Some remarks on trends in HPSG

Erhard Hinrichs and Detmar Meurers
Universität Tübingen

Tübingen, 3. November 2000

Background: What is an HPSG grammar?

A recent formulation (Ginzburg and Sag, in preparation, pp. 21f):

> *The grammar of a language thus consists of (minimally) a specification of:*
>
> - *The set of types that play a role in the grammar – a linguistic ontology*
> - *which features are appropriate for each type*
> - *what type of value is appropriate for each such feature, and*
> - *all the constraints that must be true of instances of particular types (These are usually referred to simply as 'type constraints'.)*

Two issues

- The role of linguistic data structures and constraints in HPSG

- The use of defaults in HPSG

Linguistic data structures and constraints

Is our ontology fine-grained enough to make the relevant distinctions?

- Example:
  - In an ontology which does not distinguish between different kinds of verbs, auxiliaries and full verbs, one will not be able to express the word order regularities of English.
  - So one needs to enrich the ontology by introducing either
    - a **new attribute** AUX (boolean-valued) of $verb$, or
    - two **new subtypes** $aux\text{-}verb$ and $full\text{-}verb$ of $verb$

- Which of the two encodings is preferable when?
  - When additional idiosyncratic properties are appropriate for the newly introduced subclass, encoding the subclass as a new type is preferable, with the additional properties as appropriate attributes.
  - Otherwise both encodings appear to be suitable: each of the two possibilities can easily be accommodated using the local characterizations captured by the appropriateness conditions.

Linguistic data structures and constraints (2)

Where in the ontology should new distinctions be introduced?

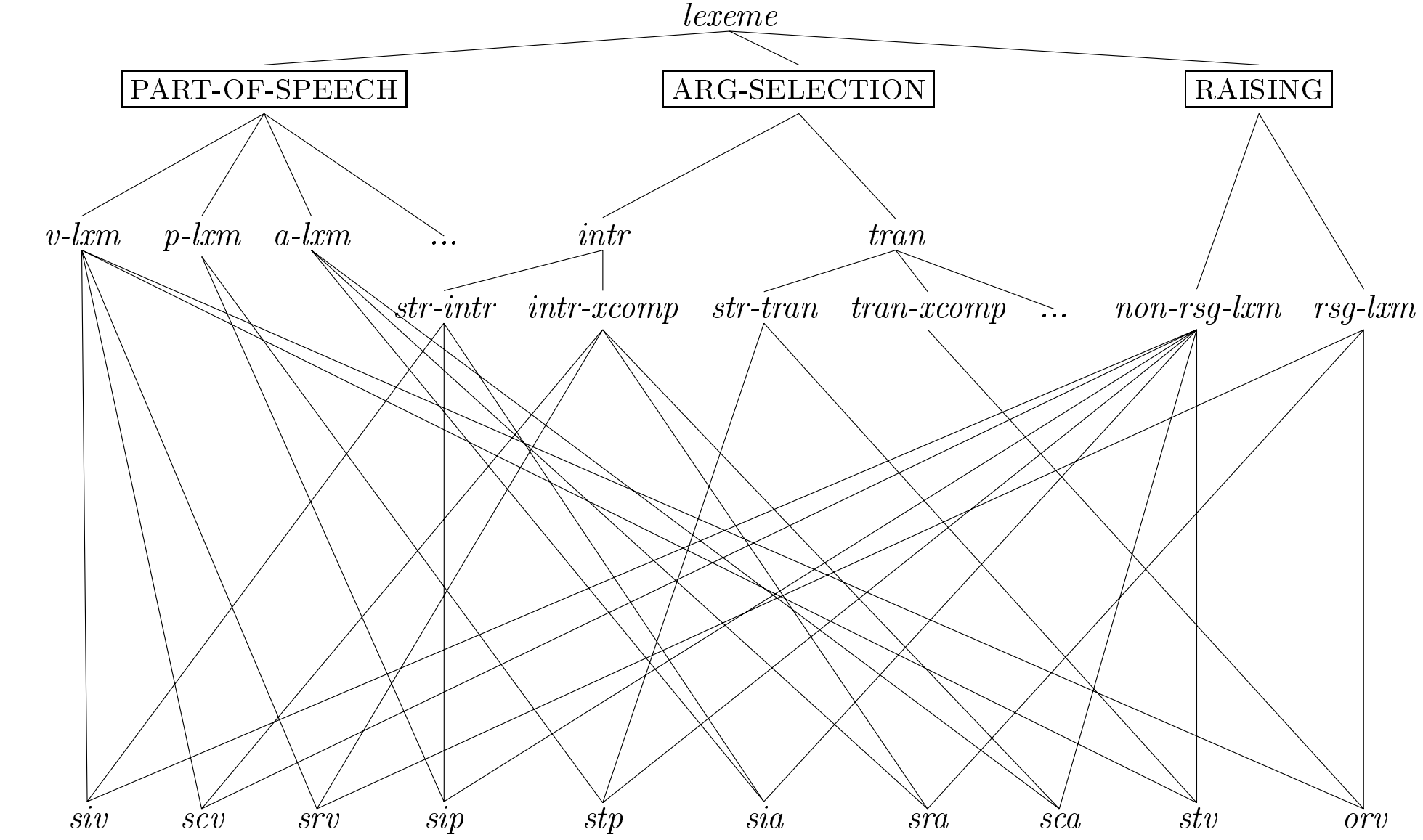
- Example: Lexeme hierarchy used as illustration in Ginzburg and Sag (1999, p. 88)

  – If types are intended to play the role of natural classes, is it sensible to reify what doesn't apply to certain entities?

    For example: the reification of "being non-raising" in the type *non-rsg-lexeme*

  – High level type distinctions coupled with "multiple inheritance along several dimensions" leads to multiplying out dimensions even if they are not appropriate.

    For example: Since the dimension RAISED is multiplied out with all parts of speech to classify prepositions as non-raising, the fact that raising is only appropriate for verbs and adjectives is lost.


lexeme
PART-OF-SPEECH
ARG-SELECTION
RAISING
v-lxm
p-lxm
a-lxm
...
intr
tran
str-intr
intr-xcomp
str-tran
tran-xcomp
...
non-rsg-lxm
rsg-lxm
siv
scv
srv
sip
stp
sia
sra
sca
stv
orv

Abbreviations

- – $v\text{-}lxm$: verb-lexeme
  - – $p\text{-}lxm$: preposition-lexeme
  - – $a\text{-}lxm$: adjective-lexeme
  - – $intr$: intransitive-lexeme
  - – $tran$: transitive-lexeme
  - – $str\text{-}int$: strict-intransitive-lexeme
  - – $intr\text{-}xcomp$: intransitive-xcomp-lexeme
  - – $tran\text{-}xcomp$: transitive-xcomp-lexeme
  - – $str\text{-}trn$: strict-transitive-lexeme
  - – $non\text{-}rsg\text{-}lxm$: non-raising-lexeme
  - – $rsg\text{-}lxm$: raising-lexeme
- – siv: strict-intransitive-verb-lexeme (e.g. die)
  - – srv: subject-raising-verb-lexeme (e.g. $seem$)
  - – scv: subject-control-verb-lexeme (e.g. try)
  - – sip: strict-intransitive-preposition-lexeme (e.g. of)
  - – stp: strict-transitive-preposition-lexeme (e.g. in)
  - – sia: strict-intransitive-adjective-lexeme (e.g. big)
  - – sra: subject-raising-adjective-lexeme (e.g. $likely$)
  - – sca: subject-control-adjective-lexeme (e.g. $eager$)
  - – stv: strict-transitive-verb-lexeme (e.g. $prove$)
  - – orv: object-raising-verb-lexeme (e.g. $believe$)

Some of the constraints

- *v-lxm* $\Rightarrow \left[\text{SYNSEM}|\text{LOC}|\text{CAT} \begin{bmatrix}\text{HEAD} & \textit{verb}\\ \text{SPR} & \langle\rangle \\ \text{SUBJ} & \left\langle [] \right\rangle\end{bmatrix}\right]$
- *p-lxm* $\Rightarrow \left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD} \ \textit{prep}\right]$
- *str-intr* $\Rightarrow \left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{ARG-ST} \left\langle \text{NP} \right\rangle\right]$
- *rsg-lxm* $\Rightarrow \left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{ARG-ST} \left\langle \left([],\right) \left[\text{LOC} \ \boxed{1}\right], \left[\text{LOC}|\text{CAT}|\text{SUBJ} \left\langle \left[\text{LOC} \ \boxed{1}\right] \right\rangle\right] \right\rangle\right]$
- *intr-xcomp* $\Rightarrow \left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{ARG-ST} \left\langle \text{NP}, \left[\text{LOC}|\text{CAT}|\text{SUBJ} \left\langle [] \right\rangle\right] \right\rangle\right]$
- *s-ctrl* $\Rightarrow \left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{ARG-ST} \left\langle \text{NP}_i, \left[\text{LOC}|\text{CAT}|\text{SUBJ} \left\langle \text{NP}_i \right\rangle\right] \right\rangle\right]$
- *str-tran* $\Rightarrow \left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{ARG-ST} \left\langle \text{NP}, \text{NP} \right\rangle\right]$

Linguistic data structures and constraints (3)

How do the new distinctions relate to the distinctions already present in the signature?

- A type constraint relates $p\text{-}lxm$ to the head value $prep$:

  $$p\text{-}lxm \Rightarrow \left[\textsc{synsem}|\textsc{local}|\textsc{category}|\textsc{head}\ \ prep\right]$$

  What is strange about such a constraint?

  - duplication of already present distinctions (redundancy)
  - The new type does not carry any appropriateness conditions in the classical, local sense. Instead, a type constraint relates the new type to an existing type as the "appropriate" value of a long feature path.
  - What rules out lexemes which are not of subtype $p\text{-}lxm$ but have $prep$ as HEAD value?

  So is this type constraint actually a linguistic generalization? What constitutes a linguistic principle?

- Is there an alternative signature which could be defined?
  - the part-of-speech dimension could be eliminated
  - transitivity and raising could be introduced as subtype of the lists which are possible values for argument-structure

Linguistic data structures and constraints (4)

What constitutes a linguistic principle?

- A principle expressing a linguistic generalization captures the covariation of two independently motivated linguistic properties.

A constraint defined by Ginzburg and Sag (1999, p. 99):

$$
sai\text{-}ph \Rightarrow \left[\begin{array}{l}
\textsc{ss|loc|cat|subj} \langle\rangle \\
\textsc{hd-dtr} \left[\begin{array}{l}
word \\
\textsc{ss|loc|cat} \left[\begin{array}{ll}
\textsc{head} & \left[\begin{array}{ll} verb & \\ \textsc{aux} & + \\ \textsc{inv} & + \end{array}\right] \\
\textsc{subj} & \left\langle \boxed{1} \right\rangle \\
\textsc{comps} & \left\langle \boxed{2}, \ldots, \boxed{n} \right\rangle
\end{array}\right]
\end{array}\right] \\
\textsc{non-hd-dtrs} \left\langle \left[\textsc{ss}\ \boxed{1}\right], \left[\textsc{ss}\ \boxed{2}\right], \ldots, \left[\textsc{ss}\ \boxed{n}\right] \right\rangle
\end{array}\right]
$$

Linguistic data structures and constraints (5)

A principle assigning accusative case in German (Meurers, 1999):

a) In an utterance,
b) each of the more oblique arguments with structural case of each verb
c) unless that argument is raised (= appears on the same subcat list as the verb)
d) is assigned accusative case.

a) $unembedded\text{-}sign \wedge$

b) $\forall\ \boxed{1}\ \forall\ \boxed{3}$

$$\boxed{1}\left[\text{L}|\text{C}\ \left[\begin{array}{l}\text{HEAD}\ \boxed{2}verb \\ \text{SUBCAT}|\text{REST}\ \ member\Big(\boxed{3}\Big[\text{L}|\text{C}|\text{HEAD}|\text{CASE}\ \ struc\Big]\Big)\end{array}\right]\right] \wedge$$

c) $\neg\exists\ \boxed{4}$

$$\boxed{4}\Big[\text{SUBCAT}\ \ member\Big(\boxed{3}\Big) \wedge member\Big(\Big[\text{L}|\text{C}|\text{HEAD}\ \ \boxed{2}\Big]\Big)\Big]$$

$\rightarrow$

d) $\boxed{3}\Big[\text{L}|\text{C}|\text{HEAD}|\text{CASE}\ \ acc\Big]$

Defaults

Thus the lexical descriptions we posit are fully consistent with a logic like SRL (King 1989) or RSRL (Richter 1999, 2000). However, the latter foundations provide no means for expressing **default regularities** of the sort that we claim **constitute linguistically significant generalizations about lexemes, words, and constructions**.
(Ginzburg and Sag, in preparation, p. 27, fn. 8)

Note that one could replace the GHFP with a set of nondefault constraints, each of which specifies the relevant identities on particular subtypes of $hd\text{-}ph$. Our use of defaults is thus in principle abbreviatory. However, our system of constraints is conceptually quite different from one cast in a pure monotonic system. By using **defeasible constraints**, we **express generalizations about construction types that are beyond the reach of (R)SRL.** We thus achieve a significant gain in descriptive simplicity which, as noted by Lascarides and Copestake (1999), is typical of systems using default constraints.
(Ginzburg and Sag, in preparation, p. 40, fn. 26)

Defaults (2)

An example: *ung*-nominalizations in German

- Generalization (Reinhard, in preparation):
  ung-nominalization applies to all transitive change-of-state verbs

  For example:
  - *abholen (pick up someone/something)* ⇒ *Abholung*,
  - *entscheiden (decide)* ⇒ *Entscheidung*,
  - *prüfen (test)* ⇒ *Prüfung*,
  - *verabreden (make an appointment)* ⇒ *Verabredung*

- Exceptions (Reinhard, in preparation):
  - *anfangen (begin)* ⇒ **Anfangung*,
  - *besuchen (visit)* ⇒ **Besuchung*,
  - *kaufen (buy)* ⇒ **Käufung*

- Explanation for exceptions:
  Blocking by competing nominal forms *Anfang (the beginning)*, *Besuch (the visit)*, *Kauf (the buy)*

Defaults (3)

Characteristics of such default generalizations

- Generalization captures the prototypical, unmarked case and it is productive.
- Exceptions are based on idiosyncratic and unpredictable facts about a language
- An explanation for when an exception can in principle arise is often available.
  E.g. lexical blocking: a fact about language motivating in what cases a default can be violated

Default (4)

An example from syntax:

- The Generalized Head Feature Principle (GHFP) (Ginzburg and Sag, in preparation):

$$hd\text{-}ph \Rightarrow \begin{bmatrix} \text{SYNSEM} & \backslash\ \boxed{1} \\ \text{HD-DTR}|\text{SYNSEM} & \backslash\ \boxed{1} \end{bmatrix}$$

- Structure of *synsem* objects:

$$\begin{bmatrix} \text{LOCAL} & \begin{bmatrix} local & \\ \text{HEAD} & part\text{-}of\text{-}speech \\ \text{CATEGORY} & \begin{bmatrix} category & \\ \text{SUBJ} & list\Big(\, synsem \,\Big) \\ \text{COMPS} & list\Big(\, synsem \,\Big) \\ \text{SPR} & list\Big(\, synsem \,\Big) \end{bmatrix} \\ \text{CONTENT} & content \\ \text{STORE} & set\Big(\, scope\text{-}obj \,\Big) \end{bmatrix} \\ \text{SLASH} & set\Big(\, local \,\Big) \\ \text{WH} & set\Big(\, scope\text{-}obj \,\Big) \\ \text{BCKGRND} & set\Big(\, fact \,\Big) \end{bmatrix}$$

Defaults (5)

Does the GHFP fit the default profile?

- Does the default cover the prototypical, unmarked case?

  ⇒ Needs to be shown.

- Are the exceptions to the default idiosyncratic and unpredictable facts about a language?

  ⇒ No. Previously, the "exceptions" were characterized by clearly delineated principles expressing generalization about different parts of the structure:

  - valence requirements and their realization
  - head feature percolation

- Is there a general explanation (like blocking) for when an exception can in principle arise?

  ⇒ Needs to be shown.

Defaults (6)

In what sense do default constraints capture generalizations?

- How many violations of the default constraint can a default constraint have and still be called a generalization?
  Problem: Degree of violatedness is not captured.

More concretely: What predictions does the GHFP make?

- Is the principle constrained enough in light of the exceptions encoded in the theory?

  $\Rightarrow$ The way the principle is formulated, one would expect combinations of overriding values throughout the entire $synsem$ structure, e.g. bundles of overriding part-of-speech, valence, background or store specifications.

  $\Rightarrow$ Instead the "exceptions" only seem to involve unrelated single or pairs of values:
  * a valence requirement is overridden since it is realized
  * a slash requirement is overridden since there is a filler
  * ...

  $\Rightarrow$ What can be violated does not seem to be captured.

Defaults (7)

What predictions does the GHFP make? (continued)

- What predictive consequences does the scope of the default have?

  For example: Would the analysis make different predictions if one modified the GHFP to share by default even more than the SYNSEM value, i.e., the entire signs?

- Allowing part of speech specifications to be overridden jeopardizes the strong notion of endocentricity embodied in the non-default classical HFP. This endocentricity assumption is a crucial ingredient of the theory of gerunds (Pullum, 1991; Malouf, 1999).

References

Ginzburg, Jonathan and Sag, Ivan A. (1999, July). English Interrogative Constructions. unpublished Ms. Hebrew University and Stanford University.

Ginzburg, Jonathan and Sag, Ivan A. (in preparation). *Interrogative Investigations. The form, meaning and use of English interrogatives*. Stanford, CA: CSLI Publications. Version of October, 2000.

Malouf, Robert (1999). Verbal gerunds as mixed categories in HPSG. In R. D. Borsley (Ed.), *The Nature and Function of Syntactic Categories*, Volume 32 of *Syntax and Semantics*, pp. 133–166. New York, et al.: Academic Press.

Meurers, Walt Detmar (1999). *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. Ph. D. thesis, Seminar für Sprachwissenschaft, Universität Tübingen, Tübingen. Available as Volume 145 in Arbeitspapiere des SFB 340, 2000.

Pullum, Geoffrey K. (1991). English nominal gerund phrases as noun phrases with verb-phrase heads. *Linguistics 29*, 763–799.

Reinhard, Sabine (in preparation). *Deverbale Komposita an der Morphologie-Syntax-Semantik-Schnitstelle: Ein HPSG-Ansatz*. Ph. D. thesis, Universität Tübingen.