MPI-EVA, May 2007

Four kinds of technical concepts in comparative grammatical research

MARTIN HASPELMATH

1. The four concept types

(1. descriptive 2. cognitive 3. comparative 4. conceptual-semantic)

1.1. Descriptive (grammatical) concepts

= concepts describing *linguistic units* in a particular language

(1) German *Stuttgart ist Meister.* 'Stuttgart is champion.'

Concepts needed to describe this sentence:

present tense, noun, verb, copula, bilabial nasal stop, diphthong, 3rd person, singular, suffix, nominative case, ...

These concepts are concepts of **descriptive linguistics**, and as such they are relevant only to a particular language.

Descriptive concepts are needed for **observational adequacy** (Chomsky 1965), i.e. for a complete account of possible and impossible sentences of a particular language.

(descriptive concepts = observational concepts = phenomenological concepts; Haspelmath 2004: a *phenomenological description* is a description that is sufficient to know how to use the language correctly but makes no claims about the knowledge of native speakers)

• Potential problem:
– Descriptive concepts are (to some extent) arbitrary and not comparable across languages

1.2. Cognitive (grammatical) concepts

= concepts describing aspects of the human *cognitive code* for knowledge of language (= Universal Grammar; Haspelmath 2004), in particular innate grammatical categories and features.

Examples:

(2) Jakobson's, Chomsky & Halle's, generative feature geometry's *distinctive features* (±high, ±coronal, ±syllabic, etc.)

(3) Chomsky's features [±N, ±V] for syntactic categories
(noun: [+N, –V], verb: [–N, +V], adjective: [+N, +V], preposition: [–N, –V])

(4) functional categories such as C(OMP), I(NFL), T(ense), AGR(eement)

(5) in Relational Grammar and Lexical-Functional Grammar:
grammatical relations/functions such as *1, 2, 3, chômeur*; *SUBJ, OBJ, OBJ2, OBL*

(6) higher-level concepts such as *head, specifier, constituent, c-command, transformation, deep structure, f-structure, merge*

These concepts are concepts of **cognitive linguistics**, and as such they are potentially relevant to any language.

Cognitive concepts are needed for **descriptive adequacy** (Chomsky 1965), i.e. for an account of the grammatical system that expresses the speakers' knowledge of their language. (In other words, cognitive concepts are needed for *cognitive descriptions*, Haspelmath 2004.)

• Potential problems:
– We do not know what the right cognitive concepts are.
(We are not even close.)
– We do not know whether grammar-specific concepts are needed at all.
– We do not even know whether different speakers of the same language have the same internalized grammar.

1.3. Conceptual-semantic (grammatical) concepts

= concepts for describing the meanings of linguistic expressions, e.g.

(7) Wierzbickian semantic primitives (*I, you, someone, something, this, the same, other, think, know, want, feel, say, do, happen, good, bad, part, kind, can, ...*)

(8) Jackendoff's ontological categories (State, Event, Thing, Place, Path) and functions (BE, GO, STAY, BECOME, CAUSE, TO, FROM, VIA, ...)

(9) Semantic role labels such as *agent, patient, recipient, theme, instrument, location, goal, ...*

(10) Reichenbach's E, R, S and their combinations

These concepts are concepts of **conceptual semantics**, which is a universal aspect of languages. They are needed for grammatical research because grammatical elements and constructions typically have content (in addition to indicating the way in which lexical meanings should be composed).

• Potential problem:
– The meanings which are expressed differ across languages.
(***Solution***: Consider meanings only at a more abstract level, where meanings are commensurable across languages. That meanings must be commensurable at some level is clear from the fact that translation is possible to a very large extent. Where translation is impossible, comparing language structure is impossible, too.)

1.4. Comparative (grammatical) concepts

= concepts which are created specifically to allow cross-linguistic generalizations, e.g.

(11) *ergative case* (= a case of A, when A ≠ S, P)

Generalization: An ergative case always has an overt marker (Dixon 1979).

(12) *subject* (= the argument with the greatest number of subject properties, Keenan 1976)

Generalization: The subject normally precedes the object (Greenberg 1963).

(13) *future tense* (= a tense form that has 'future time reference, E > S' as a prominent meaning)

Generalization: Future tenses tend to be more analytic than past tenses (Dahl 1985?).

(14) *reflexive pronoun* (= a pronoun that is used for local coreference, but not for local disjoint reference)

Generalization: Reflexive pronouns are at least as long as disjoint-reference pronouns (Faltz 1985, Haspelmath 2005).

(15) *wh-movement* (= movement of a *wh*-word to a peripheral position in the clause)

Generalization: *Wh*-movement is always to the left.

(16) *wh-word* (= a word that is used (possibly among other uses) as a question pronoun)

(17) *question pronoun* (= a word that is used in parametric questions (= special questions) to represent the questioned content)

(18) *affix* (= a morpheme that can never occur on its own)

Generalization: Affixes show a strong tendency to be postposed (as "suffxes") rather than preposed (as "prefixes").

Cf. the pairs: *proposition/sentence, question/interrogative, participant/argument, time/tense, multiple/plural*

These are concepts of **(general-)comparative linguistics**, and as such they must be applicable to any language.

If we knew what the right cognitive concepts are, and if there were only a small number of them, we could compare languages in terms of the cognitive concepts.

This is generally attempted in generative linguistics, but with fairly limited success (cf. Haspelmath 2007c+).
Comparative concepts cannot be defined in terms of descriptive concepts (because these are language-specific), so they must be defined in terms of conceptual-semantic concepts, other comparative concepts, and/or highly general relational concepts ("precedes", "is part of", "expresses", etc.).

•Problems:
– ***Linguists (even comparative linguists) are often unaware that comparative concepts are a special kind of concept.***
– ***Terms for comparative concepts are often homonymous with terms for descriptive concepts***
(one solution: capitalize descriptive concepts, because they are like proper names: *Ergative case, Future tense, Subject*, etc.; Comrie 1976, Bybee 1985, Croft 2001)

2. What happens when the concept types are not distinguished

2.1. Conflating descriptive concepts and cognitive concepts

Generative linguists generally attempt to describe (= "analyze") particular languages with cognitive concepts, assuming that the set of required cognitive concepts is within our reach, and that they bear a strong resemblance to the kinds of concepts that Western linguists have been working with for centuries.

In practice, this means that they do not describe each language in its own terms, but rather describe all languages in terms of a currently accepted list of hypothesized cognitive concepts.

• Problems:

– The assumptions about cognitive concepts are very unstable, so the descriptions are not cumulative (= cannot build on each other).

– The assumptions about cognitive concepts are typically based on widely studied languages (such as English and Japanese), and/or on a few languages that have been prominently described in the literature (e.g. Icelandic, Mohawk, Chichewa, Warlpiri), so that a strong bias is unavoidable.

Newmeyer (1998) concludes from the uncertainty about cognitive concepts that comparison of languages is difficult (Ch. 6: "Language typology and its difficulties"):

> "Assigning category membership is often no easy task... Is Inflection the head of the category Sentence, thus transforming the latter into a[n] Inflection Phrase (IP)? ... Is every Noun Phrase dominated by a Determiner Phrase (DP)? ... There are no settled answers to these questions. Given the fact that we are unsure precisely what the inventory of categories for any language is, it is clearly premature to make sweeping claims about their semantic or discourse roots. Yet much functionalist-based typological work does just that." (Newmeyer 1998: 338)

Here Newmeyer makes two wrong presuppositions:
– that one needs to know "precisely what the inventory of [cognitive? descriptive?] categories for any language is" in order to compare languages
– that the discussion in generative syntax about IP, DP etc. is about the categories of particular languages (in fact, it is about the cognitive concepts of UG, and it is driven primarily by speculative assumptions about UG)

2.2. Conflating cognitive concepts and comparative concepts

There is no well-established methodology for identifying cognitive concepts. In practice, rough similarities between categories of Latin/English and other languages are taken as sufficient reason to equate the categories (= to assume that the Latin/English category and the category of the other language represent the same cognitive concept). This leads to problems when similarity relations are more complex, e.g.

– Tagalog *ang*-phrases are similar to subjects and to topics
– Spanish *que* is similar to a relative pronoun and to a complementizer
– Chinese property words are similar to adjectives and to verbs
– the Turkish plural marker *-ler/-lar* is similar to an affix and to a clitic

In such cases, the "rough similarity" approach fails. (Category-assignment controversies are unresolvable; see Haspelmath 2007a+.)

In order to establish correspondences across languages in a consistent way, the approach adopted by typologists for comparative concepts (§1.4) must be employed. But this approach yields comparative concepts, not concepts that can be used for language-particular description.

A glaring example of the confusion between cognitive concepts and comparative concepts:

> "[F]ormal analysis of language is a logical and temporal prerequisite to language typology. That is, if one's goal is to describe and explain the typological distribution of linguistic elements, then one's first task should be to develop a formal theory...
>
> [T]he only question is how much formal analysis is a prerequisite [to functional analysis]. I will suggest that the answer is a great deal more than many functionally oriented linguists would acknowledge.
>
> To read the literature of the functional-typological approach, one gets the impression that the task of identifying the grammatical elements in a particular language is considered to be fairly trivial... it seems to be based, more often than not, on some loose semantic or pragmatic criteria" (Newmeyer 1998a: 337-338)

But Newmeyer wrongly presupposes
– that comparison of languages must be based on cognitive concepts ("formal analysis"),
– **or** that the descriptive concepts used by descriptive linguists will be readily comparable (which is impossible)
(He completely overlooks the possibility of comparative concepts.)

2.3. Conflating descriptive concepts and comparative concepts

A common objection when presenting *WALS* maps, e.g. the map on Adjective-Noun order:

> "But not all languages have adjectives!"

In what sense is *adjective* used here?

– descriptive? NO: nonsensical (cf. Not all countries have Warsaw.)

– cognitive? maybe:
But unclear how to identify *adjectives* (cf. §2.2),
and this objection tends to come from non-generativists.

– comparative!
(= the category that contains words expressing property concepts such as 'large, small, old, new, good, bad, red, blue, green, yellow')

2.4. Conflating comparative concepts and conceptual-semantic concepts

Sometimes it is said that cross-linguistic comparison is based on conceptual-semantic concepts:

> "In recent universalist studies.., the common strategy has been to employ *semantic* (or perhaps better: *cognitive*) definitions for the parameter of the typology. That is, the feature upon which the typology is to be based is defined in terms of its semantic content or function, rather than in terms of its alleged structural characteristics." (Stassen 1985:14)

> "typologists generally use definitions of a grammatical phenomenon for cross-linguistic comparison that are "external" to the linguistic structural system, that is, semantic, pragmatic, or discourse-based definitions for morphosyntactic phenomena and phonetic definitions for phonological phenomena." (Croft 1995:88)

> "In view of the foregoing objections to formal definitions, linguistic typologists opt for functional [i.e. semantic, pragmatic and/or cognitive] definitions for purposes of cross-linguistic identification." (Song 2001:11)

But in reality, typologists rarely compare everything that expresses a particular meaning. Instead, they confine their attention to **comparable constructions** expressing particular meanings. Cf.:

> "Our solution to [the problem of cross-linguistic identification] is to use an essentially semantically based definition of RC [=relative clause]. We consider any syntactic object to be an RC if it specifies a set of objects (perhaps a one-member set) in two steps: a larger set is specified, called the *domain* of relativization, and then restricted to some subset of which a certain sentence, the *restricting sentence*, is true. The domain of relativization is expressed in surface structure by the *head NP*, and the restricting sentence by the *restricting clause*, which may look more or less like a surface sentence depending on the language." (Keenan & Comrie 1977:63)

So are the following three expressions all examples of relative clauses? (Presumably not.)

a. *The hat* ***which is white*** *is my sister's.*
b. *The* ***white*** *hat is my sister's.*
c. *One of the hats is my sister's.* ***It is white****.*

3. Some concepts and generalizations about reciprocal constructions (cf. Haspelmath 2007b+)

(19) *Aisha and Pedro pinched **each other**.*

3.1. Concepts/terms

***reciprocal** construction* (comparative concept) vs.
***mutual** situation/configuration* (conceptual-semantic concept)

cf. König & Kokutani 2006:272-273: ("symmetric" instead of "mutual")

> The core area of reciprocal constructions is easy to identify in terms of prototypical examples, but there are problems at the periphery, and the task of providing explicit criteria for their identification across languages and their delimitation from other constructions is certainly not a trivial one. In contrast to most other studies, we will therefore try to formulate such criteria. To begin with, we draw a strict distinction between **symmetric** as a semantic property and **reciprocal** as a syntactic property. The class of symmetric predicates and the set of reciprocal constructions can now be defined as follows:
>
> (i) **Symmetric** predicates are basic predicates with at least two argument (valency) positions which denote binary (or ternary) relations R among members of a set A with the following semantic property: $\forall x, y \in A\ (x \neq y \longrightarrow R(x, y))$, that is, for specific substitutions of values a and b ($a, b \in A$) for the variables x and y: aRb <—> bRa.
>
> (ii) **Reciprocal** constructions are grammatical means for the expression of symmetrical relations for any n-ary predicate and for at least one set of arguments A, with $|A| \geq 2$; it is a typical feature of such constructions that one of the arguments denotes a set A as specified above and that the basic argument structure of the relevant predicate is reduced or changed in such a way that not all argument positions are filled by referential expressions.

Nedjalkov 2007: a *reciprocal construction* expresses a *reciprocal situation*; participant occurring in a reciprocal situation: *reciprocant* (Haspelmath 2007: ***mutuant***)

Mutual configurations can be left **implicit**, as in (20) and (21):

(20) a. *Hector and Achilles fought obsessively.*
b. *Lisi and Aisha are in love.*

(21) To'aba'ita (Lichtenberk, Ch. 36, p. 1554)
Kero *musu-a* *babali-daro'a.*
3DU.FACT kiss-3.OBJ cheek-3DU.POSS
'The two of them kissed them/themselves/each other on the cheek.'
(Lit. '...kissed their cheeks.')

They can be expressed in a **non-specialized** and **fully compositional** way:

(22) *Aisha pinched Pedro, and Pedro pinched Aisha.*
(23) *Each participant knows all the others well.*

Specialized constructions for expressing mutual situations are called **reciprocal constructions**. Most, but not all reciprocal constructions are monoclausal:

(24) *Aisha pinched Pedro, and Pedro pinched her **back**.*

Lexical reciprocals (= *allelic predicates*) can be defined as predicates that express a mutual configuration by themselves, without necessary grammatical marking ('marry', 'quarrel', 'friend', 'adjoin', 'next to', 'same as', 'different from', 'resemble').

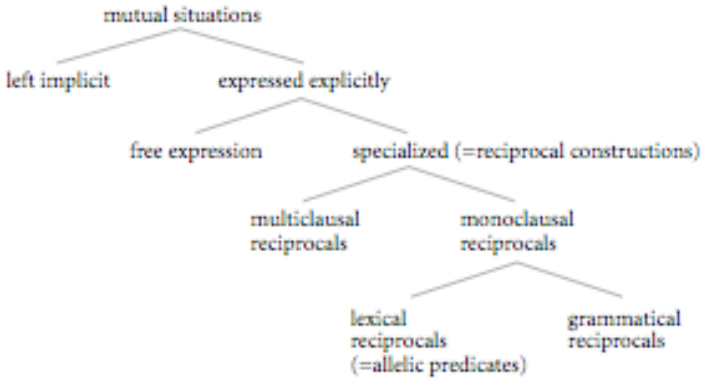

Figure 1. Ways of expressing mutual situations

Languages typically express the set of mutuants as a single argument occupying one of the two syntactic positions in which the mutuants are in the corresponding non-reciprocal clause pair. This argument is called the **reciprocator** here. The other syntactic position is called the **reciprocee**.

The reciprocee position is either not filled or is filled by a **reciprocal anaphor**, i.e. a noun phrase whose reference is dependent on and derivative of a nearby referential expression.

(25) reciprocee position is not filled: Cashinahua (Camargo, Ch. 45, p. 1869)

a. *paku-n haidu dɨt-ai.*
Paco-ERG Jairo hit-PROGR
'Paco is hitting Jairo, Jairo is hitting Paco.'

b. *paku inun haidu dɨt-**nanan**-ai-bu.*
Paco and Jairo hit-REC-PROGR-PL
'Paco and Jairo are hitting each other.'

(In such cases, the verb is typically marked as reciprocal: **verb-marked reciprocal**.)

(26) reciprocee position is filled by an anaphor:
Vietnamese (Bystrov & Stankevich, Ch. 47, p. 1943)

a. *Lan yêu Hống. Hống yêu Lan.*
Lan love Hong Hong love Lan
'Lan loves Hong. Hong loves Lan.'

b. *Lan và Hống yêu **nhau**.*
Lan and Hong love each.other
'Lan and Hong love each other.'

3.2. Universals

Universal 1:
In all languages, monoclausal reciprocal constructions are at least as complex formally as the corresponding non-reciprocal constructions denoting simple events.

Universal 2:
In all languages with reciprocal constructions, there are constructions in which the mutuants are expressed in a single (nonsingular) argument of the predicate.

i.e. all languages have simple reciprocal constructions (Nedjalkov Ch. 1, §7), but not all languages have **discontinuous reciprocal constructions** (as in 27c):

(27) German (Wiemer & Nedjalkov, Ch. 10, p. 47)
a. *Hans schlägt Paul, und Paul schlägt Hans.*
'Hans hits Paul, and Paul hits Hans.'

b. *Hans und Paul schlagen* ***sich****.*
'Hans and Paul hit each other.'

c. *Hans schlägt* ***sich*** *mit Paul.*
'Hans and Paul hit each other.' (Lit. 'Hans hits each other with Paul.')

Discontinuous reciprocal constructions = constructions in which the mutuants are expressed by two different arguments.

Universal 3:
No language has a reciprocal construction in which there are two mutuant-expressing arguments that are coded like the A (most agent-like argument) and the P (most patient-like argument) of a typical transitive clause.

Thus, we do not in general find reciprocals of the sort shown schematically in (28):

(28) *The girl-NOM kissed-****REC*** *the boy-ACC.*
'The girl and the boy kissed (each other).'

In Iwaidja (Evans 2007), there is a monoclausal construction violating Universal 3:

(29) *anb-uku-n* ***lda*** ***wamin*** *a-ngurnaj*
3PL.A>3PL.P-give-NPST and 3PL.REC 3PL-name
'They used to give each other their (clan) names.'
(Lit. 'They gave them, and mutually, their names.')

Universal 4:
Only verb-marked reciprocals allow a discontinuous reciprocal construction.

Universal 5:
All reciprocal constructions with two arguments that both refer to the set of mutuants are anaphoric reciprocal constructions.

That is, reciprocal constructions where the noun phrase denoting the set of mutuants is simply repeated, as in (18), are excluded by this universal.

(30) **Taro and Jiro phoned Taro and Jiro.*

Universal 7:
The more clearly two arguments differ in prominence, the easier it is for the more prominent argument to antecede the anaphor. Less prominent arguments cannot antecede more prominent arguments.

(31) a. English
*****Each other*** *pinched Aisha and Pedro.*

b. Basque (Hualde & Ortiz de Urbina 2003:608)
*****Elkarr-ek*** *Epi eta Blas maite ditu.*
each.other-ERG Epi and Blas love AUX
'Each other love(s) Epi and Blas.'

What exactly counts as "prominence" is a difficult question, cf. Shkarban & Rachkov (Ch. 22, p. 922) on Tagalog:

(32) *Nag-pasalamat-an ang* ***isa't isa.***
AG.REC.PFV-thank-REC NOM one.and one
'[They] thanked each other.' (lit. 'Each other was thanked by them.')

Universal 11:
In verbal reciprocals, the reciprocator is always the subject, and the reciprocee can only be the direct object, the indirect object, the possessor of a co-argument, or an adverbial.

(33) a. I love you. You love me. —> *Ø Love-REC us.
(OK: We love-REC Ø.)

b. I gave you it. You gave me it. —> *Ø Gave-REC us it.
(OK: We gave-REC Ø it.)

c. I hold your hand. You hold my hand.
—> *Ø Hold-REC our hand.
(OK: We hold-REC Ø's hand.)

d. I come to you. You come to me. —> *Ø Come-REC to us.
(OK: We come-REC Ø.)

References

Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form.* Amsterdam: Benjamins.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.

Croft, William. 1995. Modern syntactic typology. In: Shibatani, Masayoshi & Bynon, Theodora (eds.) 1995. *Approaches to language typology.* Oxford: Clarendon Press, 85-144.

Croft, William. 2001. *Radical Construction Grammar.* Oxford: Oxford University Press.

Dahl, Östen. 1985. *Tense and aspect systems.* Oxford: Basil Blackwell.

Dixon, R.M.W. 1979. Ergativity. *Language* 55:59-138.

Faltz, Leonard M. 1985. *Reflexivization: A Study in Universal Syntax*. New York: Garland.

Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements." In: Greenberg, Joseph H. (eds.) *Universals of grammar*, 73-113. Cambridge, Mass.: MIT Press.

Haspelmath, Martin. 2004. "Does linguistic explanation presuppose linguistic description?" *Studies in Language* 28.3: 554-579

Haspelmath, Martin. 2005. "A frequentist explanation of some universals of reflexive marking". Paper presented at SLE Annual Conference in Valencia, September 2005.

Haspelmath, Martin. 2007a+. "Pre-established categories don't exist: consequences for language description and typology." To appear in *Linguistic Typology* 11

Haspelmath, Martin. 2007b+. "Further remarks on reciprocal constructions." To appear in: Nedjalkov, Vladimir P. (ed.) *Reciprocal constructions.* Amsterdam: Benjamins.

Haspelmath, Martin. 2007c+. "Parametric versus functional explanations of syntactic universals." To appear in: Theresa Biberauer & Anders Holmberg (eds.) *The limits of syntactic variation.* Amsterdam: Benjamins.

Keenan, Edward L. & Comrie, Bernard. 1977. "Noun phrase accessibility and universal grammar." *Linguistic Inquiry* 8: 63-99.

Keenan, Edward. 1976. "Toward a universal definition of 'subject'." In: Li, Charles N. (ed.) *Subject and topic.* New York: Academic Press, 303-333.

König, Ekkehard & Kokutani, Shigehiro. 2006. Towards a typology of reciprocal constructions: focus on German and Japanese. *Linguistics* 44.2: 271-302.

Newmeyer, Frederick J. 1998. *Language form and language function.* Cambridge/MA: MIT Press.

Song, Jae Jung. 2001. *Linguistic typology: Morphology and syntax*. Harlow & London: Pearson Education.

Stassen, Leon. 1985. *Comparison and universal grammar.* Oxford: Blackwell.