Synthese (2021) 199:14459–14490
https://doi.org/10.1007/s11229-021-03429-5

ORIGINAL RESEARCH

Semantic theories, linguistic essences, and knowledge of meaning

Nick Haverkamp[1] · **Miguel Hoeltje**[2]

Received: 20 March 2021 / Accepted: 22 September 2021 / Published online: 19 November 2021

Abstract
This paper argues, first, that *the information problem* poses a foundational challenge to mainstream semantics. It proposes, second, to address this problem by drawing on notions from Kit Fine's essentialist framework. More specifically, it claims that the information problem can be avoided by strengthening standard truth theories, employing an operator expressing the notion of a *relative constitutive semantic requirement*. As a result, the paper proposes to construe semantic theories as theories of semantic requirements, and semantic knowledge as knowledge of such requirements.

Keywords Truth conditional semantics · Kit fine · Semantic requirements · Constitutive essence · Information problem

1 Introduction

A semantic theory should exactly capture meaning, telling us exactly what the meanings of the object language expressions are. This entails that the theory should characterise meaning uniquely and conspicuously; it needs to provide characterisations of object language expressions that no non-synonymous expressions satisfy, and it needs to do so in a transparent way. One way of getting at the notion of exactly capturing meaning is in terms of knowledge: if a theory exactly captures meaning, then knowing the theory would suffice for knowledge of meaning. Or, to put it the other way around, if knowing the theory was insufficient for knowledge of meaning—if you could know the theory while failing to know what the expressions of the object language mean—then the theory does not manage to exactly capture meaning. Thus,

✉ Miguel Hoeltje
miguel.hoeltje@uni-due.de

Nick Haverkamp
nick.haverkamp@uni-mainz.de

1 Johannes Gutenberg-Universität Mainz, Mainz, Germany

2 Universität Duisburg-Essen, Duisburg, Germany

we can also say that a semantic theory should be such that knowledge of the theory would suffice for knowledge of meaning.[1]

A large swath of mainstream approaches fail to exactly capture meaning. This goes not only for simple extensional truth-theoretic accounts along broadly Davidsonian lines and their semantic value-based cousins, but also for intensional and model-theoretic approaches which relativise truth to possible worlds, situations, or interpretations. All of these approaches suffer from the *information problem*: they are too weak to exactly capture meaning and thus unable to ground knowledge of meaning. This is our negative goal in the current paper: to hammer home that the information problem remains an unsolved foundational problem.

Nevertheless, it is hard not be impressed by the progress that has been made within formal semantics. If we had to abandon this body of work, that would be a great loss. Can the foundations be repaired without jeopardising the building? This is our positive goal: to show that the information problem can be addressed in a way which satisfies the worrying philosopher without disrupting the daily business of working semanticists. To this end, we will suggest that semantic theories (and linguistic theories more broadly) should be construed as theories of *constitutive linguistic essences*, and that semantic knowledge (and linguistic knowledge more broadly) should be construed as knowledge of the linguistic essences of one's language.

Before going into the details, we present a bird's-eye view of our overall argument. Let *S* be the German sentence 'Sokrates ist weise', which means that Socrates is wise. In its simplest incarnation, the information problem applies to the proposal that a proper semantic theory merely says such things as: *S* is true iff Socrates is wise. Note that you could have this knowledge about sentences that do *not* mean that Socrates is wise, e.g. about a sentence which means that Plato is a philosopher (or any other true sentence). Hence, knowing that *S* is true iff Socrates is wise cannot suffice to know that it means that Socrates is wise. This, roughly, is the information problem as it applies to simple extensional truth theories. Moving to theories that employ semantic values, or that relativise truth to possible worlds, situations, or interpretations does not eliminate the basic problem. We develop these considerations in more detail in Sect. 2.

The only way out is to go beyond the information provided by standard semantic theories. But this does not mean that we have to abandon the whole framework. For it does seem right that the information that *S* is true iff Socrates is wise captures something essential about the meaning of the German sentence. It is just that this information by itself is not enough. How should the theory be strengthened? An analogy can provide guidance. Someone who knows merely that all moving objects continue to move unless acted upon by a force knows less than someone who knows that *it is a law of physics* that moving objects behave in this way. Similarly, knowing that it is essential to *S* to be true iff Socrates is wise goes beyond merely knowing that it is true iff Socrates is wise. This, we suggest, points the way to a solution of the information

[1] In our view, thinking about the relevant issues in terms of knowledge of meaning is helpful and apt. However, much of what we will say could be rephrased without appeal to this notion. Thus, as long as you agree that semantic theories should aim at exactly capturing meaning, you should be able to take on board much of the following discussion, even if you would prefer to frame the relevant issues in a non-epistemic way.

problem: linguistic theories should explicitly be construed as theories about linguistic essences. More specifically, we will argue that the information problem can be solved by embedding the axioms of standard semantic theories under an operator expressing (something closely related to) Kit Fine's notion of *relative constitutive essence*. We develop our positive proposal in Sect. 3, focussing on simple predication and sentential connectives. Section 4 concludes the paper with a brief outloook, and a technical appendix shows how our approach can be applied to first-order quantification.

2 The information problem

This section introduces the information problem and illustrates how far reaching it is. We will frame the problem in terms of a condition of adequacy on semantic theories: they should exactly capture meaning so that knowledge of such a theory suffices for knowledge of meaning. The information problem consists in the fact that mainstream approaches in current semantics are unable to satisfy this condition. Section 2.1 introduces the condition and illustrates how we can show that a given theory fails to meet it. Section 2.2 indicates the scope of the problem. Section 2.3 distinguishes the information problem from two related problems: overgeneration and granularity. Finally, Sect. 2.4 prepares the development of our positive proposal by briefly addressing existing responses to the information problem.

2.1 What is the information problem?

This section develops three claims: first, that semantic theories need to *exactly capture* the meanings of object language expressions. Second, that in order to do so, they need to be *uniquely characterising* as to these meanings. And third, that the notion of exactly capturing meaning can be fruitfully construed in terms of information that would suffice for *knowledge of meaning*. Theories that do not provide information sufficient for knowledge of meaning suffer from what we will call *the information problem*.[2] Apart from making the information problem more precise by linking it to unique characterisation, one point of our paper is to emphasise that the problem extends far beyond the Davidsonian programme and afflicts a wide swath of mainstream approaches (Sect. 2.2). Another point is that, even with respect to the Davidsonian programme, it has never been addressed in a satisfying manner (Sect. 2.4).

Let $\mathcal{T}$ be a candidate theory for some language $\mathcal{L}$, and let α be some $\mathcal{L}$-expression. We take $\mathcal{T}$ to be adequate only if $\mathcal{T}$ *exactly captures* the meaning of α. What does it mean to exactly capture the meaning of an expression? Roughly speaking, it means that the theory should tell us exactly and transparently what the meaning of the expression is. In order to be able to show that specific theories fail to meet this demand, we point out a necessary condition for exactly capturing meaning on which we will rely in subsequent arguments. Let $\Phi(\alpha, \mathcal{L})$ be the totality of information $\mathcal{T}$ provides about α.

[2] What we, roughly following Segal (2008), call the information problem has been discussed with respect to simple extensional truth theories employed within the Davidsonian programme; see, for instance, Foster (1976), Davidson (1976), Larson and Segal (1995, §2.2.2), Lepore and Ludwig (2005, §8) and Lepore and Ludwig (2013).

If $\Phi(\ldots)$ could also hold of a non-synonymous expression β of some language $\mathcal{L}^*$—i.e. if $\Phi(\beta, \mathcal{L}^*)$ is true for some β that differs in meaning from α—then the theory leaves open what α means; for all the theory says, it could mean what β means. We will say that such a theory fails to be *uniquely characterising* as to meaning. In order to exactly capture the meaning of an expression, a theory has to be at least uniquely characterising.

What unique characterisation amounts to and how it can best be tested depends on the underlying conception of languages and expressions and the format of semantic theories.[3] We roughly sketch one approach, relying on an abundant conception of languages, the view that simple expressions receive individual axioms, and the view that the information provided by a theory is fully determined by the theory's axioms. For any finitely axiomatised theory $\mathcal{T}$, let $\bigwedge\mathcal{T}$ be the conjunction of $\mathcal{T}$'s axioms. For any singular terms $t_1, \ldots, t_n$ and $u_1, \ldots, u_n$, let $\bigwedge\mathcal{T}_{t_1/u_1,\ldots,t_n/u_n}$ be the result of replacing every occurrence of t_i in $\bigwedge\mathcal{T}$ by an occurrence of u_i. We assume some canonical way of denoting expressions and languages, and we will write $\overline{\alpha}$ and $\overline{\mathcal{L}}$ for the canonical designator of an expression α and language $\mathcal{L}$, respectively. Given this approach, a theory $\mathcal{T}$ for a language $\mathcal{L}$ will be uniquely characterising for a *simple* expression α iff there is no language $\mathcal{L}^*$ containing an expression β such that α and β are non-synonymous while $\bigwedge\mathcal{T}_{\overline{\alpha}/\overline{\beta},\overline{\mathcal{L}}/\overline{\mathcal{L}^*}}$ is true. Correspondingly, $\mathcal{T}$ will be uniquely characterising for a *complex* expression $\alpha = \mathcal{S}(\alpha_1, \ldots, \alpha_n)$ built from the simple expressions $\alpha_1, \ldots, \alpha_n$ by a syntactic operation $\mathcal{S}(\ldots)$ iff there is no language $\mathcal{L}^*$ containing an expression $\beta = \mathcal{S}(\beta_1, \ldots, \beta_n)$ such that α and β are non-synonymous while $\bigwedge\mathcal{T}_{\overline{\alpha_1}/\overline{\beta_1},\ldots,\overline{\alpha_n}/\overline{\beta_n},\overline{\mathcal{L}}/\overline{\mathcal{L}^*}}$ is true.

In Sect. 2.3 we will show that a unique but non-transparent characterisation might fail to exactly capture meaning. To sharpen the latter notion, we add the assumption that a central goal of linguistic semantics is to contribute to an explanation of knowledge of meaning. In particular, we assume that knowledge of an adequate semantics $\mathcal{T}$ would suffice to know, of every $\mathcal{L}$-expression, what it means, and that the notion of exactly capturing meaning can be fruitfully construed in terms of knowledge of meaning: $\mathcal{T}$ exactly captures the meaning of α iff knowledge of $\mathcal{T}$ would suffice to know what α means. Accordingly, we will freely switch between talk of exactly capturing meaning and talk of providing information sufficient for knowledge of meaning.

To sum up: a theory $\mathcal{T}$ is an adequate semantic theory for a language $\mathcal{L}$ only if $\mathcal{T}$ exactly captures the meanings of $\mathcal{L}$-expressions. This, in our view, comes to the same thing as saying that knowledge of $\mathcal{T}$ would suffice for knowing what $\mathcal{L}$-expressions mean. Theories that fail this condition suffer from the information problem. One way of showing that a theory fails this condition is to show that it is not uniquely characterising.

[3] In this paper, we do not take a stand on whether an expression belonging to a language is essential for the language and/or for the expression. However, we would like to note that while nothing depends on it here, we will later (3.2) adopt a *thick conception of expressions* according to which they have their meanings essentially.

2.2 How far reaching is the problem?

Given the previous section, we can establish the inadequacy of a candidate theory $\mathcal{T}$ for some language $\mathcal{L}_1$ by showing that, for some $\mathcal{L}_1$-expression α_1, there is an expression α_2 of some language $\mathcal{L}_2$ such that, while α_1 and α_2 share *all* properties that $\mathcal{T}$ ascribes to α_1, they are non-synonymous. In order to apply this strategy, we need to add some assumptions about the availability of non-synonymous expressions. Semantic theories combine information about simple expressions, about syntactic structure of complex expressions, and about the semantic effects of syntactic composition. Thus, if α_1 is complex, it will not suffice that α_2 instantiates some central semantic property that the theory ascribes to α_1 (e.g. contribution to truth conditions). Rather, α_2 will have to be semantically isomorphic with α_1. For instance, let α_1 = 'equilateral triangle' and α_2 = 'equiangular triangle'. Then α_1 and α_2 are intensionally equivalent, but non-synonymous. Nevertheless, an intensional semantic theory will still manage to provide a characterisation of α_1 that is not satisfied by α_2, given that it must construe α_1, roughly, as built up of morphemes 'equi', 'lateral', 'tri', 'angle', each with its own semantic properties. Hence, the easiest way of applying the above strategy is in cases where α_1 and α_2 are both simple. Accordingly, we want assumptions that ensure the availability of non-synonymous simple expressions.

We make three assumptions. First, we assume that, for every language $\mathcal{L}_1$ and for every complex predicate α_1 of $\mathcal{L}_1$, there is a synonymous *simple* predicate α_2 of some language $\mathcal{L}_2$. For instance, 'unmarried man' is complex and synonymous with the simple 'bachelor'. Our assumption says that such pairings can always be found, even if not within the same language. Second, we assume that for all languages $\mathcal{L}_1$ and $\mathcal{L}_2$, containing the simple predicates α_1 and α_2, respectively, there is a language $\mathcal{L}_3$ which differs from $\mathcal{L}_1$ only in containing α_2 (meaning in $\mathcal{L}_3$ what it means in $\mathcal{L}_2$) instead of α_1. For instance, if $\mathcal{L}_1$ contains 'cordate' (meaning *creature with a heart*), then there is a language which differs from $\mathcal{L}_1$ only in containing 'renate' (meaning *creature with a kidney*) instead of 'cordate'. Finally, we assume a fine-grained notion of linguistic meaning. In particular, we assume that for every predicate α_1, there is a logically equivalent yet non-synonymous predicate α_2. For instance, we take it that the predicates 'x is wise' and 'x is wise and if x is a molehill, then x is a mountain or a molehill' are logically equivalent, yet not synonymous.

The first two assumptions rely on an abundant conception of languages. They could be weakened by rephrasing the following in terms of which languages *could have* existed, but we take them to express the plausible default position.[4] The third assumption is more substantial. We will not try to defend it here, however. If you do not agree, you may conditionalise the rest of the paper on this assumption, and if you prefer restricted versions of our three assumptions, you might then want to check whether you could accept the particular consequences we make use of in the remainder of this subsection.

[4] Perhaps the range of languages that can be acquired by humans is restricted by Universal Grammar. Our abundance assumption is not in conflict with this. Any such restriction would have to be specified by the complete theory, which still would have to be uniquely characterising with respect to the full set of possibilities.

On a standard notion of logical equivalence, synonymous expressions need not be logically equivalent (for instance, the predicates ‘x is a snake’ and ‘x is a serpent’ are synonymous; yet, since they belong to the non-logical vocabulary, they can receive distinct interpretations in a model). But let us say that two expressions are logically equivalent* iff they have synonyms that are logically equivalent. Now, our three assumptions entail that, for every language $\mathcal{L}_1$ containing a simple predicate α_1, there is another language $\mathcal{L}_2$ that differs from $\mathcal{L}_1$ only in containing the simple predicate α_2 instead of α_1, where α_1 and α_2 are logically equivalent* yet non-synonymous. Hence, in order to be uniquely characterising, a theory has to provide a semantic characterisation of a given simple predicate that is not *eo ipso* satisfied by every simple logically equivalent* predicate. As can easily be checked, standard mainstream approaches fail in this respect.

First, consider clausal extensional truth theories familiar from work within the Davidsonian tradition, employing axioms such as the following:[5]

(A1) $\forall N\ (T(\ulcorner N \text{ ist weise}\urcorner) \leftrightarrow R(N) \text{ is wise})$

Note that (A1) is not only *true*, but *interpretive*: the object language expression mentioned is synonymous with the metalanguage expression used. Following Lepore and Ludwig (2007, p. 32) we call a truth theory all of whose semantic axioms are interpretive an *interpretive truth theory*. It is easy to see that such theories are not uniquely characterising. Any simple expression that is logically equivalent* with ‘ist weise’ is true of all and only the wise things; but many of these expressions are not synonymous with ‘ist weise’. Thus, clausal extensional truth theories do not exactly capture meaning, and suffer from the information problem.

As we pointed out above, the information problem has received some discussion with regard to clausal extensional truth theories. However, it extends much further. For instance, it is not restricted to the clausal approach, but carries directly over to an extensional semantic value-based approach as employed in Heim and Kratzer (1998). After all, when it comes to unique characterisation, nothing is gained by moving from (A1) to a lexical entry which assigns the corresponding function from objects to truth values.

Moving from extensional approaches to theories that work with possible worlds or situations significantly increases the information encapsulated in the theory. But whether we do this in the clausal format (by relativising truth, reference, and satisfaction) or in the semantic value-based framework (by adding places to our functions), this move does not yield theories that are uniquely characterising. Given that two simple predicates α_1 and α_2 are logically equivalent*, they share all properties that can be expressed in the vocabulary of possible-worlds semantics or situation semantics. Finally, augmenting extensional and intensional information with information about non-intended interpretations along model-theoretic lines will not help either.[6] For, while it is true that, under a standard notion of ‘permitted interpretation’, we can say

[5] $T(x)$ stands for ‘x is true’; $R(N)$ for ‘the referent of N’. For purposes of illustration, we have chosen a predicate axiom that quantifies over names. Nothing in our argument hinges on this choice.

[6] Lepore (1983) makes a similar point.

that there are some interpretations under which α_1 and α_2 come apart extensionally, this does not allow us to conclude that they are non-synonymous; for synonymous expressions routinely come apart in this way. Thus, the upshot so far is the following: mainstream approaches along broadly truth-theoretic lines fail to be uniquely characterising, fail to exactly capture meaning, and suffer from the information problem.

The recent literature contains proposals, arguably located at the edge of truth-theoretic semantics, that go beyond the accounts considered above. Prominent examples include various impossible-worlds semantics, truthmaker semantics, procedural semantics, and semantics in terms of structured propositions.[7] Some of these are clearly not uniquely characterising, but some of them arguably are (see e.g. the account mentioned in Berto and Jago (2019)). We are sceptical that these latter accounts exactly capture meaning, but we will not enter into this discussion here. Rather, our goal in this paper is to investigate a different route that is more conservative in one way (by sticking to standard truth-theoretic machinery), but more exploratory in another way (by augmenting the language of the theory with a relativised hyperintensional operator expressing an essentialist concept). A comparison of the resulting approach to other non-standard approaches has to be left for later work.

2.3 Overgeneration and granularity

We want to distinguish the information problem from two related problems: overgeneration and granularity. Let us turn to overgeneration first. Any formal semantics aims to derive central theorems that characterise object-language sentences. In extensional theories, these will be the interpretive T-sentences ascribing (extensional) truth conditions, in possible-worlds semantics they ascribe intensional truth conditions, and so on. For simplicity, we stick to extensional theories, but the discussion carries over to other approaches.

Via the interpretive T-sentences, a truth theory relates any object-language sentence *S* to a metalanguage sentence that could be used to explicitly give *S*'s meaning. This fact lies at the heart of why truth theories have seemed promising for natural-language semantics. However, once we consider the theory's full set of theorems, it becomes apparent that the theory by itself does nothing to *identify* any metalanguage sentence as the one that gives *S*'s meaning. Call a truth theory which employs classical logic a *classical truth theory*. Any classical truth theory yields many *non-interpretive* T-sentences such as

(1) T('Sokrates ist weise') $\leftrightarrow$ (Socrates is wise, and if Socrates is a mountain, then Socrates is a mountain or a molehill).

We call this the problem of *overgeneration*.[8]

Just like the information problem, the problem of overgeneration has been discussed with respect to the Davidsonian programme; but again, it arises for semantic theories

[7] See, for instance, Duí et al. (2010), Nolan (2013), Jago (2014) and Fine (2017).

[8] For discussion, see Foster (1976), Davidson (1976), Callaway (1988), Soames (1992), Soames (2008), Lepore and Ludwig (2007) and Larson and Segal (1995, §2.2.1).

in general. The standard response within the Davidsonian programme is to modify the inferential component of a classical truth theory. With respect to first-order languages, it is known that an appropriate modification to the inferential component is available. In such cases, it is possible to give what we will call a *canonical truth theory*: a truth theory whose deductive apparatus does not enable us to derive any non-interpretive T-sentences such as (1).[9]

Note that solving overgeneration does not by itself solve the information problem. As we argued above, (classical) truth theories are not uniquely characterising. But canonical theories fail to be uniquely characterising for precisely the same reason. A canonical theory and the corresponding classical theory share their axioms, and the set of canonical theorems is a subset of the classical theorems. Thus, everything the canonical theory says, the classical theory says, too. Since what the classical theory says is not sufficient, neither is what the canonical theory says. You cannot increase informativity by saying *less*.

In a sense, unique characterisation is a matter of granularity: a theory that fails to be uniquely characterising fails to provide sufficiently fine-grained characterisations of object language expressions. The need for a fine-grained semantics becomes especially apparent when the object language contains sensitive constructions, such as operators ascribing propositional attitudes and explanatory vocabulary. However, the information problem should be distinguished from the problem of how to treat such constructions, and it should be acknowledged to go beyond mere matters of granularity/unique characterisation.

That the problem of how to treat sensitive constructions and the information problem are distinct should be evident from the fact that the former arises only for languages which contain such operators, while the latter arises even for fully extensional languages. More importantly, even a maximally fine-grained theory can suffer from the information problem. Consider a theory that assigns full-fledged meanings as semantic values (whatever *meanings* turn out to be). The theory might do so via axioms such as the following:

(A2) $[\![ist\ weise]\!] =$ the meaning of ‘ist weise’.

Or it might do so via

(A3) $[\![ist\ weise]\!] =$ Mickey,

where ‘Mickey’ is a directly referring name for the meaning of ‘ist weise’.[10] By definition, such a theory will be sufficiently fine-grained and uniquely characterising. But it would not solve the information problem. In particular, knowing such a theory would add next to nothing to an understanding of the object language.[11]

[9] For the use of canonical truth theories, see Davies (1981), Larson and Segal (1995, §2.2.1), Kölbel (2001), Kölbel (2002) and Lepore and Ludwig (2007, p. 36).

[10] This move echoes Ludwig (2017, p. 35), who points out that knowing a meaning ascription of the form ‘S in L means Bob’ employing a directly referring name ‘Bob’ of a proposition will not be sufficient for understanding S.

[11] This recapitulates one aspect of Davidson’s argument for the ‘inutility of meanings’. See Davidson (1967) and Davidson (1963, 319f.). For further discussion, cp. Lepore and Ludwig (2005) and

2.4 Responding to the information problem

Before developing our own response to the information problem, we want to address other existing approaches. While remaining within a broadly truth-theoretic framework, there are at least three kinds of responses. The first is to modify the axiomatic base of a truth theory; the second is to modify the inferential component; the third is to move to a meta level, construing meaning theories as theories *about* truth theories.

In our view, the first kind of response seems the most promising. The information a theory provides is most naturally viewed as determined by its axioms. Hence, if a theory is not informative enough, it seems natural to include more information in its axiomatic base. One way to do this is by strengthening the axioms with some operator. Versions of this approach have been considered by Higginbotham (1992), who proposes the operator *it is common knowledge among the speakers of the object language that*, and by Wallace (1978), who proposes *it is a matter of meaning alone that*. Our own proposal falls into this category, and we will briefly indicate why we take Wallace's closely related proposal to be insufficient.[12]

Let us abbreviate Wallace's operator as $\Box_W$. Prefixing $\Box_W$ to ordinary truth-theoretic axioms, we arrive at:

(W1) $\Box_W \forall N\ (T(\ulcorner N \text{ ist weise}\urcorner) \leftrightarrow R(N) \text{ is wise})$.

(W2) $\Box_W R(\text{‘Sokrates’}) = \text{Socrates}$.

Such axioms should enable us to derive strengthened T-sentences:

(2) $\Box_W (T(\text{‘Sokrates ist weise’}) \leftrightarrow \text{Socrates is wise})$.

However, as Davies (1981, p. 246) rightly complains, the principles governing $\Box_W$ ‘are not specified by Wallace’. That specifying such principles may not be straightforward becomes apparent when one notices that $\Box_W$ cannot be closed under logical consequence. We adapt an example from Lepore and Ludwig (2005, 108f.). Let χ be a predicate which means *is wise and LT*, where *LT* is some logical truth. Hence the following holds:

(3) $\Box_W \forall N\ (T(\ulcorner N \text{ ist } \chi\urcorner) \leftrightarrow (R(N) \text{ is wise} \wedge LT))$.

Assuming closure under logical consequence for $\Box_W$, (3) entails:

(4) $\Box_W \forall N\ (T(\ulcorner N \text{ ist } \chi\urcorner) \leftrightarrow R(N) \text{ is wise})$.

Footnote 11 continued
(Lepore and Ludwig (2006), p. 9). Our considerations in this section draw on similar points made in Ludwig (2014) and Ludwig (2017). Essentially the same point is illustrated by compositional translation manuals; even if they are uniquely characterising, they do not provide the right kind of information.

[12] Higginbotham's approach is discussed in Soames (1992), Soames (2008), Richard (1992). Wallace's suggestion has not received much attention in the literature. It is briefly discussed in Davies (1981, 1982) and Lepore and Loewer (1989) mention it in a footnote; more recently, Lepore and Ludwig (2005) provide a short critique of it, and Fine (2010) mentions, but does not discuss it. As far as we are aware, that's it.

Given that χ and 'ist weise' are non-synonymous, the truth of (4) illustrates that a theory treating 'ist weise' solely by (W1) fails to be uniquely characterising. Lepore and Ludwig seem to assume that $\Box_W$ *has* to be closed under logical consequence. More moderately, one could take the above consideration to issue a challenge: Wallace needs to amend his proposal with a plausible specification of the principles governing $\Box_W$ which ensures both the derivability of intended theorems and unique characterisation.

There are, however, reasons to think that this challenge cannot be met, even if closure under logical consequence is taken off the table. Consider the following schema:

$$(5) \quad \Box_W(\Phi \leftrightarrow (\Psi \wedge \Omega)) \wedge \Box_W \Omega \quad \rightarrow \quad \Box_W(\Phi \leftrightarrow \Psi).$$

If it is a matter of meaning alone that a condition holds iff a conjunction of two other conditions holds, while one of those two conditions is itself a matter of meaning alone, then we should take it to be a matter of meaning alone that the first condition only depends on the remaining one. But now let Ω be any condition that holds as a matter of meaning alone, and let χ be a predicate which means *is wise and* Ω. Since we have $\Box_W \Omega$, (3) entails (4) via (5), and hence a failure to be uniquely characterising.

The second response to the information problem takes its cue from the fact that, once overgeneration has been addressed, we can introduce a rule which allows the derivation of explicit meaning ascriptions.[13] Given that $(\mathcal{A}, \mathcal{R})$ is a canonical theory comprising a set $\mathcal{A}$ of axioms and a set $\mathcal{R}$ of inference rules, and given that a T-sentence '$T(S) \leftrightarrow p$' is derivable from $\mathcal{A}$ via $\mathcal{R}$, the corresponding meaning ascription 'S means that p' is guaranteed to be true. We could hence introduce a new rule, r_m, which allows the derivation of 'S means that p' from the corresponding canonically derived T-sentence. The resulting theory, $(\mathcal{A}, \mathcal{R} \cup \{r_m\})$, thus produces true meaning ascriptions. However, there is something awkward about this inferential approach. Note that r_m is not truth preserving. Let $\mathcal{A}_*$ be a set of true but non-interpretive axioms. Combining $\mathcal{A}_*$ with $\mathcal{R}$ results in a theory $(\mathcal{A}_*, \mathcal{R})$ which yields only true, albeit non-interpretive, T-sentences. Hence, combining $\mathcal{A}_*$ with $\mathcal{R} \cup \{r_m\}$ allows the derivation of *false* meaning ascriptions from *true* premisses. Thus, this approach essentially models the speaker as systematically inferring conclusions from premisses that do not warrant these conclusions. The fact that a given set of truth-theoretic axioms is interpretive cannot be gleaned from these axioms themselves; yet interpretiveness of the axioms is strictly required if the suggested inference is to be legitimate. It seems more natural to think that the lesson of the information problem is that the content of the axioms themselves must be enriched.

Versions of the third, 'meta-level' approach have been endorsed by Davidson and more recently by Lepore and Ludwig. We turn to Davidson first. Let 'p' be the conjunction of the axioms of some interpretive truth theory for a language $\mathcal{L}$. In *Reply to Foster*, Davidson suggests that what suffices for understanding $\mathcal{L}$ is the information that some truth theory (satisfying conditions such as charity) for $\mathcal{L}$ states that p. The thought here is that, by explicitly characterising the content of a truth theory *as* the content of a truth theory, we say more than if we just give the truth theory itself. For

13 Versions of this approach have been considered by Larson and Segal (1995, p. 560), Kölbel (2001), Ludwig (2002), and Hoeltje (2012, §2.4).

instance, when we say about ‘Snow is white’ that it is true iff snow is white, this also holds for ‘Grass is green’ (or any other true sentence). However, given that no truth theory satisfying the pertinent conditions entails that ‘Grass is green’ is true iff snow is white, explicit recourse to truth theories allows us to differentiate between the two. However, it is hard to see how Davidson’s proposal can avoid the overgeneration problem. The ‘p’-slot in ‘some truth theory for $\mathcal{L}$ states that p’ has to be filled by the truth-theoretic axioms, and there is no explicit reference to an inferential component. Hence, this approach cannot appeal to canonical proofs. But without addressing overgeneration, there is no hope of solving the information problem.

Partly motivated by considerations along these lines, Lepore and Ludwig have developed a different version of the meta-level approach.[14] On their proposal, what suffices for understanding is the meta-linguistic information expressed by ‘“p” is an interpretative truth theory for $\mathcal{L}$’ combined with: (i) a specification of how to derive canonical theorems; (ii) a statement of the fact that canonically derived theorems are interpretive; and (iii) information that, they hold, suffices for understanding the meta-language in which the truth theory is formulated. In contrast to Davidson’s proposal, which directly targets the content of a truth theory, Lepore and Ludwig’s approach aims at truth theories construed as syntactic objects. While this allows them to make use of canonical proofs, it introduces an additional language (the language of the truth theory), semantic competence with which has to be ensured in order for the meta theory to do its work. We are sceptical that this can be done in a satisfying manner.[15] Moreover, even if it can be done, the resulting theories seem somewhat awkward, generating understanding of the object language by first generating understanding of another language, then giving a truth theory in this second language, and finally using a third language to talk about this truth theory. One wonders why there shouldn’t be a more straightforward route.

Obviously, both the inferential response and the meta-level response deserve a fuller discussion. But our main focus lies elsewhere. We merely hope that the remarks in this section suffice to motivate a fresh look at a response along the lines of the first approach: strengthening truth-theoretic axioms by embedding them under a new operator.

3 Semantic theories as theories of constitutive semantic requirements

This section presents our positive proposal. Let α_1 and α_2 be two simple predicates, the first meaning *is wise*, the second *is wise and self-identical*. Since the wise things are the self-identical wise things, vocabulary such as ‘applies to’ does not by itself enable us to express any semantic difference between α_1 and α_2, and the situation is not improved by adding possible worlds, situations, or models. But there *is* a crucial semantic difference between the two. Roughly speaking, for α_1, it is *primary* to apply

[14] Lepore and Ludwig (2005, §9), Lepore and Ludwig (2007, §1.9) and Lepore and Ludwig (2011, pp. 265–268). See also Kirk-Giannini and Lepore (2017, §7).

[15] For discussion, see Hoeltje (2013) and Ludwig (2017).

to all and only the *wise* things, while applying to all and only the *self-identical* wise things is *secondary*—it is merely a consequence of the primary fact. In the case of α_2, it is precisely the other way around: for α_2, it is primary to apply to all and only the self-identical wise things, while applying to all and only the wise things is a secondary consequence. In a nutshell, our proposal is that a semantic theory should explicitly state what the primary semantic facts governing object language expressions are. By doing so, it will be uniquely characterising, and by doing it transparently, it will be exactly capturing and able to ground knowledge of meaning.

Obviously, this talk of primacy is in need of explication. We turn to Kit Fine's essentialist framework for help. Section 3.1 introduces the framework, in particular the distinction between constitutive and consequential essence, and the notion of a semantic requirement. While we will lean quite heavily on Fine's work, we will also suggest two novel modifications. Section 3.2 argues that we should acknowledge a notion of a *constitutive* semantic requirement in addition to Fine's own consequentialist construal of semantic requirements. Moreover, Sect. 3.3 argues that Fine's *Chaining* principle, which allows him to combine semantic requirements, should be rejected for constitutive requirements. Instead, we propose a new principle, *Replacement*, and illustrate how it can be used to derive constitutive requirements on complex expressions from the constitutive requirements on their parts. Our proposal then is that semantic theories should be construed as theories that explicitly state the constitutive semantic requirements on the object language expressions, and that these requirements can be stated by strengthening interpretive truth-theoretic axioms with an operator that expresses the notion of a constitutive semantic requirement on the expression in question.

3.1 Varieties of essence and semantic requirements

Following Fine (1994), we assume that there is a robust notion of *essence*, to be thought of in terms of *real definitions*: to give the essence of an object is to say *what the object is*, and to say that a certain property is essential to the object is to say that 'the object must have that property if it is to be the object that it is' (Fine 1994, p. 4). This notion of essence allows fine-grained distinctions. For instance, while Socrates and singleton Socrates necessarily coexist, the first is involved in the essence of the latter, but not *vice versa*.

There are two distinctions between different notions of essence that will be central to our proposal: the distinction between *constitutive* and *consequential* notions and between *immediate* and *mediate* notions. While a constitutive notion covers only what is part of the definition of the object, the corresponding consequential notion in addition covers all the logical consequences of the definition; and an immediate notion is only concerned with what is directly part of the definition of the object, while the corresponding mediate notion additionally takes into account what is part of the definition by virtue of being part of a definition of something that occurs in the original definition. Since both distinctions cut across each other, we end up with four notions: immediate and mediate constitutive essence; immediate and mediate consequential essence.[16]

[16] Fine (2012) makes an analogous distinction between immediate and mediate ground.

Let us illustrate these distinctions by way of examples. Suppose that being the successor of 1 is directly part of the definition of the number 2, while being the successor of 0 is directly part of the definition of 1; i.e. the following are immediately constitutively essential to 2 and 1 respectively:

(6) $2 = S(1)$.

(7) $1 = S(0)$.

Now consider:

(8) $\exists x\ 2 = S(x)$.

(9) $2 = S(S(0))$.

(10) $\exists x\ 2 = S(S(x))$.

First, since (8) is a logical consequence of (6), (8) is part of 2's *immediate consequential* essence. Second, since 2 is defined as the successor of 1, while 1 is defined as the successor of 0, (9) is the result of substituting definiens for definiendum. Hence, (9) is part of 2's *mediate constitutive* essence. Finally, since (10) is a logical consequence of (9), it is part of 2's *mediate consequential* essence.

Focussing now on semantics, we follow Fine in drawing on a concept of *semantic requirement*. Fine (2007, p. 43) distinguishes between facts that are 'semantic as to topic' and facts that are 'semantic as to status'. While the first class includes all facts that 'pertain to the exemplification of semantic properties or relations', the second class contains only facts that are semantic in a more narrow sense by belonging 'to the semantics of a given language'.[17] For instance, the fact that 'Schnee ist weiß' is a true sentence of German is merely semantic as to topic, while the fact that the sentence is true in German iff snow is white may be considered to be semantic as to status. Semantic requirements are facts that are semantic as to status; as Fine (2010, p. 66) suggests, they can be thought of as *laws* governing the relevant language.

Some semantic requirements are requirements *simpliciter*, while others are *relative* requirements *on* certain expressions, holding '*in virtue of* their meaning or semantic features', concerning them 'as *source* rather than simply as *subject*', as Fine (2007, p. 123) puts it. An example: if 'he' functions as an anaphor for 'John', then the requirement for these two expressions to be coreferential has 'he' as its source; it is a requirement on the anaphor, not on the antecedent. The relative/simpliciter distinction can thus be understood in analogy to the distinction between essence and metaphysical modality given in Fine (1994): just as it is necessary that {Socrates} contains Socrates, while this is essential to {Socrates} without being essential to Socrates, it is a semantic requirement simpliciter that 'he' and 'John' are coreferential, but only a semantic requirement on the anaphor.

Fine conceives of semantic requirements in a consequentialist spirit, taking them to be closed under a restricted form of logical consequence, what Fine (2007, p. 48) calls *manifest consequence*. Informally, manifest consequences can be characterised

[17] Fine's distinction is closely related to Salmon's distinction between *pure* and *applied semantics*, see (Salmon 2007, p. 179).

as classical consequences that would be evident to an ideal cogniser.[18] We use $\Box$ to express consequentialist semantic requirements simpliciter, so that $\Box\Phi$ can be read as *it is a semantic requirement that* Φ. Correspondingly, we use $\Box_\alpha$ to express relative consequentialist semantic requirements, so that $\Box_\alpha\Phi$ can be read as *it is a semantic requirement on the expression* α *that* Φ. If $\vdash^{\mathrm{m}}$ stands for the relation of manifest consequence, the closure principle on relative semantic requirements can be expressed as follows:

$$\mathrm{C}^{\mathrm{m}}_{\Box_\alpha} \quad \text{If } \Phi_1, \ldots, \Phi_n \vdash^{\mathrm{m}} \Psi, \text{ and if } \Box_\alpha\Phi_1, \ldots, \Box_\alpha\Phi_n, \text{ then } \Box_\alpha\Psi.$$

A main reason for Fine to focus on a *consequentialist* notion of semantic requirements seems to be a concern about how to derive semantic requirements concerning complex expressions. Fine (2007, p. 47) observes that 'the compositional character of semantics requires that we should derive semantic facts concerning complex expressions from the semantic facts concerning simpler expressions.'

Fine needs, however, an additional principle to account for compositionality. For example, let us grant that the following two semantic requirements are obtainable in a Finean consequentialist approach:

(11) $\Box_{\text{‘es regnet nicht’}} \big(T(\text{‘es regnet nicht’}) \leftrightarrow \neg T(\text{‘es regnet’})\big).$

(12) $\Box_{\text{‘es regnet’}} \big(T(\text{‘es regnet’}) \leftrightarrow \text{it is raining}\big).$

That is, 'es regnet nicht' is semantically required to be true iff 'es regnet' is not true, and 'es regnet' is semantically required to be true iff it is raining. Then, 'es regnet nicht' should also be semantically required to be true iff it is not raining:

(13) $\Box_{\text{‘es regnet nicht’}} \big(T(\text{‘es regnet nicht’}) \leftrightarrow \neg\, \text{it is raining}\big).$

The question is how (13) can be obtained from (11) and (12) by means of plausible principles for the notion of a relative semantic requirement.

Note that (13) does not follow from (11) and (12) by means of the closure principle $\mathrm{C}^{\mathrm{m}}_{\Box_\alpha}$, since (11) and (12) give semantic requirements for *different* expressions, 'es regnet nicht' and 'es regnet', respectively. What is needed is a principle which allows one to infer that something is a semantic requirement on an expression α from claims (at least some of) which do not concern semantic requirements on α. For this reason, Fine (2007, p. 123) introduces a principle he calls 'Chaining':

[18] There is a debate about how to give a more precise definition of manifest consequence; see (Fine 2007, p. 48) and (Fine 2007, 136n14) for two proposals, and Salmon (2012), Fine (2014), Weiss (2014), and Salmon (2015) for discussion. For our purposes, the informal characterisation suffices.

$$\frac{\Box_\alpha\, \Phi(\alpha, \beta) \quad \Box_\beta\, \Psi(\beta)}{\Box_\alpha\, (\Phi(\alpha, \beta) \wedge \Psi(\beta))}$$

This principle allows one to import the content of one requirement into the content of another. Applying it to (11) and (12) leads to

(14) $\Box_{\text{‘es regnet nicht’}} \big((T(\text{‘es regnet nicht’}) \leftrightarrow \neg T(\text{‘es regnet’})) \wedge (T(\text{‘es regnet’}) \leftrightarrow \text{it is raining}) \big).$

And from this, (13) follows by $C^{m}_{\Box_\alpha}$.

3.2 The first idea: constitutive instead of consequential semantic requirements

To obtain a semantics that exactly captures the meanings of the object-language expressions, we suggest that the truth-theoretic axioms should employ an operator expressing the notion of a semantic requirement.[19] More precisely, if the truth-theoretic axiom mentions an object language expression α, then it should employ an operator expressing the notion of a relative semantic requirement on α. However, although this Finean idea points to a solution of the information problem, it does not quite get us there. The problem is that by employing a *consequentialist* notion of semantic requirement, one ends up with a theory that is not uniquely characterising (as is witnessed, for example, by two simple predicates corresponding to the manifestly equivalent ‘is wise’ and ‘is wise and self-identical’).

Now, corresponding to the distinction between a consequentialist and a constitutive notion of essence, it seems natural to draw an analogous distinction within the domain of semantics, distinguishing merely consequentialist and truly constitutive requirements. Take Fine’s example of the anaphor. Surely, the difference between the following requirements

(15) $\Box_{\text{‘He’}}$ ‘He’ is coreferential with ‘John’.

(16) $\Box_{\text{‘He’}}$ ‘He’ refers to John.

(17) $\Box_{\text{‘He’}}$ ‘He’ refers to John or molehills are mountains.

quite neatly mirrors the distinction between immediate constitutive, mediate constitutive, and consequential essence. We suggest to explicitly acknowledge constitutive semantic requirements, and to investigate how they can be put to work. More strongly, we suggest that a truth-theoretic axiom for α should employ an operator expressing the notion of a relative *constitutive* semantic requirement on α. Only such a semantically fine-grained notion allows one to formulate an axiom that exactly captures the meaning of α and therefore provides all the basic semantic information about this expression.

We will use $\boxdot$ to express constitutive semantic requirements simpliciter, so that $\boxdot\Phi$ can be read as *it is a constitutive semantic requirement that* Φ. Correspondingly,

[19] This idea is proposed in Haverkamp (2015).

we will use $\boxdot_\alpha$ to express relative constitutive semantic requirements, so that $\boxdot_\alpha\Phi$ can be read as *it is a constitutive semantic requirement on the expression* α *that* Φ. We assume that $\boxdot_\alpha\Phi$ can only be true if α is mentioned in Φ, for which we write $\alpha \in \Phi$. How are the notion of a (constitutively) essential property and the notion of a (constitutively) semantic requirement on some expression related? Like Fine we think that these notions come down to the same thing: a semantic requirement on some expression is nothing but an essential semantic property of this expression.

Note that this claim presupposes a *thick* conception of linguistic expressions, according to which their meanings are essential to them (cp. Fine 2007, 135). For example, one would say that it is an essential property of 'house' to mean *house*, rather than *dog*. On the alternative 'thin' conception of linguistic expressions, one would say that a word could have a different meaning—still being the same word.[20]

Many linguistic phenomena can be adequately accounted for on both conceptions, though in slightly different terms. Consider lexical change as an example:

> (Thin) In Middle High German the word 'witzig' meant *smart*, in Standard German this same word means *funny*.
>
> (Thick) Middle High German had a word spelled 'witzig' meaning *smart*, Standard German has a (different) word spelled 'witzig' which means *funny*.

A thick conception of expressions is compatible with the observation that it is arbitrary, perhaps conventional in the sense of Lewis (1975), which sequences of sounds/marks go with which meanings in some linguistic community. For example, instead of saying that there is a regularity among speakers of Standard German to use a (thinly conceived) expression only if they have certain attitudes towards a possible meaning—e.g. to use 'Peter ist witzig' only when they believe that Peter is funny—a Lewisian might alternatively say that there is a regularity among these speakers to use certain sequences of sounds/marks only if they have such attitudes, thereby bringing it about that a certain (thickly conceived) expression is part of the language of their community (where this could, but need not, be understood as involving reference to a *thickly conceived language*, i.e. one that has its expressions essentially, see note 3). A semantics based on thick expressions thus does not preclude a metasemantics based on thin expressions.

For many purposes, it is thus insignificant which metaphysics of expressions one adopts. Adopting a thick conception of expressions allows us to literally understand relative semantic requirements as essential properties of the relevant expressions. Without defending it any further, we will thus adopt this thick account of linguistic expressions and this essentialist understanding of relative semantic requirements.

Our proposal, then, is the following. A linguistic theory for some object language $\mathcal{L}$ should be concerned with the *real definitions* of the expressions of $\mathcal{L}$—it should state what their *constitutive essences* are. This will at least involve specifying their *syntactic* essences and their *semantic* essences, but here we will focus on their semantic

[20] Similarly for syntactic properties: on a thin conception, one might say that it is an accidental property of the expression 'red house' that is has the adjective 'red' and the noun 'house' as its immediate constituents—that this very expression could also have been, for example, a simple noun. On a thick conception, one would say that such syntactic properties belong to the essence of 'red house'.

essences.[21] Adopting a truth-theoretic approach (but see Sect. 4), we suggest that the constitutive semantic essence of a lexical item can be given by an interpretive specification of the truth conditional contribution it makes to sentences.[22] Hence, we take the axioms of an interpretive truth theory to express the content of the semantic part of a *real definition* of the object language expression mentioned in the axiom. To strengthen simple truth theories, we propose to attach indexed sentential operators for constitutive essence to truth-theoretic axioms, so that the resulting strengthened axioms explicitly convey their definitional character. Such axioms explicitly state what the constitutive semantic requirements on the lexical items of the object language are. In the following, when we use 'requirement', we mean *constitutive semantic requirement*, unless stated otherwise.

3.3 The second idea: replacement instead of chaining

To illustrate how semantic requirements can be employed in giving a compositional semantics, we now turn to the principles governing this notion. We replace Fine's *Chaining* principle, which merely allows the importation of the content of one (consequentialist) requirement into that of another, with a different principle about how (possibly mediate, but constitutive) requirements interact. As a result, Fine's closure principle, which is not plausible for a constitutive notion, can be dispensed with; compositionally deriving requirements for complex expressions can be done more directly via our new principle.

Recall that requirements can be thought of as (the semantic component of) real definitions of expressions. Central to our proposal is the idea of the *expansion* of a given definition. Suppose that Φ is defined in terms of Ψ, which is itself defined in some way. In such a case, we can expand Φ's definition by replacing Ψ with its definiens. The result of this expansion will itself have definitional status. Here is a non-linguistic example. If knowledge is defined in terms of truth:

x knows $p \Leftrightarrow_{\text{def}} p$ is true, and …,

and if truth is defined as correspondence:

p is true $\Leftrightarrow_{\text{def}} p$ corresponds to a fact,

[21] The syntactic part of the theory would identify the syntactic categories of all the lexical items as well as the constituent structure of complex $\mathcal{L}$-expressions. On our account we would say that the theory would explicitly state that these syntactic properties are parts of their constitutive essences (or, in other words, that they are *constitutive syntactic requirements* on the expressions in question). For instance, the syntax will not only tell us that 'Sokrates ist weise' is a sentence, but that it is part of the constitutive essence of this expression that it is a sentence consisting of (let us say) a proper noun 'Sokrates' and a verb phrase 'ist weise'. In the following, we will leave the syntactic part of linguistic theories tacit. Moreover, instead of full-fledged phrase structure trees or some analogous device, we will simply use quotation expressions to refer to object language expressions. These should be understood as placeholders for whatever the correct canonical structural descriptions are that the syntax provides.

[22] We do not want to claim that every aspect of lexical meaning can be treated by such truth-theoretic means. We will indicate how some problematic cases might be dealt with (Sect. 3.4), but we are aware that there are many types of lexical items and complex expressions (pejoratives, honorifics, expressives, etc.) a complete semantic treatment of which requires a supplementation and/or revision of our approach. We think, however, that such a supplementation/revision would not undermine the idea of using constitutive semantic essences to solve the information problem.

then knowledge is also defined in terms of correspondence:

x knows $p \Leftrightarrow_{\text{def}} p$ corresponds to a fact, and ....

Turning to linguistic cases, suppose a real definition of one expression links it to a second expression. By replacing the definiendum by the definiens of a real definition of the second expression, one will likewise obtain an expanded version of the initial definition. This points the way towards a compositional account of requirements. For a complex expression α, there will be a requirement on α in terms of α's constituents, while these constituents will in turn come with their own requirements. Expanding the initial requirement on α by filling in the requirements of α's parts, we obtain an interpretive requirement on α which exactly captures α's meaning.

For purposes of illustration, we again work with clausal extensional truth theories. Moreover, we restrict ourselves to some core truth-theoretic treatments: sentential operators, predicates and singular terms (in this section), and quantifiers (in the appendix). Even given these restrictions, there is still a wide variety of techniques that have been employed in treating such constructions, and there are different possibilities for how to apply our approach. We consider these differences to be of some semantic importance, but in this paper we focus on the main idea, leaving a discussion of various options of implementation for the future.[23]

The informal idea of expanding definitions leads to different formal implementations, depending on the form of the definition. Suppose an expression β is a constituent of an expression α, and a definition of α, $\boxdot_\alpha \varphi$, contains the definiendum of a definition of β. One can then replace definiendum by definiens to obtain an expanded definition of α, $\boxdot_\alpha \varphi^*$. Depending on the form the definition of β takes, there will be different implementations, which we all call 'versions of *Replacement*'. As examples, we will treat definitions based on simple biconditionals, identities, and quantified biconditionals. Ultimately, one wants a comprehensive account (cp. Sect. 4), but here we will only present those formal implementations that are relevant to the linguistic constructions at issue.

For our first example, we consider $\boxdot_\beta$-biconditionals in which only one side mentions the pertinent expression β. For specificity, we focus on cases where the expression is mentioned on the left hand side:

$$\boxdot_\beta(\psi \leftrightarrow \chi), \text{ where } \beta \notin \chi$$

We call such claims *constitutive equivalences (for β)*. In such a constitutive equivalence, ψ expresses a condition on β, and χ expresses what β's satisfying this condition consists in. An example is

$$\boxdot_{\text{'Sokrates ist weise'}}(T(\text{'Sokrates ist weise'}) \leftrightarrow \text{Socrates is wise}),$$

[23] If some natural language constructions are amendable to truth-theoretic treatment in principle, but only in a way that excludes the strengthening with our proposed operator, we would consider that to be a significant problem for our approach. However, we do not at present see reasons to expect this. On the other hand, we do think it is possible that our approach will turn out to be a better fit with some of a variety of competing truth-theoretic approaches for a given construction. In such cases, our approach provides novel motivation for choosing among alternative proposals. We take quantification to be one candidate here; see the appendix for details.

stating that it is a requirement on the sentence ‘Sokrates ist weise’ to be true iff Socrates is wise. We take such constitutive equivalences to be prime examples of *real definitions*. In line with the idea of definition expansion, we propose the following first version of *Replacement*:[24]

$$\begin{array}{ll} & \boxdot_{\alpha}\varphi \\ & \boxdot_{\beta}(\psi \leftrightarrow \chi) \quad \alpha \notin \psi, \beta \notin \chi \\ \Rightarrow & \boxdot_{\alpha}\varphi[\psi/\chi] \end{array}$$

In the relevant instances of this rule, the first line will state a requirement on α in which a condition on β, namely ψ, occurs, and we have a requirement on β in which it is stated what this condition on β consists in, namely χ. One can then replace the condition on β in the original requirement by what it consists in.[25]

Let us first turn to negation. The negated sentence ‘Sokrates ist nicht weise’ ($= \alpha$) means that Socrates is not wise. Our goal is to illustrate how to derive the relevant interpretive constitutive equivalence[26]

(18) $\boxdot_{\text{‘SNW’}}$ (T(‘Sokrates ist nicht weise’) $\leftrightarrow \neg$ Socrates is wise).

from the interpretive constitutive equivalence for the contained sentence ‘Sokrates ist weise’ ($= \beta$)

(19) $\boxdot_{\text{‘Sokrates ist weise’}}$ (T(‘Sokrates ist weise’) $\leftrightarrow$ Socrates is wise).

and an appropriate axiom for ‘nicht’. We propose the following axiom, stating that every German sentence S is such that the result of applying ‘nicht’ to S is defined to be true iff S is not true:[27]

(A4) $\forall S[\text{Sent}(S) \rightarrow \boxdot_{\text{‘nicht’}^\frown S}(T(\text{‘nicht’}^\frown S) \leftrightarrow \neg T(S))]$.

Instantiating for S,[28] and relying on the syntactic information that ‘Sokrates ist weise’ is a German sentence, *Modus ponens* yields:

24 $\varphi[\psi/\chi]$ is the result of replacing all occurrences of ψ in φ by occurrences of χ.

25 Note that *Replacement* allows one to combine requirements for different expressions without introducing additional material, such as the notion of conjunction. In this regard it differs from *Chaining* as Fine understands it.

26 We will occasionally use obvious abbreviations in the index of $\boxdot$, e.g. ‘SNW’ for ‘Sokrates ist nicht weise’.

27 The syntax of negation is a complex issue that we cannot address here. In order to sidestep this issue and to illustrate the basic semantic idea, we simply use ‘the result of applying “nicht” to S’ and ‘“nicht”$^\frown S$’ to denote the negation of S.

28 Similar to the case of quantified modal languages, the rule of universal instantiation has to be restricted here; if a universal quantifier binds a variable inside the scope of $\boxdot$, one cannot in general instantiate the variable with an arbitrary term. Here, we assume that syntax specifies *canonical terms* which can be used in this context; cp. the appendix.

(20) $\boxdot_{\text{'SNW'}}\ (T(\text{'Sokrates ist nicht weise'}) \leftrightarrow \neg T(\text{'Sokrates ist weise'}))$.

We can now apply *Replacement* to (20) and (19) to arrive at (18), our target theorem.[29] Sentential operators of arity higher than 1 can be dealt with analogously.

Next, let us turn to predication. There are various options for a truth-theoretic treatment of predication. For purposes of illustration, we consider employing a term-forming operator 'R' (for 'the referent of'). This allows us to illustrate that the informal idea of expanding definitions leads to a second version of *Replacement* if the relevant definitions are *identities* rather than biconditionals:

$$\boxdot_{\beta}(s = t), \text{ where } \beta \notin t$$

We also call these *constitutive equivalences*. Here, t discloses the identity of the object which s, a term mentioning β, denotes. For example:

(A5) $\boxdot_{\text{'Sokrates'}}\ (R(\text{'Sokrates'}) = \text{Socrates})$.

that is, it is a requirement on the name 'Sokrates' to have Socrates as its referent. For the predicate, we here use the following axiom, analogous to the one for the sentential operators:

(A6) $\forall N[\text{Name}(N) \rightarrow \boxdot_{\text{'ist weise'}^\frown N}(T(\text{'ist weise'}^\frown N) \leftrightarrow R(N) \text{ is wise})]$.

That is, every German name N is such that the application of the predicate 'ist weise' to N is defined to be true iff the referent of N is wise. Instantiating for N, relying on the syntactic information that 'Sokrates' is a German name, *Modus ponens* yields:

(21) $\boxdot_{\text{'Sokrates ist weise'}}\ (T(\text{'Sokrates ist weise'}) \leftrightarrow R(\text{'Sokrates'}) \text{ is wise})$.

Now, we have to combine this with the constitutive equivalence (A5). To this end, we use a suitable version of *Replacement*:

$$\begin{array}{lll} & \boxdot_{\alpha}\varphi & \\ & \boxdot_{\beta}(s = t) & \alpha \notin s, \beta \notin t \\ \Rightarrow & \boxdot_{\alpha}\varphi[s/t] & \end{array}$$

Applying this rule to (A5) and (21) yields the desired interpretive theorem. Predicates of arity higher than 1 can be dealt with analogously.

Note that the axioms we have considered for sentential operators as well as for predicates do not simply specify a requirement on the expressions in question. Rather, they state a requirement on the result of combining that expression with some input (e.g. a sentence or a name). This is witnessed by the fact that the $\boxdot$ in these axioms has a complex index (for instance $\boxdot_{\text{'nicht'}^\frown S}$ rather than simply $\boxdot_{\text{'nicht'}}$). However, one might

[29] So in this instance of *Replacement*, we have φ = '$(T(\text{'Sokrates ist nicht weise'}) \leftrightarrow \neg T(\text{'Sokrates ist weise'}))$', ψ = '$T(\text{'Sokrates ist weise'})$', χ = 'Socrates is wise'.

prefer that a lexical axiom for some expression α should state a requirement solely on α, and hence be governed by $\boxdot_\alpha$. We briefly indicate how this can be achieved for a sentential operator such as negation (the case of predication is analogous). Again, our focus here is on the main idea rather than on the details of different ways of implementation.

A simple index axiom for ‘nicht’ would state that it is a requirement on ‘nicht’ that the result of applying ‘nicht’ to a German sentence is true iff that sentence is not true:

$$\boxdot_{\text{‘nicht’}}\forall S[\text{Sent}(S) \to (T(\text{‘nicht’}^\frown S) \leftrightarrow \neg T(S))].$$

Note that this would be a different kind of real definition, requiring a rule that is not simply a *definiens-for-definiendum* version of *Replacement*. Restricting attention here to the case of a unary operator/predicate α, building a complex expression $\alpha^\frown\beta$ if applied to a suitable expression β, we propose a rule which combines restricted universal quantifier elimination with the move from a simple to a complex index:[30]

$$\begin{array}{ll} & \boxdot_\alpha \forall x\,(F(x) \to \psi) \\ & \boxdot_\beta F(\beta) \\ \Rightarrow & \boxdot_{\alpha^\frown\beta}\psi[x/\beta] \end{array}$$

Take an example. Since it is a requirement on ‘nicht’ that the result of applying ‘nicht’ to a sentence is true iff that sentence is not true, and since syntax tells us that it is a requirement on ‘Sokrates ist weise’ to be a sentence, we can apply this rule to derive that it is a requirement on the result of combining ‘nicht’ and ‘Sokrates ist weise’ to be true iff ‘Sokrates ist weise’ is not true. Given the appropriate theorem for ‘Sokrates ist weise’, the first version of *Replacement* now yields the desired interpretive theorem for the negated sentence ‘Sokrates ist nicht weise’. We take this extension of the aforementioned approach, which was based exclusively on *definiens-for-definiendum* versions of *Replacement*, to be especially interesting, because it observes an interplay of semantic and syntactic requirements—a topic we will not go into here.

Let us take stock. Our proposal is to employ the notion of relative constitutive semantic requirements to strengthen existing semantic approaches. In order to derive requirements for complex expressions on the basis of the requirements of their constituents, we have suggested drawing on the idea of definition expansion. Since there typically are various proposals for how a given object language construction should be treated in a broadly truth-theoretic framework, there is a corresponding range of possible strengthened theories. Moreover, even starting from a single approach for a certain construction, there will often be various options of how to implement our proposal (corresponding, for instance, to various possible placements of the $\boxdot_\alpha$-operator). The goal of this section was to illustrate the feasibility of the main idea rather than to settle the details.[31]

[30] In the first premiss, x and α are only allowed to occur in the context $\alpha^\frown x$.

[31] We are aware of the recent work in Fine (2015) which might inspire a different way to present both axioms and theorems: using a special two-place, variable-binding operator, merging the effect of $\boxdot$, a prologue such as ‘$\forall S[\text{Sent}(S) \to$’, and the embedded $\leftrightarrow$ in axioms such as (A4). Here we do not enter into the pros and cons of this alternative, but simply stick to the more conservative route.

3.4 Constitutive requirements, unique characterisation, and the information problem

This section puts forward three claims of increasing strength: *(i)* that our approach does not fall victim to the examples we used to show that other approaches fail to be uniquely characterising; *(ii)* that our approach can in fact yield theories that are uniquely characterising; *(iii)* that it can yield theories that exactly capture meaning and avoid the information problem.

We cannot give a rigorous proof of these claims, as this would require a substantive account of the notion of exactly capturing and a theory of synonymy; tasks that are beyond the scope of this paper. However, we do suggest that there are strong reasons for believing these claims. To see this, it is useful to make one further wrinkle explicit that we have so far left tacit: the distinction between three levels of constitutive semantic requirements. The first two, *immediate* and *mediate*, are familiar from Fine's work. We add a third level of *semantically fully expanded* constitutive semantic requirements. Roughly speaking, the three levels correspond to three stages in a derivation of truth conditions: we start from immediate requirements, move through mediate ones, and the derivation terminates once every semantic definiendum has been fully expanded.

Consider a conjunction 'S_1 und S_2'. For simplicity, assume that the constituent sentences S_1 and S_2 are treated as simple, receiving axioms stating that S_1 is required to be true iff p and that S_2 is required to be true iff q. A derivation starts with the axiom for 'und', expressing the immediate requirement that 'und' forms true sentences iff combined with two true sentences. By instantiating for S_1 and S_2, we derive an immediate requirement on the complex expression 'S_1 und S_2', stating that it is required to be true iff both S_1 and S_2 are true. We now appeal to the requirement on S_1 to be true iff p and derive the mediate requirement on 'S_1 und S_2' to be true iff p and $T(S_2)$. Finally, appealing to the requirement on S_2 to be true iff q, we derive the semantically fully expanded requirement on 'S_1 und S_2' to be true iff p and q. This last requirement no longer defines the truth of 'S_1 und S_2' in semantic terms, and hence our derivation terminates. We use $\circledcirc_\alpha \Phi$ to express that Φ is an immediate requirement on α. We continue to use $\boxdot_\alpha \Phi$ to express that Φ is a requirement on α, where this is neutral between the levels of requirements.

Ad *(i)*. Our proposal does not fall victim to logically equivalent* but non-synonymous expressions, since it is not based on a *consequentialist* notion. Roughly speaking, while a consequentialist notion is closed under logical (or perhaps manifest) consequence, our constitutive notion is merely closed under definition expansion that results from replacing definienda by definentia. Consider, for example, a simple expression α_1 meaning *is wise* and a simple expression α_2 meaning *is wise and self-identical*. Now, since $\boxdot_\alpha$ expresses a *constitutive* (and not just a consequential) semantic requirement on α, it is correct to say

$$\boxdot_{\alpha_1} \forall N[\text{Name}(N) \rightarrow (T(\alpha_1^\frown N) \leftrightarrow R(N) \text{ is wise})],$$

but incorrect to say

$$\boxdot_{\alpha_2} \forall N[\text{Name}(N) \rightarrow (T(\alpha_2^\frown N) \leftrightarrow R(N) \text{ is wise})].$$

It is *not* a constitutive semantic requirement on a predicate meaning *is wise and self-identical* that it forms true sentences iff combined with a name of a wise thing; you cannot arrive at such a characterisation of α_2 by merely expanding definienda. Thus, the shift to a constitutive characterisation excludes formerly problematic cases.

Ad *(ii)*. We suggest that our proposal provides the tools to formulate uniquely characterising theories. In a sharpening of the broad approach set out in the previous sections, we add that the lexical axioms should state immediate requirements, and hence be formulated by employing $\boxdot_\alpha$. We claim that, in the case of a simple expression α, the $\boxdot_\alpha$-axiom by itself is already uniquely characterising. Recall that a requirement on α is immediate if it does not hold in virtue of other requirements. In our view, the only way to state immediate requirements for lexical items in broadly truth-theoretic terms is by being *interpretive*—that is, by stating the truth conditional contribution using a metalanguage expression synonymous with the pertinent lexical item. Assume, for simplicity, that we treat monadic predicates by axioms of the following form:

$$\ldots \leftrightarrow R(N) \text{ is } \phi.$$

Now let α be a predicate meaning *is wise*. If there are *any* constitutive requirements of the above kind on α, then surely one of them must be the one obtained by replacing ϕ with ‘wise’:

$$\boxdot_\alpha \ \ldots \leftrightarrow R(N) \text{ is wise.}$$

But it also seems clear that this is the most basic of all the constitutive requirements on α—any other constitutive requirement of this form will only hold in virtue of this one. Hence, true $\boxdot_\alpha$-axioms must be interpretive.

Turning to complex expressions, note that $\boxdot_\alpha$-theories will, by virtue of their syntactic component, provide information about the syntactic structure of a given complex expression, and, by virtue of their semantic component, information about the immediate requirements governing the simple parts of the complex expression. We claim that, taken together, this provides uniquely characterising information about complex expressions. Take a simple predication, ‘Sokrates ist weise’. Assume that syntax tells us (very roughly) that ‘Sokrates ist weise’ = ‘ist weise’$^\frown$‘Sokrates’, while semantics tells us what the immediate requirements on ‘ist weise’ and ‘Sokrates’ are. Now, if β_1 is synonymous with ‘Sokrates’ while β_2 is synonymous with ‘ist weise’, then combining β_1 and β_2 in the way in which ‘ist weise’ and ‘Sokrates’ are combined in the relevant sentence will result in a sentence synonymous with ‘Sokrates ist weise’. Hence, building on the unique characterisation for simple expressions, the theory will also be uniquely characterising for complex ones.

In this paper, we have left the notion of meaning mostly intuitive. There are arguably various possible sharpenings of ‘(linguistic) meaning’, and the conditions of adequacy on semantic theories will depend on which sharpening the theory aims at. We think there is one natural and relevant notion of linguistic meaning that lies at the intersection of truth conditional contribution and what is definitional. Relying on such a notion, it is plausible to assume that an expression α is synonymous with ‘und’, for example, if there is an immediate constitutive requirement on α to form a true sentence iff it is applied to two true sentences. Arguably, this would yield a notion of synonymy according to which ‘und’ and ‘aber’ (like ‘and’ and ‘but’) are synonymous. To be sure, there are also more fine-grained notions of linguistic meaning. We think that our approach could be fleshed out in a way that allows it to discriminate between ‘und’ and ‘aber’ by explicitly incorporating information about the *completeness* of the requirements provided by the theory. Suppose that a theory not only provides the immediate requirement on ‘und’, but also states that this is the *only* requirement governing ‘und’. Since it seems plausible that ‘aber’ is governed by additional requirements regarding, for instance, felicity, such a theory could be taken to be uniquely characterising even under the assumption that ‘und’ and ‘aber’ diverge in meaning.

Ad *(iii)*. Do theories employing $\boxdot_{\alpha}$-axioms exactly capture meaning? Would knowing such a theory suffice for knowing the object language? We have argued that true $\boxdot_{\alpha}$-axiom are guaranteed to be interpretive. Moreover, recall that the rules governing constitutive requirements only allow for definitional expansion. A fully expanded constitutive equivalence derivable from interpretive $\boxdot_{\alpha}$-axioms is itself guaranteed to be interpretive, where a constitutive equivalence is *fully expanded* if we cannot replace any more definienda by definentia via applications of *Replacement*. Thus, by employing an operator in the language of the theory which, when applied to lexical axioms, forces interpretiveness, and which is governed by rules that ensure an interpretive end result, we not only obtain a theory that is uniquely characterising. Rather, since these facts transparently flow from the notion expressed by the operator, the theory will be exactly capturing, and knowledge of the theory will suffice for knowledge of meaning.

4 Conclusion

What does it take to know the meaning of ‘Sokrates ist weise’? What does it take to exactly capture its meaning? It has been the guiding thought of the mainstream approach to semantics that knowledge of meaning involves knowledge of truth conditions. In its simplest form, the claim would be that knowing what ‘Sokrates ist weise’ means comes down to knowing that this sentence is true iff Socrates is wise; more sophisticated truth-theoretic machinery allows for more sophisticated versions of the claim. As the information problem illustrates, this will not do. One goal of our paper was to emphasise that the information problem is central, and remains unsolved. The second goal of this paper was to illustrate how the information problem can be overcome. To bring out what is correct about the motivating thought, we need to make explicit the *definitional connection* between truth and truth condition. We have proposed to do this in terms of relative constitutive requirements, and we have argued that there is a central notion of linguistic meaning which is fruitfully explicated in these terms.

Depending on which aspect of our proposal one focusses on, it can be viewed as a conservative continuation of the truth-theoretic approach, or as a radical departure from the commonly accepted path. It is conservative in the sense that the core machinery of the semantic theory is left untouched (this might be a slight simplification; see, for instance, the appendix). Given a truth theory of your choice, our proposal can more or less piggyback on this theory to produce a strengthened theory explicitly stating what the constitutive semantic requirements governing the object language are. The more radical aspect of the proposal lies in the employment of a relativised hyperintensional operator in the metalanguage. This move sets our proposal apart from other non-standard approaches such as truthmaker semantics or impossible-worlds semantics, which stick with an extensional metalanguage and instead inflate the ontology of the theory.

In this paper, we have introduced the main idea of semantic theories based on the notion of constitutive requirements, and we have begun to illustrate how such theories can be formulated. However, we have merely touched upon a number of issues that deserve a fuller investigation. In closing, we want to indicate five related areas in which further work would seem to be promising and/or necessary.

(i) In this paper, we confined our approach to fairly simple constructions (basic predication, sentential operators, first-order quantification). While we are optimistic that our approach can be combined with truth-theoretic treatments of more complicated constructions (e.g. modal operators, adverbial modification, gradable adjectives, and so on), this optimism should be substantiated by providing detailed accounts of the relevant constructions in terms of constitutive requirements.

(ii) For purposes of illustration, we have assumed a clausal format for truth theories and illustrated how, based on the respective axioms of a clausal truth theory, we can bring in the notion of constitutive requirement to produce a strengthened theory. However, as this exercise has already illustrated, we will sometimes have various options of how exactly to proceed, e.g. by choosing different possible placements for the semantic requirement operator. These various options deserve a fuller investigation, as do the options that arise when applying our approach to semantic value-based truth theories.

(iii) Our main aim in this paper was to argue that ordinary semantic theories can be strengthened with $\boxdot_\alpha$ in order to solve the information problem. On the resulting picture, $\boxdot_\alpha$ could be perceived as a mere add-on; the first-order semantic work of coming up with treatments for natural language constructions proceeds just as before, the results can then be strengthened with $\boxdot_\alpha$. But once we have admitted $\boxdot_\alpha$ into our tool kit, it is natural to ask whether it could contribute to the first-order work as well. For instance, it might be worth considering whether we can construe propositional attitude operators or other non-extensional constructions as sensitive to the constitutive requirements governing their inputs, rather than merely in terms of the truth conditional properties of their inputs. We suggest this as another potentially fruitful area for further research.

(iv) As we have briefly indicated with respect to the contrast between ‘und’ and ‘aber’, there may be reason to acknowledge additional kinds of constitutive requirements besides those concerning truth conditional contribution. Perhaps there is work to be done for a notion of constitutive requirements concerning what a standard use of a given expression conveys over and above its truth conditional contribution. Giving a linguistic theory of a particular language could then be construed as the investigation into the essences of the expressions of that language, where this involves specifying syntactic, semantic/truth conditional, and semantic/conveying requirements governing these expressions, and the way these requirements interact.

(v) Finally, while we have focussed on truth-theoretic accounts of meaning, the underlying idea might be applied to other frameworks aswell. Broadly speaking, we suggest a semantics in terms of constitutive semantic requirements on the object language expressions: immediate such requirements on the simple expressions, mediate ones on the complex expressions. It calls for further work to develop a general account of mediate constitutive requirements on expressions and, ultimately, mediate constitutive essence in general, which can be applied to alternative semantic frameworks. In particular, we assume that the proposed variants of *Replacement* have to be considerably supplemented if one aims at a general account of the informal idea of expanding (real) definitions.

Acknowledgements We are grateful to two anonymous referees of this journal for detailed comments that helped improve this paper. We have discussed the themes of this paper with many colleagues and students. We would like to thank: Singa Behrens, Samuel Boardman, Vanessa Carr, Scott Dixon, Lars Dänzer, Alexander Dinges, Katharina Felka, Max Goetsch, David Heering, Dirk Kindermann, Christian Nimtz, Raphael van Riel, Stefan Roski, Thorsten Sander, Benjamin Schnieder, Alex Steinberg, Maik Sühr, Hamid Taieb, Barbara Vetter, Emanuel Viebahn, Lisa Vogt, David Wörner, Julia Zakkou, the members of the philangling group, and the participants of Ralf Busse's colloquium in Mainz, Christian Nimtz' colloquium in Bielefeld, Benjamin Schnieder's colloquium in Hamburg, Barbara Vetter's colloquium in Berlin, and Emanuel Viebahn's research seminar in Berlin. We are grateful to Jan Barbanus for his thorough comments on stylistic and grammatical issues. Finally, we are especially grateful to Ralf Busse for providing many insightful comments on various previous versions of this paper.

Funding Open Access funding enabled and organized by Projekt DEAL. Funding was provided by Volkswagen Foundation (Grant No: 88316-2)

5 Appendix

We have presented our approach with respect to names, predicates, and sentential operators. In this appendix we will indicate how it can be applied to quantifiers. To begin with, we will explain the main obstacle quantification presents for a theory of constitutive semantic requirements.

To facilitate focussing on the main point, quantification, we here present our approach with respect to an object language and a metalanguage that are both formal first-order languages. We consider an *interpreted* object language $\mathcal{L}$ with a single binary predicate symbol, assumed to have a fixed interpretation. The alphabet of $\mathcal{L}$ consists of over-lined symbols:

$$\overline{(},\quad \overline{)},\quad \overline{x_1},\quad \overline{x_2},\quad \ldots,\quad \overline{\neg},\quad \overline{\wedge},\quad \overline{\forall},\quad \overline{R}.$$

The metalanguage contains expressions synonymous with these ($\neg, \wedge, \forall, R$), as well as syntactic and semantic vocabulary. In particular, it contains individual constants for the symbols of $\mathcal{L}$—we use boldfaced symbols here—and a two-place function symbol for building sequences—we use concatenations to indicate sequences. A canonical name of a sentence, in short *s-name*, is a sequence of boldfaced symbols denoting a sentence, for example $\overline{\boldsymbol{\forall x_2 \neg \forall x_1 R x_2 x_1}}$. The semantic vocabulary comprises $\boxdot$ and a truth predicate T. We present a semantics with theorems such as

$$(\sharp) \qquad \boxdot_{\overline{\boldsymbol{\forall x_2 \neg \forall x_1 R x_2 x_1}}} \left(T(\overline{\boldsymbol{\forall x_2 \neg \forall x_1 R x_2 x_1}}) \leftrightarrow \forall x \neg \forall y R x y \right).$$

Say that a metalanguage sentence P *corresponds to an s-name* t iff P can be obtained from the object-language sentence that t denotes by erasing the lines on top of the denoted symbols and renaming bound variables. We then make the following claim:

(C_1) For every s-name t and corresponding metalanguage sentence P, our theory will enable us to derive the theorem $\boxdot_t \; (T(t) \leftrightarrow P)$.

Our theory thus yields fully expanded constitutive equivalences for all object-language sentences, thereby exactly capturing their meanings.[32]

The main difficulty in applying our approach to quantification consists in the fact that standard semantics for quantification uses *mathematical* vocabulary whose interaction with a notion of a constitutive semantic requirement is not obvious. For example, on a standard approach to quantification, one would expect an axiom such as

$$\forall \nu \forall \varphi \forall f \left(\mathrm{Var}(\nu) \wedge \mathrm{Form}(\varphi) \wedge \mathrm{Assign}(f) \rightarrow \boxdot_{\overline{\boldsymbol{\forall}}\nu\varphi, f}(\mathrm{Satis}(f, \overline{\boldsymbol{\forall}}\nu\varphi) \right.$$
$$\left. \leftrightarrow \quad \forall f^*(\nu\text{-}\mathrm{Var}(f^*, f) \rightarrow \mathrm{Satis}(f^*, \varphi)))\right).$$

We do not object to such an extended notion of semantic requirement, which is relative to an expression and an assignment. However, the difficulty for using such an axiom in a derivation of ($\sharp$), for example, is to get rid of the notions of an assignment and of a ν-variant *inside the scope of* $\boxdot$ without turning $\boxdot$ into a consequentialist operator.

Without taking a stance here on whether this indicates an actual deficiency of the standard treatment of quantification, we indicate a slightly different semantics which does not require the use of these mathematical notions. Instead of open formulas, which are interpreted relative to assignments to their free variables, we use *quasi-sentences*. These are either sentences or like sentences except for containing arbitrary objects in some (or all) of the argument places of the predicate symbol $\overline{R}$; e.g. the sequence consisting of $\overline{\forall x_2 R}$ followed by Mary (the person), followed by $\overline{x_2}$.[33] A quasi-sentence that is not a sentence thus consists of a sentential frame saturated by (in general) non-linguistic objects.

[32] We would also like to mention the following complementary claim:

(C_2) For every theorem of the form $\boxdot_t \; (T(t) \leftrightarrow P)$ with an s-name t and a metalanguage sentence P without boldfaced symbols, P corresponds to t.

To defend this second claim, we would have to precisely specify the syntactic component of the proposed theory, which we will not do here.

[33] To be precise, the objects are not allowed to be object language variables. If the object language were to contain singular terms, these objects would also not be allowed to be such terms.

Our fundamental semantic predicate T (for which we write 'is true' in our informal prose) is meant to apply to something iff it is either a true sentence or a quasi-sentence whose sentential frame is true of the objects saturating it. For example, T applies to the quasi-sentence consisting of $\overline{R}$ followed by Mary and Paul, iff $\overline{R}$ is true of Mary and Paul (in that order); that is, this quasi-sentence is true iff Mary is R-related to Paul. The semantic axiom for the universal quantifier can then be the following:

$$\forall x \forall \nu \left(Q(\overline{\boldsymbol{\forall}} \nu x) \rightarrow \boxdot_{\overline{\boldsymbol{\forall}} \nu x}(T(\overline{\boldsymbol{\forall}} \nu x) \leftrightarrow \forall y\, T(x^{y}_{\nu}))\right),$$

where Q stands for being a quasi-sentence. It thus states, roughly speaking, that a universal quantification is constitutively required to be true iff the embedded expression is true of every object, x^{y}_{ν} indicating the result of substituting y for the free occurrences of the variable ν in x.[34]

Our approach to quantification assumes that there are semantic requirements on quasi-sentences. For example, we would say that the concatenation of $\overline{R}$, Mary, and Paul is constitutively required to be true iff Mary is R-related to Paul.[35] One might consider separating the sentential frame from the saturating objects, thus employing a doubly relativised operator (roughly corresponding to 'it's a constitutive semantic requirement on sentential frame s, relative to objects xx'), but then one has to set up additional machinery for relating the two, linking the objects to the empty places of the sentential frame. To avoid such complications, here, we stick to the simpler approach.

We will call a sequence of boldfaced symbols and metalanguage variables a *q-name*. If a q-name stands for a quasi-sentence, then we will speak of a *qs-name*; e.g. $\overline{\boldsymbol{\neg\forall x_1 R}z\boldsymbol{x_1}}$. With respect to the syntactic part of the theory, we merely note two (classes of) theorems it delivers. First, if t is a *qs-name* containing precisely the distinct metalanguage variables $y_1, \ldots, y_n$, then Syntax implies the theorem

$$\forall y_1 \ldots \forall y_n \; Q(t),$$

thereby capturing that a qs-name denotes what becomes a quasi-sentence no matter which objects are taken to replace its free (metalanguage) variables. Second, Syntax produces theorems like:

$$\forall y_1 \forall y_2 \quad \overline{\boldsymbol{R}} y_1 \overline{\boldsymbol{x_1}}\,{}^{y_2}_{\boldsymbol{x_1}} \;=\; \overline{\boldsymbol{R}} y_1 y_2,$$

thus correctly evaluating the subsitution function for q-names. We will pretend that there is a single rule, SYN, producing these theorems.

It remains to specify the semantic part of the theory and the logical rules for the metalanguage. There are only four semantic axioms:

(T_1) $\forall x \forall y \boxdot_{\overline{\boldsymbol{R}} xy} \left(T(\overline{\boldsymbol{R}}\, xy) \leftrightarrow R\, xy\right)$.

34 This is only a rough indication because the universal quantification does not have to be a sentence, but can itself contain (non-linguistic) objects. We assume that the object language universal quantifier does not have object language variables in its range. One may think of ν as ranging over object language variables like $\overline{x_2}$ and x as ranging over sequences like $\overline{R}$Paul$\overline{x_2}$ such that the concatenation of $\overline{\forall}$, the object language variable, and the sequence is a quasi-sentence like $\overline{\forall x_2 R \text{Paul} x_2}$.

35 We do not believe that there are semantic requirements on purely non-linguistic objects.

(T_2) $\forall x\ \big(Q(x) \rightarrow \boxdot_{\bar{\neg}x}(T(\bar{\neg}x) \leftrightarrow \neg T(x))\big)$.
(T_3) $\forall x \forall y\ \big(Q(\bar{(}x \bar{\wedge} y\bar{)}) \rightarrow \boxdot_{\bar{(}x\bar{\wedge}y\bar{)}}(T(\bar{(}x \bar{\wedge} y\bar{)}) \leftrightarrow (T(x) \wedge T(y)))\big)$.
(T_4) $\forall x \forall v\ \big(Q(\bar{\forall} v x) \rightarrow \boxdot_{\bar{\forall} v x}(T(\bar{\forall} v x) \leftrightarrow \forall y\, T(x^y_v))\big)$.

As an example, consider (T_2): the negation of a quasi-sentence x is constitutively required to be true iff x is not true.

We will now turn to the rules for the metalanguage. As before, we will employ a version of *Replacement* for $\boxdot$, but now we have to allow for parameters in the second premiss:

$$\Rightarrow \quad \forall x_1 \forall x_2 \ldots \quad \begin{array}{l} \boxdot_\alpha \varphi \\ \boxdot_\beta(\psi \leftrightarrow \chi) \\ \boxdot_\alpha \varphi[\psi/\chi] \end{array} \qquad \alpha \notin \psi, \beta \notin \chi$$

Here $x_1, x_2, \ldots$ are precisely the (distinct) open metalanguage variables of the following formula. We refer to this rule as REP.

Finally, we have to say something about the logical rules of the metalanguage, given that it contains a hyperintensional operator. In particular, we have to specify how these rules have to be restricted in applications in which they interact with $\boxdot$.

First, we assume a rule OL for *ordinary logic*. It allows for three things. *(i)* Bound (metalanguage) variables can be renamed. *(ii)* A sentence with two initial universal quantifiers implies the corresponding 'reflexive' sentence:

$$\forall x \forall y \psi(x, y) \quad \stackrel{\text{OL}}{\Rightarrow} \quad \forall x \psi(x, x).$$

(iii) Standard inferences of first-order logic can be drawn if they are 'independent' of contained $\boxdot$-statements. We mention the case with occurrences of a single $\boxdot$-statement (the other cases are analogous):

$$\varphi(\boxdot_t \psi(y_1, \ldots, y_n)) \quad \stackrel{\text{OL}}{\Rightarrow} \quad \chi(\boxdot_t \psi(y_1, \ldots, y_n)),$$

if $\varphi(\ldots)$ implies $\chi(\ldots)$ in first-order logic.[36]

Second, we assume a rule UI for the combination of a universal instantiation by means of a q-name into a $\boxdot$-context and a universal generalisation of the result with respect to introduced free metalanguage variables. More precisely, if t is a q-name with exactly the (distinct) metalanguage variables $y_1, \ldots, y_n$, where these do not occur in ψ, then

$$\forall x \psi(x) \quad \stackrel{\text{UI}}{\Rightarrow} \quad \forall y_1 \ldots \forall y_n \psi(t).$$

Third and finally, we assume a rule ES, namely the evaluation of the substitution function for q-names in an arbitrary context. That is, suppose t is a q-name, ξ is the

[36] That is, if $\varphi(F y_1 \ldots y_n)$ implies $\chi(F y_1 \ldots y_n)$ in first-order logic for some new predicate symbol F.

canonical name of an object language variable, and y a metalanguage variable not in t. If replacing the free occurrences of ξ in t by y yields the q-name s, then

$$\psi(t^{y}_{\xi}) \quad \overset{\text{ES}}{\Rightarrow} \quad \psi(s).$$ [37]

These axioms and rules allow us to derive the theorem $\boxdot_t\ (T(t) \leftrightarrow P)$ for every s-name t and corresponding metalanguage sentence P. For reasons of space, we illustrate this for a specific example. (It is tedious but not difficult to turn this into an inductive proof of the general claim.) We will consider the object language sentence '$\overline{\forall x_2 \neg \forall x_3\ (Rx_3x_2 \wedge Rx_3x_3)}$'. The derivation will use the following qs-names:

$$t_1 := \overline{\boldsymbol{\forall x_2 \neg \forall x_3\ (Rx_3x_2 \wedge Rx_3x_3)}}, \quad t_2 := \overline{\boldsymbol{\neg \forall x_3\ (Rx_3}}u\overline{\boldsymbol{\wedge Rx_3x_3)}},$$
$$t_3 := \overline{\boldsymbol{\forall x_3\ (Rx_3}}u\overline{\boldsymbol{\wedge Rx_3x_3)}}, \quad t_4 := \overline{\boldsymbol{(R}}zu\overline{\boldsymbol{\wedge R}}zz\overline{\boldsymbol{)}}, \quad t_5 := \overline{\boldsymbol{R}}zu, \quad t_6 := \overline{\boldsymbol{R}}zz.$$

The derivation then runs as follows:

(a_1)	$\boxdot_{t_1}\big(T(t_1) \leftrightarrow \forall u T(t_2)\big)$	[T_4, UI, SYN, OL, ES]
(b)	$\forall u \boxdot_{t_2}\big(T(t_2) \leftrightarrow \neg T(t_3)\big)$	[T_2, UI, SYN, OL]
(a_2)	$\boxdot_{t_1}\big(T(t_1) \leftrightarrow \forall u \neg T(t_3)\big)$	[REP(a_1, b)]
(c)	$\forall u \boxdot_{t_3}\big(T(t_3) \leftrightarrow \forall z T(t_4)\big)$	[T_4, UI, SYN, OL, ES]
(a_3)	$\boxdot_{t_1}\big(T(t_1) \leftrightarrow \forall u \neg \forall z T(t_4)\big)$	[REP(a_2, c)]
(d)	$\forall z \forall u \boxdot_{t_4}\big(T(t_4) \leftrightarrow (T(t_5) \wedge T(t_6))\big)$	[T_3, UI, SYN, OL]
(a_4)	$\boxdot_{t_1}\big(T(t_1) \leftrightarrow \forall u \neg \forall z (T(t_5) \wedge T t_6))\big)$	[REP(a_3, d)]
(e)	$\forall z \forall u \boxdot_{t_5}\big(T(t_5) \leftrightarrow Rzu\big)$	[T_1, OL]
(a_5)	$\boxdot_{t_1}\big(T(t_1) \leftrightarrow \forall u \neg \forall z (Rzu \wedge T(t_6))\big)$	[REP(a_4, e)]
(f)	$\forall z \boxdot_{t_6}\big(T(t_6) \leftrightarrow Rzz\big)$	[T_1, OL]
(a_6)	$\boxdot_{t_1}\big(T(t_1) \leftrightarrow \forall u \neg \forall z (Rzu \wedge Rzz)\big)$	[REP(a_5, f)]

For (e) and (f), OL has been used to change the bound variables in the semantic axiom (T_1) for $\overline{R}$. For (b), UI has been applied to the semantic axiom (T_2) for $\overline{\neg}$ with respect to the q-name t_3:

$$\forall u\big(Q(t_3) \rightarrow \boxdot_{t_2} \ldots\big).$$

In addition, SYN has been used to derive $\forall u Q(t_3)$, and OL has been applied to the two results.

For (d), UI has been applied to the semantic axiom (T_3) for $\overline{\wedge}$ with respect to the q-name $\overline{\boldsymbol{R}}zu$, OL has been used to switch the initial quantifiers (from $\forall z \forall u \forall y \ldots$ to $\forall y \forall z \forall u \ldots$), UI has then been applied with respect to the q-name $\overline{\boldsymbol{R}}vv$, and OL has

[37] This loose talk of 'free occurrences of *names* of variables' is to be understood in the obvious way. An instance of ES would be

$$T(\overline{\boldsymbol{\neg \forall x_1 R x_2 x_1}}{}^{y}_{\overline{\boldsymbol{x_2}}}) \quad \overset{\text{ES}}{\Rightarrow} \quad T(\overline{\boldsymbol{\neg \forall x_1 R}}y\overline{\boldsymbol{x_1}}).$$

Note that we assume SYN to deliver such results about how to correctly evaluate the substitution function for q-names.

again been used to change the initial quantifiers (changing $\forall v \forall z \psi(v, z)$ to $\forall z \psi(z, z)$), producing:

$$\forall z \forall u \big(Q(t_4) \rightarrow \boxdot_{t_4} \ldots \big).$$

In addition, SYN has been used to derive $\forall z \forall u \, Q(t_4)$, and OL has been applied to the two results.

For (a_1) UI has been applied to the semantic axiom (T_4) for $\overline{\forall}$ with respect to the q-name $\overline{\boldsymbol{\neg \forall x_3 \, (Rx_3x_2 \wedge Rx_3x_3)}}$, and UI has then been applied with respect to q-name $\overline{\boldsymbol{x_2}}$. The rules SYN and OL (*Modus ponens*) then yield:

$$\boxdot_{t_1} \big(T(t_1) \leftrightarrow \forall y \, T(\overline{\boldsymbol{\neg \forall x_3 \, (Rx_3x_2 \wedge Rx_3x_3)}} \tfrac{y}{\overline{\boldsymbol{x_2}}}) \big).$$

Finally, OL is used to rename 'y' to 'u', and ES is used to evaluate the substitution function. ((c) is similar to (a_1).) The other steps are all effected by the replacement rule REP.

References

Berto, F., & Jago, M. (2019). *Impossible worlds*. Oxford University Press.

Callaway, H. G. (1988). Semantic competence and truth-conditional semantics. *Erkenntnis, 28*(1), 3–27.

Davidson, D. (1963). The method of intension and extension. In P. A. Schilpp & L. Salle (Eds.), *The philosophy of Rudolf Carnap*. Ill, Open Court.

Davidson, D. (1967). Truth and meaning. *Synthese, 17,* 304–323.

Davidson, D. (1976). Reply to Foster. In G. Evans & J. McDowell (Eds.), *Truth and meaning* (pp. 33–41). Clarendon.

Davies, M. (1981). *Meaning, quantification, necessity: Themes in philosophical logic*. Routledge & Kegan Paul.

Davies, M. (1982). Sentence modifiers and semantic theories: A note on a conjecture. *Analysis, 42*(1), 7–11.

Duí, M., Jespersen, B., & Materna, P. (2010). Procedural semantics for hyperintensional logic: Foundations and applications of transparent intensional logic. Vol. v. 17. Logic, epistemology and the unity of science. Springer Science + Business Media.

Fine, K. (1994). Essence and modality: The second philosophical perspectives lecture. *Philosophical Perspectives, 8,* 1–16.

Fine, K. (2007). Semantic relationism. Blackwell/Brown lectures in philosophy. Wiley.

Fine, K. (2010). Semantic necessity. In B. Hale & A. Hoffmann (eds.) Modality: Metaphysics, logic, and epistemology. Oxford University Press.

Fine, K. (2012). Guide to ground. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding* (pp. 37–80). Cambridge University Press.

Fine, K. (2014). Recurrence: A rejoinder. *Philosophical Studies, 169*(3), 425–428.

Fine, K. (2015). Unified foundations for essence and ground. *Journal of the American Philosophical Association, 1*(02), 296–311.

Fine, K. (2017). Truthmaker semantics. In B. Hale, C. Wright, & A. Miller (Eds.), *A companion to the philosophy of language*. Wiley Blackwell, 556–577.

Foster, J. (1976). Meaning and truth theory. In G. Evans & J. McDowell (Eds.), *Truth and meaning*. Clarendon.

Haverkamp, N. (2015). Intuitionism vs. classicism: A mathematical attack on classical logic. Vol. 2. Studies in theoretical philosophy. Klostermann.

Heim, I., & Kratzer, A. (1998). Semantics in generative grammar. Vol. 13. Blackwell textbooks in linguistics. Blackwell.

Higginbotham, J. (1992). Truth and understanding. *Philosophical Studies, 65*(1), 3–16.

Hoeltje, M. (2012). *Wahrheit, Bedeutung und Form: Eine Auseinandersetzung mit dem Davidson'schen Programm*. Mentis.
Hoeltje, M. (2013). Lepore and Ludwig on 'explicit meaning theories'. *Philosophical Studies, 165*(3), 831–839.
Jago, M. (2014). *The impossible: An essay on hyperintensionality*. Oxford: University Press.
Kirk-Giannini, C. D., & Lepore, E. (2017). De Ray: On the boundaries of the Davidsonian semantic programme. *Mind, 126*(503), 697–714.
Kölbel, M. (2001). Two dogmas of Davidsonian semantics. *The Journal of Philosophy, 98*(12), 613.
Kölbel, M. (2002). *Truth without objectivity*. International library of philosophy Routledge.
Larson, R. K., & Segal, G. (1995). *Knowledge of meaning: An introduction to semantic theory*. MIT Press.
Lepore, E. (1983). What model theoretic semantics cannot do? *Synthese, 54*(2), 167–187.
Lepore, E., & Loewer, B. (1989). What Davidson should have said. In J. Brandl (Ed.), *The mind of Donald Davidson. Grazer philosophische Studien*. Rodopi.
Lepore, E., & Ludwig, K. (2005). *Donald Davidson: Meaning, truth, language, and reality*. Clarendon Press and Oxford University Press.
Lepore, E., & Ludwig, K. (2006). Ontology in the Theory of Meaning. *International Journal of Philosophical Studies, 14*(3), 325–335.
Lepore, E., & Ludwig, K. (2007). *Donald Davidson's truth-theoretic semantics*. Oxford University Press.
Lepore, E., & Ludwig, K. (2011). Truth and meaning redux. *Philosophical Studies, 154*(2), 251–277.
Lepore, E., & Ludwig, K. (2013). Truth in the theory of meaning. In E. Lepore, & K. Ludwig (Eds.) A companion to Donald Davidson. Blackwell companions to philosophy (pp. 175–190).
Lewis, D. (1975). Languages and language. In K. Gunderson (Ed.), *Minnesota studies in the philosophy of science* (pp. 3–35). University of Minnesota Press.
Ludwig, K. (2002). What is the role of a truth theory in a meaning theory? In J. K. Campbell, M. O'Rourke, & D. Shier (Eds.), *Meaning and truth* (pp. 142–163). Seven Bridges Press.
Ludwig, K. (2014). Propositions and higher-order attitude attributions. *Canadian Journal of Philosophy, 43*(5–6), 741–765.
Ludwig, K. (2017). Truth-theoretic semantics and its limits. *Argumenta, 3,* 22–38.
Nolan, D. (2013). Impossible worlds. *Philosophy Compass, 8*(4), 360–372.
Richard, M. (1992). Semantic competence and disquotational knowledge. *Philosophical Studies, 65*(1–2), 37–52.
Salmon, N. (2007). Relative and absolue apriority. In Content, cognition, and communication. Oxford University Press, pp. 169–182.
Salmon, N. (2012). Recurrence. *Philosophical Studies, 159*(3), 407–441.
Salmon, N. (2015). Recurrence again. *Philosophical Studies, 172*(2), 445–457.
Segal, G. (2008). Truth and meaning. In E. Lepore & B. C. Smith (Eds.), *The oxford handbook of philosophy of language*. Oxford University Press.
Soames, S. (1992). Truth, meaning, and understanding. *Philosophical Studies, 65*(1–2), 17–35.
Soames, S. (2008). Truth and meaning: In perspective. *Midwest Studies in Philosophy, 32*(1), 1–19.
Wallace, J. (1978). Logical form, meaning, translation. In F. Guenthner & M. Guenthner-Reutter (Eds.), *Meaning and translation*. Duckworth.
Weiss, M. (2014). A closer look at manifest consequence. *Journal of Philosophical Logic, 43*(2), 471–498.