Lexical typology: introduction

Maria Koptjevskaja Tamm

Dept. of linguistics, Stockholm university
tamm@ling.su.se,
http://www.ling.su.se/staff/tamm

1

Linguistic typology vs. linguistic universals

- > Theoretical questions I:
- What is universal / frequent and what is language particular in a specific phenomenon, what phenomena are frequent / rare?
- What generalizations can be made about attested vs. possible patterns?
- How can the attested cross-linguistic patterns / generalisations be explained?

3

Modern linguistic typology vs. linguistic universals

"[T]he study of linguistics patterns that are found cross-linguistically, in particular, patterns that can be discovered solely by cross-linguistic comparison" (Croft 1990: 1)

"In the past century, typology was mostly used an alternative method in pursuing one of the same goals as generative grammar: to determine the limits of possible human languages and, thereby, to contribute to a universal theory of grammar. The paradigm result was the absolute universal law that would rule out as linguistically impossible what would see logically imaginable, e.g., a language with a gender distinction exclusively in the 1st person singular" (Bickel 2007: 238).

2

The Universals Archive in Konstanz The Universals Archive Browse Search ← of 16 <<< >>> 517 (used to be 519 in the old version) Original If a language has gender distinctions in the 1st person, it always has gender distinctions in the 2nd or 3rd person, or in both. IF there are gender distinctions in the 1st person [of the personal pronoun], THEN there are gender distinctions in the 2nd and/or 3rd person. gender (1 person) → gender (2 &, V. 3 person) pronoun, gender, person, 1, 2, 3 inflection implication achronic absolute 30 languages of Greenberg 1963 sample Greenhern 1963: 96 #44 By Frans Plank 🖾 03.08.2006. 09:49 sy r-rais манк <u>ш</u>. In Thai (Daic, Austroasiatic) (Uspensky 1965: 214, Uspensky 1968: 11, Uspensky 1972: 63) gender (sex) can be distinguished only in the 1st person. Tocharian A (IE) has gender contrast in personal pronouns exclusively in 1st person singuity. By Frans Plank M 03.08.2006. 09:49 Supported by Corbett 1991: 131, on the basis of some 200 lgs.

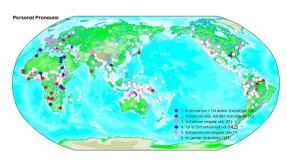
Linguistic typology as population science

"Over the past decade...[i]nstead of asking "what is possible?", more and more typologists ask "what's where why"? Asking "what's where?" target universal preferences as much as geographical or genealogical skewings, and results in probabilistic theories stated over properly sampled distributions. Asking "why?" is based on premises that (i) typological distributions are historically grown and (ii) that they are interrelated with other distributions" (Bickel 2007:238)

5

Gender distinctions in independent personal pronouns (Anna Siewierska)

Haspelmath, M., M. Dryer, D. Gil & B. Comrie (2005). The World Atlas of Language Structures (WALS). Oxford: Oxford University Press



7

Linguistic typology as population science

- > Theoretical questions II:
- Which phenomena are genetically stable and which are subject to contact-induced change?
- How are the various linguistic phenomena distributed across the world's languages?
- How can the attested distribution of the different patterns across languages be explained?

6

Kinds of typology

Grammatical typology, syntactic typology, morphological typology, phonetic typology, phonological typology...

Lexical typology

- the "characteristic ways in which language [...] packages semantic material into words" (Lehrer 1992: 249),
- typologically relevant features in the grammatical structure of the lexicon (Lehmann 1990: 163) / typologically relevant vs. language-specific patterns of lexicon-grammar interaction (Behrens & Sasse 1997).
- the cross-linguistic and typological dimension of lexicology (Maria Koptjevskaja Tamm 2007) [lexicology ≈ the characterization of words and vocabularies both as wholes and as units]

Lexical typology: major foci

- FOCUS 1 (onomasiology): what meanings can(not) be encoded as single word in different languages, what distinctions are made in lexical systems for encoding a particular cognitive domain and what factors underlie them?
- FOCUS 2 (semasiology): what different meanings can be expressed by one and the same lexeme or by lexemes synchronically and historically derived from each other (polysemy, semantic shifts etc.)?
- FOCUS 3: how does the lexicon interact with the grammar?

Illustrations: BODY, AQUA-Motion, TEMPERATURE

9

The general issues of the course

- 1. Questions asked in lexico-typological studies
- 2. Methods of data collection and their positive and negative sides
- 3. Cross-linguistic identification of the studied phenomena: how do we know that we compare like with like, and what actually counts as like for the purposes of a particular investigation (again, if possible, what is gained / lost in this approach)? What kinds of meaning does the investigation aim at: e.g., denotation / extension vs. sense / descriptive meaning / intension, vague / approximate vs. precise (incl. the problem of polysemy / semantic generality).

11

The structure of the course

- 1: Maria Koptjevskaja-Tamm (introduction)
- 2-4: Peter Koch (onomasiology, lexicon-grammar interactionh)
- 5-6: Martine Vanhove (semasiology)
- 7: Maria Koptjevskaja-Tamm (methodology, theoretical implications)
- 8: Martine Vanhove, Maria Koptjevskaja-Tamm (wrapping up)

We will regularly introduce small tasks that will be used for discussions in the

10

The general issues of the course (cont.)

- 4. Methods of data analysis, ways of representing meanings, incl. metalanguage used for semantic explications, and of formulating generalizations relevant for a particular investigation, with their positive and negative sides
- 5. Possible explanations for the attested phenomena/generalizations e.g., environmental (rooted in the properties of the real world), biological (shaped by human perceptual and cognitive predispositions or simply innate), social or cultural, historical, etc.
- Further theoretical contribution of lexical typology to both theoretical linguistics, and to other disciplines; interaction between lexical typology and related disciplines (cognitive science, computer sciences, language acquisition, anthropology etc.).

Case 1: BODY

The most universal, basic and crucial human domain. Extensive research:

Andersen 1978, Brown 2001, 2005 a/b, Wilkins 1996, ajid et al. (eds.) 2006,

ajid et al. (eds.) 2006, Koch (DECOLAR, LexiTypeDia, num. Publications) Enfield & Wierzbicka (eds.) 2002, Heine 1997



13

FOCUS 1: onomasiology (stratification of lexical fields)

English	Italian	Ruman.	Eston.	Japan. Russ	
hand	mano	mina	käsi	ude	ruka
arm	braccio	brat	käsi(vars) te	
foot	piede	picior	jalg	ashi	noga
leg	gamba				
finger	dito	deget	sõrm	yubi	palec
toe			varvas		

15

Lexico-typological questions wrt to the BODY

- Focus 1: What body-part concepts are encoded as words across languages, what distinctions are made in the systems of body-part terms and what factors underlie them?
- Focus 2: How are body-part concepts lexicalized across languages in terms of word classes? Are there morphological peculiarities characteristic for body-part terms? What syntactic constructions are used for talking about body parts?
- Focus 3: What are the possible extensions of body-part terms to other domains? Where from do the body-part terms come? How can their meanings change?

14

A few statistical generalisations (Brown 2001, 2005)

- In a sample of 118 languages, 49 languages (42%) link 'eye' and 'face' either through polysemy or overt marking.
- In a sample of 617 languages, 228 languages use the same word for 'hand'/'arm'
- In a sample of 593 languages, 72 languages use the same word for 'finger'/ 'hand'.

A few absolute generalisations

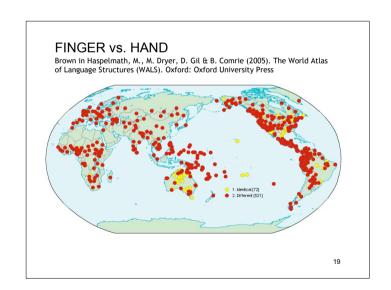
Andersen 1978, Brown 2001, Wilkins 1996

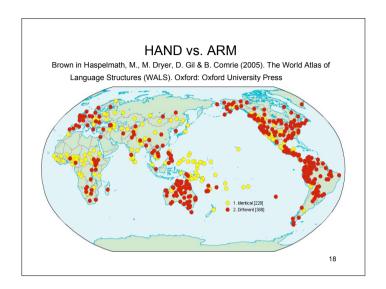
- If both hand and foot are labelled, they are labelled differently.
- If there is a distinct term for foot, then there will be a distinct term for hand
- There are never more than six levels of depth in the partonomy relating to body-part terminology.

Recent research on the BODY (Majid et al. 2006): many of the earlier generalizations have proved to be wrong.

Lavukaleve (Papua) – tau 'arm / leg', fe 'foot'

17





Which factors shape the body-term system in a language / across languages?

- Biology-rooted explanations: perceptual discontinuities, different functions (cf. Majid et al. 2006)
- Socio-cultural practices (partly related to climatic conditions, cf. Brown 2005)

FOCUS 2 (semasiology): Polysemy, semantic associations and semantic shifts with respect to the BODY

Semantic shift – a pair of meanings A and B linked by some genetic relation, either synchronically (e.g., via polysemy or derivation), or diachronically

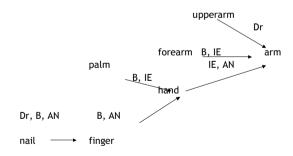
- Intrafield semantic shifts: A and B belong to the same semantic domain
- Interfield, or transfield semantic shift: A and B belong to different semantic domains

(Matisoff 1978, Wilkins 1996, DECOLAR, LexiTypeDia, Enfield & Wierzbicka (eds.) 2002, Heine 1997, Ojutkangas 2000, Kovács 2000, Schladt 2000)

21

23

Intrafield semantic shifts, body parts => body parts: visible parts and wholes (Wilkins 1996)



Also: Matisoff 1978, DECOLAR, LexiTypeDia

Deriving body-part terms, tendencies: Wilkins 1996

- A visible person-part term > the visible whole, but not vice versa
- A person-part term > a spatially contiguous person-part
- Terms for parts of the upper body <> terms for parts of the lower body
- An animal-part term > person-part
- · A term for verbal action > person-part involved in that action

22

Around the eye: LexiTypeDia (Tübingen, Koch

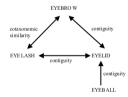


Figure 1. Attested semantic shifts within the frame EYE (Koch forthc.)

Transfield semantic shifts with body parts as source (mainly metaphor)

- 'Mapping' onto parts of other things
- Emotions
- · Spatial relations
- Numbers
- Possession
- · Reflexive-reciprocal-middles
- Etc.

25

Focus 3: Interaction between vocabulary and grammar and body-part terms

Body-part terms are normally nouns, but they can show various peculiarities:

- Special constructions for expressing possession (inalienable possession, constructions with external possessors, body-part-incorporation) – Chappell & McGregor 1996
- · Body-part terms and number
- · Body-part terms and gender

27

Transfield semantic shifts with body parts as source: universal vs. specific

- Emotions: e.g., 'heart', 'liver', 'stomach', 'throat'
- Spatial relations: e.g. anthropomorphic vs. zoomorphic patterns, 'head' vs. 'trunc' (49% each in Oceania, 38% vs. 60% in Africa)
- Australian Aboriginal languages: animals and plants named after their most salient part, e.g. 'tooth' => 'dog' and 'wild asparagus'

26

Case 2: Aqua-motion verbs





Majsak & Rakhilina eds. 2007, http://aquamotion.narod.ru; Koptjevskaja-Tamm, Divijak & Rakhilina 2010

Aqua-motion (motion and location of a nonliquid Figure in liquid medium)

Fairly universal and basic activity, but also dependent on the environmental factors and socio-cultural practices

Sub-domains of aqua-motion

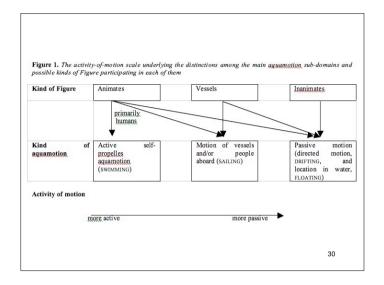
- SWIMMING prototypically self-propelled motion of animate figure
- SAILING prototypically controlled motion of vessels and people aboard
- DRIFTING motion of (in)animate figure driven by current
- FLOATING containment of (in)animate figure in confined space

29

Lexico-typological questions wrt to AQUAmotion

- Focus 1: What AQUA-motion concepts are encoded as words across languages, what distinctions are made in the systems of AQUA-motion terms and what factors underlie them?
- Focus 2: How are AQUA-motion concepts lexicalized across languages in terms of word classes? Are there morphological peculiarities characteristic for AQUA-motion terms? What syntactic constructions are used for talking about AQUA-motion?
- Focus 3: What are the possible extensions of AQUA-motion terms to other domains? Where from do the AQUA-motion terms come? How can their meanings change?

31



FOCUS 1: stratification of the AQUA-motion domain – types of system

Rich systems: more than three verbs (e.g. Swedish). Indonesian – 14 (!) verbs.

Middle systems (rare): two dedicated aqua-motion verbs for active vs.
passive motion and a general motion verb for motion on vessels
(Hindi, Persian, Maninka)

Poor systems: Russian (plyt'/plavat'), Polish, Bulgarian; Turkish

Table 1. The Swedish system of aguamotion verbs Active Motion of vessels and people aboard Passive motion motion of an Sailing Motor-driven animate boats vessels boats out of / Location Figure control simma segla (no specific flyta aquamotion (indirectly) (indirectly)



33

FOCUS 2 (semasiology): Polysemy, semantic associations and semantic shifts with respect to the AQUA-motion

Cross-linguistically recurrent "AQUA-motion" based metaphors:

SWIM => abundance, muchness (e.g., to "swim" in money, tears)
SAIL, FLOAT, DRIFT => non-inhibited motion (e.g., a political party is drifting in a certain direction)

FLOAT => absence of stability, formlessness (e.g., when I am reading this text, the letters are floating in my eyes)

35

Table 2. The English and Dutch systems of aquamotion verbs Active Motion of vessels and people aboard Passive motion Sailing Motor-Rowing Motion Neutral motion of an boats driven out of / Location boats animate vessels control Figure paddle drift float swim. sail row. (indirectly (indirectly) Dutch zwemmen. paddelen roeien drijven (indirectly) (indirectly)

34

FOCUS 3 (semasiology): lexicon-grammar interaction with respect to AQUA-motion

AQUA-motion is a typical activity (or state, in the case of FLOATING) and is normally encoded by verbs. A couple of interesting cases:

- SAILING verbs can take different subjects (humans or vessels) and behave both as intransitive and transitive verbs;
- In languages with poor lexical AQUA-motion systems (e.g., Russian) constructional distinctions can partly encode the differences among the different sub-domains

CASE 3: Linguistics of temperature

http://ling-asv.ling.su.se/mediawiki/index.php/Main_Page



Sutrop 1998, 1999; Plank 2003, Koptjevskaja-Tamm & Rakhilina 2006,

Koptjevskaja-Tamm & Sahlgren (work in progress) se/mediawiki/index.php/Main Page

37

Temperature perception in humans: physiology

- Temperature sensation/evaluation of the temperature of other entities, based on perception received by the skin, and
- thermal comfort, having to do with maintaining heat balance between the heat produced by metabolism and the heat lost or gained as a result of other physiological processes.

Different temperature scales with their own reference points:

- the physiological zero and the neutral zone ($\approx 31-36$ °C): subject to sustained thermal adaptation, and
- the comfort zone (3–4°C within 17.5°C and 31°C), dependent on the habitual temperature in the group's environment.

39

Linguistics of temperature

Sutrop 1998, 1999; Plank 2003, Koptjevskaja-Tamm & Rakhilina 2006, Koptjevskaja-Tamm & Sahlgren (work in progress). Why temperature?

- •Temperature phenomena are universal, relatively easily perceptible by humans and crucial for them:
- •Their conceptualisation involves a complex interplay between external reality, bodily experience and evaluation of the relevant properties with regard to their functions in the human life in a particular cultural setting. Koptjevskaja-Tamm & Rakhilina (2006): temperature terms are anthropocentrical. Goddard & Wierzbicka (2007): descriptors such as hot and cold, hard and soft, rough and smooth, and heavy and light, that attribute qualities to things, can be analysed as "touching something with a part of the body, feeling something in that part, knowing something about that thing because of it, and thinking about that thing in a certain way because of it."

Temperature perception vs. other physiological processes

- The same nociceptor, VR1 (sensory neurons involved in the detection of pain-producing stimuli) is activated by painful heat (≥45°C) and by chilli peppers (i.e., by capsaicin found in them) (Clapham 1997).
- Certain emotional responses involve temperature regulation: e.g. anger leads to increased body heat.

Subdomains within the linguistic temperature domain

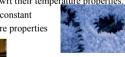
- EXPERIENCE-BASED (whereby an entity has a temperature that can be verified from "outside")
 - tactile-temperature, or "touch-temperature"
- (1) The floor / the kettle / the soup is COLD
 - non-tactile temperature: ambient (2) and clothing (3) temperature
- (2) It is cold here; the weather / Moscow /the winter is cold.
- (3) This sweater is cold.
- EXPERIENCER-BASED (whereby an entity, i.e. a living being, experiences a certain state):
 - personal-feeling temperature
- (4) I am cold (because the room is not heated, because I am ill...)

41

Temperature properties of entities: variable

Entities differ wrt their temperature properties

· some have constant temperature properties







- some have preferable temperature properties
- · some are primarily used for thermal comfort





• some have irrelevant temperature properties (e.g., books)

43

Physical environment: variable

Humans live under amazingly different climatic conditions and differ in their exposure to high vs. low temperatures, both under natural and human-made conditions (cf. a sauna, a refrigerator).





42

Social-cultural practices: variable

Cultures differ in their "temperature-related" practices and in the symbolic value of temperature evaluation:

- sauna once again
- a division of the world into 'hot' and 'cold' entities in Mesoamerica
- a division of consumables and body conditions into 'hot', 'cool', 'cold' etc. in South-East Asia

What's there to be studied wrt to linguistics of temperature?

- Focus 1: What temperature concepts are encoded as words across languages, what distinctions are made in the systems of temperature terms and what factors underlie them?
- Focus 2: How are temperature concepts lexicalized across languages in terms of word classes? What syntactic constructions are used for talking about temperature perception?
- Focus 3: What are the possible extensions of temperature terms to other domains? Where from do the temperature terms come? How can their meanings change?

45

FOCUS 1: Lexicalization of temperature concepts

- English, Russian, Swedish: elaborated systems with six or more terms (e.g., hot, warm, lukewarm, chilly, cool, cold; freeze...)
- Many languages have only two temperature terms e.g., the Oceanic languages spoken on Vanuatu (Alex François p.c.) – or three – e.g., Yucatec Maya (Olivier Le Guin p.c.).
- Even languages with a comparable degree of elaboration within their temperature systems can vary considerably as to what distinctions are relevant. E.g., what distinguishes 'warm' and 'hot'; whether there is a distinction between tactile and non-tactile perception, etc. (cf. Russian vs. Swedish, Koptjevskaja-Tamm & Rakhilina 2006)

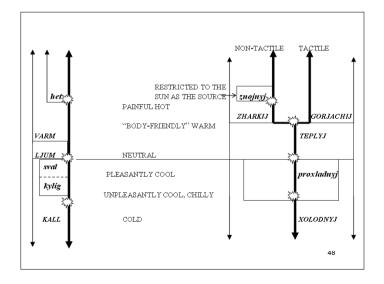
47

FOCUS 1: Lexicalization of temperature concepts, universal vs. language-specific

Do all languages lexicalize temperature concepts?

The null hypothesis: yes, but this has to be checked. However, we know already:

- Many (most?) languages lack the word for the functional concept 'temperature'
- Languages differ as to how many temperature terms they have and how these categorize the temperature domain in general.



Categorization within the temperature domain: talking about water in Ewe (Felix Ameka p.c.)

- (1) Tsi-a fa

 'The water is cool/cold'
- (2) Tsi-a gblo

'The water is lukewarm' (e.g., for medicinal purposes)

- (3) Tsi-a x3 dzo (e.g., for medicinal purposes)

 'The water is hot'
- (4) Tsi-a vé
 - 'The water is painfully hot'
- (5) Tsi-a fie (e.g., for bathing dead corpses)

'The water has boiled'

49

FOCUS 3: Lexicon-grammar interaction within the temperature domain

Languages vary considerably as to

- how temperature concepts are lexicalized across languages in terms of word classes (cf. hot - heat, cold - freeze), and
- what syntactic constructions are used for talking about temperature perception

51

FOCUS 2: Semantic shifts relevant for the temperature domain

- Temperature meanings are often semantically related to other meanings, either synchronically (within a polysemantic lexeme) or diachronically.
 - Temperature concepts are often used for talking about emotions ('hot temperament', 'warm feelings') and for referring to other perceptional modalities ('hot spices').
 - Temperature meanings can develop from others, e.g., 'burn, fire'
 'hot', or 'ice' > 'cold'.
 - The meanings of temperature terms can also change within the temperature domain itself (e.g., 'warm, hot' > 'lukewarm', as in Lat. tep- 'warm', Sanskrit tápas 'heat' vs. English tepid 'lukewarm', or Swed. sval 'cool' vs. German schwül 'stiffy, unpleasantly warm').

50

Distinguishing among the temperature subdomains: syntactic constructions vs. lexical means

- Syntactic constructions
 - English: The stones are cold, It is cold here, I am cold
 - German: Die Steine sind kalt, Es is kalt hier, Mir ist kalt
 - French: Les pierres sont froides, Il fait froid, J'ai froid
- Lexical means
 - English: I am freezing
 - German: Ich friere
- Syntactic constructions and lexical means
 - Japanese: Ishi ga (SUB) tsumetai, Kyô wa (TOP) samui, (Watashi wa ((TOP) samui

Which factors shape the linguistic temperature system in a language / across languages?

 Are there temperature-related "universals" that can be explained by neurophysiology and cognition, rather than by the more external and accidental factors?

53

Other possible generalizations: temperature words used in other domains

- Universal metaphors suggested within cognitive linguistics:
 - 'AFFECTION IS WARMTH' (Lakoff & Johnson 1997:50)
 - 'ANGER IS HEAT' (Kövecses 1995, also Goossens 1998; cf. also Shindo 1998-99).
- Cross-modal perception: 'hot day' vs. 'hot pepper'

55

Possible generalizations? Universal 1

Plank 2003: "The number of basic temperature terms a language can maximally have is probably quite limited. Probably there are only 2-term, 3-term, or 4-term systems of basic terms.

The **2-term** system only distinguishes **warm** and **cold**, as an equipollent opposition.

The **3-term** system distinguishes **warm** (pleasant for the human perceiver/experiencer, unmarked), **cold** (unpleasantly non-warm, marked relative to warm), and **hot** (unpleasantly, even dangerously very-warm, also marked, forming the opposite of cold in terms of extremes).

The **4-term** system adds a neutral term for the absence of either a pleasant or an unpleasant perception/experience of temperature, **luke**. Luke can probably not be added to equipollent 2-term systems."

54

Cross-linguistic differences

- Not all languages use temperature expressions in other domains.
 Oceanic languages (Alex François), Pirahā (Dan Everett), Nganasan (Valentin Goussev & Maria Brykina): no uses of temperature words in addition to the straightforward temperature situations.
- Languages differ as to whether they can use 'hot' for cross-modal perception: e.g. peppers can be sharp (Russian, German) and strong (Swedish) (cf. Rakova 2003).

Cross-linguistic differences (cont.)

- Extensions from the temperature domain are dependent on categorization within it. Palula (Henrik Liljegren & Naseem Haider): due to the lack of lexical elaboration within the warmer part of the temperature domain, both metaphors 'anger is heat' and 'affection is warmth' are realized by one and the same adjective taatu, cf. so taatu miish 'He is an angry person' (lit. 'He is a hot/warm person') vs. and taatu hiRu 'generous' (lit. 'a hot/warm heart').
- Languages can show "unexpected" extensions. Swahili (II'ja Grountov), Wolof (Michel-Loïc Perrin: 'nice' extensions from 'cool' rather than 'warm' (due to climatic conditions)

57

Which factors shape the linguistic temperature system in a language / across languages (cont.)?

- Language contact: cf. cool; hot line; 'cold' for 'slow' in several African languages (e.g., Bambara and Bozo – Mande, Cerma and Nateni – Gur)? The Komi-Zyryan tactile vs. nontactile distinction.
- Environmental (climatic) factors: can the paucity of the temperature terms in the Oceanic be related to the relatively narrow range of temperature fluctuations in the environment?
- Social-cultural practices: cf. the abundance of temperature terms for qualifying water in Ewe related to food preparation, bathing, washing, medicinal and ritual purposes.

59

Which factors shape the linguistic temperature system in a language / across languages?

• Genetic factors: cf. Hot, heiß, het; warm, varm; cold, kalt. Temperature terms are very stable! But even closely related languages can show amazing differences.

Cf. Icelandic: *heitur* has been generalized to cover almost the whole warming zone (the cognate of warm has been more or less lost)

Turkic: Uzbek sovuk 'cold; chilly, cool' vs. Turkish soğuk 'cold' /serin 'cool'

The tactile/ non-tactile distinction: Russian vs. the other Slavic; Komi Zyrjan (πöcь / жар) vs. Udmurt (πöcь)

58

Coda: more on lexical and semantic typology

The three foci do not exhaust all the possible issues within lexical typology. E.g., stability vs. borrowability of lexical concepts belongs also to lexical typology.

Semantic typology – "the systematic cross-linguistic study of how languages express meaning by way of signs" (Nick Evans, in press).

E.g., semantics of tense-aspect systems belongs to semantic typology but not to lexical typology.

Conversely, lexical typology is not necessarily restricted to semantic questions: e.g., cross-linguistic studies of "word".

Task 1

Imagine that you are going to conduct a lexico-typological study of (one of) the three domains introduced in the lecture. How will you attack the task? More precisely, what will you do for collecting data?