On Idioms and Compositionality

Stefan Keine
LING 720 · 12/10/2013

1 Introduction

• The overwhelming majority of theories of semantics presuppose the Principle of Compositionality in (1), standardly attributed to Gottlob Frege.

(1) Principle of Compositionality

The meaning of a complex expression is a 'function of'

(i) the meanings of its component expressions and

(ii) their mode of syntactic composition.

• Within Montague Grammar, this principle is implemented as a homomorphism between the syntax and the semantics (Montague 1970). The definition in (2) is from Westerståhl (1998: 635ff).

(2) Compositionality as a Homomorphism

Given the syntactic algebra \( A = (A, (F_\gamma)_{\gamma \in \Gamma}) \) and a meaning function \( m \) from \( A \) to a set of meanings \( M \). Let \( F \) be a \( k \)-ary operation of \( A \). \( m \) is \( F \)-compositional if there is a \( k \)-ary partial function \( G \) on \( M \) such that whenever \( F(a_1, \ldots, a_k) \) is defined,

\[
m(F(a_1, \ldots, a_k)) = G(m(a_1), \ldots, m(a_k)).
\]

With \( G \) as above we say that \( m \) is \( F \)-compositional with \( G \), and we say that \( m \) is compositional if it is \( F \)-compositional for all operations \( F \) of \( A \).

• The Problem with Idioms:

Idioms seem to be in direct contrast with the requirement in (2). The meaning of, e.g., kick the bucket seems to be unrelated to its parts and their combination.

• Central Question:

Do idioms force us to abandon (2) (and perhaps the basic idea that natural languages are compositional)?

• A Terminological Point:

Given the definition in (2), it does not make sense to ask whether certain expressions are compositional. Only entire interpretations can be. What we can ask is whether we can capture idioms in a compositional semantics.

• Outlook:

Westerståhl (2002) is a detailed study of idioms and compositionality. He argues that idioms are unproblematic for compositionality.

2 The Compositionality of Idioms: Westerståhl (2002)

• Westerståhl (2002) develops three ways in which idioms can be implemented in a compositional system. His first develops a system and then shows three ways this system has to be amended to include idioms. We will go through them and then step back to draw some general conclusions about compositionality.

2.1 The Basic System

(3) Grammar

A grammar

\[
E = (E, A, g)_{\alpha \in \Sigma}
\]

consists of a set \( E \) of expressions, a set \( A \subseteq E \) of atomic expressions and for each function symbol \( \alpha \in \Sigma \) a corresponding syntactic rule: a partial map \( g \) from \( E^n \) to \( E \), for some \( n \).

(4) Grammatical Terms

The set \( GT(E) \) of grammatical terms and the function \( val : GT(E) \to E \) are given by:

a. \( \alpha \in A \) is an atomic grammatical term, and \( val(\alpha) = \alpha \).

b. Suppose \( \alpha \in \Sigma \) in an \( n \)-ary function symbol, and \( p_1, \ldots, p_n \in GT(E) \) with \( val(p_i) = e_i \). If \( g(e_1, \ldots, e_n) \) is defined, say \( g(e_1, \ldots, e_n) = e \), the term \( \alpha(p_1, \ldots, p_n) \) is in \( GT(E) \), and \( val(\alpha(p_1, \ldots, p_n)) = e \).
• Note:
  NB: grammatical terms are analysis trees/derivations; val gives back the string associated with an analysis tree; E is the set of surface expressions

(5) **Semantics**
  A semantics for E is a function μ whose domain is a subset of GT(E). p ∈ GT(E) is μ-meaningful if p ∈ dom(μ).

(6) **Compositionality**
  μ is compositional if dom(μ) is closed under subterms and for each α ∈ Σ there is a function rα such that, whenever α(p1,...,pn) is μ-meaningful,
  \[ \mu(\alpha(p_1,\ldots,p_n)) = r_\alpha(\mu(p_1),\ldots,\mu(p_n)) \]

2.2 Treating Idioms 1: Atomic Extensions

• Intuition:
  Idioms are syntactically unanalyzable atoms.

(7) **Grammar**
  Let e0 ∈ E - A. The new grammar E^a is defined as
  \[ (E, A \cup \{e_0\}, \alpha)_{\alpha \in \Sigma} \]
  E^a is the atomic extension of E.

(8) **Example**
  a. e_0 = kick-the-bucket = val(α_0(kick, β(the, bucket)))
  b. 'John kicked the bucket.'
     (i) α(John, α_0(kick, β(the, bucket))) \sim \text{literal}
     (ii) α(John, e_0) \sim \text{idiomatic}

• Remarks:
  This treatment is clearly unsatisfactory for the majority of idioms. It leaves entirely unaccounted for the fact that idioms overwhelmingly are syntactically regular, and that parts of them may be targeted by syntactic operations (e.g., Strings were pulled to secure Henry his position; John kick+ed the bucket).

• Possible candidate:
  by and large

2.3 Treating Idioms 2: Indexed Operations

(9) **Grammar**
  We define a new grammar
  \[ E^i = (E, A, \alpha)_{\alpha \in \Sigma}, \]
  where \( \Sigma' = \Sigma \cup \{\alpha_0^i\} \), and \( \alpha_0^i \) is a new k-ary function symbol such that \( \alpha_0^i = \alpha_0 \).
  \( E^i \) is called a duplicated rule extension of E. Let \( q_0^i = \alpha_0^i(q_0,\ldots,q_k) \).

(10) **Example**
  'John kicked the bucket.'
  a. \( \alpha(\text{John}, \alpha_0(\text{kick}, \beta(\text{the, bucket}))) \sim \text{literal} \)
  b. \( \alpha(\text{John}, \alpha_0^i(\text{kick}, \beta(\text{the, bucket}))) \sim \text{idiomatic} \)

(11) \( s^- \) is the result of deleting all superscripts \( i \) in \( s \).

(12) **Semantics**
  The semantics for \( E^l, \mu^l \), is defined as follows:
  a. Its domain is \( K^l = \{p : p^- \in \text{dom}(\mu)\} \);
  b. \( \mu^l(a) = \mu(a) \) for all \( a \in A \) (whenever defined);
  c. Let \( p = \alpha(p_1,\ldots,p_n) \) be a complex term in \( GT(E^l) \). \( p^- \) is of the form \( \beta(p_{i_1},\ldots,p_{i_k}) \), where \( \beta = \alpha \) if \( \alpha \in \Sigma \), and \( \beta = \alpha_0^i \) if \( \alpha = \alpha_0^i \). If \( p^- \) is in \( \text{dom}(\mu) \) then so is each \( p_j \), so \( \mu^l(p_j) \) is defined, by induction hypothesis, and
  \[ \mu^l(p) = r_\alpha(\mu^l(p_1),\ldots,\mu^l(p_n)) \]

(13) **Definition of \( r_{\alpha_0} \)**
  \[ r_{\alpha_0}(m_1,\ldots,m_k) = \begin{cases} m_0 & \text{if } m_j = \mu(q_0), 1 \leq j \leq k \\ r_{\alpha_0}(m_1,\ldots,m_k) & \text{otherwise} \end{cases} \]
  where each \( q_0j \) is a specified meaning (e.g., \( \mu(\text{kick}), \mu(\text{the bucket}) \)).

• Consequence:
  Nothing bad happens if \( \alpha_0^i \) applies to elements that lack an idiomatic interpretation (e.g., see John). In this case \( r_{\alpha_0} = r_{\alpha} \).

• A Drawback:
  Because the definition of \( r_{\alpha_0} \) is stated in terms of the output of \( \mu \), i.e., in meaning, it follows that substitution of synonymous expressions should preserve the idiomatic interpretation. If \( q \) and \( q_01 \) are synonymous, then
  \[ \mu^l(\alpha_0^i(q_01,q_02,\ldots,q_0k)) = \mu^l(\alpha_0^i(q,q_02,\ldots,q_0k)) = m_0 \]
• If, e.g., bucket and pale are synonymous, then kick the pale should mean die. Likewise, cemetery shift or burial ground shift should mean what graveyard shift means.

2.4 Treating Idioms 3: Indexed Terminals

• Many idioms, like pull strings, may appear in various syntactic constructions:

(14) a. Strings were pulled to get John his position.
b. Mary tried to pull a lot of strings.
c. We could pull yet more strings.
d. Those strings, he wouldn't pull for you.

• Analysis:
There are special homophones of pull and string—pull\textsuperscript{i} and strings\textsuperscript{i}, respectively—that give rise to the idiomatic meaning.

• Similar accounts have been argued for by Nunberg et al. (1994) for some and by Marantz (1996, 1997) for all idioms.

(15) Grammar
We first adopt a new general definition of a grammar

\[ \left( (E, A, g)_{\alpha \in \Sigma}, \nu \right) \]

where \( E \) and \( \Sigma \) are as before but \( A \) no longer has to be a subset of \( E \). Instead there is function \( \nu \) from \( A \) to \( E \).

We now slightly revise our definition of \( GT(E) \) (only the clause for simplex expressions changes):

(16) Grammatical Terms
The set \( GT(E) \) of grammatical terms and the function \( \text{val} : GT(E) \to E \) are given by:

a. \( \alpha \in A \) is an atomic grammatical term, and \( \text{val}(\alpha) = \nu(\alpha) \).
b. Suppose \( \alpha \in \Sigma \) in an \( n \)-ary function symbol, and \( p_1, \ldots, p_n \in GT(E) \) with
\[ \text{val}(p_1) = e_1, \ldots, \text{val}(p_n) = e_n \]
say \( \alpha(e_1, \ldots, e_n) \) is defined, then \( \text{val}(\alpha(p_1, \ldots, p_n)) = e \).

NB: grammatical terms are analysis trees; \( \text{val} \) gives back the string associated with an analysis tree

(17) Idioms
We then add new, idiomatic, elements \( a_1, \ldots, a_k \) to \( A \) such that

\[ \nu(a_j) = \nu(a_j), 1 \leq j \leq k \]

That is, superscripted elements are syntactically identical to their non-indexed counterpart.

(18) The final grammar has the form

\[ E^* = ((E, A, \{ a_1, \ldots, a_k \}, g)_{\alpha \in \Sigma}, \nu^*) \]

where \( \nu^* \supseteq \nu \) and \( \nu^*(a_j) = \nu(a_j) \).

(19) Example
'Mary tried to pull several strings.'
a. \( \alpha(Mary, a_0(\delta(\text{try-to}, pull), \gamma(\text{several}, strings)))) \to \text{literal} 
b. \( \alpha(Mary, a_0(\delta(\text{try-to}, pull\textsuperscript{i})), \gamma(\text{several}, strings\textsuperscript{i}))) \to \text{idiomatic} 

(20) Semantics
The semantics for \( E^* \) is \( \mu^* \), which has the following properties:

a. Its domain is \( K^* = \{ p : p \in \text{dom}(\mu) \} \),
b. \( \mu^*(a_j) = m_j \)
c. \( \mu^*(a(q_1, \ldots, q_n)) = \mu_r^*(q_1, \ldots, \mu^*(q_n)) \)

• A Problem:
How do we constrain the distribution of indexed expressions? In particular, how we make sure that strings in Mary tried to tie several strings does not get an idiomatic interpretation.

• Westerstål’s Solution:
Make the occurrence of strings\textsuperscript{i} dependent on the occurrence of pull\textsuperscript{i} and vice versa.

(21) A term in \( GT(E^*) \) is meaningful iff it belongs to \( K^* \) and for each subterm of the form \( \alpha(p, a_0(q_1, q_2)) \) it holds that pull\textsuperscript{i} occurs in \( q_1 \) iff strings\textsuperscript{i} occurs in either \( q_2 \) or \( p \).

• What about Passives?
Passives are treated as underlyingly active and then derived by applying a passive rule \( \alpha_P \). Thus, Several strings were pulled by Mary would have the structure

\[ \alpha_P(\alpha(Mary, a_0(\delta(pull\textsuperscript{i}), \gamma(\text{several}, strings\textsuperscript{i})))) \]

and (21) is respected.
• More Serious Problem

(21) is arguably too lax as it allows pull\textsuperscript{i} to be licensed by strings\textsuperscript{i} in subject position. Moreover, (22), from Nunberg et al. (1994), shows that elements can be interpreted as strings\textsuperscript{i} even if they are not syntactically identical to strings and even in the absence of pull\textsuperscript{i} in the required position.

(22) a. Pat got the job by pulling strings that weren’t available to anyone else.
   b. Kim’s family pulled some strings on her behalf, but they weren’t enough to get her the job.
   c. I would not want you to think that we are proud of our ability to pull strings, such as the ones we pulled to get you down here.

3 What This Tells Us About Compositionality

• Question:
   Is there anything we could not do in a compositional system?

• Answer:
   No. Compositionality is formally vacuous: Any semantics can be implemented as a compositional one, as proven by Janssen (1986), Groenendijk & Stokhof (1991) and Zadrożny (1994). If we have unlimited freedom in our syntax, compositionality is also empirically vacuous.

• As a result, it is not surprising that idioms are unproblematic for compositionality. In fact, since compositionality is empirically vacuous, nothing could ever be an empirical problem.

• Strengthening Compositionality:
   In light of this, one might impose constraints on the syntactic rules (Partee 1979, Szabó 2013, and also Dowty 2007).

• Problem:
   Constraints on syntactic rules do not seem to rule out any of the accounts discussed here.

• Consequences of Homomorphism:
   Requiring every semantic operation to be mirrored by a syntactic one arguably inflates our syntax beyond what is syntactically reasonable. It also renders this framework incompatible with a wide range of declarative or representational theories of syntax, like HPSG (Pollard & Sag 1994), LFG (Bresnan 2001) or variants of MP (Brody 2002). Furthermore, if movement reduces to Merge (Chomsky 2004), it is unclear how to implement it.

• Conclusion:
   Implementing compositionality as a homomorphism between the syntax and the semantics is empirically vacuous. It is also at least not directly compatible with several types of syntactic frameworks.

Marantz, Alec (1996). ‘Cat’ as a phrasal idiom: Consequences of late insertion in Distributed Morphology, Ms., MIT, Cambridge, MA.