists x_i)P(x_i)$ is true if and only if there is a true sentence of the form $P(o)$ where object o is a part of the substitution class of the language. As an example, 'There is something that is white' would be considered true if there is a term such as 'snow' and the sentence 'snow is white' is true. The substitutional quantifier, by not quantifying over objects but over phrases, is said to avoid Quinean ontological commitment and thus to be ontologically lighter. On the other hand, it is argued that by not ranging over objects the substitutional quantifier does not maintain the connection between language and the world.

Wallace ("Frame of Reference"; "Convention T") and Tharp argued for referential quantification in a few influential papers. They write that the claimed benefits of substitutional quantification are not actually present because one cannot give a correct truth-definition for a substitutional quantifier without covert appeal to a notion of satisfaction. Instead, satisfaction can be found in these definitions, even if it is not immediately obvious. Because of this, both theories have the same ontological commitments.

As a response to this, Saul Kripke published "Is There a Problem about Substitutional Quantification". Kripke argues that there is no problem at all with giving a truth-definition of a substitutional language that does not make use of satisfaction. He shows that both interpretations can have correct truth-definitions and that they are, properly speaking, two independent concepts. The quantifiers can even be combined or used in the same language. Kripke argues that the choice of which quantifier to use depends on the system in which we use it, and that in many instances the choice of quantifier is unimportant. This paper virtually ended the discussion on substitutional and referential quantification.

While the discussion has ended, this does not mean that all questions surrounding referential and substitutional quantification have been answered. While Kripke has convincingly argued that both quantifiers can be given a definition, this does not explain wherein exactly the difference between the two lies, nor where the difference originates from. This thesis will attempt to answer what the differences between substitutional and referential quantification are and what causes these differences.

It will do this by providing an overview of the discussion so far, by first of all giving an account of Wallace's argument that the substitutional quantifier needs to make use of satisfaction, and secondly by giving an account of the arguments made by Kripke in "Is There a Problem about Substitutional Quantification". After this, in the second part of this thesis, an analysis will be given of the reason for and the importance of satisfaction, and the question whether terms must denote.

This thesis will argue that proper analysis shows that satisfaction cannot be avoided for a significant class of substitutional languages, namely those with an infinite amount of atomic sentences, and that the difference between the two quantifiers instead lies in at which point of the construction of the language the connection between language and the world is made. Furthermore, it will argue that many of the points raised as the philosophical significance of satisfaction are instead already present in the notions of 'language' and 'truth'.

Chapter 2

Wallace, Satisfaction and Substitutional Quantification

As stated in the introduction, one of the main opponents of the idea that substitutional quantification provides a reasonable alternative to referential quantification is John Wallace. Wallace argues that for a truth definition of a substitutional language to fulfil the requirements of a truth definition, axioms need to be added beyond a formalisation of the definition of substitutional quantification. However, he argues, these axioms would introduce a notion of satisfaction to the substitutional language and thus eradicate any relevant differences between the two quantifiers. This chapter will analyse the argument put forward by Wallace by first looking at the goals and aims that he sets for a theory of truth, and his justifications for this. Then, it will look at his definitions of the referential and the substitutional quantifier. Thirdly, it will look at the problems that Wallace argues the substitutional quantifier has, and finally it will look at the relation between the new axioms and satisfaction. As a whole, this chapter aims to give a comprehensive overview of the argument put forward by Wallace, rather than critically assess the arguments themselves, which will take place in later chapters.

2.1 Conditions of Truth-Theories

Before we can discuss different theories of truth, it is first necessary to discuss what we view as successful theories of truth, and what conditions we impose on these theories. For Wallace, these conditions are largely determined by Tarski. Wallace, on the basis of the works of Tarski, places four requirements on any theory of truth.

First of all, the concept of truth has to be relative to a language. Truth is taken to be a predicate of sentences, which in turn are a specific combination of symbols. It is very possible, likely even, that a specific combination of symbols has different meanings in different languages, and that because of this different truth-values are ascribed to it for different languages.

Secondly, a language cannot contain its own truth predicate. This is done to avoid situations like the Liar's paradox, where for a sentence like 'This sentence is false' we would need to construct a partial definition of truth that states that 'This sentence is true if and only if this sentence is false', which would result in a paradox. Because of this, we have to make use both of an object language, which is the language for which we aim to give a truth-definition, and a metalanguage, in which we intend to give the truth-definition of the object language.

Thirdly, only complete sentences of a language can be true or false. In cases in which the sentence is not closed and still contains open variables, it cannot be given a truth-value and must remain undetermined.

Finally, Wallace argues that any definition of truth has to fulfil Tarski's convention T, which states that for any sentence such as 'snow is white', we are able to construct a sentence of the form:

'Snow is white' is true if and only if snow is white.

where the first half of the biconditional is the sentence for which we are defining truth, and the second half is a translation of this sentence into the language in which we are defining truth. Wallace writes that

> *Tarski's next point concerns the role of partial definitions of truth, i.e., sentences of the form of 'Lizzy is playful' is true if and only if Lizzy is playful, in setting adequacy conditions for a theory of truth. Roughly, an adequate theory of truth for a given object language must explain each partial definition of truth.* ("Frame of Reference", 120)

Taken together, these four conditions place several limitations on the metalanguage. It needs to have sufficient expressive power to be able to both name and translate every sentence of the object language. It also needs to include a notion of logical consequence, as logical consequence is necessary to form the sentences of convention T. Furthermore, it needs to include every sentence of the form '$T(\overline{p})$ if and only if p', where $\overline{p}$ is the canonical name of a sentence of the object language and p is its translation (Wallace, "Convention T" 201).

The notion of translation has not been explained. Wallace avoids the question in the following way:

> *What is a good or correct translation between two languages is of course a difficult question on which there is little substantive agreement and less hard theory. In the case of truth theories taken up in this paper the general problem*

> *of translation is bypassed in two ways: (1) we shall usually be interested in cases where the object language is included in the metalanguage, e.g., ML = OL + a bit of recursive apparatus; in these cases we study the identity translation, which seems clearly to be a good one; (2) the main results of the paper extend to large classes of translations, but these classes can be characterized in terms of abstract structural features of translations, e.g., translation such that every sentence in a certain class of metalanguage sentences is a translation of something.* ("Frame of Reference", 122-123)

Wallace argues that even without properly defining translation we can look at theories of truth by letting the example definition be given in a metalanguage that includes the object language itself. If the object language is included, translation becomes trivial. This result can then afterwards be extended to other types of translation.

Wallace further states the condition that a theory of truth should be finitely axiomized. He gives several reason for this. In *Convention T and Substitutional Quantification* he states that this condition is necessary to avoid trivial definitions of truth.

> *As described so far, Convention T would be satisfied by a theory that took all target conditionals as axioms. To guard against this trivialization one may want to add the condition that an adequate theory of truth be finitely axiomatized.* (201)

In "On the Frame of Reference" he further justifies why trivial theories of truth should be rejected.

> *We want a theory of truth to be analytic or recursive, picking out a finite*

> *number of compoundable features of sentences whose effect on truth conditions are uniform. On the other hand, we want not to say in advance what the compoundable features and their effects are. A recursion we want; the shape or strategy of the recursion we want to leave open. The demand for a finite theory seems a – perhaps crude – way of satisfying both desires. To demand only that the theory be recursively axiomatizable is too weak: the set of partial definitions is itself recursive. But obviously no analysis is achieved by a theory that takes all "target" biconditionals as axioms.* (122)

If a trivial theory of truth can be given simply by taking all the target biconditionals, Wallace argues, this would not reveal any structure to truth, and it would not teach us anything about the nature of truth.

2.2 Satisfaction Definition of Truth

The theory of truth of which Wallace himself is a proponent defines truth via a notion of satisfaction. This approach defines a set of atomic sentential functions, which are similar to sentences, with the exception that they contain open variables. For example, we could have the sentential function

x is white.

These atomic sentential functions can be used to build more complex sentential functions through conjunction, disjunction, negation and quantification.

We can define a satisfaction-predicate for sentential functions by stating for each atomic sentential function what kind of objects would satisfy it. To come back to the previous example, 'x is white' would be satisfied by all white things. Satisfaction can then be recursively defined for all sentential functions by saying that a negation of a sentential function is satisfied by a sequence if and only if the sentential function is not, a conjunction is satisfied if both parts are satisfied, and an existential quantification is satisfied by a sequence if a sequence that differs at most in the relevant term would satisfy the sentence. Sentences are defined as those sentential functions without open variables. The truth of a sentence is defined on the basis of satisfaction. Wallace takes a similar approach to this informal definition of satisfaction ("Frame of Reference", 125) Languages defined this way are called referential languages, and the quantifiers used is called the referential quantifier.

Wallace formalises this definition for a metalanguage which not only includes the predicates and the logical apparatus of the object language, but also arithmetic, a syntactical apparatus for structural-descriptive names, a two-place predicate 'Sat', a one-place predicate 'Seq' and a two-place function sign 'Val'. ("Frame of Reference", 125) The two-place predicate 'Sat' refers to the satisfaction of a sentential function by a sequence s. The one-place predicate 'Seq' says of an object that it is a sequence. As for 'Val', Val(s, n) refers to the nth member of a sequence s.

Furthermore, Wallace uses the expression '$s \approx_n s'$' as an abbreviation for '$(m)(m \neq n \Rightarrow Val(s, m) = Val(s', m))$', which says of a sequence s that it differs from s' in at most the nth place. After this, he defines formally satisfaction as

(1′) s Sat neg(f) $\iff$ $\neg$ *(s Sat f)*

(2′) s Sat conj(f, g) $\iff$ *(s Sat f & s Sat g)*

(3′) s Sat exquant (f, x_n) $\iff$ *(Es′)(s $\approx$ s′ & s′ Sat f)*

(4′) s Sat pred ($\overline{before}$, x_m, x_n) $\iff$ *Val(s, m) is before Val(s, n)* ("Frame of Reference", 125)

To round off the theory of truth, it is necessary to have two further axioms. First of all there needs to be an axiom that ensures that there are sequences.

(5′) ($\exists$x)(Seq x) ("Frame of Reference", 125)

Secondly, sequences need to be able to vary freely over a certain range.

(6′) ($\exists$s′)(s $\approx$ s′ & Val(s′, n) = x) ("Frame of Reference", 126)

Wallace states that with these axioms for a sentence $Fx_1...x_n$with only those $x_1, \ldots\ x_n$ variables free and a sentence '$F\ Val(s,k_1)...Val(s,k_n)$' that is the result of replacing all those variables with a term from the sequence, we can prove every sentence of the form:

(A) s Sat $\overline{Fx_1...x_n} \iff F\ Val(s,k_1)\ ...\ Val(s,k_n)$ ("Frame of Reference", 126)

This shows that satisfaction is closely related to the truth of the sentences that are a result of replacing the variables in a sentential functions with terms. Wallace writes:

> *'The present theory gives 's Sat f' the intuitive sense: f comes out true when the reference of each of its free variables is fixed to be the value of that variable under s.*

("Frame of Reference", 126)

As outlined at the start of this chapter, the truth of a sentence is defined on the basis of satisfaction, in that a sentence is considered true if and only if all sequences satisfy it, and false if and only if no sequences satisfy it.

> *Evidently, whether a sequence satisfies a sentence depends only on what the sequence assigns to numbers corresponding to variables free in the sentence. If the sentence is closed, whether it is satisfied by a sequence does not depend at all on what the members of the sequence are. A closed sentence is true if it is satisfied by every sequence, and false otherwise.'* ("Frame of Reference", 126)

While this alone is sufficient to express a complete recursive theory of truth on the basis of satisfaction, Wallace adds two further axioms.

> Hilbert and Bernays have observed that every first-order language with a finite primitive vocabulary of predicates contains an open sentence *'Rxy'* with two free variables such that
>
> *(x)Rxx*
>
> and all sentences of the form
>
> $(x)(y)(Rxy \Rightarrow (Fx \iff Fy))$ (Hilbert & Bernays, pp. 381ff, via Wallace, "Frame of Reference", 126)

where we can take *Rxy* to mean that *x* and *y* are indistinguishable relative to the ontology of the language. Wallace takes 'Eq' as the structural-descriptive name of this predicate, which is assumed to range over the first two variables of the language, and 'Eq′' as the result of identifying the variables in Eq. He then formalises the results of Hilbert and Bernays as

> *(C) s Sat Eq′* ("Frame of Reference", 127)

and from the existence of Eq proves

> *(D) <x, y> Sat Eq ⇒ (<...x...> Sat f ⟺ <...y...> Sat f)* (Wallace, "Frame of Reference", 127)

This states that if two objects are indistinguishable by a language, a sequence containing the one will satisfy a sentential function of the language if and only if a sequence containing the other in its place would do so as well.

Wallace defines the totally of his theory of truth on the basis of satisfaction as follows:

> *Five formulas now sum up the effects of a satisfactional truth theory that produces the homophonic partial definitions of truth.*
>
> *(I)* $True(\overline{F}) \iff F$
>
> *(II) (x)(True(x) ⟺ x is a closed sentence & (s)(Seq s ⇒ s Sat x))*
>
> *(III)* $s\ Sat\ \overline{Fx_1...x_n} \iff F\ Val(s, k_1)\ ...\ Val(s, k_n)$
>
> *(IV) s Sat Eq′*
>
> *(V) <x, y> Sat Eq ⇒ (<...x...> Sat f ⟺ <...y...> Sat f)* ("Frame of Reference", 127)

These five formulae express the result of the theory of truth defined through satisfaction that Wallace puts forward. It includes all partial definitions of truth, states that a sentence is true if it is satisfied by every sequence, that if a sequence s satisfies F, the result of replacing the variables of F with the terms of the s will result in a true sentence, and that two objects that are indistinguishable satisfy the same things.

2.3 Substitutional Theory of Truth

We have seen that the referential theory of truth functions through the definition of a satisfaction relation, which says of a predicate what objects would make it true. A substitutional theory of truth, on the other hand, does not define truth via a recursive definition of satisfaction, but aims to give a direct recursive definition of truth. Instead of using a universe of objects to quantify over, quantification in this theory happens via substitution. This means that rather than looking at whether an object satisfies the sentential function, we simply take an existentially quantified sentence such as '*(∃x) x is white*' to be true if there is a sentence that is the result of replacing the variable with a term of a specifically-defined substitution class of the language and this sentence is true, such as '*snow is white*'.

Wallace gives three different formalisations of this notion, namely the Naive substitution interpretation, the Hilbert-Bernays substitution interpretation and the McKinsey substitution interpretation ("Frame of Reference"). The Naive interpretation will be focussed on here, as the other two are mostly variants of the first.

The recursion of the Naive substitution interpretation of quantification languages can be given as follows: a negation is true if and only if what is being negated is not true, a conjunction is true if and only if both of the conjuncts are true, and an existential quantification $(\exists x_n)(Fx_n)$ is true if and only if some substitution instance of Fx_n is true. Wallace expresses this recursion as follows:

> *[7] True (neg(f))* $\Longleftrightarrow$ *¬ True(f)*
>
> *[8] True (conj (f, g))* $\Longleftrightarrow$ *(True(f) & True(g))*
>
> *[9] True (exquant (f,* v_n*)* $\Longleftrightarrow$ *(∃a)(S-class(a) & True (sub (a,* v_n*, f))).*
>
> ("Frame of Reference", 128)

Here, '*sub (a,* v_n*, f)*' stands for the substitution of all free occurrences of v_n in f with a. These three definitions do not provide a recursive definition, as there are no base cases: no use is made of atomic predicates. To give a full definition, it is necessary to give the truth for a set of atomic sentences.

2.4 Consequences of Satisfaction

While the question of how we should categorise truth seems like a natural one, in that it is an often-used term and we might want to know when it applies and how we should use it, this is not Wallace's core intention for his paper. For Wallace, the question whether truth has to be defined via satisfaction is not just a technical detail of the construction of a successful truth-definition, but itself tells us something about the notion of true: if it were to be successfully shown whether truth has to be defined via satisfaction or not, this teaches us something meaningful and concrete about truth. He writes that:

> *From a semantical point of view, quantification, ontology, predication and extensionality form a single structure. The semantical interpretation of predication is that predicates are satisfied by object, independently of how objects are described; satisfiers of quantified sentences are determined by satisfiers of their predicate parts, relative to a universe of discourse. This semantical point*

> *of view is Tarski's; the referential structure it brings to light is exactly represented in Tarski's theory of truth for quantificational languages.* ("Frame of Reference", 117)

He argues that a definition of truth that occurs via satisfaction is one that shows that there is a relation between predication and ontology, independent of how we describe this ontology. If all definitions of truth would go via satisfaction, Wallace argues that this would imply a strong form of the thesis of extensionality, which says that once all obscurities and confusions are removed from an ordinary language, the entire notion of truth for that language can be expressed through extensional language.

An absence of satisfaction would, according to Wallace, break this relationship between truth and ontology.

> *In making no appeal to a range of quantification the substitutional recursion seems to break the traditionally held connection between truth and extra-linguistic things. And coordinately, in making no appeal to a relation of satisfaction or denotation, it seems to undermine the importance traditionally attached to predication.* ("Convention T", 200)

Wallace argues that without satisfaction, we can talk about the truth of a sentence without talking about the world, and we can define the truth of sentences without looking at the relation between predicates and objects.

The question then is for Wallace, is it in fact necessary that any definition of truth moves via satisfaction? He states:

> *The question naturally arises: is the referential structure Tarski's theory represents*

categorial in character? Is it contained in every reasonable theory of truth? Or is it an artifact of Tarski's approach? ("Frame of Reference", 118)

Wallace argues that if the notion of substitution is one that is implicitly present in any definition of truth, this says something important about the nature of truth. Because of this, it is necessary to analyse whether alternatives are possible.

2.5 Does Substitution Entail Satisfaction?

In previous sections we have seen the standard definition of truth via satisfaction and its main rival, a direct definition of truth with substitutional quantification and a set of base-cases. We have also seen why Wallace assigns such an importance to the existence of a need for satisfaction. Wallace aims for his paper to prove that the substitutional interpretation does in fact contain a satisfaction notion. He states that

> *In essence, my claim will be this: if a finite theory puts enough conditions on 'true' to entail all instances of [I] (i.e., if it meets Tarski's Convention T) then it puts enough conditions on some relation 'satisfies' and on some name 'Eq' to entail [III] and all instances of [II]. Tarski showed that satisfaction is a way to truth; I hope to persuade you that it is the way.* ("Frame of Reference", 118)

This section will outline his argument for the presence of satisfaction in the substitutional interpretation.

We saw previously that beyond defining the effect of the operators of truth as done in (7)-(9), there was also need for a base case for the recursion.

Wallace argues that this base case cannot be provided by simply adding as axioms a partial definition of truth for each atomic sentence. He argues that if we have axioms of the form

$$True(\overline{\textit{Able is a man}}) \iff \textit{Able is a man.}$$
$$True(\overline{\textit{Baker is a man}}) \iff \textit{Baker is a man.}$$
$$True(\overline{\textit{Cain is a man}}) \iff \textit{Cain is a man.}$$

Then we cannot derive the partial truth definition

$$True(\overline{(\exists x)(x \textit{ is a man})}) \iff (\exists x)(x \textit{ is a man})$$

He argues that if we assume

$$True(\overline{(\exists x)(x \textit{ is a man})})$$

we can from (9) deduce that

$$(\exists a)(\textit{S-class}(a) \mathbin{\&} True\ (sub\ (a, \overline{x}, \overline{x \textit{ is a man}})))$$

but that from here we cannot reach the wanted partial definition. Considering the conditional from right to left, Wallace states that if we assume

$$(\exists x)(x \textit{ is a man})$$

we can use quantifier elimination to get

m *is a man*

for a singular term '*m*' that is new for the theory, and thus is not one of the terms previously introduces in the substitution class. No further conclusions about individual sentences can be made, Wallace argues, and once again no partial definition of truth can be reached ("Frame of Reference", 129).

If this argument is correct, it shows that a set axioms similar to those introduced by (7)-(9) together with some atomic base cases does not fulfil convention T, and should, if we accept convention T as a requirement, either be strengthened or rejected.

Wallace proposes that the substitutional theory can be strengthened to fulfil convention T by adding the following further axioms:

$[10]\ True\ (pred\ (\overline{man}, a)) \iff den\ (a)\ is\ a\ man.$

$[11]\ den\ (\overline{Able}) = Able$

$den\ (\overline{Baker}) = Baker$

$den\ (\overline{Cain}) = Cain$

$[12]\ (\exists a)(den\ (a) = x)$ ("Frame of Reference", 130)

Where (10) says that a predication of something in the substitution class is only true if what is being denoted has that predicate, (11) defines the denotation of each term of the substitution class, and (12) states that everything is denoted by some member of the substitution class.

A theory including (7)-(12) and a set of atomic definitions fulfills Convention T and only assigns truth to closed sentences. It thus fulfils the conditions

(I) and (II) of Wallace's definition via satisfaction. However, when these axioms are added, satisfaction can also be found. Wallace states that:

> *... it is easy to find 'Seq', 'Sat', and 'Val' in it. A sequence is a sequence of members of the substitution class. 's Sat ($Fx_1...x_n$)' means that the result of substituting the k_1th, ...k_nth member of s for the corresponding variables in $Fx_1...x_n$ is True. Val(s, n) is the denotation of the* nth *member of* s. *Eq is obtained by the same method as the satisfaction theory. With these definitions, it is straightforward to verify that (III), (IV), and (V) are provable."* ("Frame of Reference", 131)

If Wallace's argument is correct, then the substitutional definition covertly contains a notion of satisfaction, and because of this, truth functions in exactly the same way as in the substitutional definition, namely with truth being determined by predicates that are true of objects. If substitutional theories of truth make use of satisfaction, for Wallace this means that there can be no question whether it is possible to define truth without a commitment to a universe of objects.

According to Wallace, this result also holds for the Hilbert-Bernays and the McKinsey interpretation. The Hilbert-Bernays is a more finely structured version of the Naive interpretation that uses Gödel numbering as structural-descriptive names, and the McKinsey interpretation defines truth of sentences on the basis of truth of their set-theoretic analogues. Wallace argues that his argument applies to both.

2.6 Conclusion

This chapter has presented the argument put forward by Wallace. Wallace argues that the axioms that are provided to give a truth-definition of a substitutional language do not fulfill the conditions set for truth-definitions, specifically Convention T, as there are certain partial definitions that cannot be reached. To entail all partial definitions, further axioms would need to be added that describe the denotation of the terms, ensure that a term is only ascribed a predicate if its denotation has that quality, and ensure that everything is denoted by something. However, once this is done, it is possible to find satisfaction in the definition. Because of this, Wallace argues that either a substitutional truth-definition fails the essential demands of a truth-definition, or covertly appeals to a notion of satisfaction.

Chapter 3

Kripke On Quantification

As a response to the two papers written by Wallace and, to a lesser extent, the paper written by Tharp which reject the idea that one can have a general definition of truth for a substitutional language without also containing in one way or another a notion of satisfaction, Kripke wrote the paper "Is There a Problem of Substitutional Quantification?". In this paper, he responds to Wallace's arguments and aims to show that it is in fact possible to give a definition of a substitutional language that does not make use of satisfaction.

Kripke summarises the central point of the papers by Wallace and Tharp as follows:

> *The claim seems to be that, contrary to the usual impression, a careful examination of truth definitions for the substitutional quantifier will show that these definitions, if they succeed at all, must make a covert appeal to some range for the variables.* (326)

This is also the argument that this paper is primarily aimed at. This chapter will outline the arguments put forward by Kripke and see how they relate to Wallace's arguments.

Kripke states that much of the confusion between substitutional and referential quantification is a result of typographical issues, namely that both quantifiers are denoted by the same symbol. To avoid this mistake, Kripke introduces separate notation for the referential and the substitutional quantifier, with the referential quantifier and variables using the traditional notation, and the substitutional quantifier being written as (Σx) for the existential quantifier and (Πx) for the universal quantifier, with substitutional variables being unitalisized. This thesis will follow this typographical decision.

3.1 Definition of Truth for Substitutional Languages

Kripke defines a substitutional theory of truth as follows: first, we take a language L_0, the sentences of which are assumed to be effectively specified syntactically. This language will provide the set of atomic sentences for the larger language *L*. We also assume that L_0 contains a non-empty class *C* of expressions which will be the substitution class. The members of *C* are called terms and can be any class of expressions of L_0, be it individual words, sentences, or even parentheses. Kripke explicitly states that we do not assume that the terms of *C* denote, or are syntactically similar to the terms of a referential language.

The result of replacing zero or more terms of a sentence *A* of L_0 with variables not contained in L_0 results in an (atomic) preformula or preform. If the result of replacing a variable with a term is itself a sentence of L_0, the preformula is called an (atomic) form.

If there is an effective test of formhood, all forms are taken to be atomic formulae of *L*. If such a test is not available, a specific set of forms of L_0 is taken as the set of atomic formulae. We assume that these include all the sentences of L_o and that if $\phi(x_{i_1}...x_{i_n})$ is an atomic formula, so is the result of replacing the listed variables with any other variables.

Kripke then recursively defines the set of formulas of L. He states as a base case that atomic formulae are formulae. Then, he states that if ϕ and ψ are formulae, so are $\phi \wedge \psi$, $\neg\phi$ and $\Sigma x_i \phi$, where the truth-functions and existential substitutional quantifiers are new notations not already present in L_o. Sentences of L are defined on the basis of formulae of L, by saying that the sentences of L are those formulae that do not have any free variables.

Having defined the language of L, Kripke presents a definition of truth of sentences of L. He assumes that truth has already been defined for sentences of L_o. This truth-definition of L_o can then be extended to the entirety of L by saying that:

> *[13]* $\neg\phi$ *is true iff* ϕ *is not;*
> *[14]* $\phi \wedge \psi$ *is true iff* ϕ *is and* ψ *is;*
> *[15]* $\Sigma \mathrm{x}_i \phi$ *is true iff there is a term* t *such that* ϕ' *is true, where* ϕ' *comes from* ϕ *by replacing all free occurrences of* x_i *by* t. (Kripke, 330)

Kripke then proves that given any set S of sentences of L_o which we take to be true sentences, there is a unique set S' of truths of all of L satisfying (13)-(15) and coinciding with S on L_o.

Kripke takes this as a proof that truth of L has been uniquely characterised. He states that:

> *I would have thought that any mathematical logician at this point would conclude that truth for L has been characterized uniquely. If someone asserts a*

> *formula ϕ of L, we know precisely under what conditions his assertion would be true. (Namely, that ϕ is a member of a set satisfying [13]-[15] and coinciding with S on L_o.)* (333)

Furthermore, he states that "There is no hidden fallacy in the proof that conditions [13] – [15] uniquely extend a truth concept for Lo to one for L. (If Wallace means to deny this, he is wrong)" (333).

Thus, while Wallace argues that no truth definition for substitutional quantification can exist without collapsing into referential quantification, Kripke gives a truth-definition that seem to do exactly that. Kripke writes:

> *There is one point which, indeed, deserves strong emphasis. This is that the use of substitutional quantifiers cannot per se be thought of as guaranteeing freedom from 'ontological commitment' (other than to expressions). It is true that the clauses [13]-[15] can be stated without mentioning entities other than expressions and the entities mentioned in characterizing truth for L_o. If other entities are mentioned in characterizing truth for L_o, they still are used when the notion is extended to L.* (1976, p. 333)

However, Kripke states that this shows nothing about the collapse of the substitutional definition into the referential one, for "... there is no reason to think that other entities are used in every truth characterization for every L_o; such an assumption would obviously be false" (333).

Kripke also argues that we reach conceptual problems if we take the terms of the substitution class to necessarily denote, as the substitution class can not only consist of words, but also sentences, connectives, or parentheses. He states that it is questionable whether we can sensibly say that such things denote.

Similarly, substitutional quantification into opaque contexts is possible because the truth of opaque statements will already be contained in L_o. Extension of this to L is then unproblematic (Kripke, 334).

The first section of Kripke's paper aims to show that it is possible to give a recursive definition of a substitutional language, and his method of showing this is by giving such a definition. He is at this point not occupied yet by how this relates to the arguments of Wallace and Tharp:

Without looking any further at Wallace's and Tharp's arguments, I find myself saddled with a complex dilemma. Perhaps they do not really mean to deny that truth for L is intelligibly characterized, given truth for L_0, regardless of whether L_0 has opacities or any denoting terms. But it is hard for me to interpret them otherwise. Alternatively, they mean to impose additional requirements ... But then, since it is a theorem that truth has been characterized for L, it would seem that either (i) the additional criteria are unjustified, or (ii) they are directed towards some problem other than the intelligibility of 'true in L' ... or (iii) the claims that truth in L … fails to satisfy the additional criteria are incorrect. (335)

After formalising the theory just presented [1], Kripke discusses how Wallace could have missed the adequacy of these axioms. He presents three possible explanations, namely that Wallace missed the necessity of an atomic truth condition, that Wallace focuses too much on finding a homophonic truth theory, which this formalisation cannot present as the object language is not included in the metalanguage, or that he objects to the use of truth predicates on the right side of the partial definitions of truth. The last objection is

[1] The details of which can be found in (Kripke, 337-349)

either wrong, if he objects to the use of 'True-in-L' on the right side, as these sentences do not contain the predicate 'True-in-L', or solvable, if he objects to the use of 'True-in-L_o', as there is no reason why this predicate has to be taken as primitive, or that "*if it is defined, its (explicit) definition* must *involve 'semantical terms'*". (Kripke, 347)

3.2 Satisfaction and Pseudo-Satisfaction

Wallace observes that just as we can characterize truth through Convention T, the concept of satisfaction can be paradigmatically characterized through the formula

$$(x_1)(R(x_1, \overline{\phi(x_1)}) \iff \phi(x_1))$$

Where any two-place predicate $R(x_1, \alpha)$ satisfying the above formula for each formula $\phi(x_1)$ will have the satisfaction relation as its extension. This formula essentially says that for all possible values of x_1, the predicate R holds of it and a specific sentence α with one free variable, only if the replacement of this free variable with the specific value of x_1 is true.

Kripke argues that such a predicate cannot be found in his homophonic theory of truth, and that such a conclusion would be a result of a confusion between the satisfaction predicate and the predicate 'pseudo-satisfaction', which is present in his theory. For this homophonic theory, Kripke defines P-Sat(x_1, α_1) as short for ($\exists\alpha_2$)($\exists\alpha_3$)(Q(x_1, α_2) $\wedge$ T(α_3) $\wedge$ Subst (α_3, α_1, α_2, x_1)) where Q(t, α) functions as a device like quotation which for each term t represents a formula satisfied by t alone. Then it can be shown that

$(\Pi\, x_1)(\text{P-Sat}(x_1, \overline{\phi(x_1)}) \iff \phi(x_1))$

is a theorem for each formula $\phi(x_1)$ with one free substitutional variable (Kripke, 369). Kripke writes that:

> *We can conclude that it induces any relation at all only if (i)* all *the terms are assigned denotations and (ii) it is transparent. As in the case of the Q-formulae, it is easy to see that even in the exceptional case where (i) holds, (ii) will not; the reason is that the definition of P-Sat involves the opaque form Q(x_1, α_2).* (369)

Kripke thus argues that satisfaction cannot be found in his theory, as pseudo-satisfaction involves opaque contexts.

3.3 The Choice Between Quantifiers

As we have seen, Kripke regards both the substitutional and the referential quantifier as intelligible. This does not mean that there are no limits on which quantifier we can use. In the case of the construction of a truth-definition in a metalanguage, the decision is made by the object language. Kripke writes that:

> *In particular, if certain variables in the object language are substitutional, they must remain so in the metalanguage. Therefore, the interpretation we gave for Q(*x*, α) is* forced *on us, since its first variable is substitutional. It is not just that we need not read it as Wallace does; we are prohibited from his referential reading.* (377)

Beyond the question of interpretation of variables for the construction of theories of truth, Kripke challenges the intelligibility of the question which quantifier is right. "... the query goes: 'What is the proper interpretation of the quantifier, referential or substitutional?' What can these queries mean?" (377).

Rather than viewing it as two interpretations of the same concept, Kripke views the referential and substitutional quantifiers as two different (but related) concepts. The choice which of these concepts we want to use for a language is mostly determined by the language.

> *If [the queries] refer to uninterpreted first-order quantification theory (the pure predicate calculus) the answer has already been given: both the substitutional and the standard interpretation make all theorems valid. ... The point is that an uninterpreted formal system is just that – uninterpreted; and it is impossible to ask for the 'right' interpretation.* (Kripke, 377)

The question which quantifier is right for uninterpreted language becomes a purely technical question. Kripke argues that there are some situations in which we might give preference to a referential quantifier, such as situations where there are objects for which no terms are available, or situations where we might give preference to a substitutional quantifier, such as when the quantifier would have to range over sentences (378).

In the case where the system is interpreted, these questions are unnecessary, Kripke states. If the language is interpreted, the symbols are already given meaning. In this case, the quantifier that is used is determined by the interpretation.

3.4 Conclusion

In the conclusion to "Is There a Problem about Substitutional Quantification", Kripke concludes that there was never a problem with substitutional quantification, and that for any class of expressions C of a language L_0, we can extend the language with substitutional quantification, independent of whether these expressions denote. For Kripke, "The issue of whether truth conditions have been given for substitutional languages is one of mathematical fact, not philosophical opinion" (406).

The main conclusion of Kripke's paper, that it is possible to give a definition of truth for substitutional languages, is already obtained in its first section. Kripke shows that if we assume the truth of an atomic language L_0, some part of which is a substitution class C, that we can extend this atomic truth predicate to the entirety of a language L, and that it is not necessary that the terms of C denote. This method can also be formalised and used to create a homophonic theory of truth.

Kripke argues that his truth-definition does not include satisfaction, as the equivalent in his theory, Pseudo-satisfaction (P-Sat) includes the predicate $Q(\mathrm{x}_1, \alpha)$ which is opaque. Because of this, he concludes, satisfaction cannot be present in his theory.

Finally, Kripke argues that referential and substitutional quantification are not two interpretations of the same concept, but rather two different concepts, which can even be used together in the same language. The choice of which quantifier we use is primarily determined by the language. If the language is uninterpreted, we can pick any interpretation that makes all theorems come out as true. If the system is interpreted, the choice is determined by the interpretation.

Chapter 4

Tarski and Satisfaction

The first part of this paper showed the arguments presented by Wallace and Tharp that argue that every definition of truth of a language (covertly) makes use of satisfaction, and the arguments by Kripke that reject these conclusions. To put this discussion into context and see what the substantial difference would be between referential quantification defined through satisfaction and substitutional quantification defined through the methods stated by Kripke, it is helpful to first answer the question of why one would need to use satisfaction in the first place. This question is best answered by looking at the works of Tarski. This section will show that the way in which Tarski builds up his languages, combined with his notion of truth, seems to necessitate a notion of satisfaction. It will do this through the analysis of Tarski's works *The Concept of Truth in Formalized Language*, in which Tarski first introduces the method of satisfaction, and "Truth and Proof" in which Tarski most clearly states his conception of truth.

4.1 Tarski and Truth

In both *The Concept of Truth and Formalized Language* and "Truth and Proof", one of Tarski's primary aims is to explain what we mean by the concept of truth, both with respect to natural and formal languages. He aims to give a concrete and precise definition that conforms with our intuitive understanding of the term.

Before we can determine when something is or is not true, it is first important to determine what we ascribe truth to. Traditionally, truth is seen in one of three ways. It can be seen as a quality of a judgement, a quality of a proposition, or as a redundant linguistic element. These are however not the views that Tarski bases his definition on: for Tarski, truth is a semantic notion, and is a quality that is had or not had by sentences. Tarski writes that:

> *In this article, however, we are interested only in what might be called the logical notion of truth. More specifically, we concern ourselves exclusively with the meaning of the term 'true' when this term is used to refer to sentences. Presumably this was the original use of the term 'true' in human language.* ("Truth and Proof", 63)

Thus, the term that Tarski wants to define is the quality of a sentence, and only that, as opposed to the use of true in contexts like 'true friend'.

Secondly, Tarski is not interested in the analysis of truth as a random metamathematical concept, but instead aims to capture the meaning of the concept of 'truth' that we use in our everyday life and seem to have an intuitive grasp of. This intuition, Tarski holds, is captured by the traditional philosophical explanation of truth, for example the one given by Aristotle.

> *Our understanding of the notion of truth seems to agree essentially with various explanations of this notion that have been given in philosophical literature. What may be the earliest explanation can be found in Aristotle's metaphysics: 'to say of what is that it is not, or of what is not that it is, is*

> *false, while to say of what is that it is or of what is not that it is not, is true.'* ("Truth and Proof", 63)

Of this quotation he says that while in modern philosophy alternative formulations have been offered, they more or less aim to capture the same idea, whether it states that a sentence is true if it denotes a state of affairs, or that a sentence is true if it corresponds to reality. At its core, then, Tarski holds a correspondence theory of truth, in which the truth of a sentence in one way or another depends on the content of the sentence matching the way the world is. Tarski prefers the description by Aristotle over some of the modern equivalents due to its technical simplicity, and the goal of a definition of truth should be to capture the basic intentions of this formulation whilst being more precise. ("Truth and Proof", 64)

Tarski has defined truth as a whole as a quality that sentences have if their content matches a state of affairs in the world. An individual sentence, then, is true if what it says is the case, and false if what it says is not the case.

We ask ourselves the question: What do we mean by saying that S is true or that it is false? The answer to this question is simple: in the spirit of Aristotelian explanation ... we arrive at the following formulations: [16] 'snow is white' is true if and only if snow is white. [16′] 'snow is white' is false if and only if snow is not white. Thus [16] and [16′] provide satisfactory explanations of the meaning of the terms 'true' and 'false' when these terms are referred to the sentence 'snow is white'. ("Truth and Proof", 64)

An important remark to make here is that according to the present definition, the truth-definition does not need to be homophonic. In the example of 'snow is white' given above, the truth-definition is homophonic, in that the sentence for which we define truth is repeated exactly in the second half of the definition. However, the definition would work equally well if this was not the case. Thus, we could have an alternative version of (16) that defines not the truth of a particular English sentence but of a particular German sentence as "'Schnee ist weiß' is true if and only if snow is white."

A definition such as was given above for the sentence 'snow is white' can also be made for any other sentence that we come across. For each sentence 'p' we can construct a definition of the form

(17) 'p' is true if and only if p

where 'p' is to be replaced on the left side by the sentence for which we are constructing the definition and on the right side by its translation into the language that we are using.

As a whole, we want to find a semantic definition, which for each sentence of a language says that it is true if it describes a state of affairs in the world, and that state of affairs actually is the case. Furthermore, the semantic definition has to be formally correct and materially adequate, that is to say, for each sentence 'p' of the language, it should entail a definition of the form given by (17). This second requirement is called Convention T.

4.2 Constructing a Definition of Truth

In *The Concept of Truth in Formalized Languages*, Tarski attempts to give a definition along these lines for truth of natural languages. However, this quickly runs into problems due to the fact that natural languages are not semantically closed: they contain semantic terms that refer to the language itself.

For example, it is possible to talk about the truth of English sentences in English. Because of this, we can construct problems such as the Liar Paradox. We construct a sentence 'p' that states "'p' is false.". If we follow the methods previously given, for this sentence we would build the following partial truth-definition for the sentence 'p':

(18) 'p' is true if and only if 'p' is false

This partial truth-definition quite straightforwardly results in a paradox.

Because of this problem, Tarski concludes that truth cannot be defined for natural languages, and he continues with finding a truth definition for formal languages. Formal languages are semantically closed, and because of this cannot lead to versions of the Liar paradox. Tarski defines these languages as languages which "can be roughly characterized as artificially constructed languages in which the sense of every expression is unambiguously determined by its form" (*Concept of Truth*, 166). It is important to note that Tarski does not view formal languages as lacking meaning. He writes that:

> *It remains perhaps to add that we are not interested here in 'formal' languages and sciences in one special sense of the word 'formal', namely sciences to the signs and expressions of which no material sense is attached. For such sciences the problem here discussed has no relevance, it is not even meaningful. We shall always ascribe quite concrete and, for us, intelligible meanings to the signs which occur in the languages we shall consider. The expressions which we call sentences still remain sentences after the signs which occur in them have been translated into colloquial language.* (*Concept of Truth*, 167)

While the language is formally defined and unambiguously determined by its form, it should still be seen as an actual part of language, with concrete and intelligible meaning, not as an abstract object of study. Because of this, the concepts of truth of this formal language should also be intelligible and natural to us.

> *The sentences which are distinguished as axioms seem to us to be materially true, and in choosing rules of inference we are always guided by the principle that when such rules are applied to true sentences, the sentences obtained by their use should also be true.* (Tarski, *Concept of Truth*, 167)

This means that 'true' and 'false' cannot be seen as a more-or-less arbitrary division of sentences into two sets, but should conform to the intuitive correspondence notion that was defined above.

As the formal language itself is semantically closed, it cannot not contain its own truth predicate. Because of this, we cannot give the truth definition in the language for which we want to find a truth definition. Instead, we distinguish between object language, the language for which we aim to give a truth definition, and the metalanguage, the language in which this truth definition is given.

If the language for which we want to define truth is finite, giving a definition is entirely unproblematic. We can for each sentence build a partial definition of the form (17). We can then define the entire notion of truth for the language through the conjunction of all these partial definitions.

However, if the language for which we are trying to construct a truth definition is infinite, this method cannot be successful, as it is impossible to finish such a definition:

> *But the situation is not like this. Whenever a language contains infinitely many sentences, the definition constructed automatically according to the above scheme would have to consist of infinitely many words, and such sentences cannot be formulated either in the metalanguage or in any other language. Our task is thus greatly complicated.* (Tarski, *Concept of Truth*, 188)

Almost all of the formal languages which we are interested in will contain infinitely many sentences, not only owing to the wanted inclusion of standard logical operators such as conjunction, negation and quantification, but also because many of these formal languages require the possibility of talking about infinite or expandable lists of objects. Because of this, another method of giving truth definitions is required.

4.3 Recursive Truth Definitions

The preferred method of giving truth definitions would be through recursion, where some atomic sentences would be determined and the truth-definition of these simple sentences would be explicitly given in the style described above. Then, a list would be given of the operations through which simpler sentences are combined into composite ones, and how the truth of these composite sentences depends on the truth of the simpler sentences. Through this, all sentences of the language could be constructed. However, Tarski argues that this method cannot succeed.

His demonstration of this impossibility starts with an illustration of a language of the calculus of classes. For the calculus of classes, he defines a primitive predicate, namely inclusion, and three fundamental operations by means of which compound expressions are formed from simpler ones, namely negation, logical addition and universal quantification. Then, if we begin with the inclusion *i* of the variables v_k and v_l., written as $i_{k,l}$, and apply the fundamental operations any number of times, we obtain a class of expressions which Tarski calls *sentential functions*. Tarski defines the list of sentential functions by stating that x is a sentential function if and only if (a) k and l are natural numbers and $x = i_{k,l}$, (b) x is the negation of y and y is a sentential function, (c) x is the conjunction of y and z, where y and z are sentential functions, or (d) there is a natural number k and a sentential function y, and x states that for all k, y.

Tarski defines the free variables v_k of a sentential function x as the variables for which k is a number distinct from 0, and for which either (a) $x = i_{k,l}$, (b) x is the negation of y and k is a free variable in y (c) x is the conjunction of y and z, and k is free in either y or z, or (d) there is a natural number l different from k and a sentential function y, and x states that for all l, y, and k is free in y.

Then, he defines sentences as just those sentential functions for which none of the variables are free. Sentences are thus special cases of sentential functions.

However, if we accept this as the structure of the language for which we want to give a truth-definition, the recursive method runs into problems. Tarski writes that:

> *In attempting to realize this idea we are however confronted with a serious obstacle. Even a superficial analysis of Defs 10-12 of Par 2 shows that in general composite sentences are in no way compounds of simple sentences.*

> *Sentential functions do in fact arise in this way from elementary functions, i.e. from inclusions; sentences on the contrary are certain special cases of sentential functions. In view of this fact, no method can be given which would enable us to define the required concept directly by recursive means.* (*Concept of Truth*, 189)

The recursive method cannot work, because the recursive method would need to be recursively defined over sentences, but sentences are not built out of simpler sentences, but out of sentential functions.

4.4 Satisfaction

As we have seen in the previous paragraph, for a language of infinite size, it is impossible to define truth via a recursive method over sentences, as sentences are not made up out of simpler sentences, but out of sentential functions. If a recursive method is still desired, it will have to make use of recursion over sentential functions. However, this cannot be a recursive definition of truth because, as we have seen, truth is a property of sentences. Because of this, a recursive method over sentential functions will have to make use of another predicate, which can then in turn be related to the truth of the specific sentential functions that are sentences. This is also the approach taken by Tarski. He writes that:

> *The possibility suggests itself, however, of introducing a more general concept which is applicable to any sentential function, can be recursively defined, and, when applied to sentences, leads us directly to the concept of truth. These requirements are met by the notion of the* satisfaction of a given sentential

> function *by given objects, and in the present case by given classes of individuals* (*Concept of Truth*, 189).

This is the reason that Tarski introduces a notion of satisfaction: unlike truth, satisfaction as a predicate can range over sentential functions. From this, we can recursively define satisfaction over all sentential functions in the way that was shown by Wallace in the first part of this paper. From satisfaction we can define the truth of sentences, by saying that a true sentence is one that is satisfied by all sequences and a false sentence one that is satisfied by no sequences. From this truth definition we can in turn derive all the partial truth definitions required by convention T.

4.5 Conclusion

In this chapter, we have seen that Tarski viewed truth not as an abstract mathematical notion, but as a notion of natural language which ranges over sentences and of which we have an intuitive grasp, namely that a true sentence is one whose content corresponds with reality. For finite languages we can straightforwardly define truth by saying of each sentence whether it is true or not through an explicit definition. For an infinite language, Tarski argues that such a method cannot work, because we would never be able to finish this explicit definition. Because of this, we need to create a recursive definition.

The recursive definition that would be most preferable is one in which we define the truth of some set of atomic sentences and define several truth-functional operators through which all other sentences are constructed. However, Tarski argues that this is problematic, as sentences are not made up out of simpler sentences, but out of sentential functions, for which truth cannot be defined. As a result, Tarski introduces the notion of satisfaction, which we can recursively define for all sentential functions. From satisfaction, we can derive the truth of a sentence.

Chapter 5

Kripke's Definition of Substitutional Truth

As we have seen in chapter 3, Kripke argues that satisfaction is not required to define truth for a language with substitutional quantification, and gives a recursive truth definition of the basis of a pre-given truth definition of an atomic language L_0. He then argues that this definition clearly shows that satisfaction is not required to define truth. However, in the previous chapter we have seen that Tarski introduced the notion of satisfaction because he was unable to create a recursive definition for truth defined over sentences. This chapter will look more into the recursive definition of truth given by Kripke, and analyse whether, and if so how, Kripke avoids the issues that motivated Tarski to define truth via satisfaction. It will do this by first looking in more detail at the definition given by Kripke, and then by analysing how the atomic truth predicate is defined.

5.1 Kripke's Truth Definition

The truth definition that Kripke proposes appears multiple times throughout "Is There a Problem of Substitutional Quantification", once in section 1, where it is informally defined, once in section 2, where it is given a formal definition, and once in section 5, where a specifically homophonic version is given.

Furthermore, he gives some variants of the definition, such as one where the metalanguage itself also makes use of substitution quantification. However, for this thesis, a focus on the initial informal definition will be enough.

As we have seen, Kripke first introduces a language L_o, the sentences of which will serve as the atomic sentences of the language L. This language L_o is assumed to be effectively specified syntactically. It is also assumed that L_o has a non-empty class C of expressions, the elements of which will be called terms. These terms will form the substitution class for the substitutional quantification. Kripke specifically states that we do not assume that the terms of L_o are syntactically similar to the terms of a referential language, and that terms could be any class of expressions of L_o, such as sentences, connectives or even parentheses.

The language L is an extension of L_o. Kripke defines the formulae of L inductively, taking L_o as the atomic formulae of L, and stating that if ϕ and ψ are formulae, so are $\phi \wedge \psi$, $\neg\phi$ and $(\Sigma x_i)\phi$. It is assumed that the truth-functions and substitutional quantifier are new notions that are not yet present in L_o. If ϕ and ψ are sentences of L, so are $\phi \wedge \psi$ and $\neg\phi$. $(\Sigma x_i)\phi$ is a sentence of L iff ϕ is a formula with at most x_i free.

Kripke then assumes that truth has been defined for the sentences of L_o. Then, he states that the extended truth conditions for sentences of L are:

> *[13] $\neg\phi$ is true iff ϕ is not;*
>
> *[14] $\phi \wedge \psi$ is true iff ϕ is and ψ is;*
>
> *[15] $\Sigma x_i \phi$ is true iff there is a term* t *such that ϕ' is true, where ϕ' comes from ϕ by replacing all free occurrences of x_i by* t. (330)

Important to note here is that the formulae of L, the sentences of L, and the truth of L, are all defined recursively on the basis of the pre-defined language L_0.

5.2 On the Use of L_0

This leads us to an interesting point where Tarski believes that satisfaction is necessary because it is impossible to give a recursive definition of truth, and Kripke states that he has given a recursive definition of truth. How does Kripke avoid the problems raised by Tarski?

A key part of Kripke's proof of the unique existence of a recursive definition of truth for a language L is the existence of a definition of truth for atomic language L_0. To gain more insight into Kripke's definition of truth, it is useful to look more at this notion of atomic truth.

Kripke says several things about the truth definition of the atomic language L_0. First of all, he is of the opinion that it is obvious and unproblematic that one needs to make use of a concept 'true in L_0.' Wallace states of this recursive definition that "I think that some philosophers will resist the idea that [13]-[15] are not already an adequate account of truth conditions" ("Frame of Reference", 233). To this, Kripke replies that philosophers most certainly should resist this idea, because there is nothing wrong with the proof that the conditions (13)-(15) uniquely extend a truth concept for L_0 to one for L. He then states:

> *Perhaps Wallace means that [13]-[15] are not enough by themselves; we must have the concept 'true in L_0' (the set S) already. This is true enough, but no lengthy disquisitions are needed to establish the point. It has nothing*

> *to do with substitutional quantification; it would remain true even if [15] and* (Σx_i) *were dropped from the language and only the truth-functions remained. Of course an inductive definition requires a basis; we can only use [13]-[15] to extend truth for* L_o *to L. Who ever thought otherwise? The phenomenon is characteristic of all recursive definitions.* (333)

Kripke seems to take this to show that there is no reason to worry about the use of a pre-defined truth definition for L_o, however, if this is indeed his argument, he is wrong: while a recursive definition does indeed require a basis, and there is nothing wrong with using a basis in a recursive definition, giving an (attempted) recursive definition does mean that this basis also unproblematically exist. Instead, it means that the basis is required for such a recursive definition, and that if the basis does not exist, the definition fails. Because of this, the above quotation says nothing concrete about the existence and content of the required L_o.

Kripke shows more of how he views L_o when he discusses whether this truth definition and substitutional quantification guarantee freedom from 'ontological commitment'. Here, he states that:

> *It is true that the clauses [13]-[15] can be stated without mentioning entities other than expressions and the entities mentioned in characterizing truth for* L_o*. If other entities are mentioned in characterizing truth for* L_o*, they are still used when the notion is extended to L. Note that, however, there is no reason to think that other entities are used in every truth characterization for every* L_o*; such an assumption would be obviously false.* (333)

Kripke clearly views the truth-definition of L_o as something that can, at least possibly, be defined without reference to objects. This is further confirmed by a similar point he makes in the discussion of the formal truth theory, where the atomic truth-definition of L_o is called $R(\alpha)$:

> *Does the meta-theory involve an ontology of expressions alone? As stated, it does, since $R(\alpha)$ was taken as primitive ... Suppose, however, we wish not to take $R(\alpha)$ as primitive but to define it in terms of more basic primitives. After all, $R(\alpha)$ is itself a truth predicate for a primitive language Lo, and for each particular language L we may wish this predicate to be explained. The definition of $R(\alpha)$ in more basic terms depends heavily on the choice of L_o; we can in general say nothing specific about $R(\alpha)$ other than what has been said above. Whether the definition of $R(\alpha)$ requires us to extend the ontology of the metalanguage beyond an ontology of expressions alone depends both on the choice of Lo and on the question of what predicates of expressions (or, alternatively, via Gödel numbering, of numbers) we are willing to take as primitive.* (343-344)

Here Kripke once again argues that while it is possible that the truth of L_o refers to objects, this is by no means a requirement.

5.3 Truth-Definitions For L_o

For the construction of truth-definitions of an atomic language L_o, we can distinguish between the definitions for a L_o with a finite amount of sentences and a L_o with an infinite amount of sentences.

In the case of a L_o with a finite amount of sentences, it is relatively easy to construct a truth-definition that does not require any objects beyond the expressions of the language itself. Like with a language L with a finite amount of sentences, we can simply for each sentence say whether it is true or not.

This results in a complete definition of True-In-L_o. This definition of L_o can then be recursively expanded to a language L with a substitutional quantifier, as all the sentences of L are made up out of previous sentences.

An explicit definition of a language L_o of finite length can through (13)-(15) provide the truth for the language L without any reference to objects or sentential functions. This method is not open to languages with a referential quantifier, as per definition, the referential quantifier makes use of the sentential function. The substitutional language can avoid the use of satisfaction in cases in which there exists a L_o of finite length, and in these cases it significantly differs in truth-definition from referential languages, for which this possibility does not exist.

Does this result also hold for languages L whose atomic language L_o has an infinite amount of sentences? Such a language is introduced by Kripke himself as an example of a language that is syntactic and does not require the use of any objects or denotation. He states that:

> *Let me give an example of what I just called the 'most favoured case'. The terms of L_0 are arbitrary strings of the letters 'a' and 'b'. A formula of L_0 is anything of the form M(t_1, t_2) where t_1 and t_2 are two terms. Such a formula counts as true whenever the first string is the mirror image of the second; for example, M(abb, bba) is true while M(abb,bab) is false. This truth characterization for L_0 is indeed syntactic, and the truth characterization can then be extended to the larger language L by means of [4]-[6)]. Note that no question arises of any 'denotations' for the 'terms' – the terms are merely uninterpreted strings of the letters 'a' and 'b'. Nor do the substitutional variables have any 'range'.* (345)

There is indeed no doubt that this truth characterization does not add an ontology of objects beyond the ontology of expressions. However, we can wonder about the importance of that finding. The language based on the predicate M(t_1, t_2) is a language about expressions, and because of this, a language with referential quantifier and truth definition would also be limited to an ontology of expressions. This, then, is not enough to show that the two languages are fundamentally different.

It is also not obvious whether it is really the case that the terms of this language do not denote. To be a language, the sentences have to be about *something*. If it is not, we are not talking about a language and not talking about truth, we are merely looking at an abstract mathematical construct. This language is about something, namely whether two strings mirror each other. The terms, then, do denote something, namely these strings. The string is a syntactical object about which the language is, but at the same time each string is the name of itself. This also, then, does not distinguish between the substitutional and referential languages, as both deal with the same things, namely strings made up out of the letters ’a’ and ’b’. Kripke does discuss this autonymous interpretation, in which each term designates itself, but says of it that:

> *I assume, however, that this trivial interpretation is not all that Wallace is claiming when he argues that satisfaction invariably lurks in the background of the substitutional quantifier. Indeed, even without the autonymous interpretation, it is clear that every substitutional formula is equivalent to one involving referential quantifiers over expressions.* (353)

However, denotation is a requirement for satisfaction, be it trivial denotation or otherwise. Even if the denotation is trivial, it seems unwise to decide the question of satisfaction by rejecting autonymous denotation for being trivial.

Kripke states that we can give a truth characterization by saying that a formula of the form M(t_1, t_2) counts as true whenever the first string is the mirror image of the second string. It seems that Kripke assumes that through this, we can effectively give a truth characterization of the entirety of L_o. It is, however, not entirely obvious how this would work. We can see that the language L_o is one with an infinite amount of sentences. The formulas of L_o are all those of the form M(t_1, t_2), where the terms are arbitrary strings of the letters 'a' and 'b'. As these strings can be of any length, there is an infinite amount of strings, and thus an infinite amount of formulas of the form M(t_1, t_2). Because of this, we cannot simply look at each formula of L_o and for each one see whether the terms mirror each other, as this task could never be completed. We can also not say for each individual formula M(a, b) that it is true if the terms mirror each other and false otherwise, because, like with the explicit definition, we would need to create an infinite amount of these partial definitions. Instead, we need to talk about all sentences of the form M(t_1, t_2), and define the truth for all these sentences. The truth characterization, then, does not deal directly with the individual sentences of L_o, but rather with the infinite set of all sentences of the form M(t_1, t_2), or alternatively the sentential function M(t_1, t_2) which determines the membership of the set of all sentences of the form M(t_1, t_2). We have seen in the previous chapter that for Tarski, truth was a quality of sentences, and sentences alone. Here, what we define for the infinite set of all sentences of the form M(t_1, t_2) or the sentential function M(t_1, t_2) cannot be truth, as neither the set nor the sentential function are sentences. Instead, it needs to say of the form M(t_1, t_2) that it *would be true* in certain situations.

A method using satisfaction would here, following Tarski, create the truth definition in the following way. The form M(t_1, t_2) would be taken as the sentential function. Satisfaction would be defined as *for every* a *and* b*, we have* a *and* b *satisfy M(t_1, t_2) if and only if a is the mirror image of* b. This could be extended into satisfaction by infinite sequences to deal with longer sentences. Then, truth of a sentence would be defined as a sentence satisfied by every sequence.

Here we can wonder whether the method using satisfaction substantially differs from the one informally used by Kripke, and I do not think it does. *'A formula of L_o is anything of the form M(t_1, t_2)'* functions exactly the same as taking M(t_1, t_2) as a sentential function that is later instantiated by turning the variables into concrete individual terms. The statement *'Such a formula counts as true whenever the first string is the mirror image of the second'* seems to informally at the same time both play the same role as defining something similar to satisfaction for M(t_1, t_2) and define truth for instantiated sentential functions on the basis of this satisfaction-like concept. If we assume that we have the formula M(abb, t_2) it would make sense to say that this formula would count as true if we were to take t_2 to mean 'bba', but not if we took t_2 to mean 'bab'. We can say that an instantiation of M(t_1, t_2) is true iff t_1 is a mirror instance of t_2. This concept that says of sentential functions when it would or would not be true essentially fulfils the exact same role as satisfaction. It does not seem that we can avoid seeing it as a satisfaction-predicate, just because it lacks the explicit formal definition and the proper terminology. We could equally well say that a pair t_1, t_2 satisfies M(t_1, t_2) iff t_1 is a mirror instance of t_2, and the rest of the definition would stay exactly the same.

As we have seen previously, Wallace observed that just as we can paradigmatically characterize truth through convention T, we can paradigmatically characterize satisfaction through the formula

$$(x_1)(R(x_1, \overline{\phi(x_1)}) \iff \phi(x_1))$$

Where any two-place predicate $R(x_1, \alpha)$ satisfying the above formula for each formula $\phi(x_1)$ will have the satisfaction relation as its extension. This specific definition holds for formulae with one free variable, but we can easily construct an alternative for formulae with two free variables, namely

$$(x_1)(x_2)(R(x_1, x_2, \overline{\phi(x_1, x_2)}) \iff \phi(x_1, x_2))$$

Where R is a three-place predicate with satisfaction by pairs of objects as its extension. If the truth-definition for the atomic language L_o fulfils this definition, per Kripke himself, satisfaction is present in the definition. If we take the class of objects over which the quantification ranges to be the set of all expressions of arbitrary length made up out of 'a' and 'b', then per the definition given by Kripke of truth of L_o, for all x_1 and x_2, and all formulas $\phi(x_1, x_2)$, which in this case is only the formula M(x_1, x_2), we can per the above suggestion construct a predicate R(x_1, x_2, α) such that all partial definitions of satisfaction hold. As a result of this, satisfaction is present in the truth-characterisation for language L_o.

In this specific case, the P-Sat predicate that Kripke showed was present in his theories can be expanded to satisfaction. Kripke stated that it induces a relation if all the terms are assigned denotations and it is transparent. He states that it was possible that all the terms had denotations, but not that it was transparent, as P-Sat included the opaque predicate Q(x_1, α_2). Here, all terms have denotations, as we gave an autonymous interpretation in which each term denotes itself. The first condition is thus fulfilled. As for the opaqueness of Q, as it ranges over terms, it is unproblematic to quantify into it with a referential quantifier if the range of objects over which the quantification happens is a range of terms. In this specific case then, P-Sat does equate to normal satisfaction.

As we have seen, in this specific case the satisfaction-class ranges over autonymous denotations, which each term denoting itself. The next chapter will discuss whether this will always be the required satisfaction-class. It will also discuss whether it will always be the case that satisfaction can be found in any theory with an infinite amount of atomic sentences.

The reasons that such a predicate like satisfaction also has to be present in Kripke's definition are the same that motivated Tarski to define truth via satisfaction to begin with: first of all, for a language with an infinite amount of sentences, we cannot define truth explicitly. Secondly, sentences are not made up out of simpler sentences but out of sentential functions.

If Kripke's definition of this specific language L_o contains satisfaction, then the language L which is an extension of L_o through conjunction, negation and quantification also contains satisfaction. Assume that the instantiation of a variable t_i or t_j of M(t_i, t_j) with terms has been defined as true iff all sequences replacing the *i*th or *j*th place with the specific term satisfies M(t_1, t_2). Then we can prove that the language L as a whole also has a truth definition via satisfaction.

Proof. Assume that ϕ and ψ are satisfied or not by n and m terms in an infinite sequence s and that ϕ and ψ are true iff all sequences satisfy them and false if it is satisfied by no sequence.

(19) Negation

A sentence $\neg\phi$ is true iff ϕ is not.

A sentence ϕ is not true if no sequence satisfies it.

A sequence does not satisfy ϕ iff it satisfies $\neg\phi$.

A sentence ϕ is not true if all sequences satisfy $\neg\phi$.

A sentence $\neg\phi$ is true iff all sequences satisfies $\neg\phi$.

(20) Conjunction

A sentence $\phi \wedge \psi$ is true iff ϕ is true and ψ is true.

A sentence ϕ is true iff the n relevant terms of all sequences satisfy ϕ.

A sentence ψ is true iff the m relevant terms of all sequences satisfy ψ.

A sentence $\phi \wedge \psi$ is true if all sequences satisfy ϕ and all sequences satisfy ψ.

A sequences satisfies ϕ and ψ iff it satisfies $\phi \wedge \psi$.

A sentence $\phi \wedge \psi$ is true if all sequences satisfy $\phi \wedge \psi$.

(21) Existential Substitutional Quantification

A sentence $(\Sigma x_i)\phi$ is true iff there is a sentence ϕ' that is true where ϕ' is the result of replacing a variable x_i with a term in ϕ.

A sentence ϕ' where a variable x_i is replaced by a term t is satisfied by a sequence iff this sequence would satisfy ϕ if its ith member would be replaced by t.

A sentence ϕ' is true iff it is satisfied by all sequences.

A sentence $(\Sigma x_i)\phi$ is true iff all sequences satisfy ϕ if their ith members would be replaced by t.

Per assumption, the truth of all atomic sentences of L can be defined via satisfaction. Per induction, the truth of all sentences of L can be defined via satisfaction. This proof shows that if we accept that the truth definition of L_0 contains satisfaction, the truth definition of the entirety of the language L

contains satisfaction.

Does this conclusion that Kripke here covertly appeals to satisfaction hold for any pair of languages L_0 and L? Clearly not. The step towards satisfaction was only required because we were dealing with an atomic language with an infinite amount of sentences. Had the language L_0 been of finite size, it would be possible to define the truth of each formula of L_0 explicitely and build L from that. L would in this situation still be a language with an infinite amount of sentences. However, if the language L_0 is of infinite size, it will always be necessary to define the truth of L_0 in a way that does not directly deal with the formulae of L_0, but instead discusses more general forms and what would make such forms true. This is exactly what satisfaction does.

5.4 Conclusion

In this chapter, we have seen that Kripke aims to give a truth-definition of the language L by recursively defining it over a pre-given truth-definition of the atomic language L_0. However, further analysis has shown that this is not unproblematic. If L_0 is a language with an infinite amount of sentences, it is impossible to define the truth of the sentences of L_0 explicitly. The method that Kripke uses to define truth of a language with an infinite amount of sentences seems to work exactly the same as Tarski's method using satisfaction, with the difference being in wording and less detailed explanation in the case of Kripke. If the truth of a language L_0 is defined through satisfaction, then the truth of language L which is an extension of L_0 through the use of conjunction, negation and quantification is also defined through satisfaction.

While there are certainly languages with a substitutional quantifier and an infinite amount of sentences that can be defined without appealing to satisfaction, which does not hold for similar languages with a referential quantifier, there also seem to be languages with a substitutional quantifier that do have to (covertly) appeal to satisfaction. In these cases, however, the differences between the substitutional and referential quantifiers are still present. Because of this, the differences between the two quantifiers cannot be explained through their respective use of satisfaction.

Chapter 6

Terms, Denotation and Ontology

The previous chapter has shown that to provide a truth-definition for a language L through a recursive definition of truth, it is necessary to assume a pre-given truth of L_0. However, if this language L_0 is itself also infinite, it is necessary to use a notion that is equivalent to satisfaction for the truth-definition of L_0, which in turn leads to the truth of L also being determined by satisfaction. This holds both for languages using referential quantifiers and substitutional quantifiers. However, while it is clear that sentential formulae of a referential language are satisfied by objects, it is yet unclear what satisfies the sentential formulae of substitutional languages. In the example in the previous language, the terms had an autonymous denotation, but it is unclear whether this always works. This chapter will look at the objects used by a substitutional language, and the relation between the terms of a substitutional language and the objects that a language makes use of.

6.1 Terms and Denotation

Kripke rejects the idea that the terms of language L must denote. He writes that:

> *Wallace and Tharp, especially the former, seem to argue that the 'terms' must denote or a determination of truth conditions for L will be impossible. Since*

> *any class of expressions of L_o can be the substitution class, and since L_o can be any language, and since we can extend a truth concept for L_o to L automatically, such a result would seem to mean that any expressions of any language must denote, a rather astounding result.* (334)

Any class of expressions of L_o can be the terms of the substitution class, so if we hold that all terms of the substitution class must denote, then any expression of L_o must denote. The idea that all expressions must denote has, according to Kripke, ridiculous consequences:

> *True, for injudicious choices of the substitution class there will be very few forms. It is entirely feasible, however, to let the substitution class consist of the sentences of L_o (see Section 5(a)). Does this provide an a priori proof of Frege's view that sentences denote? Suppose L_o contains connectives, say primitive truth-functions $\wedge, \neg, \vee, \Rightarrow, \iff$. The substitution class can consist of expressions where a sentence of L_o is followed by a binary connective. Then $x\psi$, where ψ is a sentence, is a form, and $(\Sigma x_i)x\psi$ is implied by each of its instances $\phi_1 \Rightarrow \psi$, $\phi \wedge \psi$, $\phi \iff \psi$, $\phi_2 \Rightarrow \psi$, etc. Is this supposed to mean that expressions such as '$\phi \Rightarrow$' denote? Usually they are not even regarded as significant units!* (334)

The reason that Wallace argues that the 'terms' must denote is his belief that without denotation, we cannot fulfill the requirement of convention T. He specifically argues that it is impossible to derive the sentence

$$\text{True}(\overline{(\exists x)(\text{x is a man})} \iff (\exists x)(x \text{ is a man})$$

from his equivalent of (13)-(15) together with some partial definition of truth for atomic sentences. He argues that to do this, it is necessary to add the following axioms:

> *[22] a predication of* $\overline{man}$ *with a member of the substitution class, say* a*, is true if and only if the denotation of* a *is a man.*
> *[23] the denotation of* $\overline{Able}$ *= Able*
> *the denotation of* $\overline{Baker}$ *= Baker*
> *the denotation of* $\overline{Cain}$ *= Cain*
> *and so on for each member of the substitution class.*
> *[24] everything is denoted by some member of the substitution class.* ("Frame of Reference", 129-130)

This critique fails, as the sentence that Wallace claims needs to be derived for convention T is not in fact a partial truth definition. Wallace interprets the first quantifier as substitutional and the second quantifier as referential. The formula that he tries to derive can, following Kripke's typology, better be written as True($\overline{(\exists x)(\text{x is a man})}$ $\iff$ (Σx)(x is a man). As the two parts mean something different, there is no need for this sentence to be provable to fulfill convention T.

However, while Wallace's argument is wrong, that does not mean that his conclusions are entirely wrong as well. The need for denotation is not a result of convention T, but a result of the concept of truth. As we have seen, Tarski views truth as something that describes a relationship between a sentence and the world, and it was this notion that we tried to capture. However, this means that we cannot view the object language for which we try to define truth as uninterpreted: the object language is not a mathematical object but an actual *language*, with meaning. Meaning is required for us to find truth. If a sentence does not have meaning, it cannot say anything about the world, and as such we can say nothing about its relation to the world. However, the thing that must denote is not necessarily the terms of the substitution class, but rather the words of the language of L_o. These are the entities that determine the meaning of the sentences of L_o, even in the cases where the substitution class is made up of things like parentheses and connectives.

If the language for which we are constructing a truth-definition deals entirely with syntactic issues, such as the language containing M(t_1, t_2) it is sufficient to have an ontology that contains just syntactical entities, as it is a language that speaks only of syntactical issues. However, if we have a language that deals with non-syntactic objects, references to objects beyond the syntactic are necessary. This holds not just for languages with an infinite amount of atomic sentences: even if we have a completely finite language for which we can give an explicit definition of truth, this language still has to have defined denotations to the objects about which it speaks. If we took the sentence 'snow is white' as nothing but an uninterpreted string, what could justify us in saying that this sentence is true if snow is white?

We have seen that some reference to objects is necessary. At the same time, however, it seems that interpreting the substitutional quantifier as range over denotations of terms leads to equally big conceptual trouble. One of the main advantages of the use of a substitutional quantifier is that it can quantify into opaque contexts. Assume a pair of true sentences 'John believes that he can see the Morning Star' and 'John does not believe that he can see the Evening Star', where the Morning Star and the Evening Star co-designate one object, namely the planet Venus. The referential quantifier cannot be used to quantify into this opaque context. Doing so would result in the sentence 'There is an x so that John believes that he can see x', but which object would this x be? If it is the Morning Star, then it is the Evening Star, of which John did not believe that he could see it. If, however, we quantify over terms, it is perfectly unproblematic. This would result in the sentence 'There is a term x so that John believes that he can see x', which is true, as there exists the term 'Morning Star', which does not equal 'Evening Star', of which it is true that John believes he can see it. If we were to claim that the terms over which a substitutional quantifier ranges must denote and that this denotation must determine the satisfaction and truth of formulae, this benefit would be lost. In fact, not only this benefit would be lost, but the substitutional quantifier would collapse into the referential quantifier, as we shall see in the next section. Kripke writes about the problems of opacity and denotation that:

> *The same puzzlement arises with respect to opacity. Suppose the terms of L_o do denote; the truth characterization for L_o somehow uses the concept of denotation for all the terms. Then L_o may contain opacities; codesignative terms may not be intersubstitutable in L_o salva veritate. Yet this phenomenon, if it exists, will be entirely irrelevant to the extension of L_o to L, and of truth for L_o to truth for L. L will then contain 'substitutional quantification into opaque contexts', this is why the conventional view asserts that such quantification is unproblematic. Wallace challenges this view. Does he challenge the theorem on which it is based (that a truth characterization for L_o uniquely extends to L)? If not, there can be no trouble with truth for L which was not already trouble with truth for L_o.* (334)

Wallace challenges this view precisely because he believes that the truth and satisfaction of a sentence are determined entirely by its denotation. Because of this, he argues that co-designative terms in L_o are already problematic, and cannot be expanded to L.

6.2 Equivalence between Referential and Substitutional Quantifiers

So far, we have seen that it is necessary that a substitutional language gives denotations to its terms, because if it does not there is nothing that makes a sentence true or false, but that at the same time, this denotation cannot be taken as determining the entire satisfaction and truth of a sentence, as this would disable some of the benefits of substitutional quantification. To get further insight into the relations between the referential and substitutional quantifier, it is useful to look at the situations in which they can be said to be equivalent. Kripke writes that:

> *In the most favoured case we can drop $R(\alpha)$ and define the class of true atomic strings without introducing any new vocabulary or ontology not already in the metalanguage; in this case, no-one could conceivably quarrel with the claim that our axiomatic theory involves an ontology of expressions alone. Otherwise, we can add either new predicates true of strings, or new ontology, or both, to the metalanguage M. If we extend the ontology of the metalanguage, and add a new binary predicate of denotation to the metalanguage, relating the terms to some of the new entities, and postulate that every term denotes one of the new entities, and certain additional conditions to be detailed in the first half of Section 3 are satisfied, then we can say, as we will in Section 3, that the substitutional quantifiers of L are in a sense equivalent to Referential quantifiers. Otherwise, the new ontology may have nothing to do with the terms or alleged denotata for them.* (344)

To be able to say that the referential and the substitutional quantifier are equivalent it is necessary that (a) both quantifiers extend the ontology of the metalanguage with the same objects, (b) all the terms are taken to denote, and (c) each of the objects added in (a) has a term. Later, Kripke states the condition that (d) "all formulae in L are transparent" (353).

Of these (a) has been shown to be the case earlier in this chapter. If two languages, one with a substitutional quantifier and one with a referential quantifier, deal with the same topics, then it is necessary that they add the same objects to the pre-given ontology of expressions. Without this ontology of objects, the referential language would not function, and the substitutional language would not be a 'language' in the meaning that we intend. (b) has also been fulfilled. Earlier in this chapter, we saw that to be a sentence, each part of the sentence had to say something, and as such each term had to denote. Where this leads to problems, such as in a language which talks about strings, or where the terms are expressions that are not usually taken to have meaning such as '$\phi \Rightarrow$' or punctuation marks, we can offer an autonymous interpretation, where these expressions denote themselves. (c) is normally raised as one of the key problems with substitutional quantification (Wallace, "Frame of Reference", 132). This objection raises the problem that in the situations where there is an ontology with an unnamed object, it is possible that there is a true sentence $(\exists x)(Px)$ but no single sentence Pa is true. However, this problem is easily avoided if we do not let the substitutional quantifier range over the terms of a language L_o, but rather over the terms of any finite extension of L_o. This seems philosophically justified: if we encounter a situation in which we have unnamed objects, we can solve this situation simply by naming the unnamed objects.[1] (c), then, is also satisfied. This leaves us with (d). As (a)-(c) have been shown to be fulfilled for each substitutional language with its respective referential equivalent, it seems that the difference between the substitutional quantification and referential quantification

[1]Such an approach of defining the truth of quantifiers via an extension of language that names every object is also taken by Shoenfield (18-19).

lies entirely in how they deal with opaque contexts. Substitutional quantification can happen into opaque contexts, while referential quantification over objects cannot.

Is it possible to have a language with only referential quantifiers that can quantify in and outside of opaque context, and thus has the same expressive power as a language with a substitutional quantifier? We can create a standard referential language with the standard referential quantifier $\exists x$ that ranges over objects, and add to that a second referential quantifier $\exists' x$ that ranges over terms of L_o. This language can quantify into transparent contexts using the standard $\exists x$ quantifier, and into opaque contexts using the $\exists' x$ quantifier.

This pair of referential quantifiers can function in the exact same way as a substitutional quantifiers when it comes to quantifying with respect to atomic sentences. However, when it comes to quantification with respect to any sentence, it becomes more problematic. The referential quantifiers can have the same outward effect by quantifying over either terms or objects depending on whether the predicate in the atomic sentence says something about the term or its non-syntactic denotation, this is not possible in the case of conjunctions where one part of the conjunction says something about a syntactic element and the other part about its denotation. For example, a referential language with two quantifiers could make sense of, following a previous example, the formula '*John believes that he can see x*' where John believes he can see the Morning Star but does not believe that he can see the Evening Star and '*x is a planet*' by in the first instance quantifying over terms and in the second instance over denotations of terms. This solution does not work for the formula '*x is a planet and John believes that he can see x*', as we would either have to quantify over terms, in which case '*x is a planet*' would be false, or we would have to quantify over non-syntactic objects, in which case '*John believes that he can see x*' would be problematic. Substitutional quantification seems to avoid this problem.

This difference can be explained by how sentences using these quantifiers are built up. A sentence using a referential quantifiers is built up from a sentential function. Because of this, it is necessary to decide when a referential quantifier is used whether the truth of the sentence is based on the syntax of the term or the object denoted by the term. A substitutional quantifier, on the other hand, is defined on the basis of other sentences. The determining whether the truth of this sentence is based on the syntax of the term or the object denoted by it has already happened, namely during the truth definition for the atomic language. Co-extensive terms are given the same truth-values in cases where their denotation determines truth, but can have different truth-values in case syntax determines truth. Because of this, the substitutional quantifier can safely limit itself to ranging over terms.

This also clarifies the relation between P-Sat and satisfaction for languages with an infinite amount of atomic languages. In the case where predicates range over terms and say something about syntax, P-Sat and satisfaction describe the same relation, as seen in the previous chapter. If the predicates range over objects, denotation is necessary and a definition has to be given that uses satisfaction by objects to provide the truth-conditions of the atomic sentences. Terms that co-designate are given the same truth-values, so P-Sat and satisfaction coincide, despite the opaqueness of $Q(x_1, \alpha)$. Any sentence that is different from a true sentence only by the replacement of one term with a co-designative one will still be true. It thus does not matter that satisfaction cannot designate only one term where there are multiple that co-designate and instead designates the object, as the truth-conditions in these transparent cases will always coincide.

6.3 Conclusion

We have seen that there are four conditions necessary for the equality between a substitutional language and its referential equivalent, namely (a) both quantifiers extend the ontology of the metalanguage with the same objects (b) all the terms are taken to denote (c) each of the objects added in (a) has a term and (d) that no sentences are opaque. We have seen that (a)-(c) are always met due to what it means to be a language. A language has to be about something, so the terms must denote and the ontology must be extended with these denotations. Furthermore, a language can always be extended to include terms for yet-unnamed objects. The difference between substitutional and referential quantification, then, lies in how they deal with opaque context.

This difference can be explained by how the construction of sentences with quantifiers takes place in referential and substitutional languages. In a substitutional language, sentences are built up from other sentences. The truth-definitions have already been given at the level of the atomic sentences, and at this point it is decided whether a predicate says something about the syntax of a term or about its denotation. Co-extensive terms have the same truth-conditions for predicates that discuss the objects, but can have different truth-conditions for predicates that deal with syntax. At this point, both syntactical and ontological issues can be discussed purely through reference to terms, because these terms will refer to themselves in syntactical contexts, and to their normal denotation in ontological contexts. Because of this, in transparent context, it does not matter which one of a set of co-extensive terms is picked, and because of this, quantification over terms suffices, as these predicates had their truth-values determined on the basis of the denotation of the term. In opaque contexts, quantification over terms also works, as in these cases it is the terms themselves that determine the truth of a sentence-part. In an important sense, substitutional quantification ranges over both terms and their denotation, depending on how the predicate was defined in the atomic language.

In referential languages, the sentences are built up from sentential functions. The distinction between predicates that deal with syntax and predicates that deal with objects has not yet been made, and because of this, it is necessary to distinguish between quantification over syntax and quantification over objects. As a result, quantification can either only go over terms or only go over denotations, and because of this, a combination of transparent and opaque contexts can lead to problems.

Conclusion

This thesis has looked at the feasibility of a definition of truth that does not make use of satisfaction. The first part of this thesis looked at the discussion surrounding the covert use of satisfaction for the definition of truth of a substitutional language. It did this by first of all presenting the arguments put forth by Wallace as to why a substitutional language has to make use of satisfaction. After this, it presented Kripke's reply to Wallace, which argued that a truth definition for a substitutional language did not need to make use of satisfaction, as it is possible to recursively define truth on the basis of a pre-determined truth definition for an atomic language L_0.

In the first part of the second half of this thesis, an overview was given of the work of Tarski to gain more insight in the original motivation for the use of satisfaction. This showed that rather than being given philosophical importance, as Wallace argued, it was a tool to construct definitions of truth for languages with an infinite amount of sentences. Satisfaction, for Tarski, is not necessary for finite languages as we can simply list each partial definition in turn.

The second part of the second half of this thesis showed that while Kripke for a large class of substitutional languages gives a definition that does not make use of satisfaction, namely those with an finite amount of atomic sentences, substitutional languages with an infinite amount of atomic sentences do need to appeal to satisfaction. For these languages, satisfaction is necessary to give an atomic truth definition, for much the same reasons that Tarski uses satisfaction: we cannot give an enumerative definition for a language of infinite size. From this we concluded that if there is a difference between substitutional and referential quantification, it is not determined by their respective use of satisfaction. Instead, the difference seems to be that a substitutional definition of truth is built up out of simpler sentences, for which it is already determined whether they refer to syntax or denotation, which leads to a situation in which co-designative terms have the same truth-values in transparent contexts but not necessarily in opaque contexts. This allows the use of a single quantifier that ranges over terms yet still maintains the relation to its denotation. With a referential definition of truth, on the other hand, sentences are built up out of sentential functions, and the choice whether the sentence refers to syntax or denotation happens all at once for the entire sentence. Because of this, a referential quantifier can only refer to objects or only to terms, whereas a substitutional quantifier can refer to both objects and terms. This can create problems in situations where a sentences includes both transparent and opaque contexts.

As for the philosophical significance of satisfaction, this thesis has revealed that there is remarkably little. The philosophical significance that was given to satisfaction by Wallace, namely the importance of predicates and the connection between sentences and the world, were in fact not a result of the use of satisfaction, but rather essential qualities of language and truth. To be a language, a sentence has to say something, and to say something, the words of the language need to have meaning. Similarly, truth, following Tarski, was taken to be a relation between what a sentence expresses and what is the case. This holds equally in cases in which satisfaction cannot be found. The question of satisfaction, then, is one of technical detail, and not of philosophical importance.

Works Cited

Hilbert, David, and Paul Bernays. *Grundlagen Der Mathematik II*, 1939, pp. 329ff.

Kripke, Saul A. "Is There a Problem About Substitutional Quantification?" *Truth and Meaning*, edited by Gareth Evans and John McDowell, Oxford University Press, 1976, pp. 324–419.

Mckinsey, J. C. C. "A New Definition of Truth." *Synthese*, vol. 7, no. 6-A, 1948, pp. 428–433.

Shoenfield, Joseph R. *Mathematical Logic*. Addison-Wesley, 1967.

Tarski, Alfred. "The Concept of Truth In Formalized Languages." *Logic, Semantics, Metamathematics*, 1936, pp. 152–278.

Tarski, Alfred. "Truth and Proof." *Scientific American*, June 1969, pp. 63–77.

Tharp, Leslie H. "Truth, Quantification, and Abstract Objects." *Noûs*, vol. 5, no. 4, 1971, pp. 363–372.

Wallace, John. "Convention T and Substitutional Quantification." *Noûs*, vol. 5, no. 2, 1971, pp. 199–211.

Wallace, John. "On the Frame of Reference." *Synthese*, vol. 22, no. 1, 1970, pp. 117–150.