
Journal of Logic, Language, and Information **10:** 7–28, 2001.
© 2001 *Kluwer Academic Publishers. Printed in the Netherlands.*

Formal Features of Compositionality

WILFRID HODGES
Department of Mathematical Sciences, Queen Mary & Westfield College, Mile End Road, London E1 4NS, U.K.
E-mail: w.hodges@qmw.ac.uk

(Received 2 July 1998; in final form 17 May 2000)

Abstract. We consider two formalisations of the notion of a compositional semantics for a language, and find some equivalent statements in terms of substitutions. We prove a theorem stating necessary and sufficient conditions for the existence of a "canonical" compositional semantics extending a given partial semantics, after discussing what features one would want such an extension to have. The theorem involves some assumptions about semantical categories in the spirit of Husserl and Tarski.

Key words: Compositionality, fully abstract, semantic category, semantics

1. Introduction

Compositionality is a condition on semantics for languages. A common statement of it is that the meaning of a complex expression is a function of the meanings of its immediate constituents and the grammatical rule which is used to combine them (see, for example, Partee et al., 1990: 318). The first systematic treatment of compositional semantics seems to be the paper of Alfred Tarski (1983) on the concept of truth, though the name "compositional" is more recent.

In the current literature, most "problems" for compositionality are invitations to find a compositional semantics which extends a given partial semantics, for example one defined only on sentences. Why does one ask for the semantics to be compositional? This I confess I never understood. In this paper I try to go back to first principles and ask "What should we expect to get out of a semantics extending the given partial one?" In Section 6 I discuss one rational answer that seems to go back to Frege in some form. This answer yields a precise condition on the extension, and I refer to extension semantics that meet this condition as *Fregean*. The condition involves a notion ("Husserlian") which seems not to have been discussed at all widely since Tarski, who devoted several pages to related ideas in his Concept of Truth paper (Tarski, 1983). In fact Tarski worked with a much stronger notion that makes good sense for formal languages of logic but seems wholly inappropriate for natural languages; in Section 4 I suggest a weakening of Tarski's notion.

So there is a precise question: which partial semantics have Fregean extensions? Theorem 14 answers this question. When a partial semantics has a Fregean extension to the whole language, the Fregean extension is in fact compositional; so compositionality falls out as a corollary, not as one of the original requirements on the extension.

Of course even when we know that a Fregean extension exists, there may be a major problem of linguistics to describe such an extension explicitly.

A companion paper (Hodges, 1998) considers some of the logical and historical issues that hover in the background.

2. Grammars

We want to define compositionality. The first step is to define grammars. Formal treatments of compositionality often assume that the grammar is a term algebra, or at least a subset of one. For natural languages this is a counterintuitive assumption. But in fact we can assume much less about the language and still reduce the question to one about term algebras. This section shows how.

A *grammar* is a triple (E, A, Σ) where

- E is the set of *expressions*;
- A, the set of *atomic expressions*, is a subset of E;
- Σ is the set of *syntactic rules*; each syntactic rule α is a partial map $\alpha : E^n \to E$ for some non-negative integer n called the *arity* of α.

The *term algebra* of (E, A, Σ) is the set $T = T(E, A, \Sigma)$ of *terms*, defined inductively by:

- Every variable $\xi, \eta, \zeta, \xi_0, \ldots$ is in T, and is an *atomic term*.
- Every expression is in T, and is an *atomic term*.
- If α is a syntactic rule of arity n and $t_0, \ldots, t_{n-1}$ are all in T, then the term "$\alpha(t_0, \ldots, t_{n-1})$" is in T; it is a *complex term* and its *immediate constituents* are the occurrences of the terms $t_0, \ldots, t_{n-1}$ in it.

The variables $\xi, \eta, \zeta, \xi_0 \ldots$ are metavariables of the grammar, not symbols of the language of E. The definition of *subterm* (of a term) is as usual; every term is a subterm of itself. An occurrence of a subterm in t is called a *constituent* of t. We shall take for granted the well-known fact that each term in a term algebra has a unique parsing.

The *grammatical term algebra* of (E, A, Σ) is the set $GT = GT(E, A, \Sigma)$ of *grammatical terms*, defined inductively by:

- Every atomic expression is in GT; its *value* is itself.

– If α is a syntactic rule of arity n, the terms $t_0, \ldots, t_{n-1}$ are all in GT and have values $e_0, \ldots, e_{n-1}$ respectively, and the expression $\alpha(e_0, \ldots, e_{n-1})$ is defined, then the term "$\alpha(t_0, \ldots, t_{n-1})$" is in GT and its *value* is the expression $\alpha(e_0, \ldots, e_{n-1})$.

Since every grammatical term is a term, the definitions of subterm etc. for T carry over to GT. There is a map val : $GT \to E$ taking each term in GT to its value. *We require that val is surjective.* If t is a term in GT with value e, we say that t is a *structural analysis* of e. By assumption every expression has a structural analysis; some expressions might have more than one. The atomic expressions are those that have trivial structural analyses; they might also have other non-trivial structural analyses, for example if they are idioms. (For an application to idioms, see Westerståhl, to appear.)

3. Meanings and Synonymies

In general one expects a semantics to give different meanings to an expression under different structural analyses of the expression. So I shall assume that meanings attach to elements of the grammatical term algebra GT rather than to elements of E. In fact the set E plays only a parasitic role from this point onwards. But as a shorthand I will write for example $GT(E)$ and *grammar of* E, rather than $GT(E, A, \Sigma)$ and *grammar of* (E, A, Σ). The reader can assume without loss that T is an arbitrary term algebra whose atoms include variables, and $GT(E)$ – the set of grammatical terms – is an arbitrary subset of T that contains no variables and is closed under subterms.

Thus a *semantics* for E is a map μ whose domain is a subset of $GT(E)$. We call $\mu(t)$ the *meaning* of t, or more loosely of val(t). We say that μ is *total* if its domain is the whole of $GT(E)$. If $\mu(t)$ is defined, we say that t is *μ-meaningful*.

Likewise a *synonymy* for E is an equivalence relation $\equiv$ whose field (i.e. the set of x such that $x \equiv x$) is a subset of $GT(E)$. If μ is a semantics for E, then we define a synonymy $\equiv_\mu$ for E by

$$s \equiv_\mu t \;\Leftrightarrow\; \mu(s) = \mu(t).$$

(In particular $s \equiv_\mu s$ if and only if μ is defined at s.) We write $s \not\equiv_\mu t$ to mean that $s \equiv_\mu s$, $t \equiv_\mu t$ but not $s \equiv_\mu t$. We say that two semantics μ, ν for a grammar E are *equivalent* if $\equiv_\mu$ equals $\equiv_\nu$.

Going in the other direction, if $\equiv$ is a synonymy for E then we can define a semantics $\mu^{\equiv}$ for E by taking $\mu^{\equiv}(s)$ to be the $\equiv$-equivalence class of s if s is in the field of $\equiv$, and undefined otherwise.

LEMMA 1. *With the definitions above, if $\cong$ is a synonymy for E then*

$$\equiv_{\mu^{\cong}} = \cong,$$

and if ν is any semantics for E then

$$\mu^{\equiv_\nu}, \quad \nu \text{ are equivalent.}$$

Proof. Immediate. □

Given two synonymies $\equiv$ and $\equiv'$ for E, we say that $\equiv$ is *compatible with* $\equiv'$ if for all grammatical terms p, q in the field of both synonymies,

$$p \equiv q \;\Leftrightarrow\; p \equiv' q.$$

We say that two semantics μ, ν for E are *compatible* if their associated synonymies $\equiv_\mu, \equiv_\nu$ are compatible. Then μ is equivalent to ν, in the sense defined earlier, if and only if μ and ν are compatible and have the same domain.

If $\equiv, \equiv'$ are synonymies for E, we say that $\equiv'$ is an *extension* of $\equiv$ (or that it *extends* $\equiv$) if the field of $\equiv'$ includes that of $\equiv$ and the two synonymies are compatible. If μ, ν are semantics for E, we say that ν is an *extension* of μ (or that it *extends* μ) if $\nu(p)$ equals $\mu(p)$ whenever $\mu(p)$ is defined. The next lemma relates these notions.

LEMMA 2. *Suppose μ and ν are semantics for E. Then $\equiv_\nu$ is an extension of $\equiv_\mu$ if and only if ν is equivalent to some extension of μ.*

Proof. Immediate. □

Finally if μ and ν are semantics for E, we write $\nu \geq \mu$ to mean that $\equiv_\nu \supseteq \equiv_\mu$. This implies that the domain of ν includes that of μ, but within that domain μ may make more distinctions than ν does. The semantics μ and ν are equivalent if and only if both $\mu \geq \nu$ and $\nu \geq \mu$.

For the rest of this paper I drop the phrase "for E." Terms will always be assumed to come from the term algebra of E, and grammatical terms from $GT(E)$.

4. Substitution

If s is a term, $\xi_0, \ldots, \xi_{k-1}$ are distinct variables and $t_0, \ldots, t_{k-1}$ are terms (not necessarily distinct), we write

$$s(t_0, \ldots, t_{k-1} \mid \xi_0, \ldots, \xi_{k-1})$$

for the term formed by simultaneously replacing each occurrence of variable ξ_i by t_i $(0 \leq i < k)$. This term is not necessarily in $GT(E)$; when it is, we say that it is *grammatical*. A μ-meaningful term (for any semantics μ) is necessarily grammatical.

In his discussion of the conditions under which a compositional semantics is possible, Tarski (1983: 215ff) calls attention to a notion that we shall find important too. Namely, if E is a grammar, μ a semantics and p, q are grammatical terms, then

we say that p, q *have the same* μ*-category*, in symbols $p \sim_\mu q$, if for every term s with the single variable ξ,

$s(p \mid \xi)$ is μ-meaningful if and only if $s(q \mid \xi)$ is μ-meaningful.

Tarski notes that this is an equivalence relation on the set of grammatical terms, and he refers to its equivalence classes as *semantical categories*. (The name comes from the *Bedeutungskategorien* in Husserl's "Logische Untersuchungen" of 1900, Husserl, 1993: 318.) Then Tarski proposes a principle which, he says, is very natural "from the standpoint of the ordinary usage of language." Namely:

> First Principle of the Theory of Semantical Categories. If there is a term s with the single variable ξ, such that $s(p \mid \xi)$ and $s(q \mid \xi)$ are both μ-meaningful, then p and q have the same μ-category.

(The name is Tarski's, but I have translated the principle into the terminology of this paper. Tarski has in mind the case where the field of μ is the class of formulas of a formal language.) This principle holds for Montague-type languages. It very definitely does not hold for English. I take the following two examples from a longer list in Gazdar et al. (1985: 32):

a. The beast ate the meat.
b. The beast devoured the meat.
c. The beast ate.
d. *The beast devoured.

a. It is likely that Alex will leave.
b. It is probable that Alex will leave.
c. Alex is likely to leave.
d. *Alex is probable to leave.

So Tarski's First Principle had better go out of the window at once. Fortunately we can get many of the same advantages if we adopt instead a weaker condition. Namely, let μ and ν be semantics. We say that ν and its associated synonymy $\equiv_\nu$ are μ*-Husserlian* if

for all terms p, q, if $p \equiv_\nu q$ then $p \sim_\mu q$.

We say that μ is *Husserlian* if it is μ-Husserlian. (The reason for the two-step definition will emerge in Section 6.) Thus in a Husserlian semantics for English, "likely" and "probable" have different meanings, and likewise so do "eat" and "devour."

It is immediate that a semantics μ satisfying Tarski's First Principle of Semantic Categories is Husserlian, so that our notion is a weakening of his.

LEMMA 3. *If* $p \sim_\mu q$ *and* $s(p \mid \xi)$ *is* μ*-meaningful, then* $s(p \mid \xi) \sim_\mu s(q \mid \xi)$.

Proof. Since $p \sim_\mu q$ and $s(p \mid \xi)$ is μ-meaningful, we know that $s(q \mid \xi)$ is μ-meaningful. Let t be a term with the one variable η, and suppose $t(s(p \mid \xi) \mid \eta)$

is μ-meaningful. We have to show that $t(s(q \mid \xi) \mid \eta)$ is μ-meaningful. It suffices to observe that $t(s \mid \eta)$ is a term t' with the one variable ξ, and that for any term r,

$$t'(r \mid \xi) = t(s(r \mid \xi) \mid \eta).$$

Thus $t(s(p \mid \xi))$ is $t'(p \mid \xi)$ and $t(s(q \mid \xi))$ is $t'(q \mid \xi)$. □

5. Compositionality

One speaks of semantics being *compositional* or not. There are several definitions of "compositional" in circulation, not all of them equivalent. The two theorems of this section pick out the main contenders.

THEOREM 4. *Let μ be a semantics for the grammar E, and suppose that every subterm of a μ-meaningful term is μ-meaningful. Then the following are equivalent:*

(a) *There is a function r such that for every complex μ-meaningful term $s = \alpha(t_0, \ldots, t_{m-1})$,*

$$\mu(s) = r(\alpha, \mu(t_0), \ldots, \mu(t_{m-1})).$$

(b) *There are a function b and for each syntactic rule α a function r_α such that for every μ-meaningful term s,*

$$\mu(s) = \begin{cases} b(s), & \text{if } s \text{ is atomic,} \\ r_\alpha(\mu(t_0), \ldots, \mu(t_{m-1})), & \text{if } s = \alpha(t_0, \ldots, t_{m-1}). \end{cases}$$

(c) *If s is a term and $p_0, \ldots, p_{n-1}, q_0, \ldots, q_{n-1}$ are grammatical terms such that $s(p_0, \ldots, p_{n-1} \mid \xi_0, \ldots, \xi_{n-1})$ and $s(q_0, \ldots, q_{n-1} \mid \xi_0, \ldots, \xi_{n-1})$ are both μ-meaningful and for all $m < n$,*

$$p_m \equiv_\mu q_m,$$

then

$$s(p_0, \ldots, p_{n-1} \mid \xi_0, \ldots, \xi_{n-1}) \equiv_\mu s(q_0, \ldots, q_{n-1} \mid \xi_0, \ldots, \xi_{n-1})$$

whenever both these terms are μ-meaningful.

If moreover μ is Husserlian, then (a)–(c) *are equivalent to*

(d) *If s is a term with one variable ξ, and p, q are grammatical terms such that $s(p \mid \xi)$, $s(q \mid \xi)$ are μ-meaningful, then*

$$p \equiv_\mu q \;\Rightarrow\; s(p \mid \xi) \equiv_\mu s(q \mid \xi).$$

Proof. The equivalence of (a) and (b) is immediate.

(a) $\Rightarrow$ (c): Assuming (a), we prove (c) by induction on the complexity of s. The first case is that s is a μ-meaningful atomic term e. In this case $n = 0$ and the conclusion $s \equiv_\mu s$ is trivial. The second case is where s is a variable ξ_0. Then $s(p_0 \mid \xi_0) = p_0$ and $s(q_0 \mid \xi_0) = q_0$, so again the result is trivial. Finally there is the case where s is a term $\alpha(t_0, \ldots, t_{m-1})$. In this case

$$s(p_0, \ldots \mid \xi_0, \ldots) = \alpha(t_0(p_0, \ldots, \mid \xi_0, \ldots), \ldots)$$

and similarly with $s(q_0, \ldots \mid \xi_0, \ldots)$. Since $s(p_0, \ldots \mid \xi_0, \ldots)$ is assumed to be μ-meaningful, its subterms $t_i(p_0, \ldots \mid \xi_0, \ldots)$ are also μ-meaningful, and so by induction hypothesis

$$t_i(p_0, \ldots \mid \xi_0, \ldots) \equiv_\mu t_i(q_0, \ldots \mid \xi_0, \ldots).$$

The result now follows at once by (a).

(c) $\Rightarrow$ (a): (a) follows at once from the special case of (c) where s has the form $\alpha(\xi_0, \ldots, \xi_{m-1})$.

Finally suppose μ is Husserlian. Clause (d) is a special case of (c) even without this assumption. Assuming (d), we prove (c) as follows. For each $i \leq m$, let s_i be the term

$$s(q_0, \ldots, q_{i-1}, p_i, \ldots, p_{m-1} \mid \xi_0, \ldots, \xi_{m-1}).$$

It suffices to show for each $i < m$ that if $p_i \equiv_\mu q_i$ then s_i and s_{i+1} are both μ-meaningful and $s_i \equiv_\mu s_{i+1}$, by induction on i. By assumption s_0 is μ-meaningful. Suppose it has been shown that s_i is μ-meaningful. Since $p_i \equiv_\mu q_i$ and μ is Husserlian, s_{i+1} is also μ-meaningful, and it follows that $s_i \equiv_\mu s_{i+1}$ by (d). □

I shall read the condition (d) as: "μ is *1-compositional*" (even when the domain of μ is not closed under subterms).

It is difficult to give a convincing natural language example to show why we need the Husserlian property in case (d), since it would involve producing a non-compositional semantics that obeys (d), and knock-down examples of non-compositional semantics for natural languages are notoriously difficult to find. But suppose that the English word "fast" (as in "fast asleep") means the same as "deeply," and "asleep" means the same as "sleeping"; allow me also that both "fast asleep" and "deeply sleeping" are grammatical English, but neither "fast sleeping" nor "deeply asleep" is. So we have a failure of the Husserlian property. Compositionality as in (a)–(c) would allow us nevertheless to deduce that "fast asleep" means the same as "deeply sleeping," by making both replacements simultaneously. But with only (d) holding, it would be quite possible for "fast asleep" and "deeply sleeping" to have different meanings.

COROLLARY 5. *If a semantics μ for a grammar E is compositional, then so is any semantics equivalent to μ.*

Proof. Clause (c) of the theorem depends only on $\equiv_\mu$. □

Thanks to the corollary, we can say that a semantics μ and its associated synonymy $\equiv_\mu$ are *compositional* when every subterm of a μ-meaningful term is μ-meaningful and any of the three conditions (a), (b), (c) of the theorem hold.

Clause (b) of the theorem explains why compositional semantics are sometimes known as *recursively defined* semantics – though the name is hardly justified, since (b) would still form a definition of μ by recursion on the syntax even if $t_0, \ldots, t_{m-1}$ were also arguments of r_α. The same clause also shows that in a compositional semantics one can give separate meanings not only for the expressions but also for the syntactic rules, by taking r_α as the meaning of the rule α.

Our notion of equivalence is already very weak. Zadrozny (1994) claims (after his proposition 3) that he has shown that "any semantics is equivalent to a compositional semantics." What he shows is that for every semantics μ there is a compositional semantics ν such that $\nu(t)(t) = \mu(t)$ for each term t. (This part of Zadrozny's (1994) result in is trivial – put $\nu(t) = \mu$ for each term t.) Clearly Zadrozny's notion of equivalence is even weaker than ours. (See also after Lemma 13 below.)

We turn to the second common notion of compositionality. Clause (c), which limps a little, is meant to cover formulations such as "The meaning of an expression e is determined by the meanings of its proper constituents and the way these constituents are combined syntactically in e."

THEOREM 6. *Let μ be a semantics for the grammar E, and suppose that every atomic subterm of a μ-meaningful term is μ-meaningful. Then the following are equivalent:*

(a) *There is a function r such that for every term s containing no atomic terms except the variables $\xi_0, \ldots, \xi_{m-1}$, and all atomic terms $t_0, \ldots, t_{m-1}$, if $s(t_0, \ldots, t_{m-1} \mid \xi_0, \ldots, \xi_{m-1})$ is μ-meaningful then*

$$\mu(s(t_0, \ldots, t_{m-1} \mid \xi_0, \ldots, \xi_{m-1})) = r(s, \mu(t_0), \ldots, \mu(t_{m-1})).$$

(b) *If s is a term and $p_0, \ldots, p_{n-1}, q_0, \ldots, q_{n-1}$ are atomic grammatical terms such that $s(p_0, \ldots, p_{n-1} \mid \xi_0, \ldots, \xi_{n-1})$ and $s(q_0, \ldots, q_{n-1} \mid \xi_0, \ldots, \xi_{n-1})$ are both μ-meaningful and for all $m < n$,*

$$p_m \equiv_\mu q_m,$$

then

$$s(p_0, \ldots, p_{n-1} \mid \xi_0, \ldots, \xi_{n-1}) \equiv_\mu s(q_0, \ldots, q_{n-1} \mid \xi_0, \ldots, \xi_{n-1})$$

whenever both these terms are μ-meaningful.

If every subterm of a μ-meaningful term is μ-meaningful, then (a) *and* (b) *are each equivalent to*

(c) *There is a function r with the following property. For each μ-meaningful complex term $u = s(t_0, \ldots, t_{m-1} \mid \xi_0, \ldots, \xi_{m-1})$, where s contains no atomic terms except the variables $\xi_0, \ldots, \xi_{m-1}$, and $t_0, \ldots, t_{m-1}$ are atomic, write $s^<(t_0, \ldots, t_{m-1})$ for the function which maps each proper subterm s' of s to $\mu(s'(t_0, \ldots, t_{m-1} \mid \xi_0, \ldots, \xi_{m-1}))$. Then*

$$\mu(u) = r(s, s^<(t_0, \ldots, t_{m-1})).$$

If μ is Husserlian, then (a) *and* (b) *are each equivalent to*

(d) *If s is a term with one variable ξ, and p, q are atomic grammatical terms and $s(p \mid \xi)$ is μ-meaningful, then*

$$p \equiv_\mu q \;\Rightarrow\; s(p \mid \xi) \equiv_\mu s(q \mid \xi).$$

Proof. To deduce (b) from (a), extend both the lists $p_0, \ldots, p_{n-1}$ and $q_0, \ldots, q_{n-1}$ so as to include all the atomic subterms of the term s in (b). Conversely (a) follows from the special case of (b) where all the atomic subterms of the term s are variables.

We prove the equivalence (a) $\Leftrightarrow$ (c) by showing that a function $r = r_a$ as in (a) gives us all the information needed to compute the function $r = r_c$ as in (c), and vice versa. Given r_a and $s, s^<(t_0, \ldots, t_{m-1})$ as in (c), we can read off $\mu(t_0), \ldots, \mu(t_{m-1})$ from $s^<(t_0, \ldots, t_{m-1})$ and then apply r_a. In the other direction, given r_c and $s, \mu(t_0), \ldots, \mu(t_{m-1})$ as in (a), we apply (c) inductively to find $\mu(s'(t_0, \ldots, t_{m-1} \mid \xi_0, \ldots, \xi_{m-1}))$ for each subterm s' of s, and thus we find $s^<(t_0, \ldots, t_{m-1})$ so that we can apply r_c.

The equivalence of (a) and (d) goes like the equivalence of (a) and (d) in the previous theorem. □

We say that a semantics μ and its associated synonymy $\equiv_\mu$ for a grammar E are *weakly compositional* when every atomic subterm of a μ-meaningful term is μ-meaningful and (a), (b) of the theorem hold. Clearly every compositional semantics is weakly compositional. Weak compositionality is a very weak property, and I shall say no more about it.

6. The Extension Problem

The *extension problem* runs as follows: given a semantics μ and a set Y of grammatical terms which contains the domain of μ, find a semantics ν which is compatible with μ and defined on Y, and perhaps has some other specified properties.

This problem tends to turn up in one or other of two quite different forms. The *first form* is where we have a semantics for *sentences* of a language, but not for general expressions; the problem is to extend from sentences to the whole language. Two examples that I discussed in Hodges (1998) were Bittner's (1995) problem of extending a semantics for sentences of Greenland Eskimo so as to describe the semantic role of the various suffixes on verbs in that language; and the challenge of Hintikka and Sandu (1997) to find a compositional semantics for their Independence-Friendly logic. The *second form* is where we already have a semantics for a language, and the problem is to extend it to a larger language with expressions of a new kind. A famous example of this is Frege's (1992) discussion of how to extend the notion of Bedeutung to a language containing words for "believe" and "after."

In formal terms this distinction corresponds to a distinction that model theorists make between two kinds of extension of partially ordered structures. We can borrow the model theorists' terminology. Suppose E is a grammar, and let X and Y be subsets of $GT(E)$ with $X \subseteq Y$. Then we say that X is *cofinal in* Y, or that Y is a *cofinal extension* of X, if every term in Y is a subterm of some term in X. We say that Y is an *end-extension* of X if every term in Y that is a subterm of a term in X is already in X.

Suppose μ and ν are semantics with domains X and Y respectively, and ν extends μ. Then we say that ν is a *cofinal extension* of μ if X is cofinal in Y. We say that ν is an *end-extension* of μ if Y is an end-extension of X. The analogous definitions for $\equiv_\mu$ and $\equiv_\nu$ are obvious.

The Bittner and Hintikka–Sandu extension problems are both about finding cofinal extensions, because all the relevant terms can be found inside sentences of the language. The Frege problem is about finding an end-extension. Theorem 16 will give a little information about end-extensions. But the rest of this section, and indeed most of the remainder of this paper, is about cofinal extensions.

Suppose again that X is the domain of the semantics μ, and X is a cofinal subset of Y. We want to find a semantics ν on Y extending μ. The problem is completely trivial unless we add that ν must satisfy certain conditions. Here one has to ask oneself "What is a semantics supposed to be for?" The simplest assumption, and a reasonable starting point, is the following principle:

> **Principle F**. The meaning of a term in Y is the contribution that the term makes to the meanings of terms in X containing it.

Here "F" comes from Frege, because Frege is often credited with this principle in the case when X is the set of sentences and Y is the set of all expressions.

To cash Principle F in formal terms, we can begin by translating it into synonymies. The following statement says that $p \equiv_\nu q$ if and only if p can always be substituted for q (and vice versa) in expressions in X without affecting the $\equiv_\mu$-class of these expressions:

Terms p and q in Y have the same ν-meaning if and only if for all terms t such that either $t(p \mid \xi)$ or $t(q \mid \xi)$ is in X, both $t(p \mid \xi)$ and $t(q \mid \xi)$ have the same μ-meaning.

This says something precise. It is the conjunction of the following three conditions on μ and ν:

F(a): If $p \equiv_\nu q$ and $t(p \mid \xi)$ is in the field of $\equiv_\mu$ then $t(q \mid \xi)$ is in the field of $\equiv_\mu$.

F(b): If $p \equiv_\nu q$ and $t(p \mid \xi), t(q \mid \xi)$ are both in the field of $\equiv_\mu$, then $t(p \mid \xi) \equiv_\mu t(q \mid \xi)$.

F(c): If $p \not\equiv_\nu q$ then there is a term t such that either exactly one of $t(p \mid \xi)$ and $t(q \mid \xi)$ is in the field of $\equiv_\mu$, or both are and $t(p \mid \xi) \not\equiv_\mu t(q \mid \xi)$.

When these three conditions F(a)–F(c) hold, I shall say that ν is a *Fregean cover* of μ; if moreover ν is an extension of μ, I shall say that ν is a *Fregean extension* of μ.

Condition F(a) we already know; it says that ν is μ-Husserlian. Condition F(b) is a kind of partial 1-compositionality; we can read it as "ν is *partially 1-compositional* over μ."

I shall say that ν is *fully abstract* over μ if F(c) holds. The name "fully abstract" is stolen from work of Milner (1977) and Plotkin (1977) in the 1970s, on the semantics of computer programs. In that setting a fully abstract semantics is one that distinguishes two programs if and only if they have the different outputs in some context.

All these names and labels carry over to synonymies in the obvious way.

Roughly speaking, F(a) and F(b) together say that ν does not make too few distinctions, and F(c) says that it does not make too many. In other words:

LEMMA 7. *Let E be a grammar. Suppose X and Y are subsets of $GT(E)$ with $X \subseteq Y$, and μ, ν, σ are semantics with domains X, Y, Y respectively. If ν is fully abstract over μ and σ is μ-Husserlian and partially 1-compositional over μ, then $\sigma \leq \nu$. In particular there is, up to equivalence, at most one Fregean extension of μ with domain Y.*

Proof. Immediate. □

When does a semantics μ have a Fregean extension? We can give necessary conditions at once:

LEMMA 8. *Suppose X, Y and Z are subsets of $GT(E)$ with $X \subseteq Y \subseteq Z$, and μ, σ and ν are semantics with domains X, Y and Z respectively. Suppose ν is an extension of σ. Then:*

(a) *If ν is a Fregean cover of μ then so is σ.*

(b) *If μ is a Fregean cover of itself then μ is Husserlian and 1-compositional.*

Proof. Immediate from the definitions. □

Applying the lemma with $\sigma = \mu$, μ cannot have any Fregean extension unless μ is Husserlian and 1-compositional. Theorem 14 below will show the converse: if μ is Husserlian and 1-compositional then μ has a unique Fregean extension to each set in which the domain of μ is cofinal. But before we come to that, let me prove a few lemmas that put some structure into the landscape.

Our next lemma allows us to restrict attention to cofinal extensions when we are discussing Fregean extensions. Here and below, we say that a semantics ν with domain Y is the *one-point extension* of the semantics μ with domain X if ν is an end-extension of μ, and all the terms in $Y \setminus X$ are ν-equivalent to each other but not to any term in X. If μ, X and Y are given and Y is an end-extension of X, then up to equivalence there is a unique one-point extension of μ to Y.

Note that any extension ν to Y of a semantics μ can be factored in a unique way as a cofinal extension σ of μ, followed by an end-extension ν of σ; σ is the restriction of ν to the set of subterms of terms in the domain of μ.

LEMMA 9. *Let X, Y be subsets of $GT(E)$ with $X \subseteq Y$ and let μ, ν be semantics with domains X, Y respectively, with ν an extension of μ. Write Z for the set of terms in Y that are subterms of terms in X. Then the following are equivalent:*

(a) *ν is a Fregean extension of μ.*
(b) *Writing σ for the restriction of ν to Z, σ is a Fregean extension of μ and ν is the one-point extension of σ.*

Proof. Straightforward from the definitions. □

LEMMA 10. *Let X, Y be subsets of $GT(E)$ with X a cofinal subset of Y, and suppose Y is closed under subterms. Let μ, ν be semantics with domains X, Y respectively. Then:*

(a) *If ν is μ-Husserlian then ν is Husserlian.*
(b) *If ν is a Fregean cover of μ then ν is compositional.*

Proof. (a) Let p, q be terms of Y such that $p \equiv_\nu q$, and suppose a term $s(p \mid \xi)$ is in Y. We have to show that $s(q \mid \xi)$ is also in Y. By assumption X is cofinal in Y, and so there is a term t with the variable η, such that $t(s(p \mid \xi) \mid \eta)$ is in X. There is a term t' such that for any term r, $t'(r \mid \xi) = t(s(r \mid \xi) \mid \eta)$. Then $t'(p \mid \xi)$ is in X (since it is $t(s(p \mid \xi) \mid \eta)$), and hence $t'(q \mid \xi)$ is in X (since $p \equiv_\nu q$ and ν is μ-Husserlian), and thus $t(s(q \mid \xi) \mid \eta)$ is in X (since it is $t'(q \mid \xi)$). Hence $t(s(q \mid \xi) \mid \eta)$ is in Y, so that its subterm $s(q \mid \xi)$ is in Y as required.

(b) Assume that ν is a Fregean cover of μ. Then ν is Husserlian by (a), and so by Theorem 4 we need only prove that ν is 1-compositional. Suppose for contradiction that $p \equiv_\nu q$ and $s(p \mid \xi), s(q \mid \xi)$ are both in Y but $s(p \mid \xi) \not\equiv_\nu s(q \mid \xi)$. By full abstraction there is a term t such that either just one of $t(s(p \mid \xi) \mid \eta), t(s(q \mid \xi) \mid \eta)$ is in X, or both are in X but $t(s(p \mid \xi) \mid \eta) \not\equiv_\mu t(s(q \mid \xi) \mid \eta)$. Again we can form a term t' such that for every grammatical term r, $t(s(r \mid \xi) \mid \eta) = t'(r \mid \xi)$. Since $p \equiv_\nu q$ and ν is a Fregean cover of μ, if at least one of $t'(p \mid \xi), t'(q \mid \xi)$ is in X then $t'(p \mid \xi) \equiv_\mu t'(q \mid \xi)$. This contradicts the choice of t. So $\equiv_\nu$ is compositional. □

Note one possible reading of Lemma 10(b): *Frege's context principle implies compositionality*.

7. Horn Properties

This section consists largely of bookkeeping and could have been an appendix. It shows that a number of the concepts appearing in earlier sections are well-behaved. A convenient net to catch them is Horn logic (e.g., Hodges, 1993), or if you prefer, pure PROLOG without cut.

Let $\equiv$ be a binary relation on the set $GT(E)$ of grammatical terms of a grammar E. We shall say that a property of $\equiv$ is a *hyperstrict Horn property* if it can be expressed by a set of conditions, each of which has the form

$$(p_0 \equiv q_0) \wedge \ldots \wedge (p_{k-1} \equiv q_{k-1}) \;\Rightarrow\; (p_k \equiv q_k),$$

where $k \geq 1$ and the p_i and the q_i are grammatical terms. We shall say that a property of $\equiv$ is a *strict Horn property* if it can be expressed by a set of conditions, each of which has either the form above or the form

$$p \equiv q,$$

where p, q are grammatical terms. Finally we shall say that a property of $\equiv$ is a *Horn property* if it can be expressed by a set of conditions, each of which has one of the forms above or the form

$$\text{Not } [(p_0 \equiv q_0) \wedge \ldots \wedge (p_{k-1} \equiv q_{k-1})],$$

where the p_i and the q_i are grammatical terms. Note that every hyperstrict Horn property is strict, and every strict Horn property is Horn.

Fact 11. Let π be a property.

(a) Suppose π is hyperstrict Horn. Then the empty relation on $GT(E)$ has π.
(b) Suppose π is strict Horn. Then for every binary relation R on $GT(E)$ there is a least binary relation $\equiv \supseteq R$ on $GT(E)$ which has π.

(c) Suppose π is Horn. Then for every binary relation R on $GT(E)$, if there is a binary relation $\equiv \supseteq R$ on $GT(E)$ which has π, then there is a least one. □

Thus you can always find a smallest relation $\equiv$ with a given strict Horn property. But to find a least relation with a Horn property that is not strict, you must first prove that there exists at least one relation with that property.

PROPOSITION 12. *The following properties are hyperstrict Horn:*

- *Being a synonymy.*
- *Being a synonymy where [atomic] subterms of meaningful terms are meaningful.*
- *Being a compositional synonymy.*
- *Being a Husserlian synonymy.*

The following properties are strict Horn:

- *Being a total synonymy.*
- *Being $\supseteq$ a given synonymy.*

The following properties are Horn:

- *Being compatible with a given synonymy.*
- *Being a fully abstract extension of a given synonymy.*
- *Being a μ-Husserlian synonymy (for a fixed μ).*

Proof. For a synonymy:

$$(p \equiv q) \Rightarrow (q \equiv p).$$

$$(p \equiv q) \wedge (q \equiv r) \Rightarrow (p \equiv r).$$

For the condition on subterms we add

$$(s \equiv s) \Rightarrow (p \equiv p) \quad (p \text{ [atomic] subterm of } s).$$

For "compositional" we add

$$\begin{aligned}&(p_0 \equiv q_0) \wedge \ldots \wedge (p_{n-1} \equiv q_{n-1}) \wedge \\ &\quad (s(p_0, \ldots, p_{n-1} \mid \xi_0, \ldots, \xi_{n-1}) \equiv s(p_0, \ldots, p_{n-1} \mid \xi_0, \ldots, \xi_{n-1})) \wedge \\ &\quad (s(q_0, \ldots, q_{n-1} \mid \xi_0, \ldots, \xi_{n-1}) \equiv s(q_0, \ldots, q_{n-1} \mid \xi_0, \ldots, \xi_{n-1})) \Rightarrow \\ &\quad (s(p_0, \ldots, p_{n-1} \mid \xi_0, \ldots, \xi_{n-1}) \equiv s(q_0, \ldots, q_{n-1} \mid \xi_0, \ldots, \xi_{n-1})).\end{aligned}$$

For "Husserlian,"

$$(p \equiv q) \wedge (s(p \mid \xi) \equiv s(p \mid \xi)) \Rightarrow (s(q \mid \xi) \equiv s(q \mid \xi)).$$

For "total,"

$(p \equiv p)$ (for all grammatical terms p).

For "$\supseteq R$,"

$(p \equiv q)$ (whenever $(p, q) \in R$).

For "compatible with μ,"

$(p \equiv p) \wedge (q \equiv q) \Rightarrow (p \equiv q)$ (when $p \equiv_\mu q$)

Not: $(p \equiv q)$ (when $p \not\equiv_\mu q$).

For "fully abstract extension of μ," use the clauses for $\supseteq$ and "compatible," together with

$(p \equiv p) \wedge (q \equiv q) \rightarrow (p \equiv q)$

whenever p, q are such that for all terms s, $s(p \mid \xi) \equiv_\mu s(q \mid \xi)$ whenever at least one of these terms is μ-meaningful.

For "μ-Husserlian,"

Not: $(p \equiv q)$ (whenever $p \not\sim_\mu q$).

□

Thus if μ is any semantics, then there are, up to equivalence, a unique $\leq$-least semantics $\geq \mu$ which is compositional, and a unique $\leq$-least total semantics $\geq \mu$ which is compositional.

8. Existence and Extension Theorems

LEMMA 13 (Existence Lemma). *Let E be a grammar and let X be a set of terms of $GT(E)$ which is cofinal in $GT(E)$. Let μ be a semantics whose domain is X. Then μ has a Fregean cover ν which is total (i.e., has domain $GT(E)$); ν is unique up to equivalence, and it is compositional. If p, q are terms in X with $p \equiv_\nu q$, then $p \equiv_\mu q$.*

Proof. Define a relation $\equiv_\nu$ on the set of grammatical terms as follows:

$p \equiv_\nu q$ if and only if $p \sim_\mu q$ and for every term $s(p \mid \xi)$ in X, $s(p \mid \xi) \equiv_\mu s(q \mid \xi)$.

We show first that $\equiv_\nu$ is a total synonymy. Reflexiveness is clear; the field is $GT(E)$. For symmetry we note that if $p \sim_\mu q$ and $s(p \mid \xi)$ is in X then $s(q \mid \xi)$ is in X too. For transitivity, suppose $p \equiv_\nu q$ and $q \equiv_\nu r$. Then certainly $p \sim_\mu r$ since $p \sim_\mu q \sim_\mu r$. Suppose $s(p \mid \xi)$ is in X; then by assumption $s(p \mid \xi) \equiv_\mu s(q \mid \xi)$.

But then $s(q \mid \xi) \in X$, so by assumption $s(q \mid \xi) \equiv_\mu s(r \mid \xi)$ too, giving $s(p \mid \xi) \equiv_\mu s(r \mid \xi)$ as required. Thus $\equiv_\nu$ is a total synonymy. We take ν to be a semantics with synonymy $\equiv_\nu$; then it is immediate from the definition of $\equiv_\nu$ that ν is a Fregean cover of μ.

Uniqueness is immediate from the definition of Fregean cover, and ν is compositional by Lemma 10. For the final clause, let p and q be terms in X such that $p \equiv_\nu q$. Let s be the term ξ. Then $p = s(p \mid \xi) \in X$, and so $p \equiv_\mu s(q \mid \xi) = q$ by the definition of $\equiv_\nu$. □

When $X = GT(E)$, the existence lemma gives the Synonymy Theorem of Lappin and Zadrozny (preprint), except for their use of function application (which needs non-well-founded sets).

Our main extension result follows at once:

THEOREM 14 (Extension Theorem). *Let E be a grammar and let X be a set of terms of $GT(E)$ which is cofinal in $GT(E)$. Let μ be a semantics whose domain is X. Then the following are equivalent:*

(a) *μ is Husserlian and 1-compositional.*
(b) *μ has a total Fregean extension.*

Proof. We already have (b) ⇒ (a) by Lemma 8. To prove the converse, assume (a) and let ν be the total Fregean cover of μ given by Lemma 13. We have to show that ν is (up to equivalence) an extension of μ; i.e., that if p, q are terms in X then $p \equiv_\mu q$ if and only if $p \equiv_\nu q$. Right to left we already know from the lemma. In the other direction, suppose $p \equiv_\mu q$. Then $p \sim_\mu q$ since μ is Husserlian. If $s(p \mid \xi)$ is in X then $s(p \mid \xi) \equiv_\mu s(q \mid \xi)$ since μ is 1-compositional. Hence $p \equiv_\nu q$ as required. □

COROLLARY 15. *Let E be a grammar and let X be a set of terms of $GT(E)$ which is cofinal in $GT(E)$. Let μ be a Husserlian semantics whose domain is X. Then the following are equivalent:*

(a) *μ is 1-compositional.*
(b) *There is a 1-compositional semantics extending μ.*
(c) *There is, up to equivalence, exactly one total Husserlian and compositional Fregean extension of μ.*
(d) *There is, up to equivalence, a unique $\leq$-greatest total compositional μ-Husserlian semantics which extends μ.*
(e) *There is, up to equivalence, a unique $\leq$-least total compositional fully abstract semantics which extends μ.*
(f) *There is, up to equivalence, a unique $\leq$-least total compositional semantics which extends μ.*

Proof. The implication from (b) to (a) is trivial. Clauses (c)–(f) all imply (b). The implication (a) $\Rightarrow$ (c) comes from the theorem and Lemma 10. The implications from (c) to (d) and (e) are by Lemma 7. Finally the implication from (c) to (f) is one of several similar consequences of Fact 11 and Proposition 12. □

The equivalence between (a) and (c) is the central fact here. Clauses (d) and (e) express the semantics in (c) as a maximum and as a minimum.

I included clause (f) with some polemical intent; we can read off any number of similar results by choosing other Horn properties of the semantics in (c) and using general nonsense. The point to note is that we have no guarantee whatever that the semantics in (f) is the same as that in (c)–(e) (see Examples One and Two in Section 10). If you believe that the real problem is to find a *compositional* extension of μ, then these other semantics will serve you just as well as the Fregean one. This is what comes of ignoring the Husserlian and full abstraction properties.

In the case of the Fregean extension, what is the form of the function r in clause (a) of Theorem 4, assuming we use the equivalence-class semantics $\nu = \mu^{\equiv_\nu}$? One can check that

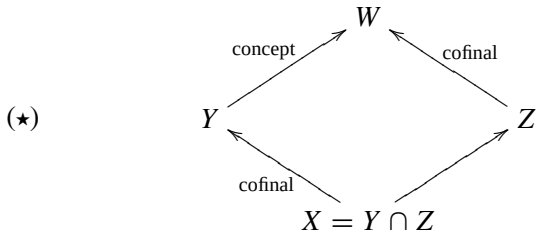
$$\begin{aligned} r(\alpha, \nu(t)) &= \{\alpha(t') : \nu(t) = \nu(t')\} \\ &= \{\alpha(t') : (\forall s)(s(t \mid \xi) \in X \vee s(t' \mid \xi) \in X \rightarrow s(t \mid \xi) \equiv_\mu s(t' \mid \xi))\}. \end{aligned}$$

9. Extending the Language

We turn to end-extensions of a semantics μ. Here the extension problem is normally trivial: the one-point extension of μ always exists, and it is always Fregean if μ was Husserlian and 1-compositional (see before Lemma 9). When Frege (1892) discussed the semantic implications of the words "believe" and "after," he was not just trying to find a semantics that incorporated them; he was trying to find one that gave the right truth-values to sentences containing these words, without abandoning the semantics that worked for expressions without these words.

Suppose W is the set of grammatical terms of a language, Y the set of grammatical terms of a sublanguage, and W "comes from Y by adding new concepts." In formal terms this says among other things that (i) $Y \subseteq W$, (ii) Y and W are both closed under subterms, and (iii) every term in $W \setminus Y$ has an atomic subterm that is not in Y. When (i)–(iii) hold I shall say that W is a *conceptual extension* of Y. A conceptual extension is necessarily an end-extension.

The following configuration describes Frege's problem. All the arrows are inclusion maps.



We have a semantics μ with domain X, a Fregean extension ν of μ with domain Y and an end-extension σ of μ with domain Z. The question is whether there is a Fregean extension τ of σ with domain W, with the property that the restriction of τ to Y is equivalent to ν. Now if there is a Fregean extension τ of σ at all, then it is unique up to equivalence, and so the question reduces to determining whether the restriction of τ to Y is Fregean.

THEOREM 16. *In the configuration* (⋆) *above, suppose τ is a semantics with domain W; write μ, ν, σ for the restrictions of τ to X, Y, Z respectively. Suppose τ is a Fregean extension of σ, and let ν' be a Fregean extension of μ to Y. Then $\nu \leq \nu'$. Moreover ν and ν' are equivalent if the following two conditions are met for all p, q in Y:*

(a) $p \sim_\mu q \;\Rightarrow\; p \sim_\sigma q$.
(b) *If $p \sim_\sigma q$ and there is a term t such that $t(p \mid \xi)$ is in Z but $t(p \mid \xi) \not\equiv_\sigma t(q \mid \xi)$, then there is such a term t with $t(p \mid \xi)$ already in X.*

Proof. Assume that p, q are in Y and $p \equiv_\nu q$. Then $p \equiv_\tau q$. Since τ is a Fregean extension of σ, it follows that if $s(p \mid \xi)$ is a term in Z then $s(q \mid \xi)$ is in Z and $s(p \mid \xi) \equiv_\sigma s(q \mid \xi)$. But if $s(p \mid \xi)$ is also in X, then q and $s(p \mid \xi)$ are in Y and $s(q \mid \xi)$ is in W, so that $s(q \mid \xi)$ is already in Y by (ii) and (iii) of "conceptual." In short, if $s(p \mid \xi)$ is in X then $s(q \mid \xi)$ is also in X and $s(p \mid \xi) \equiv_\mu s(q \mid \xi)$. This proves that $p \equiv_{\nu''} q$.

Conversely suppose that (a) and (b) hold, and that $p \equiv_{\nu'} q$. Then $p \sim_\mu q$, and for every term $s(p \mid \xi)$ in X, $s(p \mid \xi) \equiv_\mu s(q \mid \xi)$. So we have $p \sim_\sigma q$ by assumption (a). If $t(p \mid \xi)$ is a term in Z, then by (b) we deduce that $t(p \mid \xi) \equiv_\sigma t(q \mid \xi)$. So $p \equiv_\nu q$. □

We violate (a) by introducing a new meaningful term which contains p but becomes meaningless when q is put for p, though there was no such term in X. If (a) is not violated, we violate (b) by introducing a meaningful term whose meaning is changed by replacing p by q, though there was no such term in X. The second situation is exactly what Frege describes with "believe" and "after," if the meaning of a sentence is taken to be its truth-value.

Frege believed that every genuine function is defined everywhere, so that he is not likely to have been too much aware of violations of (a). But we get a rough

example by taking the phrases "adult" and "aged at least eighteen" in English without comparison of adjectives, and then adding the word "very" to the language.

10. Examples

FIRST EXAMPLE: $\leq$-LEAST COMPOSITIONAL $\neq$ FREGEAN EXTENSION

The set E of expressions consists of all the nonempty strings of either of the forms

$$\beta_0 a \beta_1 a \beta_2 \ldots, \quad a \beta_0 a \beta_1 a \ldots,$$

where each β_i is either b or c. Single letters are atoms. There is one syntactic rule; it adds an allowed letter at the lefthand end of a string. We take X to be the set of all strings in E which begin with a. The *length* $\ell(P)$ of an expression P is the number of occurrences of letters in it. We define μ on X by

$$\mu(P) = \ell(P).$$

Then μ is Husserlian and 1-compositional.

THEOREM 17.

(a) *Define ν on E by*

$$\nu(P) = \begin{cases} \ell(P), & \text{if } P \text{ begins with } a, \\ (\ell(P), 0), & \text{otherwise}. \end{cases}$$

Then ν is the Fregean extension of μ to E.

(b) *Define ρ on E by*

$$\rho(P) = \begin{cases} \ell(P), & \text{if } P \text{ begins with } a, \\ (\ell(P), b), & \text{if } P \text{ begins with } b, \\ (\ell(P), c), & \text{if } P \text{ begins with } c. \end{cases}$$

Then ρ is the $\leq$-least compositional extension of μ to E. □

SECOND EXAMPLE: FIRST-ORDER LOGIC

Suppose L is first-order logic of signature σ. We choose a grammar for L in which the grammatical terms are the formulas of L. A natural choice of semantics for the sentences of L is the map which assigns, to each sentence ϕ of L, the map $\mu(\phi)$ taking each σ-structure A to the truth-value of ϕ in A. This map $\mu(\phi)$ is a proper class, which is inconvenient. But the total number of $\equiv_\mu$-classes is at most the number of sentences of L, so the problem vanishes when we translate from μ to $\equiv_\mu$.

In fact it is immediate that $\phi \equiv_\mu \psi$ if and only if ϕ and ψ are logically equivalent sentences. Given this semantics μ for sentences, how might we extend it to a semantics ν for all formulas?

If a formula ϕ occurs as a subformula of a sentence θ, and θ' is the formula that comes from θ by replacing this occurrence by a formula ψ, then θ' is a sentence again provided that all free variables of ψ become bound in θ'. This tells us that $\phi \sim_\mu \psi$ if and only if ϕ and ψ have the same free variables. It follows at once that μ is Husserlian, since all sentences have the same number of free variables, viz. none.

It also follows that a μ-Husserlian extension of μ gives different interpretations to any two formulas with different free variables. The semantics of Tarski and Vaught (1957) fails this condition, because it gives the same interpretation to any sentence ϕ as it does to the formula $\phi \wedge (x_0 = x_0)$. This is an accidental byproduct of the fact that Tarski and Vaught interpret a formula in a structure by a set of assignments to *all* variables. We can get around it by somehow attaching the set of free variables of a formula to its "meaning." (The more appealing trick of using only satisfying assignments to the free variables of the formula does not quite work, because it fails to distinguish between $\bot$ and $\bot \wedge (x_0 = x_0)$.)

Suppose ϕ, ψ, θ are three sentences, and θ' comes from θ by replacing one or more subformulas of the form ϕ by subformulas of the form ψ. A standard metatheorem says that if ϕ and ψ are logically equivalent then so are θ and θ'. This tells us that μ is 1-compositional. Also the sentences are cofinal in the set of formulas, and we have seen that μ is Husserlian. So by Theorem 14, μ has a Fregean extension to all formulas.

Let ν be the total semantics that assigns to each formula ϕ its logical equivalence class and its set of free variables; so ν is essentially the usual semantics. I leave it to the reader to check that ν is a compositional, μ-Husserlian and fully abstract extension of μ. So ν is the Fregean extension of μ.

In this example Tarski's First Principle of Semantic Categories does hold for ν, for the trivial reason that in formulas, any formula is substitutable for any other formula. In a footnote, Tarski (1983: 217) apologises that it does not hold; but he has a less sensible grammar in mind. (And in fact his First Principle fails for μ.)

Let $\equiv_0$ be the smallest equivalence relation on the class of formulas of L such that if ϕ is any formula and θ, θ' are logically equivalent sentences, then ϕ is $\equiv_0$-equivalent to any formula got by replacing an occurrence of θ as a subformula in ϕ by an occurrence of θ'. Then $\equiv_0$ is Husserlian (for the same trivial reason as ν), μ-Husserlian and a compositional extension of $\equiv_\mu$, but it is not fully abstract over μ.

THIRD EXAMPLE: HINTIKKA'S DOUBLE-SLASH LOGIC

Hintikka (1996: 53) describes a logic L^h which (after a little inessential tweaking and some guesses about his intentions) runs as follows. For a given signature σ,

the formulas of L^h are those of the first-order language with signature σ, with two exceptions. First, the negation sign $\sim$ occurs only immediately in front of a relation symbol, or in a string of negation signs immediately in front of a relation symbol. Second, besides the usual universal quantifiers $(\forall x)$ we allow "double-slashed" quantifiers $(\forall x//y_0, \ldots, y_{k-1})$.

Each sentence ϕ describes a game on a given structure. For simplicity I assume ϕ is prenex. At an existential quantifier $(\exists x)$, player $\exists$ chooses an element of A for x to name. At a universal quantifier $(\forall x//y_0, \ldots, y_{k-1})$, player $\forall$ chooses an element for x. Then when player $\exists$ chooses the elements for any of $y_0, \ldots, y_{k-1}$ later in the play, she is not allowed to remember what element was chosen for x (unless her choice is within a smaller scope of another universal quantifier on x). When the elements have been chosen for the quantifiers, player $\exists$ wins if and only if the assignments to the variables satisfy the quantifier-free matrix of ϕ in A. Then ϕ counts as true in A if and only if player $\exists$ has a winning strategy for this game; otherwise it is false in A.

To form a grammar for L^h in the sense of this paper, take the grammatical terms to be the formulas. The game semantics defines a map μ such that if ϕ and ψ are any sentences, then $\phi \equiv_\mu \psi$ if and only if ϕ and ψ are true in exactly the same structures. The quantifier $(\forall x//y_0, \ldots, y_{k-1})$ binds only the variable x, and the occurrences of $y_0, \ldots, y_{k-1}$ in the quantifier count neither as free nor as bound. It follows as for first-order logic that μ is Husserlian and two formulas are $\sim_\mu$-equivalent if and only if they have the same free variables. Sentences are cofinal among formulas.

The following result answers an open question in Hodges (1997).

THEOREM 18. *For Hintikka's double-slash logic, let μ be Hintikka's game semantics on sentences and the classical semantics on atomic formulas. Then μ has a total fully abstract, Husserlian and compositional extension.*

Proof. We noted that μ is Husserlian. We claim that μ is also 1-compositional. For atomic formulas this is straightforward. Suppose then that θ is a sentence, the sentence ϕ is a subformula of θ, and ϕ' is μ-equivalent to ϕ. Let θ' be θ with an occurrence of ϕ replaced by ϕ'; both ϕ' and θ' are sentences. We have to show that if $\exists$ has a winning strategy for the game G of θ in a structure A, then she has one for the game G' of θ' in A too. Let her play in the game G' exactly as she would play G, until the play reaches ϕ'. She would reach ϕ at this point in G, and so by assumption she has a winning strategy for ϕ, and hence also one for ϕ'. We only have to check that her play on ϕ' is not affected by any universal quantifiers $(\forall x//y_0, \ldots, y_{k-1})$ in θ' that have ϕ' within their scope. There are two cases. First, x never occurs in ϕ'. In this case the previous value of x is irrelevant and player $\exists$ loses nothing by ignoring it. Second, x occurs in θ'. Then since ϕ' is a sentence, the value chosen for x in the game on ϕ' is chosen either by player $\exists$, or by player $\forall$ within the game on ϕ', and the previous choice for x is again irrelevant. This shows that μ is 1-compositional. The result follows by Theorem 14. □

Acknowledgements

I thank the referee for uncovering a serious confusion in an earlier draft of the paper, and Dag Westerståhl for some valuable comments.

References

Bittner, M., 1995, "Quantification in Eskimo: A challenge for compositional semantics," pp. 59–80 in *Quantification in Natural Languages*, E. Bach, E. Jelinek, A. Kratzer, and B. Partee, eds., Dordrecht: Kluwer Academic Publishers.

Frege, G., 1892, "Über Sinn und Bedeutung," *Zeitschrift für Philosophie und philosophische Kritik* **100**, 25–50. Translated as "On Sense and Reference," pp. 23–45 in *On Meaning and Reference*, A.W. Moore, ed., Oxford: Oxford University Press, 1993.

Gazdar, G., Klein, E., Pullum, G., and Sag, I., 1985, *Generalized Phrase Structure Grammar*, Cambridge, MA: Harvard University Press.

Hintikka, J., 1996, *The Principles of Mathematics Revisited*, Cambridge: Cambridge University Press.

Hintikka, J. and Sandu, G., 1997, "Game-theoretical semantics," pp. 361–410 in *Handbook of Logic and Language*, A. ter Meulen and J. van Benthem, eds., Amsterdam: Elsevier.

Hodges, W., 1993, "Logical features of Horn clauses," pp. 449–503 in *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 1: Logical Foundations*, D. Gabbay, C. Hogger, and J. Robinson, eds., Oxford: Clarendon Press.

Hodges, W., 1997, "Some strange quantifiers," pp. 51–65 in *Structures in Logic and Computer Science*, J. Mycielski, G. Rozenberg, and A. Salomaa, eds., Lecture Notes in Computer Science, Vol. 1261, Berlin: Springer-Verlag.

Hodges, W., 1998, "Compositionality is not the problem," *Logic and Logical Philosophy* **6**, 7–33.

Husserl, E., 1993, *Logische Untersuchungen II/1*, Tübingen: Max Niemeyer Verlag.

Lappin, S. and Zadrozny, A., 1999, "Compositionality, synonymy, and the systematic representation of meaning," preprint.

Milner, R., 1977, "Fully abstract models of the typed lambda-calculus," *Theoretical Computer Science* **4**, 1–22.

Partee, B.H., ter Meulen, A., and Wall, R.E., 1990, *Mathematical Methods in Linguistics*, Dordrecht: Kluwer Academic Publishers.

Plotkin, G.D., 1977, "LCF considered as a programming language," *Theoretical Computer Science* **5**, 223–255.

Tarski, A., 1983, "The concept of truth in formalized languages," pp. 152–278 in *Logic, Semantics, Metamathematics*, J. Corcoran, ed., Indianopolis: Hackett Publishing.

Tarski, A. and Vaught, R., 1957, "Arithmetical extensions of relational systems," *Compositio Mathematica* **13**, 81–102.

Westerståhl, D., "On the compositionality of idioms. An abstract approach," in *Proceedings of LLC8*, D. Barker-Plummer, D. Beaver, J. van Benthem, and P. Scotto di Luzio, eds., Stanford, CA: CSLI Publications, to appear.

Zadrozny, W., 1994, "From compositional to systematic semantics," *Linguistics and Philosophy* **17**, 329–342.