

Why is language typology possible?

Martin Haspelmath

Languages are incomparable

Each language has its own system.

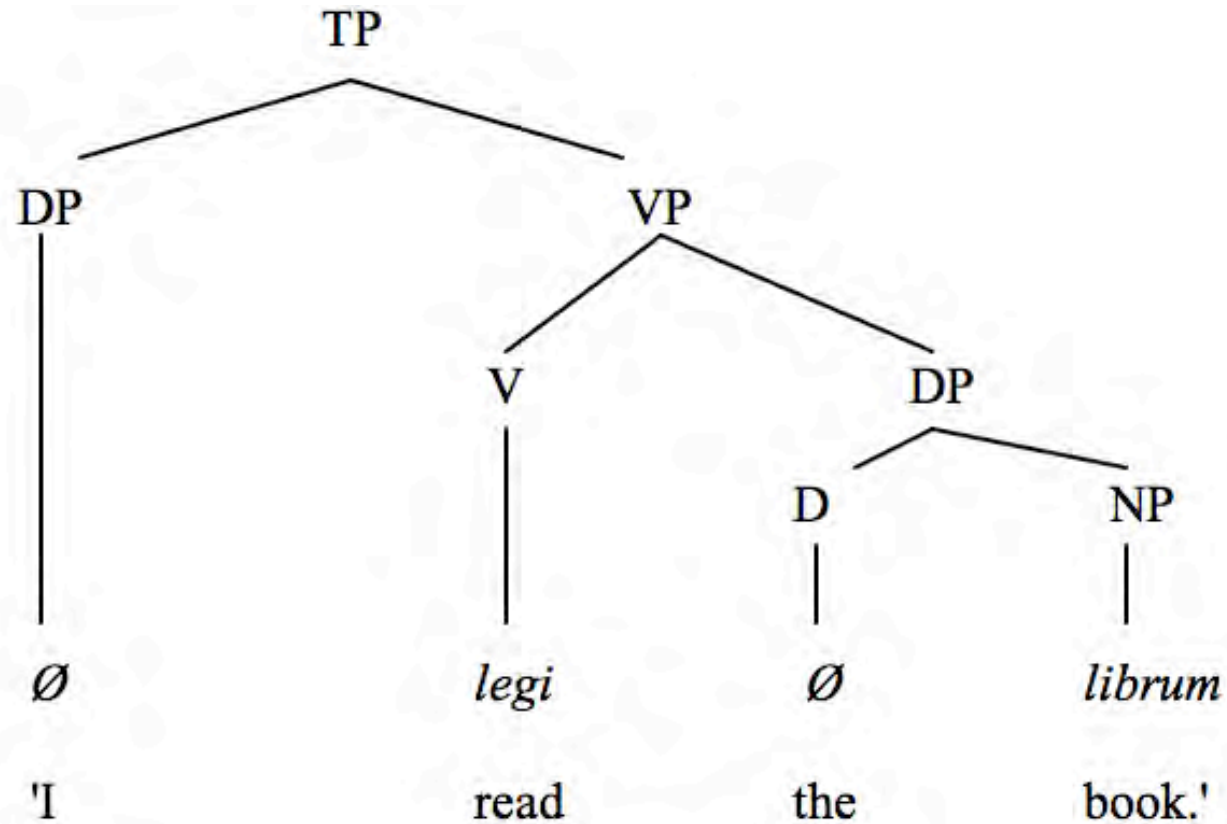
Each language has its own categories.

Each language is a world of its own.

Or are all languages like Latin?

nominative	<i>the book</i>
genitive	<i>of the book</i>
dative	<i>to the book</i>
accusative	<i>the book</i>
ablative	<i>from the book</i>

Or are all languages like English?



How could languages be compared?

If languages are so different:

What could be possible ***tertia comparationis***
(= entities that are identical across comparanda
and thus permit comparison)?

Three approaches

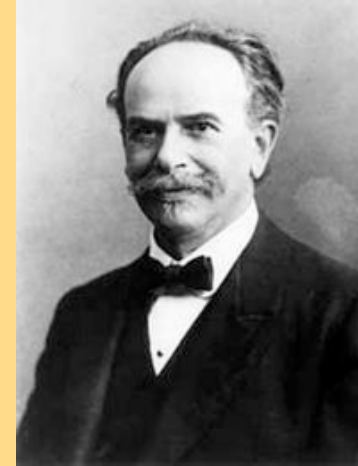
- Indeed, language typology is impossible (**non-aprioristic structuralism**)

- Typology is possible based on cross-linguistic categories (**aprioristic generativism**)

- Typology is possible without cross-linguistic categories (**non-aprioristic typology**)

Non-aprioristic structuralism:

Franz Boas (1858-1942)



The categories chosen for description in the *Handbook* **“depend entirely on the inner form of each language...”**

Boas, Franz. 1911. Introduction to *The Handbook of American Indian Languages*.

Non-aprioristic structuralism:

Ferdinand de Saussure
(1857-1913)



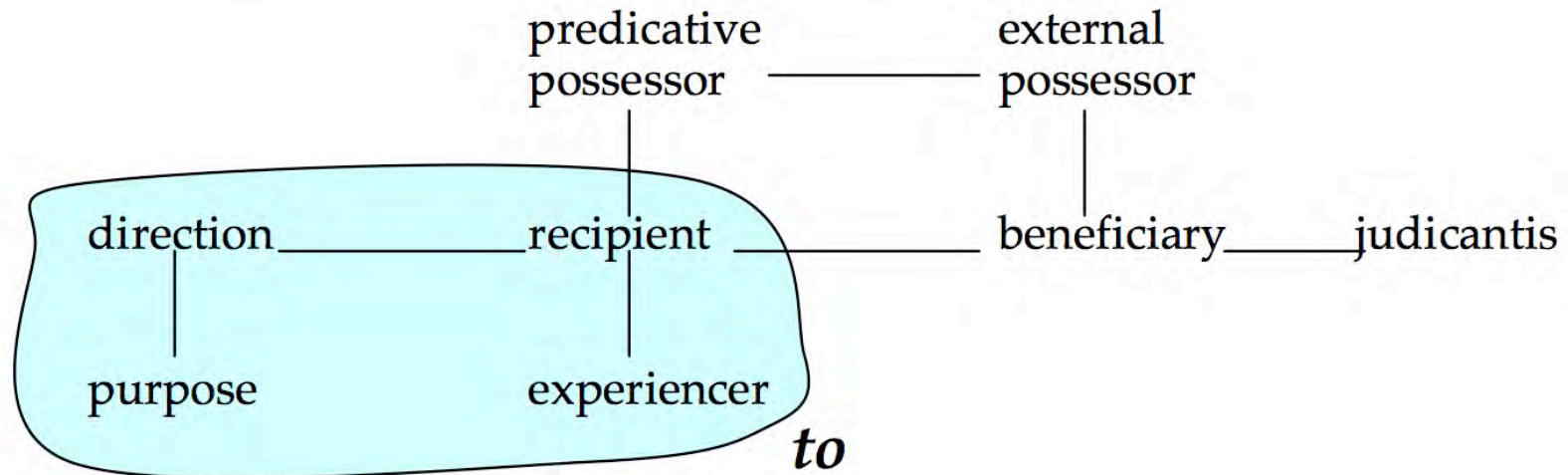
**“dans la langue il n’y a que des
différences...”** (In a language there are only
differences)

i.e. all categories are determined by the ways in
which they differ from other categories, and each
language has different ways of cutting up the sound
space and the meaning space

de Saussure, Ferdinand. 1915. *Cours de linguistique générale*.

Example: Datives across languages

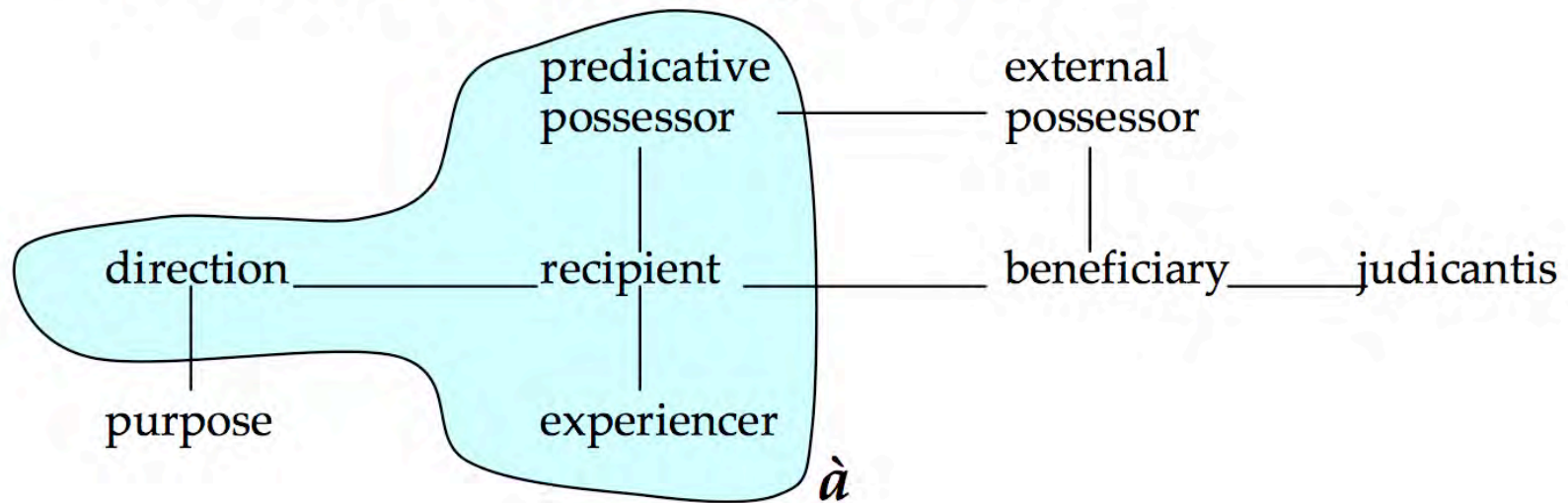
the English Dative preposition *to*



cf. Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and cross-linguistic comparison

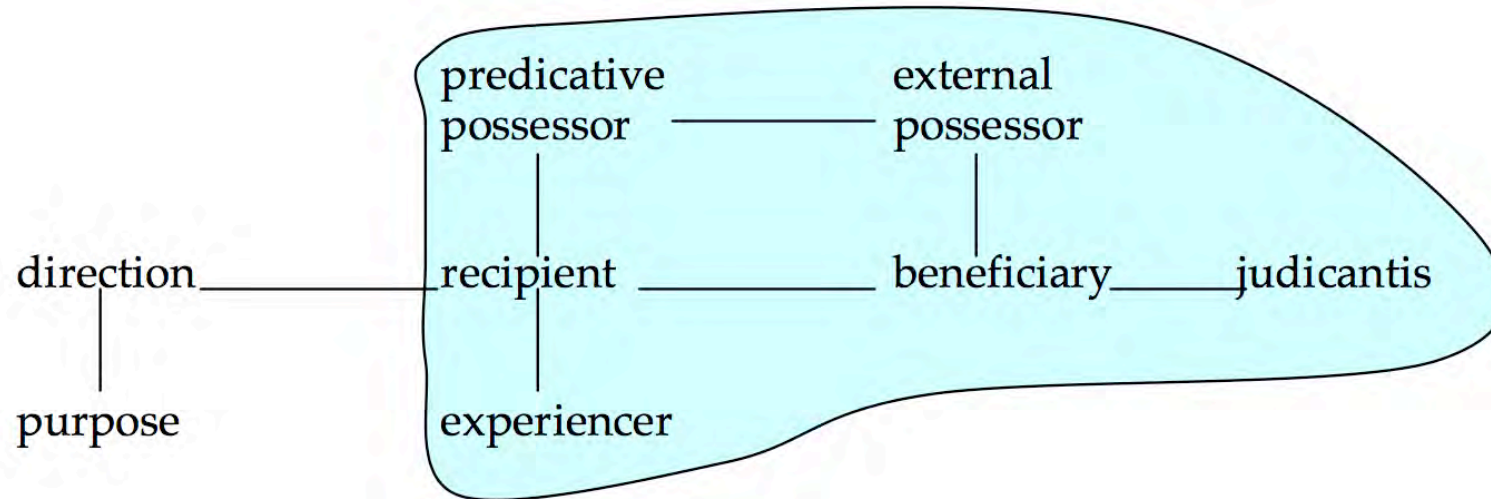
Example: Datives across languages

the French Dative preposition *à*



Example: Datives across languages

the German **Dative** case



Non-aprioristic structuralism:

Peter H. Matthews
(University of Cambridge)

Matthews 1997:199:

"To ask whether a language 'has' some category is...to ask a fairly sophisticated question... Such warnings were once commonplace... [but] in many typological studies, scholars seem to proceed as if they were irrelevant. Cross-linguistic comparison is, on the contrary, initiated independently or in advance of detailed analyses of individual systems. One approaches each language with, in effect, a checklist....

In principle, **comparisons must be based on analyses of particular systems:** but, in so many interesting cases, such analyses are lacking."

Matthews, Peter H. 1997. Structural linguistics in the 1990s. *Lingua* 100:193-203.



Non-aprioristic structuralism: Language typology is impossible

- Languages have incommensurable systems
- Each language has its own categories
- Language typology is not possible
- Hence “American structuralists largely ignored typology” (Greenberg 1974)

aprioristic generativism:
Noam A. Chomsky (1928-)



Grammatical categories are assumed to be *substantive universals*, belonging to the innate Universal Grammar

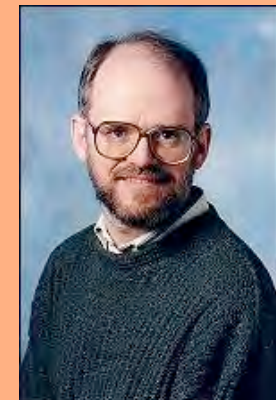
"We require that the grammar of a given language be constituted in accord with a specific theory of linguistic structure in which such terms as "phoneme" and "phrase" are defined independently of any particular language." (Chomsky 1957: 50)

aprioristic generativism:

Principles and Parameters

- Languages share a large amount of their grammatical patterns: principles and categories
- They differ in the way the open parameters are set (“a few mental switches”)

“[I]n the domain of grammar we find that languages are not so unique. Although the total number of imaginable grammatical systems is in theory quite large, in practice we find relatively few systems being used over and over again in different parts of the world.” (Baker 2001:116)



Aprioristic generativism:

Language typology is easy (in principle)

- All languages are made from the same (innate) building blocks
- There is a set of universally available categories from which languages may choose
- These are **cross-linguistic categories**
- Language typology simply requires that we match language-specific phenomena with the cross-linguistic categories, and then compare the way the categories combine across languages

Example: Datives across languages

dative = [−nominal, +verbal, +oblique]

nominative = [−nominal, −verbal, −oblique]

accusative = [−nominal, +verbal, −oblique]

genitive = [+nominal, +verbal, −oblique]

Bierwisch, Manfred. 1967. Syntactic features in morphology: general problems of so-called pronominal inflection in German. In: *To Honour Roman Jakobson*. The Hague: Mouton, 239-270.

Müller, Gereon. 2005. Syncretism and iconicity in Icelandic noun declensions. In: *Yearbook of morphology 2004*. Dordrecht: Springer, 229-271.

Aprioristic generativism:

Language typology is not straightforward in practice (I)

Identifying language-specific phenomena with cross-linguistic categories turns out to be very difficult

"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... ." (Newmeyer 1998: 338)



Aprioristic generativism:

Language typology is not straightforward in practice (2)

The idea that language-specific categories are equated with cross-linguistic categories has given rise to countless **category-assignment controversies.**

- (i) Is English *-like* a **stem** or a **suffix**? (cf. Tuggy 1992)
- (ii) Is the Romanian definite article a **clitic** or a **suffix**? (Ortmann & Popescu 2000)
- (iii) Is English *silver ring* a **phrase** or a **compound**? (e.g. Bauer 1998, Giegerich 2004)
- (iv) Are Mandarin Chinese property words **adjectives** or **verbs** (McCawley 1992, Dixon 2004)?
- (v) Is the Tagalog *ang*-phrase a **subject** or a **topic**? (Schachter 1976)
- (vi) Is German *er* a **pronoun** or a **determiner**? (Vater 2000)
- (vii) Is English *that* in relative clauses a **pronoun** or a **complementizer**? (van der Auwera 1985)
- (viii) Is the English adverbial *-ly* an **inflectional** or a **derivational** suffix?
- (ix) Are the two types of intransitive verbs in Jalonke (Mande) **unaccusatives** and **unergatives**? (or are they something else?) (Lüpke 2006)
- (x) Are French subject clitics (*je, tu, il...*) **pronouns** or **agreement markers**? (De Cat 2005)
- (xi) Is the German dative a **structural case** or an **inherent case**? (Wegener 1991, Woolford 2006)

Haspelmath, Martin. 2007. Pre-established categories don't exist—consequences for language description and typology. *Linguistic Typology* 11:119–132.

Aprioristic generativism:

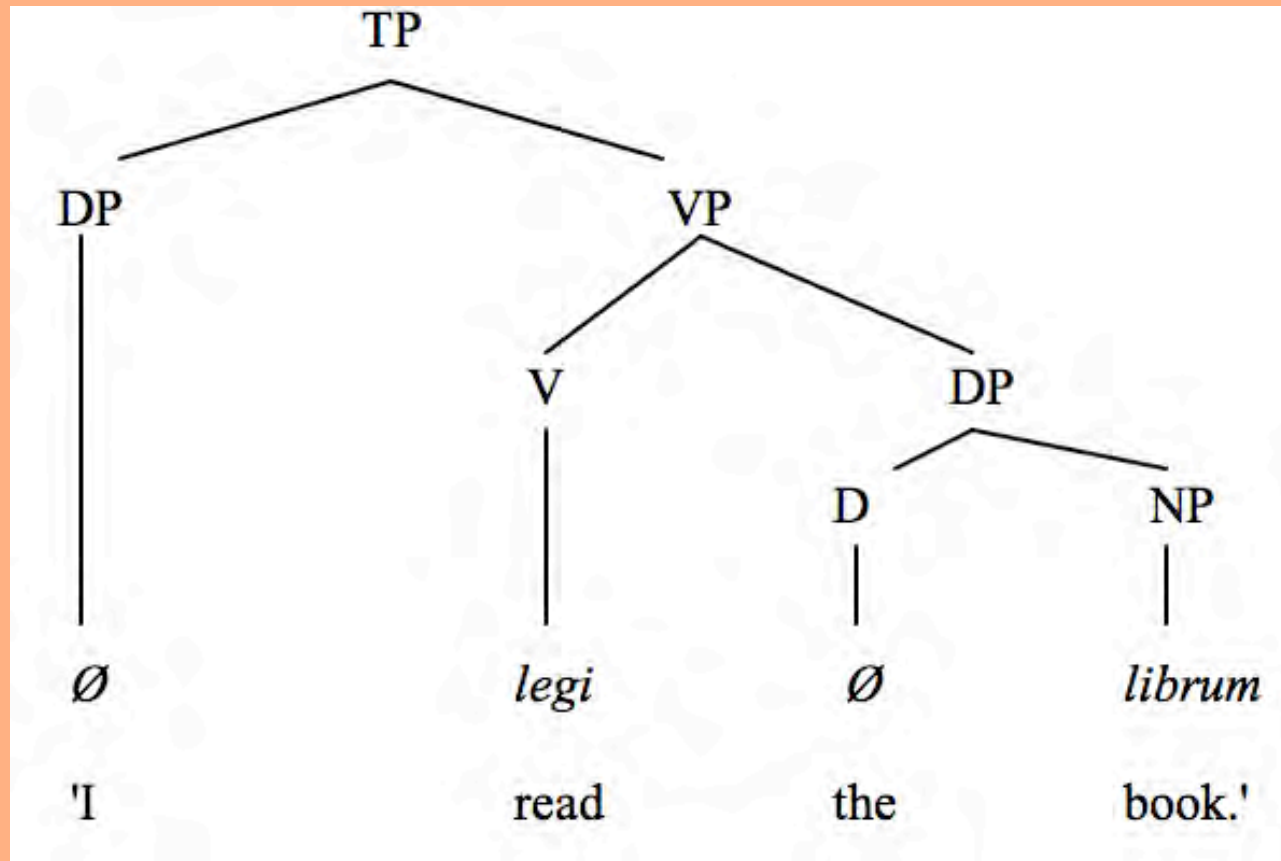
Language typology is not straightforward in practice (3)

In practice, linguists have often simply used the categories that were motivated for English and carried them over to other languages

- early generative grammar „took over with no substantial justification the categories of traditional grammar... The initial empirical base was English and as this base was broadened to include more and more different languages, these categories were naturally taken over for the “new” languages“ (Bach 2004:56-57).



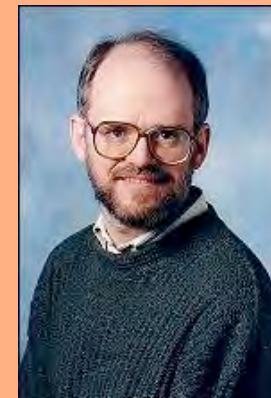
...as we saw earlier:



Aprioristic generativism:

Language typology is not straightforward in practice (4)

- “The first and most famous parameter to have been proposed is the Pro-drop Parameter (or Null Subject Parameter) of Chomsky (1981) and Rizzi (1982: ch. 4)... History has not been kind to the Pro-drop Parameter as originally stated. It is now well-known that the cluster of properties that this parameter claimed to be related fragments in various ways when one looks at a wider range of languages, either dialects or languages from other families (Gilligan 1987). Moreover, **few new proposals for parameters with the same kind of scope as the Pro-drop Parameter have gained currency since.**” (Baker, to appear)



Is a non-aprioristic typology possible?

- Perhaps surprisingly, typologists have recently tended to emphasize non-apriorism
- Matthew S. Dryer, William Croft, Gilbert Lazard
 - Haspelmath, Martin. 2007. Pre-established categories don't exist—consequences for language description and typology. *Linguistic Typology* 11:119-132.
- "A recent trend within linguistic typology ... rejects the idea that there are cross-linguistic formal grammatical categories" (Newmeyer 2007: 133)
- "It is interesting to observe that a substantial proportion of the entries in the *World Atlas of Language Structures* (Haspelmath et al. (eds.) 2005) utilize cross-linguistic formal categories, despite the position of two of the editors that such categories do not exist." (Newmeyer 2007: 138)



Matthew S. Dryer

1997. "Are grammatical relations universal?" In: Bybee, Joan & Haiman, John & Thompson, Sandra A. (eds.) *Essays on language function and language type*. Amsterdam: Benjamins, 115-143.

There are four kinds of things that might exist in the domain of grammatical relations, listed in (1).

(1) a. Grammatical relations in **particular languages**

b. **Similarities** among these language-particular grammatical relations

c. Functional, cognitive and semantic **explanations** for these similarities

d. Grammatical relations **in a cross-linguistic sense**

... the first three things in (1) exist but ... **the fourth one is unnecessary.** (p. 116)"



Language-specific grammatical relations

e.g. Tagalog **agent, topic, subject**

(Schachter & Otanes 1972:69):

*Bumabasa ng diyaryo **ang** titser.*
read.ActFoc GEN newspaper TOP teacher
'The teacher is reading a newspaper.'

*Binabasa ng titser **ang** diyaryo.*
read.ObjFoc GEN teacher TOP newspaper
'The teacher is reading the newspaper.'

The language-specific grammatical relation
“ang-relation” is sufficient to describe Tagalog.
There is no need to decide whether it is a “subject”
or not, or a “topic” and if so what kind, etc.

Language-specific parts of speech

e.g. Mandarin Chinese **verbs** vs. **adjectives**?

	verb	adjective
predicative use	<i>niaor fei</i> 'the bird flies'	<i>hua hóng</i> 'the flower is red'
attributive use	<i>fei-de niaor</i> 'bird that flies'	<i>hóng-de hua</i> 'red flower'
comparative construction	<i>A bi C xihuan B</i> 'A likes B more than C'	<i>A bi C gao</i> 'A is taller than C'
compounding	*fei-niaor 'flying bird'	<i>hóng-hua</i> 'red flower'

Language-specific parts of speech

e.g. Mandarin Chinese verbs vs. adjectives?

Whether Mandarin has adjectives as a separate part of speech depends on whether one accepts the compounding construction as a relevant criterion.

But there is **no principled basis** on which to decide which criteria should be relevant for assigning category-membership.

The criteria for grammatical categories are typically **language-specific**, and cannot be carried over from one language to the next.

→ Linguists have to make **arbitrary** decisions (“methodological opportunism”).

William Croft

2001. *Radical Construction Grammar*. Oxford: Oxford University Press.

- “There is no universal inventory of atomic primitives...
- ...the empirical argument against universal categories is not based on an esoteric fact of a single little-known language; it is based on the well-known and pervasive grammatical diversity of languages. The alternative view, **rejecting universal categories and relations**...avoids the inconsistencies of cross-linguistic methodological opportunism, which leads to undecidable disputes over questions such as whether Makah has the Noun–Verb distinction...The alternative view allows each language to be itself: **it respects the grammatical diversity of languages, and the uniqueness of each language’s grammar**. In this respect, the alternative view is little different from the view espoused by American structuralists” (2001:32-34)



Is a non-aprioristic typology possible?

- **Yes**, but without **cross-linguistic categories**, and not on the basis of language-specific **descriptive categories**.
- Non-aprioristic typology is based on **comparative concepts**:
 - i.e. concepts created by comparative linguists, for the specific purpose of cross-linguistic comparison.
 - Comparative concepts are universally applicable, and they are defined on the basis of other universally applicable concepts: **universal meanings** and **universal formal notions** (or on the basis of other comparative concepts).

Non-aprioristic typology: comparative concepts

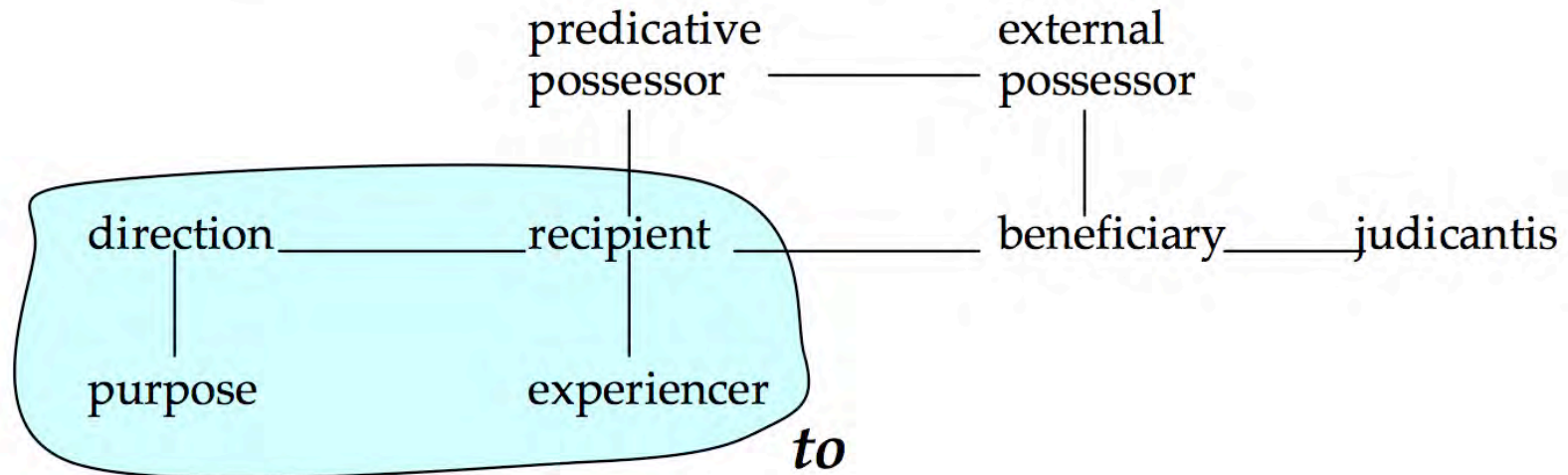
- Comparative concepts are not needed by descriptive linguists or by speakers.
- They are not psychologically real, and they cannot be right or wrong.
- They can only be more or less well-suited to the task of permitting cross-linguistic comparison; each comparative linguist can make his or her own comparative concepts.
- Comparative concepts are like measuring units in other sciences: somewhat arbitrary, but indispensable for comparing objects of study (B. Bickel)

An example of a comparative concept: dative markers

- **Generalization:** In all languages with a dative and an accusative marker, the dative marker is at least as long as the accusative case marker.
- **Definition:** A dative marker is a case marker or adposition that has among its functions the coding of the recipient argument of a physical transfer verb (such as 'give', 'lend', 'sell', 'hand'), when this is coded differently from the theme argument.

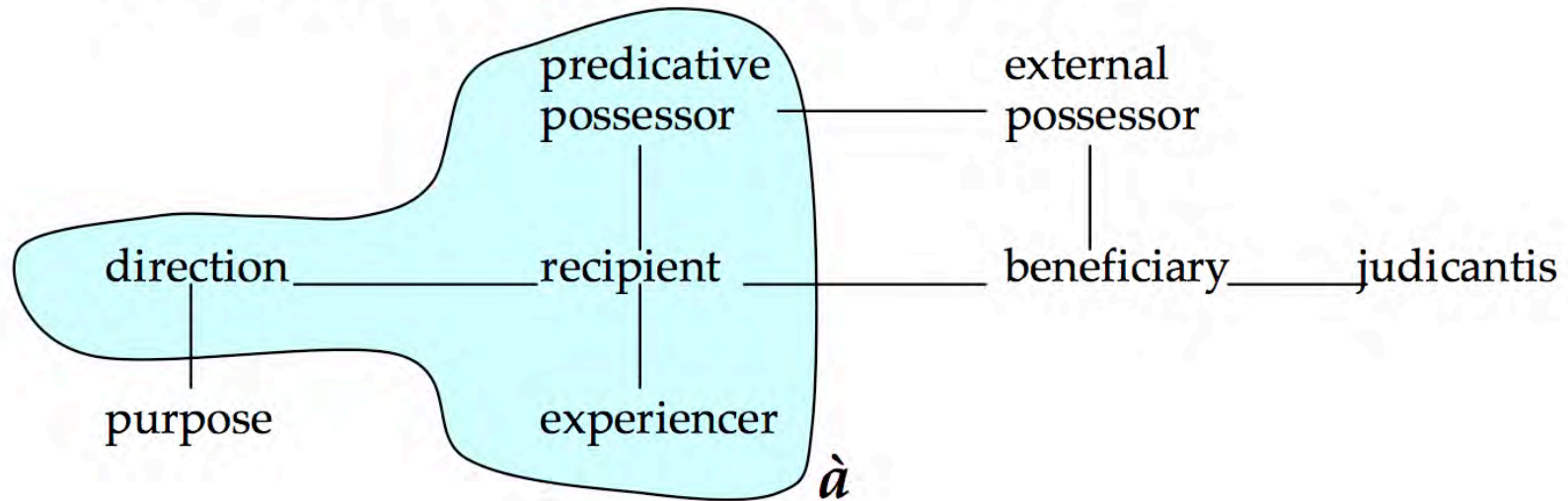
Comparative concept “dative”

the English Dative preposition *to*



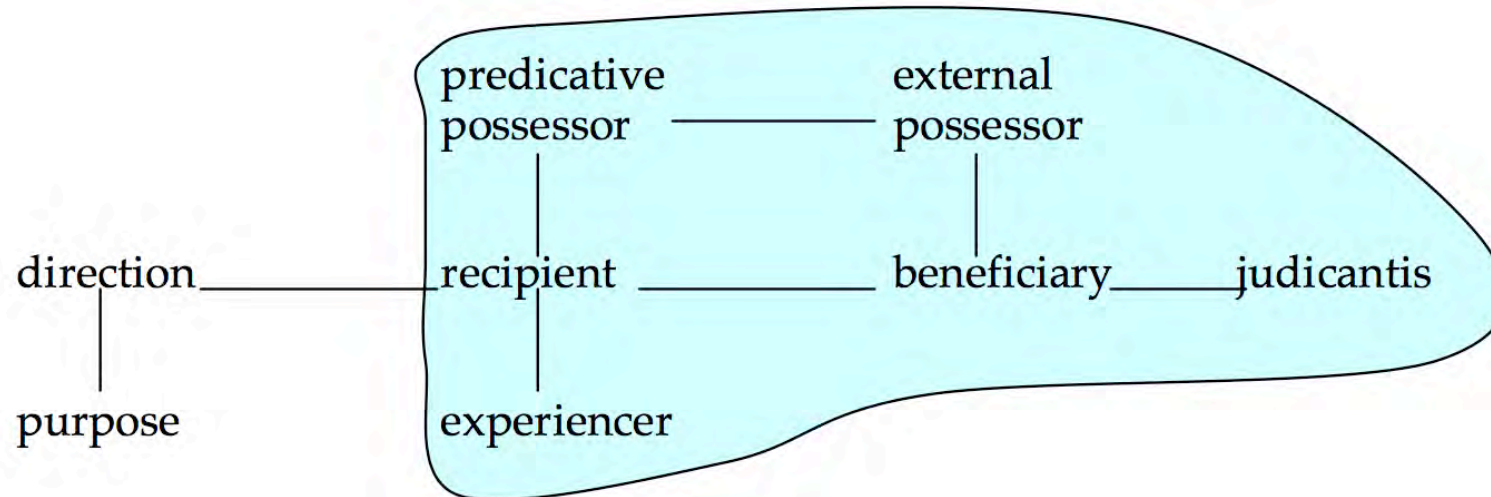
Comparative concept “dative”

the French Dative preposition *à*



Comparative concept “dative”

the German **Dative** case



Comparative concept “dative”

- **formal condition: recipient ≠ theme**
- Martuthunira (Pama-Nyungan; Dench 1995:67)

Ngunhu kanyara ngurnu jinkarn-ku yungku-lha ngurnula-ngu-u mimi-i.
that.NOM man the.ACC stick-ACC give-PST he-GEN-ACC uncle-ACC
'That man gave the digging stick to his uncle.'

- This case is **not** a dative case in the comparative sense.

More comparative concepts

- adjective
- future tense
- *wh*-movement
- relative clause
- ergative case
- for each concept, I show how it can be defined without appeal to language-specific criteria
- for each concept, I show that the descriptive categories of individual languages may be quite different

More comparative concepts (I): **adjective**

- **generalization:** If a language has dominant SOV word order and the genitive follows the governing noun, then the adjective likewise follows the noun (Greenberg 1963, universal 5).
- **definition:** An adjective is a lexeme that is primarily used to attribute a property to a nominal referent.
- **comparative concept vs. descriptive category:** a language may or may not have a separate word class Adjective. If in a language property concepts are Verbs (e.g. in Thai, Wolof) they still fall under the above definition.

More comparative concepts (2): **future tense**

- **generalization:** In all languages, future tenses are coded with at least as much material as present tenses (e.g. Latin *lauda-Ø-t* 'praises' vs. *lauda-bi-t* 'will praise').
- **definition:** A future tense is a verbal grammatical construction that has future time reference as one prominent meaning.
- **comparative concept vs. descriptive category:** the Spanish Future (*vendrá* 'will come') also has a probability sense; the Lezgian Future (*qweda* 'will come') also has a habitual sense; in the Korean Volitional Mood (*-keyss*), the future meaning is not the primary sense, but it still falls under the above definition.

More comparative concepts (3): ***wh*-movement**

- **generalization:** *Wh*-movement is always to the left.
- **definition:** *Wh*-movement is a syntactic construction in which a *wh*-word occurs in a special position in which its non-*wh*-counterpart would not normally occur.
- **definition 2:** A *wh*-word is a word that can be used as a question pronoun, i.e. to represent the questioned content in a content question.
- **comparative concept vs. descriptive category:** In many languages, *wh*-words are also used as indefinite pronouns and/or as relative pronouns. Alternative terms such as *epistememe* (Durie 1985) and *ignorative* (Wierzbicka 1980) have therefore been proposed. Still, these fall under the above definition.

More comparative concepts (4): **relative clause**

- **generalization:** If the relative clause precedes the noun either as the only construction or as an alternate construction, either the language is postpositional, or the adjective precedes the noun or both (Greenberg 1963, universal 24).
- **definition:** A relative clause is a clause that is used to narrow the reference of a referential phrase and in which the referent of the phrase plays a semantic role.
- **comparative concept vs. descriptive category:** a construction does not have to be a “relative clause” in a particular language in order to qualify as a relative clause for this definition, cf. Japanese...

More comparative concepts (4): **relative clause**

Japanese

Matsumoto (1988), Comrie (1997, 1998):

- a. gakusei ga katta hon
[student NOM bought] book
'the book that the student bought'
- b. gakusei ga hon o katta zizitu
[student NOM book ACC bought] fact
'the fact that the student bought the book'
- c. dareka ga doa o tatoku oto
[someone NOM door ACC knock] sound
'the sound of someone knocking at the door'

Japanese **Noun-Modifying Constructions** (= *descriptive category*) fall under the comparative concept of *relative clause*.

More comparative concepts (5): **ergative case**

- **generalization:** Ergative case is always overtly marked (Dixon 1994).
- **definition:** An ergative case marker is a marker of an A-argument in a language in which the S-argument is not (always) coded in the same way as the A-argument.
- **definition 2:** The A-argument is the agent of a typical transitive clause such as ‘The farmer killed the duckling’.
- **comparative concept vs. descriptive category:** in Eskimo languages, the Relative case also expresses adnominal possession, but it still falls under this definition.

Objections (I)

(Frederick Newmeyer 2007, *Linguistic Typology*)



- Typology cannot be done on the basis of purely **semantic categories**.
- A lot of existing typological research makes use of **formal** cross-linguistic categories, e.g. in *WALS*:
 - Corbett, Number of Genders
 - Iggesen, Number of Cases
 - Siewierska, Passive Constructions
 - Comrie & Kuteva, Relative Clauses

Rebuttal (I)

- The typological research that Newmeyer cites is based on **comparative concepts**, not on **cross-linguistic formal categories**
- (But Newmeyer is right that typology is often based on formal concepts – comparative concepts often contain a formal component;

this was not sufficiently emphasized in Haspelmath 2007 and Croft 2001; but cf. already Haspelmath 1997:9)

Objections (2)

(Frederick Newmeyer 2007, *Linguistic Typology*)



- **Universal semantic categories** are no less problematic than cross-linguistic formal categories:
- Just as linguists do not agree on formal categories, they do not agree on semantic categories.
- Thus, reliance on meaning is no serious alternative to reliance on grammatical form.

Rebuttal (2)

- Universal semantic categories do not have the same problems as universal/cross-linguistic formal categories:
- As we saw, **the criteria for category assignment are language-specific**, so there is no rigorous method for matching formal categories across languages.
- The criteria for assigning forms to universal meanings are not language-specific; **there is a rigorous method for determining meaning** (= semantic analysis).
- Since translation is possible from any language to any other language, it is possible to determine an expression's meaning in universal terms.

Objections (3)

“Cross-linguistic comparison is [nowadays often] initiated independently or in advance of detailed analyses of individual systems. One approaches each language with, in effect, a checklist....

In principle, **comparisons must be based on analyses of particular systems:** but, in so many interesting cases, such analyses are lacking.”

Matthews, Peter H. 1997. Structural linguistics in the 1990s. *Lingua* 100:193-203.

“**[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...”

Newmeyer, Frederick J. 1998. *Language form and language function*.

Rebuttal (3)

- Cross-linguistic comparison **is not based on language-specific analysis** (i.e. on descriptive categories, or on cross-linguistic categories).
- Instead, it is based on **comparative concepts** and the process of matching them with the facts of the language.
- For this purpose, we need to know what the facts are, but not what the best analysis is.

cf. Haspelmath, M. 2004. Does linguistic explanation presuppose linguistic description?
Studies in Language 28: 554-579.

Objections (4)

(Stephen C. Levinson, p.c.)



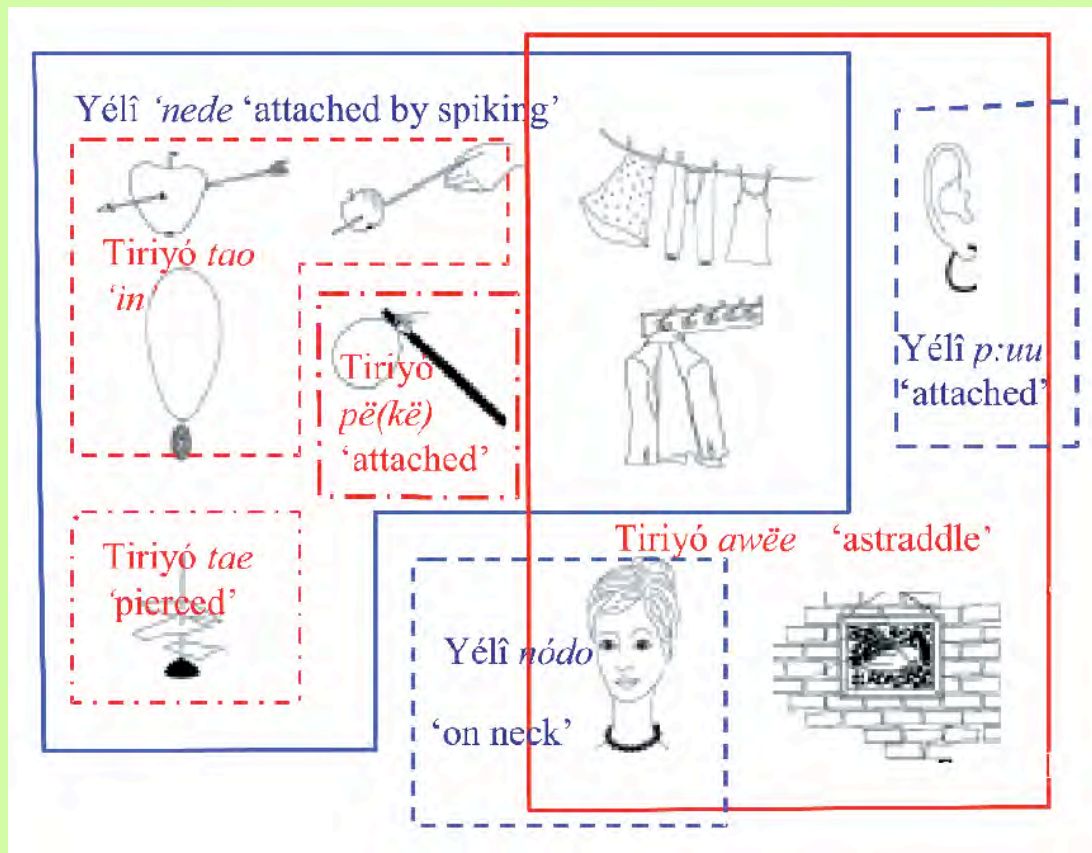
- **Universal semantic categories** do not exist;
- The possibility of true translation is illusory.
- Just as all languages have different formal categories, **all languages have different semantic categories**. There are no universal semantic primitives.
- Hence, typology cannot rely on semantics as a *tertium comparationis*.

Rebuttal (4)



- Cross-linguistic comparison can be based on non-linguistic *tertia comparationis*
- Linguists can construct **nonlinguistic stimuli** (pictures, videos, artificial social situations) and observe speakers' reaction in a systematic way

Nonlinguistic stimuli: an example



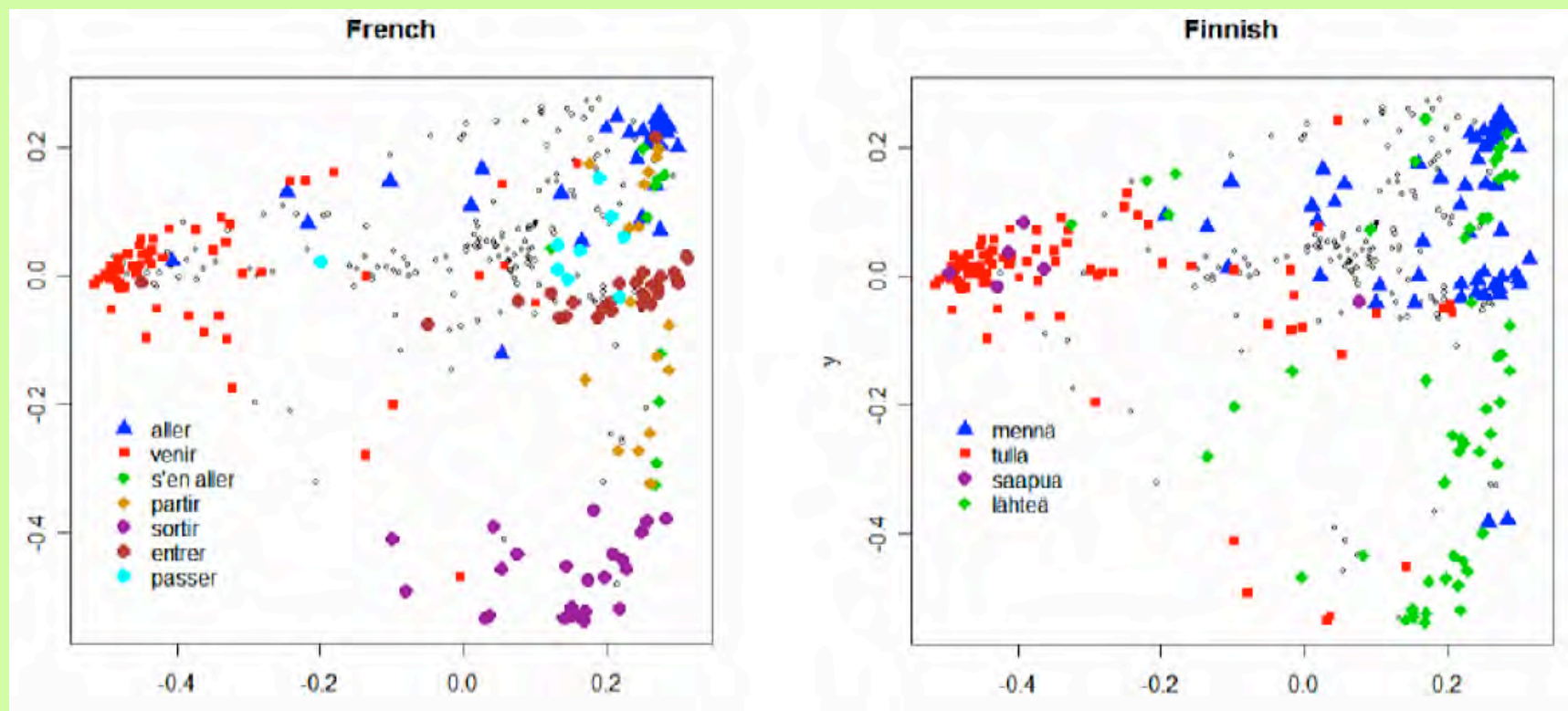
from:
Levinson &
Meira,
Language 2003

Four adpositions from Tiriyo, and three from Yelî

Nonconventional comparative concepts

- Conventional comparative concepts look a lot like descriptive categories, and the same kinds of technical terms are used for them (*adjective, ergative, relative clause,...*).
- But cross-linguistic comparison is also possible with **nonconventional comparative concepts**:
 - nonlinguistic stimuli (a la Berlin & Kay, Levinson and colleagues)
 - **tokens in parallel texts** (= texts that are translations of each other), cf. Bernhard Wälchli's work

Nonconventional comparative concepts: tokens of motion verbs in the Gospel according to Mark



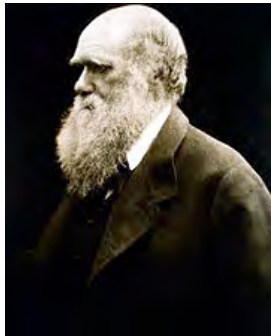
Wälchli, Bernhard. 2007. Constructing semantic maps from parallel text data. Ms., University of Berne.

Conclusion

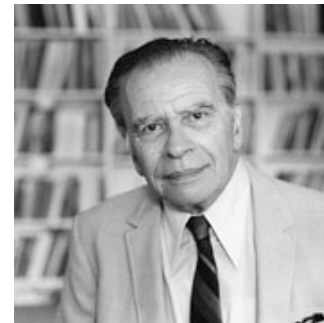
- Language typology is possible because we can construct **comparative concepts** that serve as *tertia comparationis*.
- These comparative concepts often look like **descriptive categories**, and have sometimes been confused with them.
- But comparative concepts make no claim about the best analysis of a language. **Typology can be done without deep linguistic analysis.**
- It would be nice if we could do typology and language-specific analysis at the same time, using **cross-linguistic categories**. But this is illusory.

Conclusion

- Language typology without deep linguistic analysis is somewhat like doing **comparative biology without genetics**.
- Both approaches have given rise to fruitful research results, and will continue to do so.



Charles Darwin, 1809-1882



Joseph H. Greenberg, 1915-2001

Thank you