

Compositionality

Part 1: Basic ideas and definitions

Dag Westerståhl
Stockholm University

EASLLC 2012
Southwest University
Chongqing, August 27–31, 2012

Outline

Outline of the course

Aim: To present the idea of compositionality and some of the discussion around it (for natural but also formal languages).

I start with some historical background, and then present a precise framework—or rather two related but slightly different frameworks, both due to Wilfrid Hodges—in which the idea of compositionality, and related ideas, can be expressed and studied.

Then I address some claims that compositionality is **trivial** or **empty** or simply **false**, and see what can be learnt from them.

At some point I will try to connect to Pauline Jacobson's notion of **direct** compositionality and things she does in her course.

Two important topics remain:

- how sentence meaning determines the meaning of words (Frege's **Context Principle**);
- how compositionality can deal with **context dependence**, with an application to the semantics of **quotation**.

An old idea

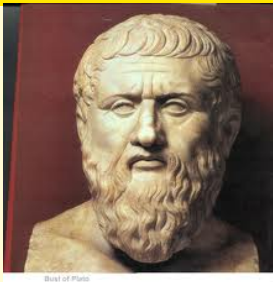
Sentences have both **structure** and **meaning**.

The meaning can be 'read off' the structure, since the **parts** of the sentence correspond to **parts** of the meaning.

That is, the structure of a sentence **reflects** its meaning.

This is an essential feature of language.

Plato



In Plato's *Sophist*, the Eleatic Stranger argues that a sentence is made up of a **noun** and a **verb**.

It is **true** if the verb attributes to what the noun signifies 'things that are':

(1) Theaetetus is sitting.
(since Theaetetus is currently sitting).

It is **false** if the verb attributes 'things that are not':

(2) Theaetetus is flying.

Thus, sentences have structure, and their parts perform different functions.

So the truth value of the sentence is a function of what the verb attributes and the noun signifies. When these are taken as meanings of sentences, verbs, and nouns, we have a compositional analysis.

Al-Fārābī



Al-Fārābī, ca. 870–950 (from Hodges):

“ . . . the imitation of the composition of meanings by the composition of expressions is by [linguistic] convention . . . ”

Linguistic expressions reflect (imitate) thoughts.

In practically all philosophy up to Frege, meanings are thoughts or ideas in the mind. And for Frege too, they are structured objects, not **in** people's minds. but **graspable** by minds.

The 'additive principle'

Middle Age philosophy is full of statements expressing some form of compositionality.

For example, in 1372, Pierre d'Ailly refers to the common view that it

belongs to the [very] notion of an expression that every expression has parts each one of which, when separated, signifies something of what is signified by the whole.

What is not clear here is **how** we put together meanings: we need to distinguish e.g. **John likes Mary** from **Mary likes John**.

In short, we need a better understanding of **structure**.

Structure is a mathematical notion, and one idea is to explicate it using concepts from **algebra**.

The characteristic universalis



Leibniz was perhaps the first to suggest an algebraic treatment of human concepts (when constructing his *universal language*).

He thought of the task as combinatorial: all concepts are combinations of a small number of simple ideas.

The simple ideas can be **coded** as **prime numbers** and concepts as **products**: with 2 coding *animal* and 3 *rational*, *man* would be coded by 6: $man = animal \times rational$ (*On Universal Synthesis and Analysis*).

This is a compositional **code**, or **translation** (not of sentences so far, but of concept words).

It is also **inversely** compositional.

6 of 34

Frege and functions

Another ingredient in a precise idea of compositionality is the notion of a **function**; the assignment of values ('meanings') to linguistic items.

Frege used a mathematical notion of function when talking about concepts and language.

Concepts for Frege are functions from objects to truth values.

And what holds sentences, and thoughts, together is function application, which Frege called **saturation**.

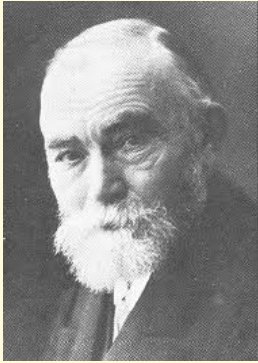
But although he was perhaps the first to express a modern idea of compositionality, he didn't use functions for the correspondence between language and meanings.

His first formulation uses **substitution**, and then one doesn't need to mention the correspondence.

His second and most explicit formulation relies on an informal idea of correspondence, or isomorphism.

7 of 34

Frege's substitution version of compositionality



Let us assume for the time being that the sentence has a reference. If we now replace one word of the sentence by another having the same reference, this can have no bearing upon the reference of the sentence. (Frege, 'On sense and reference', 1892)

This is (a special case of) the **substitution version** of compositionality. Note that the values here are *Bedeutungen* (referents), such as truth values (for sentences) and individual objects (for individual-denoting terms).

Frege is often read as having implicitly endorsed a similar principle for *Sinn* (sense) in that paper.

Frege's 'structured meanings' version of compositionality

A famous quote:

It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a terrestrial being for the very first time can be put into a form of words which will be understood by somebody to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence serves as an image of the structure of the thought. (Frege, 'Compound thoughts', 1923, opening paragraph)

Many things are notable about this passage:

Frege talks about **communication**, not only understanding.

He considers both the task for the **speaker** to find a suitable linguistic expression to convey her Thought, and the task for the **hearer** to find the appropriate Thought that corresponds to the expression uttered.

A (very) simple communicative situation

Suppose A wants to communicate a thought T to B.

She finds a sentence S that means T.

She utters S.

B hears S.

B understands S, i.e. reconstructs T.

B may never have heard S before: **compositionality** is invoked to explain how B can nevertheless understand S.

But also: A may never have uttered S before, so how did she find it?

This seems to require **inverse compositionality**: from thoughts to sentences.

Frege assumes a correspondence (isomorphism?) between thought and sentence.

He not only describes compositionality, but offers an **argument** from communication that it holds, and that it is an essential feature of language.

Carnap

Both the substitution version and the function version were explicitly stated by Rudolf Carnap (*Meaning and Necessity*, 1956) and labeled 'Frege's Principles of Interchangeability':

Let ... \mathfrak{A}_j ... be a complex name containing an occurrence of the name \mathfrak{A}_j and ... \mathfrak{A}_k ... be a corresponding expression with the name \mathfrak{A}_k instead of \mathfrak{A}_j .

*First principle. If \mathfrak{A}_j and \mathfrak{A}_k have the same nominatum, then ... \mathfrak{A}_j ... and ... \mathfrak{A}_k ... have the same nominatum. **In other words**, the nominatum of the whole expression is a function of the nominata of the names occurring in it.*

*Second principle. If \mathfrak{A}_j and \mathfrak{A}_k have the same sense, then ... \mathfrak{A}_j ... and ... \mathfrak{A}_k ... have the same sense. **In other words**, the sense of the whole expression is a function of the senses of the names occurring in it.*

The word "compositional"

It seems Hilary Putnam was the first to use this word (in its present meaning), in a talk in Oxford in 1960, later published as 'Do true assertions correspond to reality?' (*Mind, Language and Reality, Philosophical Papers, vol. 2, 1975*).

Perhaps the first appearance in print is in Jerrold Katz and Jerry Fodor's 'The structure of a semantic theory', 1964, a paper which is about adding (generative) semantics to Chomskyan syntax.

Algebraic methods in logic

Algebraic notions are implicit in the development of formal logical languages from Frege and onwards.

This aspect was in particular developed by Polish logicians, notably Tarski and Lindenbaum.

A systematic program for applying methods from algebra and topology in logic was presented by Helena Rasiowa and Roman Sikorski: *The Mathematics of Metamathematics*, 1963.



Algebraic methods for other languages

The algebraic approach was extended to

- semantics for programming languages (Dana Scott and Christopher Strachey, often called *denotational semantics*);
- semantics for **natural languages** (from the late 60's and onwards: David Lewis, Richard Montague, Barbara Partee, . . .), where compositionality was a fundamental assumption.

But aren't natural languages much too unspecific and vague to apply formal methods from logic and algebra?

Tarski thought so.

But his student Richard Montague disagreed.

Montague



There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory.

(Montague, 'Universal Grammar', 1970)

Syntactic and semantic algebras

In 'Universal grammar' Montague's approach is explicitly algebraic.

There is a (many-sorted) **syntactic algebra** which generates (disambiguated) syntactic expressions of the language: (roughly)

$$\mathbf{A} = \langle A, f_\gamma, X_\delta \rangle_{\gamma \in \Gamma, \delta \in \Delta}$$

Here the $X_\delta \subseteq A$ are of the various (disjoint) categories or **sorts**, and A is the set **generated** by the operations f_γ from the X_δ .

\mathbf{A} is **free**, so each $e \in A$ is generated in a **unique** way.

Let $\mathbf{B} = \langle B, g_\gamma \rangle_{\gamma \in \Gamma}$ a **semantic algebra** of 'meanings', **similar** to $\langle A, f_\gamma \rangle_{\gamma \in \Gamma}$.

An **interpretation** is a **homomorphism** h from $\langle A, f_\gamma \rangle_{\gamma \in \Gamma}$ to \mathbf{B} (more exactly, the unique homomorphism generated from some function h_0 from $\cup_{\delta \in \Delta} X_\delta$ to B).

Thus, for example (if f_γ is binary),

$$h(f_\gamma(e_1, e_2)) = g_\gamma(h(e_1), h(e_2))$$

So Montague **builds in** compositionality in his notion of a semantic interpretation.

Barbara Partee



The compositionality principle, in its most general form, can be expressed as follows:

The meaning of an expression is a function of the meanings of its parts and the way they are syntactically combined.

*The principle is usually attributed to Frege, and is often called 'Frege's principle' . . . In its most general form, the principle is **nearly uncontroversial**; some version of it would appear to be an **essential part of any account of how meanings are assigned to novel sentences**. (Partee, 'Compositionality', 1984)*

Jerry Fodor

So by now you may think it is fairly clear what the compositionality principle says, and that it is fundamental in semantics.

Not quite.

Several theorists think it is **false**, or **unimportant**, or **empty**. We will get back to those claims.

As to whether it is clear what the principle says . . . Here is Jerry Fodor (2001):

So not-negotiable is compositionality that I'm not even going to tell you what it is. Suffice it that the things that the expression (mutatis mutandis the concept) 'brown cow' applies to are exactly the things to which the expressions 'brown' and 'cow' apply.

(Is this a good example of compositionality?)

Fodor again (ibid.):

Nobody knows exactly what compositionality demands, but everybody knows why its demands have to be satisfied.

David Dowty

David Dowty says ('Compositionality as an empirical problem', 2007):

I believe that there is not and will not be — any time soon, if ever — a unique precise and "correct" definition of compositionality that all linguists and/or philosophers can agree upon

. . . .

I propose that we let the term NATURAL LANGUAGE COMPOSITIONALITY refer to whatever strategies and principles we discover that natural languages actually do employ to derive the meanings of sentences, on the basis of whatever aspects of syntax and whatever additional information (if any) research shows that they do in fact depend on.

This makes compositionality trivially true.

BUT:

Compositionality is not vague or mysterious

We will establish the following:

- Given a language L with a 'reasonable' syntax that identifies parts of complex expressions, and given an assignment μ of semantic values ('meanings') to expressions, the question whether μ is compositional or not is not vague.
- Indeed, although there are a few distinct notions of compositionality, each notion is precise and allows a definite answer to the question.
- Moreover, these notions are general: they don't depend on how the syntax or semantics of L is specified.

This should be ground for some optimism.

Of course, the real work lies in specifying the syntax-semantics interface, an enterprise guided by considerations which are empirical as well as theoretical.

Compositionality may be one such consideration.

If so, we should avoid mystifying or trivializing it.

Structured expressions

The first task is to define a notion of structured expression that fits most current syntactic theories.

The main thing is to have a well-defined relation x is a constituent of y .

We will look at two slightly different ways of making this precise, both due to Wilfrid Hodges.

The first is a simplification of Montague's syntactic algebra approach.

The second is even more general (and simpler), in terms of so-called constituent structures.

Both build on intuitively clear ideas of what syntactic structure is.

Syntactic and semantic categories

In both cases, we have no primitive syntactic or semantic categories.

But we will say that two expressions e and f **have the same category** iff e can be replaced by f without destroying grammaticality.

More precisely, let X be a set of expressions.

Definition

$e \sim_X f$ iff for all expressions $s[e], s[f] \in X \Leftrightarrow s[f] \in X$

This is an equivalence relation on the set of expressions, and we can think of the equivalence classes as **(X-)categories**.

X could be the set of grammatical expressions, or the set of meaningful expressions.

The notation $s = s[e]$ means that (an occurrence of) e is a constituent of s .

So the definition presupposes that we have explained 'constituent of'.

Grammars as partial algebras

Think of **grammar rules** as operating on a set E of **expressions**, e.g. strings (of signs). So they are **functions** on E ; for example,

$$\alpha(\text{Mary}, \beta(\text{like}, \text{he})) = \alpha(\text{Mary}, \text{like him}) = \text{Mary likes him}$$

These functions are **partial**: β is **defined** for (like, he) but not for e.g. $(\text{like}, \text{like})$.

The expressions and the grammar rules constitute a **partial algebra**:

$$\mathbf{E} = (E, A, \Sigma)$$

where $A \subseteq E$ is a set of **atomic expressions**, and Σ the set of partial functions. We can call such partial algebras **grammars**.

We can also use $\bar{\alpha}$ as a **function symbol** that names α , and write

$$\mathbf{E} = (E, A, \alpha)_{\bar{\alpha} \in \Sigma}$$

Then Σ is a set of function symbols, called the **signature** of \mathbf{E} .

Finally, we can think of the atoms as 0-ary functions, and write simply

$$\mathbf{E} = (E, \alpha)_{\bar{\alpha} \in \Sigma}$$

Derivations as terms

We assume \mathbf{E} is **generated**, i.e. all expressions are generated from the atoms via the grammar rules.

It can happen that the same expression (string) e can be generated in different ways, and that this matters for what it means, i.e. e is **structurally ambiguous**:

- (3) [[Old men] and women] were rescued first.
- (4) [Old [men and women]] were rescued first

This is a matter of the **order** in which the rules are applied.

$$e = \delta(\text{and}, \alpha(\text{old}, \text{men}), \text{women}) = \alpha(\text{old}, \delta(\text{and}, \text{men}, \text{women}))$$

Using the function **symbols** we can record these **derivation histories** (analysis trees) as **terms** in a corresponding partial term algebra:

$$t_1 = \bar{\delta}(\text{and}, \bar{\alpha}(\text{old}, \text{men}), \text{women}) \neq \bar{\alpha}(\text{old}, \bar{\delta}(\text{and}, \text{men}, \text{women})) = t_2$$

t_1 and t_2 are different, but have the same **(string) value**: they are spelled or pronounced the same:

$$\text{val}(t_1) = \text{val}(t_2) = e$$

Syntactic algebras

Let $GT_{\mathbf{E}}$ be the set of **grammatical terms** obtained in this way. For example,

$$\bar{\alpha}(a, \bar{\beta}(b))$$

$(a, b$ atoms) is grammatical iff α is defined for the arguments $(\text{val}(a), \text{val}(\bar{\beta}(b)))$, and then

$$\text{val}(\bar{\alpha}(a, \bar{\beta}(b))) = \alpha(\text{val}(a), \text{val}(\bar{\beta}(b))) = \alpha(\text{val}(a), \beta(\text{val}(b)))$$

We can also handle **lexical ambiguity** by introducing more atomic terms, say, \overline{bank}_1 and \overline{bank}_2 , with $\text{val}(\overline{bank}_1) = \text{val}(\overline{bank}_2) = \text{bank}$.

The symbols $\bar{\alpha}$ now also name operations on terms, so we get a **partial term algebra** $GT_{\mathbf{E}}$ (of the same signature) corresponding to \mathbf{E} .

Definition

A **syntactic algebra** is a partial term algebra obtained from a grammar \mathbf{E} , with an associated (surjective) homomorphism $\text{val}: GT_{\mathbf{E}} \rightarrow E$.

Syntactic constituents

The mapping val is a **homomorphism**, since

$$val(\bar{\alpha}(t_1, \dots, t_n)) = \alpha(val(t_1), \dots, val(t_n))$$

Among grammatical terms, the constituent relation is simply the **subterm relation**.

'Subterm' is not the same as 'substring'.

he is not a substring of **Mary likes him**,

but **he** is a subterm of $\bar{\alpha}(Mary, \bar{\beta}(like, he))$.

(Strict) **directly compositional grammar formalisms** (Jacobson 2002, 2007) essentially only allow grammar rules to perform concatenation of strings, hence not rules such as α and β above.

Then **he**, **him**, **like**, **likes** must be distinct atomic strings.

We still need terms to deal with things like

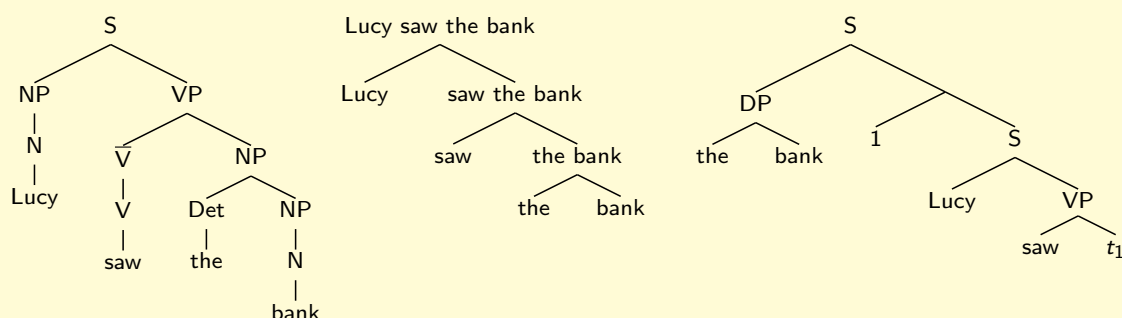
(5) Old men and women were rescued first.

Syntactic choices

The syntactic algebra format fits many semantic frameworks, such as **Categorial Grammar** (with type-shifting) and **Montague Grammar** with various extensions.

It **doesn't** fit current views of Logical Form (LF). LFs are not generated from grammar rules as above, but rather obtained from surface structure by 'movement'. Meaning is assigned to LFs, not to surface structure.

However, most tree structures, including LFs, can be represented as terms in some term algebra.



Frames

The syntactic algebra approach is **generative**: describes how complex expressions are built from simpler ones, thereby generating structure.

Now we start with another intuition: Given well-formed syntactic sentences, removing one or more parts, gives a **frame**, a structure with **holes**, that can be filled with other (appropriate) expressions. Example: From

(6) Henry knows some students.

you can get various frames, such as

(7) x knows some students

(8) x knows D students

(9) x knows Q

(10) Henry R some A

Note that this presupposes that a structural analysis is somehow given:

(11) Old men and women were rescued first.

(12) Old A were rescued first

(13) B and C were rescued first

28 of 34 Frames can be seen as **partial functions**.

Hodges' notion of a constituent structure

A **constituent structure** is a pair (\mathbb{E}, \mathbb{F}) , where \mathbb{E} is a set of expressions and \mathbb{F} a set of partial functions on \mathbb{E} , satisfying certain conditions.

(So it is very similar to grammars, but the intuition is different.)

Conditions:

(a) All functions in \mathbb{F} are non-empty (have non-empty domain).

(b) The (total) identity function 1 is in \mathbb{F} : $1(e) = e$ for all $e \in \mathbb{E}$.

(c) Nonempty Composition (NC): If $F(x_1, \dots, x_n), G(y_1, \dots, y_m) \in \mathbb{F}$, and

$$F(e_1, \dots, e_{i-1}, G(f_1, \dots, f_m), e_{i+1}, \dots, e_n) \in \mathbb{E},$$

then

$$F(x_1, \dots, x_{i-1}, G(y_1, \dots, y_m), x_{i+1}, \dots, x_n) \in \mathbb{F}$$

(d) Nonempty Substitution (NS):

$$\text{If } F(e_1, \dots, e_n) \in \mathbb{E}, \text{ then } F(x_1, \dots, x_{i-1}, e_i, x_{i+1}, \dots, x_n) \in \mathbb{F}.$$

An example

(14) If $f = F(e_1, G(e_2, e_3)) \in \mathbb{E}$, there is $H \in \mathbb{F}$ such that $f = H(e_2)$.

Proof: Since $f = F(e_1, G(e_2, e_3)) \in \mathbb{E}$, (NC) says that

$$H_1(x, y_1, y_2) = F(x, G(y_1, y_2)) \in \mathbb{F}$$

So $f = H_1(e_1, e_2, e_3) \in \mathbb{E}$, and therefore, by (NS),

$$H_2(x_1, x_2) = H_1(x_1, x_2, e_3) \in \mathbb{F}$$

So $f = H_2(e_1, e_2) \in \mathbb{E}$, and therefore, by (NS),

$$H(x_2) = H_2(e_1, x_2) \in \mathbb{F}$$

So

$$f = H(e_2)$$

□

Constituents

We now say that e is a **constituent** of f if there is a frame G such that $f = G(e)$. It is a **proper** constituent, write $e < f$, if it is distinct from f .

This is very liberal notion of constituent.

For one thing, it allows **overlapping** constituents. By contrast, if two (occurrences of) terms are subterms of another term, either they are disjoint, or one is a subterm of the other.

Also, there is no guarantee that the 'proper constituent' relation is **transitive**.

For example, nothing precludes **2-loops**: $e, f \in \mathbb{E}$ and $F, G \in \mathbb{F}$ s.t.

$$e = G(f) \text{ and } f = F(e)$$

Then $e < f$ and $f < e$. But $e \not< e$, so transitivity fails.

We can avoid this by requiring that $<$ is **wellfounded**, i.e. that there are no infinite descending chains

$$\dots e_3 < e_2 < e_1$$

A 2-loop generates such a chain: $\dots e < f < e < f$.

Exercises

Exercise 1. How should the definition of the relation $e \sim_X f$ be formulated for (a) syntactic algebras; (b) constituent structures?

Exercise 2. Suppose a sentence e , say,

Henry likes his mother.

is an expression in \mathbb{E} . One frame we may obtain from e is simply

S

(where ' S ' is a (meta)variable). Is this frame the unit frame 1 in \mathbb{F} ?

Exercise 3. Show that the relation $<$ is transitive if and only if there are no 2-loops.

Question 4. Call a frame F **rigid** if

$$F(e_1, \dots, e_n) = F(f_1, \dots, f_n) \text{ implies } e_i = f_i \text{ for } 1 \leq i \leq n,$$

i.e. if F is injective. Can you give an intuitive argument that frames should be rigid?

Semantics

Definition

A **semantics** is a partial function μ from $GT_{\mathbb{E}}$ (alternatively, from \mathbb{E}) to some set M of values ('meanings').

Thus, $dom(\mu) \subseteq GT_{\mathbb{E}}$: only grammatical terms are meaningful.

One may still want $dom(\mu)$ to be a proper subset of GT :

- Colorless green ideas . . .
- μ is a semantics for a fragment.

Might one also want some ungrammatical terms or expressions to be meaningful?

(15) John give Anna fish.

But note that (a) these are often **perceived** as ungrammatical, (b) It is because they resemble grammatical expressions that we can give them a meaning, (c) so we need the grammatical expressions anyway.

Synonymy

A semantics μ induces a sameness of meaning relation:

$$t \equiv_{\mu} u \text{ iff } \mu(t), \mu(u) \text{ are both defined and } \mu(t) = \mu(u)$$

\equiv_{μ} is a **partial equivalence relation** on $GT_{\mathbb{E}}$ (or on \mathbb{E}) — a **synonymy** — i.e. it is symmetric, transitive, and reflexive on $dom(\mu)$.

Any such synonymy \equiv in turn induces a semantics μ_{\equiv} :

$$\mu_{\equiv}(t) = [t] = \text{the equivalence class of } t \text{ (} = \emptyset \text{ if } t \notin dom(\equiv)\text{)}$$

μ_{\equiv} is the **equivalence class semantics** of \equiv .

Exercise. Show that the buck stops here, i.e., that

$$\text{if } \equiv \text{ is a synonymy, then } \equiv_{\mu_{\equiv}} = \equiv.$$