TRUTH AND HPSG (PART 1)

Seminar on Comparative Introduction to Lexicalist Syntactic Theories
PD Dr. Valia Kordoni
Department of Computational Linguistics and Phonetics
Saarland University

Peter Backes

November, Winter 2005

contents

introduction

fundamental questions re linguistic paradigms like HPSG:

1. predictions for natural language?

2. how to test these predictions?

notions of King:

truth: characterization of truth for grammar of a natural language.

verification: how to assess truth empirically.

both very interesting; without, no scientific merit

here: truth (first part)

introduction (cont)

in P&S94:

- theory about empirical domain:

- pheonomena of interest modelled by mathematical structures

- theory not about empirical, but mathematical structures

- predictive power depends on correspondence between math structures and empirical domain

ontological categories of linguistic objects in empirical domain?

rigorous modelling:

theory		empirical domain
linguistic theory	$\leftrightarrow$	natural language
	$\Updownarrow$	
theory of relativity	$\leftrightarrow$	space and time

precise definition of well-formedness for modelling structures

introduction (cont)

rigorous formalization must be possible (at least in principle)

in contrast to: obscure and intuition bound notions (Chomsky)

HPSG modelling domain: system of sorted feature structures

$\uparrow$

1:1

$\downarrow$

types of natural language expressions and subparts

role of linguistic theory:

precise specification of which feature structures are admissible

$\Updownarrow$

predictions of the theory

benefit of formalization: computational implementation

introduction (cont)

feature logics:

- Kasper and Rounds 1986
- Johnson 1988, 1991
- Gazdar et al. 1988

constraint languages:

- Moshier 1988
- Höhfeld and Smolka 1988
- Ait-Kaci and Nasr 1986

correspond in a direct way

feature logics for formalization of linguistic theories:

- Carpenter 1990
- Carpenter et al. 1991
- Carpenter and Pollard 1991
- King 1989
- Pollard, n.d.
- Pollard and Carpenter 1990
- Carpenter 1992

↑ very close to fully formalized fundament of King theory

truth

Carpenter 1992: sorted variant of Kasper and Rounds (1986) logic

- path inequalities
- define relations
- set values

truth straightforward

- grammar: set of formulae derived from augmented Kasper and Rounds logic
  - determines (in)admissibility
  - feature structures ⌉ 1:1 linguistic types ⌋
  - grammar predicts linguitic type

    $\Longleftrightarrow$

    grammar admits the feature structure corresponding to the ling. type

- truth of a grammar of a natural language:
  completeness: predicts all
  linguistic types in the natural language
  soundness: predicts only

linguistic types/knowledge

linguistic types:

- (P&S94: 14, 57, 58)
- not linguistic tokens, however relationship
- token = individual instance ("I'm sleepy")
- type = class of tokens
- type abstracts token, token instances type (iff token is member of type)

linguistic knowledge: mental faculty to decide admissibility of linguistic types

empiricality

- linguistic tokens: empirical (they are uttered and listened to)

- linguistic types?
  - suppose two distinct tokens in natural language L
  - two HPSG grammars of L assigning types to linguistic tokens of L
  - first grammar admits type containing both tokens, second not
  - which to prefer? empirical justification?
  - only by linguistic knowledge of ordinary user:
    can particular class of linguistic tokens be distinguished so?

    far from certain!

  - certain: empiricality of linguistic types

    |

    restricted by

    ↓

    empiricalitiy of linguistic knowledge

empiricality (cont)

- linguistic knowledge?
  - some can be surmised from linguistic tokens
  - recent technology: some mental processes directly observable.
  - yet: very limited, usually indirect

⇒characterization of truth of HPSG grammar wrt. natural language by P&S94:

- cannot be tested empirically
  - in current practice
  - in principle? depends on advances in cognitive technology
- this way: writing HPSG grammar ⟷ writing untestable theory

token/type relationship

- better: characterize truth in a different way, testable with current technology.

$\Rightarrow$ Clearly peferable.

But: How?

- obvious idea:

linguistic types
$\uparrow$
explicate relationship
$\downarrow$
linguistic tokens

- arguably already exploited by P&S94 sometimes, too.

criticism

however,

1. explicit relationship might violate *principle of ontological parsimony*:

if it does not make things simpler or easier, do not predict non-empirical constructs

$\Rightarrow$ suggests linguistic types be avoided if not for ease and simplicity

P&S: theory of an empirical domain interpreted by mathematical structure
that models the domain

$\Updownarrow$

most scientists: domain itself is mathematical structure
directly interprets the domain

$\Rightarrow$ formal machinery of P&S to characterize truth

- is superflous
- should be dispensed with

criticism (cont)

2. characterization poorly defined:
   - unclear aspects about feature structures
     - clear: well typed and sort resolved
     - concrete (individual graph) ⌉ ?
       abstract (isomorphism class of graphs) ⌋
     - footnote in P&S suggests concrete
     - unspecified if finite or infinite

     ⇒ not exact and counter-intuitive
   - formal language for admitting feature structures
     - not defined explicitely by P&S themselves
     - refer to feature logics, especially augmented Kasper&Rounds
       * partial information allowed
       * models partial information about linguistic entities

↑ does not fit ↓

HPSG94:

feature structures

=

models of linguistic entities themselves

criticism (cont)

2. characterization poorly defined:

   - formal language for admitting feature structures
     - refer to feature logics, especially augmented Kasper&Rounds
       - ∗ partial information

         |

         represented by

         |

         its most general satisfier/set of satisfiers
       - ∗ Kasper&Rounds logic agumented with classical implication:
         - · most general satisfier/set of satisfiers does not exist
         - · cannot indicate partial information

         (thus, used with non/restricted classical implication)
       - ∗ BUT: HPSG94 constraints are all classical implications!

     ⇒ Kasper&Rounds does not match P&S94:

not really suited for HPSG94

criticism (cont)

(1) and (2): could be rectified ⇒ Pollard 1999. But not:

3. accurateness: must capture intuitive concept it is intended to express

   grammarian

   - formulates grammar of natural language
   - has intuitions about linguistic types
   - constructs grammar accordingly

   but: with P&S94 truth, grammar reflects intuitions only if

   (a) commensurate with P&S characterization
   (b) formal machinery of P&S expresses them

   intuitions

   - barely mentioned by P&S94, nor Pollard 1999
   - no consensual concept about them shared by linguists

   ⇒ HPSG94:

big problems of intuitions↔machinery interaction left to grammarian

criticism (cont)

big problems:

> "an unreasonable, unpleasant and perverse imposition on the grammarian"

attempt already hindered by:

1. as described, formal machinery
   - ambiguous
   - incompletenely defined
   - contradictious

   Σ_∞ = "Adam thinks that Adam thinks that ..."

   - grammatical English sentence? No concensus.
   - Σ_∞ grammatical according to HPSG94? Unspecified.
   - could not be uttered or parsed
     but: many strings with this property usually judged grammatical
   - grammar should express knowledge of ideal language user
   - English user can grasp Σ_∞

criticism (cont)

1. formal machinery

   HPGSG94:

   - if infinite feature structures forbidden $\Rightarrow \Sigma_\infty$ ungrammatical
   - if allowed $\Rightarrow$ no HPSG grammar for
     - establishing grammaticality of Σ_n = Adam thinks (that Adam thinks)^n
     - at the same time not establishing grammaticality of Σ_∞

   however: infinite sentences bizarre (even if interesting)
   $\Rightarrow$ not really significant

criticism (cont)

2. well defined aspects of formal machinery

   - too far from intuitions regarding linguistic types
   - grammarian cannot be sure that intuitions captured correctly
   - feature structures:

     finite
     +
     totally well typed
     $\Rightarrow$
     problematic consequences

   - sort-resolvedness and well-typedness:
     – carries too much intuitive 'baggage'
     – baggage cannot be discovered in straight-forward way

   $\Rightarrow$ too much for grammarian

criticism (cont)

criticism shows: extending P&S94 truth to linguistic tokens most probably

- not useful
- not desirable
- not possible

solution: reformulate characterization of truth completely new from principles

- motivated by ontological simplicitly
- construal of natural language as a system of linguistic tokens
- formulation of alternative truth with
  - parsimonious ontology
  - clearly stated intuitions
  - well defined and demonstrably correct formal mechanisms

ontology

token:

- individual or particular

- can be object, event, state, location, complex and others

two problems:

1. token implies actual existence of token, however

   - any of the infinitely many Σ_n are legitimate
   - only finitely many can occur in English speech tokens

   $\Rightarrow$ modifications needed for notion of *linguistic* token

problems

intuition: grammar of natural language should account

- for linguistic knowledge of mature users of the natural language
- not for biological, physical, psychological limitations

problem: linguistic knowledge only accounted by linguistic behaviour

solution: notion of nonactual tokens

- could exist in some universe
- happen not to exist in ours

linguistic knowledge of mature user of natural language determines:

which actual and nonactual tokens are tokens of natural language

truth of a grammar:

delimits collection of tokens in sound and complete way.

$\Rightarrow$ avoids criticism of P&S94 truth

problems (cont)

2. different mature users of natural language:

   - do not agree about all tokens of the natural language
   - may use tokens that all paritcipants agree are not proper tokens

   therefore natural language =

   - family of collections of tokens
   - each contains improper tokens

   $\Rightarrow$ need idealisation

   Chomsky: ideal hearer-speaker:

$$\text{token is of natural language} \\ \Longleftrightarrow \\ \text{lingusitic knowledge of ideal user decides it to be.}$$

subsort formally

formal specification of subsort: $\preceq \in \texttt{Sort} \times \texttt{Sort}$

$\varsigma_1 \preceq \varsigma_2 \iff \varsigma_1$ is a subsort of ς_2

$\preceq$ defined by axiomatic system:

- $|\texttt{Sort}| \in \mathbb{N}$
- Reflexivity: $\varsigma \preceq \varsigma$
- Antisymmetry: $\varsigma_1 \preceq \varsigma_2 \wedge \varsigma_2 \preceq \varsigma_1 \Rightarrow \varsigma_1 = \varsigma_2$
- Transitivity: $\varsigma_1 \preceq \varsigma_2 \wedge \varsigma_2 \preceq \varsigma_3 \Rightarrow \varsigma_1 \preceq \varsigma_3$
- $\varsigma_1 \preceq \varsigma_2 \wedge \varsigma_1 \preceq \varsigma_3 \Rightarrow \varsigma_2 \preceq \varsigma_3 \vee \varsigma_3 \preceq \varsigma_2$
- $\mathit{object} \in \texttt{Sort}$
- $\varsigma \preceq \mathit{object}$

subsort formally (cont)

- ς_1 is a daughter of ς_2
  $:\Longleftrightarrow$
  - $\varsigma_1 \preceq \varsigma_2$
  - $\varsigma_1 \neq \varsigma 2$
  - $\varsigma_1 \preceq \varsigma \wedge \varsigma \preceq \varsigma_2 \Rightarrow \varsigma_1 = \varsigma \vee \varsigma = \varsigma_2$
- ς is maximal $:\Longleftrightarrow \quad \varsigma' \preceq \varsigma \Rightarrow \varsigma' = \varsigma$
- $\mathtt{Daug}(\varsigma) := \{\varsigma' \mid \varsigma'$ is a daughter of $\varsigma\}$

this axiomatic system

- explains algebraic properties
- not really semantic properties

sorts are symbols that denote collections of tokens

- `Univ`: universe of linguistic tokens the grammar deals with
- $[[\varsigma]]$: collection of tokens that sort ς denotes.

subsort semantics

yields the following formal semantics:

- $[\![object]\!] := \mathtt{Univ}$

- for each nonmaximal ς,
  - $\varsigma' \in \mathtt{Daug}(\varsigma) \Rightarrow [\![\varsigma]\!] \supseteq [\![\varsigma']\!]$
  - $\exists! \varsigma' \in \mathtt{Daug}(\varsigma) : [\![\varsigma]\!] \subseteq [\![\varsigma']\!]$

Features:

- $\mathtt{Feat} = \mathtt{Univ} \rightharpoonup \mathtt{Univ}$

- $\varphi \in \mathtt{Feat} \Rightarrow [\![\varphi]\!]$ the partial function which φ denotes

- φ: $\begin{bmatrix} \varphi_1 & \varsigma_1 \\ \vdots & \vdots \\ \varphi_n & \varsigma_n \end{bmatrix} \iff \forall v \in [\![\varsigma]\!] : [\![\varphi_i]\!](v)$ is defined $\wedge [\![\varphi_i]\!](v) \in [\![\varsigma_i]\!]$

subsort examples

examples from P&S94:

- [[PHONOLOGY]] is defined on all tokens in $[[\mathit{sign}]]$
- [[DAUGHTERS]] is defined on some tokens in $[[\mathit{sign}]]$ but not all
- [[LOCAL]] is defined on no tokens in $[[\mathit{sign}]]$

to have well-definedness in seconds case, if ς is maximal,

- if ς has declaration $[\varphi\ \varsigma']$
  $\Rightarrow$ $[[\varphi]]$ is defined for all $\upsilon \in [[\varsigma]]$
- else $[[\varphi]]$ is undefined on all $\upsilon \in [[\varsigma]]$

formalism

to avoid problems of Kasper&Rounds logic

- use King 1989 logic (SRL)
- unrestricted classical implication
- problem: formulae denote sets of entities
  not sets of feature structure representations of entities.
  $\Rightarrow$ might be an ontological advantage.
- designed for HPSG.
- sound, complete and decidable.

signature: triple $\Sigma = (\mathrm{S}, \mathrm{F}, \mathrm{A})$,

- S is a set of species
- F is a set of features
- $\mathrm{A} \in \mathrm{S} \times \mathrm{F} \to \mathit{Pow}(\mathrm{S})$

interpretation

interpretation of signature Σ: triple $I = (U, S, F)$

- U is the universe, a set of entities in I $(\upsilon \in I :\Leftrightarrow \upsilon \in U)$
  - if U is empty, I is called trivial
  - species denote sets of entities, partitioning U
- $S \in U \rightarrow \mathsf{S}$
  - maps each entity to the unique species denoting the set it belongs to
  - $\sigma \in \mathsf{S} \Rightarrow \sigma$ denotes $\{\upsilon \in U \mid S(\upsilon) = \sigma\}$
- $F \in \mathsf{F} \rightarrow U \rightharpoonup U$

such that $\forall \varphi \in \mathsf{F}, \upsilon \in U$:

- $F(\varphi)(\upsilon)$ is defined $\Leftrightarrow \mathsf{A}(S(\upsilon), \varphi) \neq \emptyset$
- $F(\varphi)(\upsilon)$ is defined $\Rightarrow S(F(\varphi)(\upsilon)) \in (A(S(\upsilon), \varphi)$.

terms and descriptions

Terms T_Σ and descriptions D_Σ defined by CFG:

$T_\Sigma \rightarrow$:
$T_\Sigma \rightarrow T_\Sigma \mathsf{F}$
$D_\Sigma \rightarrow T_\Sigma \approx T_\Sigma$
$D_\Sigma \rightarrow T_\Sigma \sim \mathsf{S}$
$D_\Sigma \rightarrow \neg D_\Sigma$
$D_\Sigma \rightarrow [D_\Sigma \wedge D_\Sigma]$
$D_\Sigma \rightarrow [D_\Sigma \vee D_\Sigma]$
$D_\Sigma \rightarrow [D_\Sigma \rightarrow D_\Sigma]$

- A subset of D_Σ is called a theory in Σ.
- $T_I \in T_\Sigma \rightarrow U \rightharpoonup U$
  assigns each term the partial function it denotes
- $D_I \in D_\Sigma \rightarrow Pow(U)$
  assigns each description the set of entities of which it is true
- $\Theta_I \in Pow(D_\Sigma) \rightarrow Pow(U)$
  assigns each theory the set of entities of which it is true

terms and descriptions (cont)

Let $\tau, \tau_1, \tau_2 \in \mathsf{T}_\Sigma, \varphi \in \mathsf{F}, \upsilon \in U, \sigma \in \mathsf{S}, \delta, \delta_1, \delta_2 \in \mathsf{D}_\Sigma, \theta \subseteq \mathsf{D}_\Sigma$

$$T_I(:)(\upsilon) := \upsilon$$
$$T_I(\tau\varphi)(\upsilon) := F(\varphi)(T_I(\tau)(\upsilon))$$
$$D_I(\tau_1 \approx \tau_2) := \{\upsilon \in U \mid T_I(\tau_1)(\upsilon) = T_I(\tau_2)(\upsilon)\}$$
$$D_I(\tau \sim \sigma) := \{\upsilon \in U \mid S(T_I(\tau)(\upsilon)) = \sigma\}$$
$$D_I(\neg\delta) := U \setminus D_I(\delta)$$
$$D_I([\delta_1 \wedge \delta_2]) := D_I(\delta_1) \cap D_I(\delta_2)$$
$$D_I([\delta_1 \vee \delta_2]) := D_I(\delta_1) \cup D_I(\delta_2)$$
$$D_I([\delta_1 \rightarrow \delta_2]) := D_I([\neg\delta_1 \vee \delta_2])$$
$$\Theta_I(\theta) := \{\upsilon \in U \mid \forall\delta \in \theta : \upsilon \in D_I(\delta)\}$$

- I satisfies θ in Σ $:\Leftrightarrow \exists \upsilon \in I : \upsilon \in \Theta_I(\theta)$

- I models θ in Σ $:\Leftrightarrow \forall \upsilon \in I : \upsilon \in \Theta_I(\theta)$

examples

Example: Anne, Adam, Beth, Bill, Cath, Carl

- maried:
  - Anne and Adam
  - Beth and Bill

- best friends:
  - Anne$\leftrightarrow$ Adam
  - Cath$\leftrightarrow$ Carl
  - Beth$\leftarrow$ Carl
  - Bill$\leftarrow$ Anne

formalization as SRL signature

- species: *single*, *wife*, *husband*

- features: SPOUSE, FRIEND

$\Rightarrow$ see p.325

examples (cont)

An interpretation I of Σ
$\Rightarrow$ see p.326

examples of

- terms
- descriptions
- a theory

$\Rightarrow$ see p.327

theory $\theta = : \sim \textit{wife}$:

- only trivial interpretations $(U = \emptyset)$ model θ
- proof easy (by contradiction: every $\textit{wife}$ has a SPOUSE which is a $\textit{husband}$)

application of SRL to HPSG

HPSG grammar:

1. sort hierarchy
   idea: express as a signature

   $\texttt{Maxi}(\varsigma) = \{\sigma \in \texttt{Sort} \mid \sigma \text{ is maximal and } \sigma \preceq \varsigma\}$

   Let $\Sigma = (\mathrm{S}, \mathrm{F}, \mathrm{A})$ such that

   - $\mathrm{S} = \{\sigma \in \texttt{Sort} \mid \sigma \text{ is maximal}\}$
   - $\mathrm{F} = \texttt{Feat}$
   - $\mathrm{A}(\sigma, \varphi) = \begin{cases} \texttt{Maxi}(\varsigma) & \text{if } \sigma \text{ has a declaration } [\varphi\ \varsigma] \\ \emptyset & \text{otherwise} \end{cases}$

   Let $I = (U, S, F)$ such that

   - $U = \texttt{Univ}$
   - $\forall v \in U : v \in [[S(v)]]$
   - $F(\varphi) = [[\varphi]]$

application of SRL to HPSG (cont)

observation:

- Σ is a signature
- I is an interpretation of Σ
- Signature has no notion of subsort. hierarchy?
  - hierarchy encodes: denotations of sorts and features

    sets of entities denoted by a sort

    |

    fully determined by

    ↓

    sets of entities denoted by maximal subsorts of the sort
  - can be already inferred from the signature
  - subsort relationships are superfluous

  $\Rightarrow$ not a problem
- handling of nonmaximal sorts: by metasyntactical convention.

  $\tau \sim \varsigma \leftrightsquigarrow [\tau \sim \sigma_1 \vee \cdots \vee \tau \sim \sigma_n]$

  if ς represents set of maximal sorts $\sigma_1, \ldots, \sigma_n$
- also solves other problems of HPSG like diamond inheritance

application of SRL to HPSG (cont)

2. set of principles
   idea: express as theory.

   examples:

   - p. 331
   - head feature principle
   - subcategorization principle
     - needs special functions
     - can be done in SRL with 'junk' features
     - technical, not here.

conclusion and outlook

- express HPSG grammar as:
  - signature Σ
  - theory $\theta \subseteq \mathrm{D}_\Sigma \Rightarrow (\Sigma, \theta)$ can be called SRL grammar
- using the SRL grammar, truth condition can be formulated: next time.

summary

HPSG94 problems:

- empiricality is questionable
- P&S truth is superfluous and should be dispensed with
- characterization of truth is poorly defined
- Kasper&Rounds logic is not suited for HPSG94
- accurateness not give, intuitive concepts not captured

solutions:

- formulation of alternative truth
- formalism of King 1989 logic (SRL)
- more appropriate
- sound, complete and decidable

END PART 1

Failure of prediction is no sin.
— *Noam Chomsky*
(Jubilee 2000)

references

The base article for this presentation is Paul John King: Towards Truth in Head-Driven Phrase Structure Grammar. *Tübingen Studies in Head-Driven Phrase Structure Grammar* (1999), Valia Kordoni (ed.), Sonderforschungsbereich 340, Arbeitspapiere des SFB 340, Bericht Nr. 132 (ISSN 0947-6954/99, in 2 volumes, 527pp.), Seminar für Sprachwissenschaft, Universität Tübingen.