Word order typology and the case of a Uralic language

Erika Asztalos

09.12.2024

- Linguistic typology, word order typology, and the application of a theory on word order typology to a particular language: Udmurt
- Languages in the world
- Language families
- Linguistic typology
- Theories in word order typology
- Application of a theory to a language: Udmurt (the material of this last point is not needed for the exam)

How many languages are there in the world?

- **7164** languages are in use today (ethnologue.com)
- But this number is constantly in flux:
  - ➢ Languages may disappear (roughly **44%** of all languages are now **endangered**, with < 1000 users!)
  - ➢ The **distinction** between **'language'** and **'dialect'** in some cases is arbitrary:
    - ▪ e.g., certain Italian dialects diverge from Standard Italian so much that they could be considered as separate languages
    - ▪ Serbian and Croatian, Bosnian and Montenegrin are mutually intelligible; they are standardised varieties of Serbo-Croatian
  - Max Weinreich: **"A language is a dialect with an army and navy"**
  - ➔ social and political conditions can impact a community's perception of language and its status

# Top 10 most spoken languages in 2024
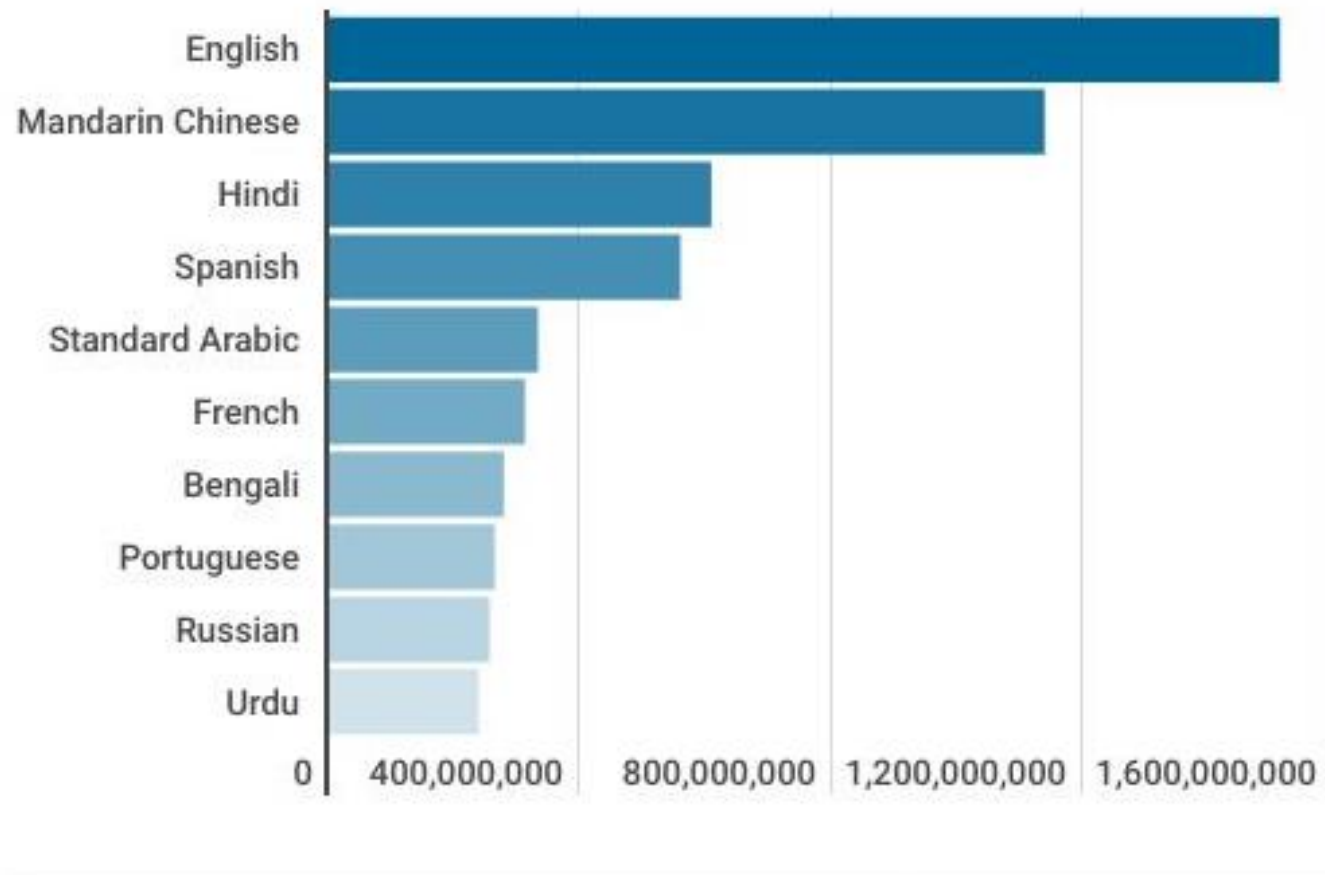
Languages differ but may also be similar to one another

- What can account for **grammatical similarities** between different languages?

  1. **Common origin** – historical descendence: when two or more languages belong to the **same language family**, e.g.: English and Dutch → *genealogical linguistics*

  2. **Typological similarities** – **structural** similarities may also exist between languages which do **not** belong to the same language family, e.g.: Hungarian and Turkish → *linguistic typology*

Language families and language isolates

- Languages that go back to a **common ancestor language** make up **language families**

- About **142** language families in the world

- Some language families in Europe:
  - **Indo-European**
  - **Uralic**

- Languages that do not belong to any language family are called **language isolates** → around 150-200 language isolates in the world. An example in Europe: Basque

The world's major language families
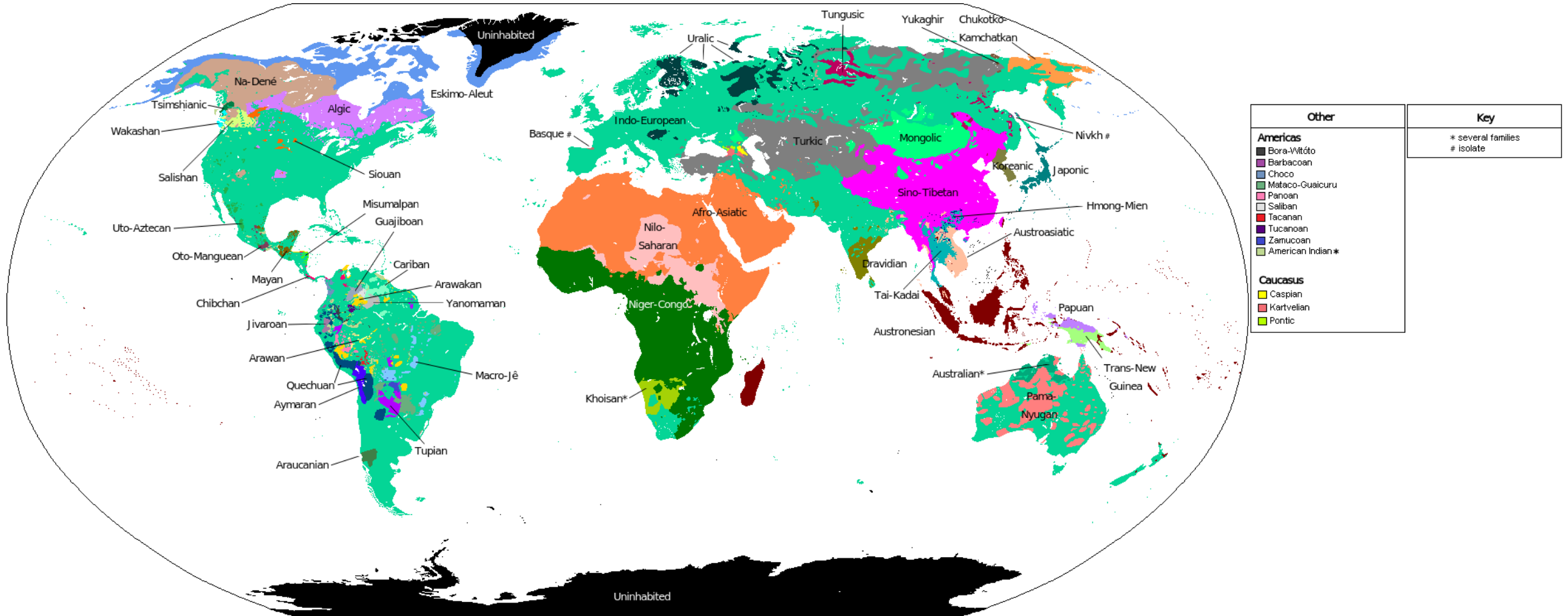
The Indo-European language family
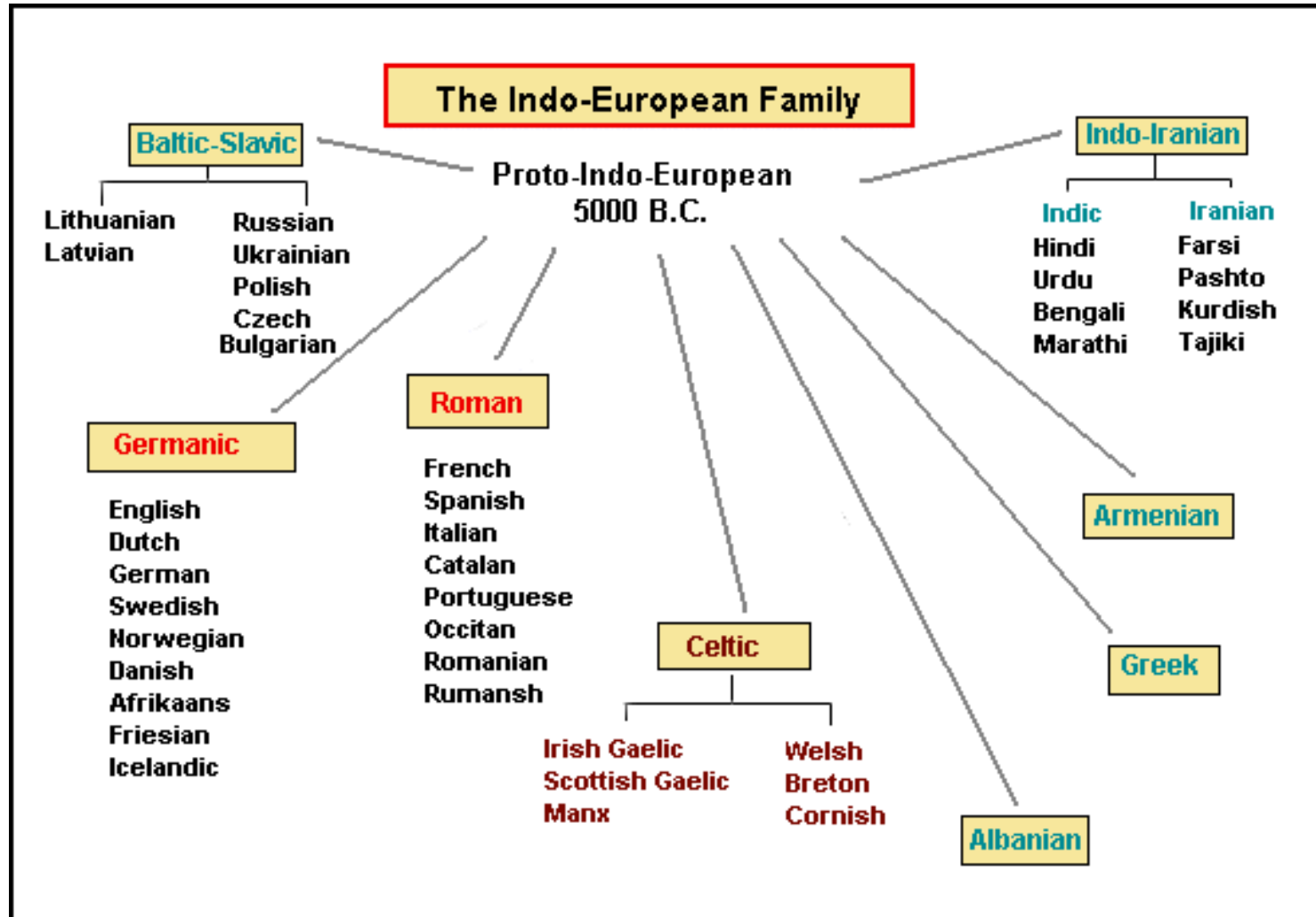
The Uralic language family
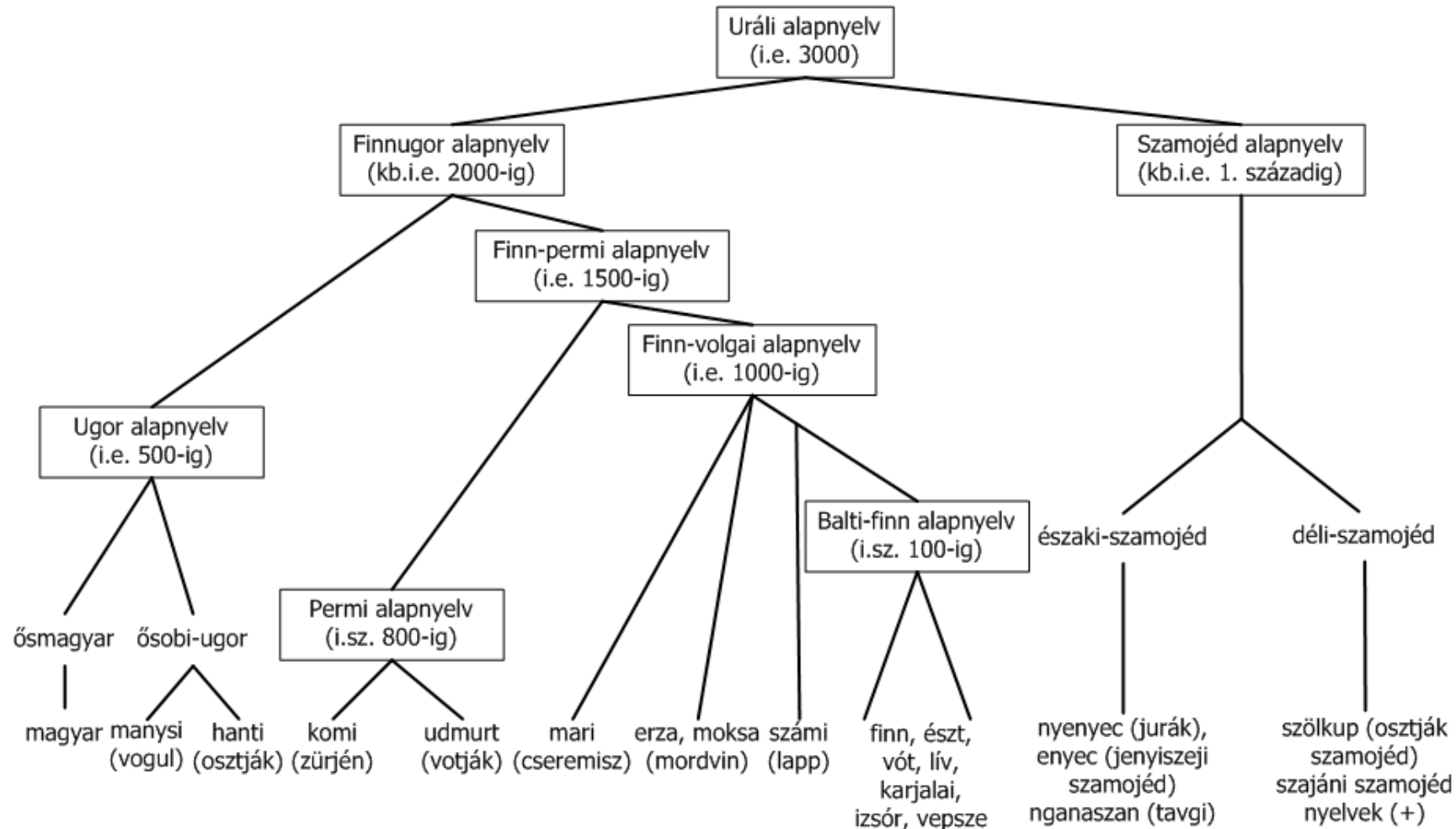
Linguistic typology

- Studies and classifies languages according to their **structural** features to allow their **comparison**
- Aim: to describe and explain the **structural diversity** and the **common properties** of the world's languages
- Contrasted with genealogical linguistics: typology groups languages based on **formal similarities rather than historic descendence**

Outline

1. Word order typology
   1. Greenberg (1963) and his word order correlations
   2. Dryer's (1992) *correlation pairs*
   3. The concept of basic word order

2. Udmurt: an ongoing (S)OV > (S)VO change?

Greenberg (1963): Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements

Relevance

- typology before Greenberg: classifying languages into morphological types: agglutinative/fusional/isolating
- Greenberg → linguistic **universals** – patterns that are potentially true for all languages of the world → revolutionising linguistic typology
- object of his study: 30 languages, roughly representing areally + geneologically the world's languages

Greenberg's (1963) language sample

Europe:
- Basque
- Serbian
- Welsh
- Norwegian
- Modern Greek
- Italian
- Finnish

America:
- Loritja
- Maya
- Zapotec
- Quechua
- Chibcha
- Guarani

Africa:
- Yoruba
- Nubian
- Swahili
- Fulani
- Masai
- Songhai
- Berber

Asia:
- Turkish
- Hebrew
- Burushaski
- Hindi
- Kannada
- Japanese,
- Thai
- Burmese
- Malay

Oceania:
- Maori

Greenberg's (1963) linguistic universals

- **45 universals** (mainly, but not only on word order) – many of them held true also in the light of more extensive later studies

- A morphological one:

Universal 36. *If a language has the category of **gender**, it always has the category of **number**.*

- *gender: masculin/feminin(/neuter)*, c.f., German, Latin, French, Italian, Spanish, Russian etc.
- *number: singular/plural*

Language types in Greenberg (1963)

1. Based on the basic (neutral) order of **subject (S)**, **verb (V)**, and **object (O)**:

   1. **SOV:** e.g., Nepali:

      Ma   *kitāb*   *paḍhchhu.*

      I   book   read

      SUBJECT   OBJECT   VERB

      'I read a book.'

Language types in Greenberg (1963)

2. SVO: e.g., English:

Frank **ate** **oranges.**

SUBJECT VERB OBJECT

3. VSO: e.g., Irish:

Itheann **Seán** **arán.**

eats Séan bread

VERB SUBJECT OBJECT

'Sean eats bread.'

\+ VOS
OSV
OVS
languages also exist,
but they are very rare

World Atlas of Language Structures (wals.info)
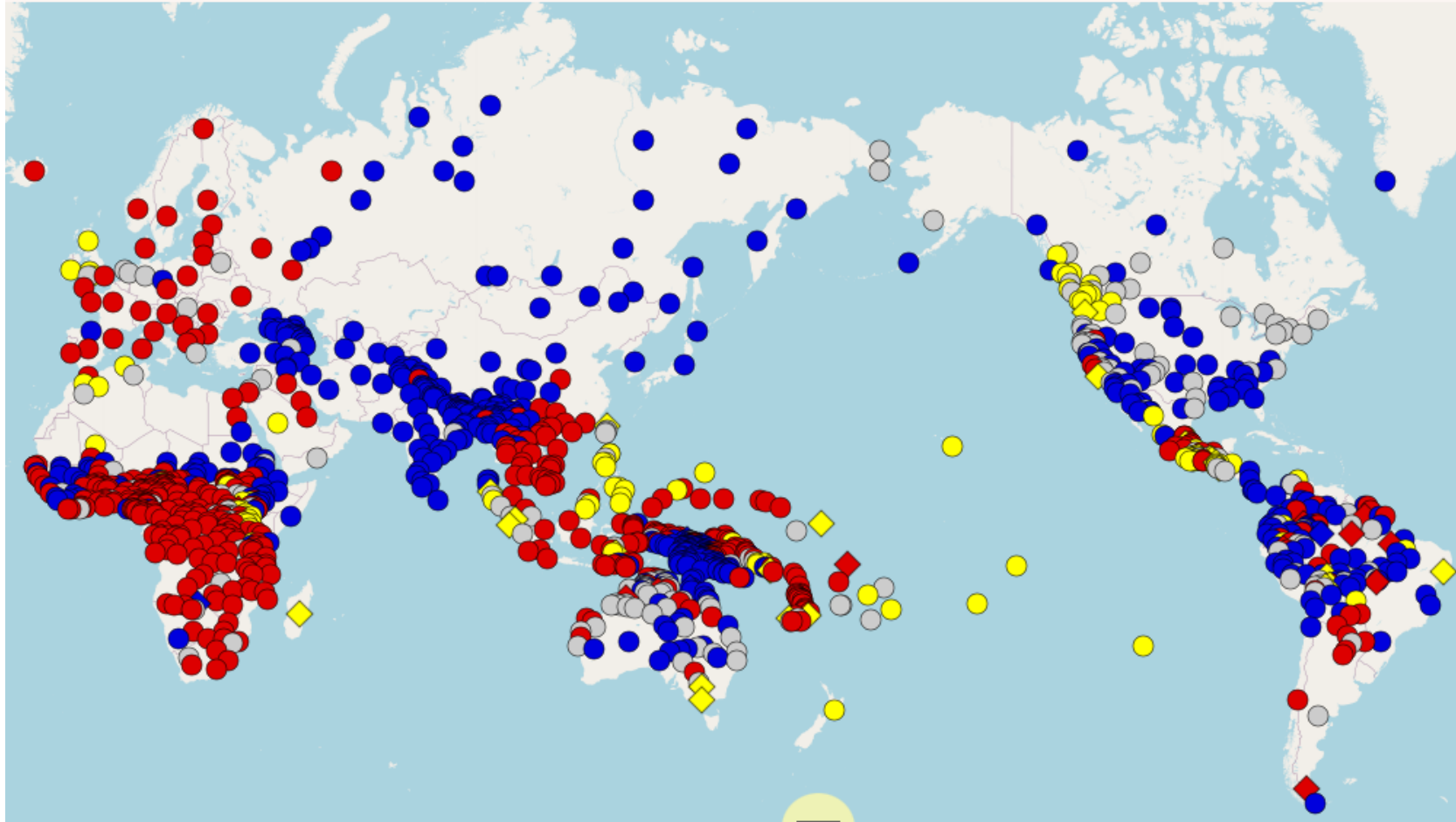
Value		Representation
●	Subject-object-verb (SOV)	564
●	Subject-verb-object (SVO)	488
●	Verb-subject-object (VSO)	95
◆	Verb-object-subject (VOS)	25
◆	Object-verb-subject (OVS)	11
◆	Object-subject-verb (OSV)	4
●	Lacking a dominant word order	189
	Total:	1376

Language types in Greenberg (1963)

2. Based on the relative order of **nouns** + **adpositions**
   → 2 types: **Prep N** / **N Postp**:

   1. *preposition – noun*
      after dinner (English)

   2. *noun – postposition*
      vacsora után (Hungarian)

Language types in Greenberg (1963)

3. Based on the relative order of **nouns** + **adjectives**

   → 2 types: **N Adj / Adj N:**

   1. *noun – adjective:*

      acqua calda (Italian)

   2. *adjective – noun:*

      hot water (English)

Some other examined parameters: the relative order of

- verb + adverb: *ran + slowly*
- auxiliary + verb: *has + run*
- nominal possessee („genitive") + possessor: *father + of John*
- demonstrative pronoun + noun: *that + man*
- numeral + noun: *five + cats*
- relative clause + noun: *movies + that we saw*
- position of question particles in a sentence
- position of interrogative words in a sentence: *What + did you eat?*

Linguistic universals (Greenberg 1963)

- Linguistic universals tell us about the existing combinations (the so-called correlations) between the possible values for these word order parameters

Types of universals (Greenberg 1963)

1. **SCOPE:**
   a) **absolute** universals: no exceptions:
      Universal 3: *Languages with dominant VSO order are always prepositional.*

   b) **statistical** universals: tendencies – they allow for some exceptions:
      Universal 5: *If a language has dominant SOV order and the genitive follows the governing noun, then the adjective **likewise** follows the noun.*

Types of universals (Greenberg 1963)

2. LOGICAL FORM:

a) **unrestricted** universals: form "All (or almost all) languages have X":

Universal 1: *In declarative sentences with nominal subject and object, the dominant order is almost always one in which **the subject precedes the object***.

Types of universals

b) implicational universals:

<u>Universal 3</u>: *Languages with dominant VSO order are* **always** *prepositional.*

- logical formula: p → q (*VSO → prepositions*)
- ! unidirectional correlations: ≠ q → p

The OV-VO typology

Dryer (1992)

- Data: database containing **625** languages (/a 543-language subset) representing **areally** and **geneologically** the world's languages

   → What are those pairs of elements whose order **statistically** correlates with that of **V** and **O**?

   → **Correlation pairs:** "If a pair of elements **X** and **Y** is such that

   X tends to precede **Y** **significantly more often** in **VO** languages

   than in **OV** languages, then

   <X, Y> is a correlation pair, and

   X is a **verb** patterner and

   Y an **object** patterner with respect to this pair"

Correlation pairs (Dryer 1992: 108)

VERB PATTERNER	OBJECT PATTERNER	EXAMPLE
verb	object	*ate + the sandwich*
verb	subject	*(there) entered + a tall man*
adposition	NP	*on + the table*
copula verb	predicate	*is + a teacher*
'want'	VP	*wants + to see Mary*
tense/aspect auxiliary verb	VP	*has + eaten dinner*
negative auxiliary	VP	cf. 7 in §4.2
complementizer	S	*that + John is sick*
question particle	S	cf. 8 in §4.4.
adverbial subordinator	S	*because + Bob has left*
article	N'	*the + tall man*
plural word	N'	cf. 9 in §4.7
noun	genitive	*father + of John*
noun	relative clause	*movies + that we saw*
adjective	standard of comparison	*taller + than Bob*
verb	PP	*slept + on the floor*
verb	manner adverb	*ran + slowly*

TABLE 39. Complete list of correlation pairs.

Non-correlation pairs (Dryer 1992: 108)

DEPENDENT	HEAD	EXAMPLE
adjective	noun	*tall + man*
demonstrative	noun	*that + man*
intensifier	adjective	*very + tall*
negative particle	verb	*not + go*
tense/aspect particle	verb	cf. examples in 6 in §3.5

TABLE 40. Noncorrelation pairs.

(Dryer 1992: 108)

The Branching Direction Theory (Dryer 1992)

- Verb patterners are **heads** and object patterners are phrasal (branching) categories (**dependents**):
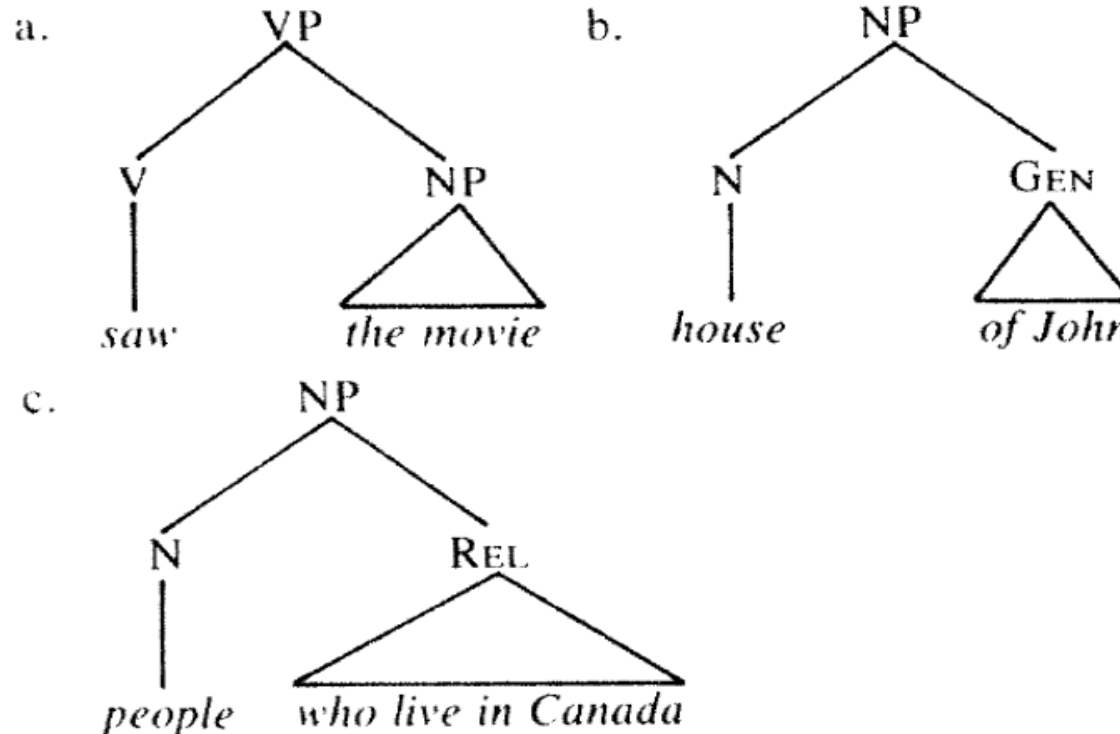
(Dryer 1992: 90)

The OV-VO typology in a nutshell

- Roughly:
  - **OV** (S**OV**, **OV**S, **O**S**V**) languages have **head-final** phrases: **dependent**-**head** order at the phrasal level

  - **VO** (S**VO**, **V**S**O**, **VO**S) languages have **head-initial** phrases: **head**-**dependent** order at the phrasal level

The notion of basic word order

Basic word order

- Word order **correlations** exist between so-called **basic** (or **neutral**) **word order** variants

- What does that mean?

- Roughly speaking:
  - Languages can have **more than one word order** both at the **sentence-level** and at the **phrasal level**. However, one of the possible word order variants is more **basic** (or **neutral**) than the others (and the other orders are marked), and word order **correlations** exist between these basic (neutral) word order variants.

    e.g., in English, **OSV** is **marked**:   *That book, I really liked.*

                                                         O   S   V

              SVO is **basic/neutral**:   *I really liked that book.*

                                                         S   V   O

Flexible word order languages also have one word order which is more basic than the others

Hungarian: all 6 permutations of **S, V** and **O** are grammatical, but **SVO** is the neutral (basic) order:

SVO: *Áron megnézte a filmet.* 'Áron watched the movie.'

SOV: *Áron a filmet megnézte.* 'The movie, Áron did watch it.'

OVS: *A filmet megnézte Áron.* 'The movie, Áron did watch it.'

OSV: *A filmet Áron megnézte.* 'The movie, Áron did watch it.'

VSO: *Megnézte Áron a filmet.* 'Áron did watch the movie.'

VOS: *Megnézte a filmet Áron.* 'Áron did watch the movie.'

Basic word order

- The word order of "stylistically neutral, independent, indicative clauses with full noun phrase (NP) participants, where the subject is definite, agentive and human, the object is a definite semantic patient, and the verb represents an action, not a state or an event" (Siewierska 1988: 8) → basic word order **at the sentence level**

  Barbara hit Matthew. (SVO)

- basic word order also exists at the **phrasal** level

- Criteria:
  1. Unmarkedness
  2. **Pragmatic neutrality**
  3. **Frequency**

Pragmatic neutrality

- neutral **context**, or

- neutral **information structure** of the sentence → no focused element in the sentence, and the topic is the subject (or no topic):

That book, **I liked**. → **OSV** → O is topicalised → not neutral

→ possible syntactic test for eliciting pragmatically neutral sentences (so-called *all-new* sentences): **- What's new? / What happened?**

- What's new? / What happened?

- Sally left New York. → SVO is the neutral order in English

Frequency

- Greenberg (1963: 67), Givón (1979: 50), Brody (1984: 717): basic word order = the most frequent order

a) **textual/statistical frequency** → a big quantity of texts of different genres is needed

b) **frequency in the grammatical system** (Hawkins 1983: 13): e.g., the number of adposition lexemes (postpositions outnumber prepositions in Finnish)

Udmurt: an ongoing (S)OV > (S)VO change?

The Uralic language family
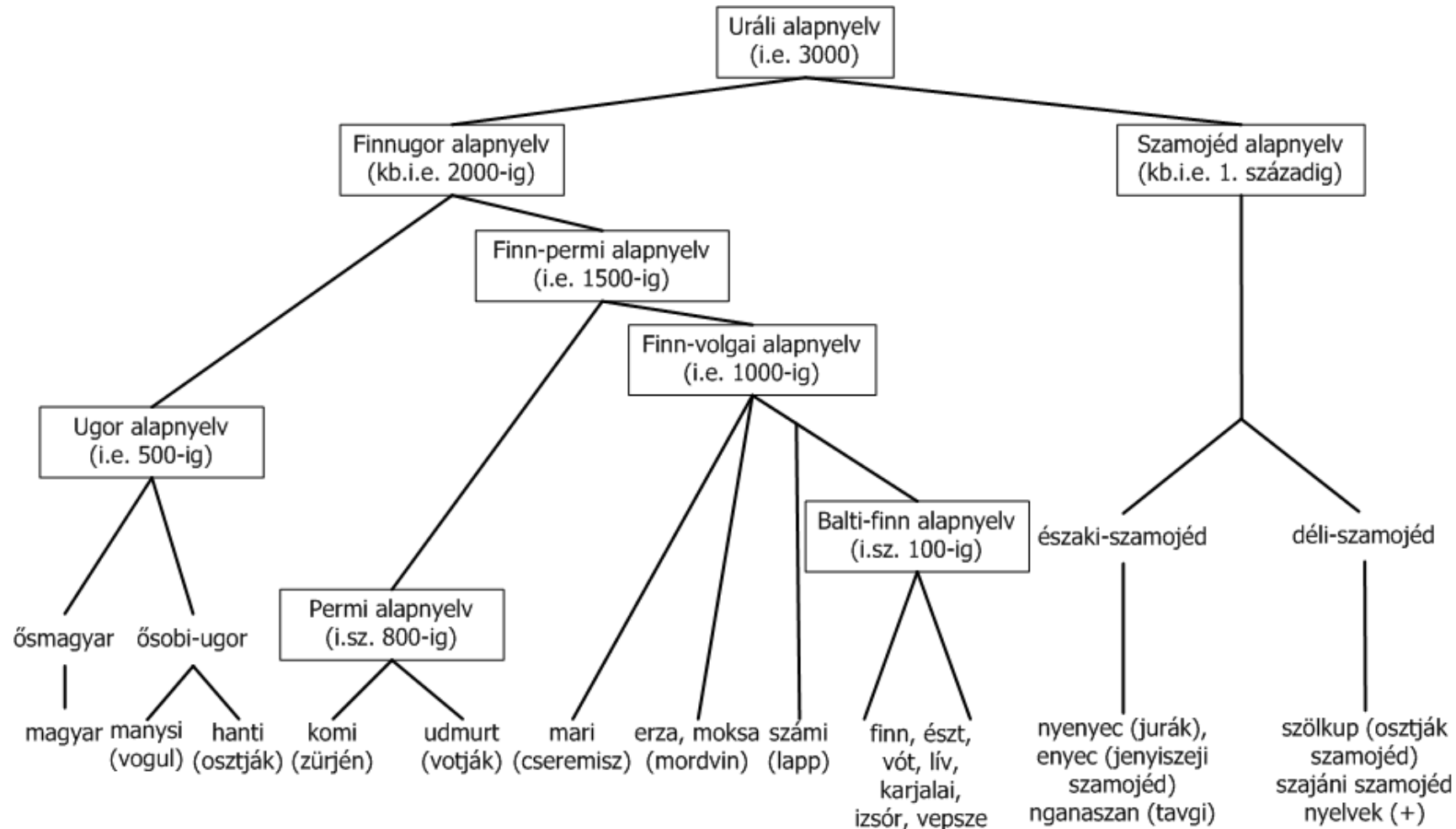
Uralic languages in the light of word order typology

- Uralic protolanguage (the common ancestor language of all Uralic languages): assumed to have been

  (S)OV: *Kimi*   *fish*   *eats*.

          Subject   **Object**   **Verb**

  with head-final (*dependent*–*head* ordered) phrases (Bereczki 2003)

- Contemporary Uralic languages:
  - **(S)OV:** Khanty, Mansi, Mari, Samoyedic
  - **(S)OV > (S)VO** change (due to the influence of SVO languages): Hungarian, Finnish, Komi, Mordvin, Saami

The Udmurt language

- Uralic / Finno-Ugric
- 255 877 native speakers in Russia (2021)
- **minority** language, strong Russian influence
- **bi-** and **trilingual** speakers (all Udmurts speak also **Russian** at a (near-) native level; Southern parts: they speak **Tatar** as well)

Word order in Udmurt according to previous studies

- Similarly to Hungarian, **all 6** permutations of S, O and V are **grammatical** (SOV, SVO, OSV, OVS, VSO, VOS)

- But **SOV** has been considered as the basic/neutral word order in most works, while other permutations, including SVO, have been considered as pragmatically marked (e.g., Zhuikov 1937, Bulychov 1947, Gavrilova 1970, Suihkonen 1990, Csúcs 1990, Timerkhanova 2011, Vilkuna 1998, Winkler 2011)

→ ... thus, we could maybe finish our class at this point?

Word order in contemporary Udmurt

- No! Actually, the picture is more complicated:
  - Frequency: (S)VO order is not rare in Udmurt (Salánki 2007)
  - (S)VO order can be both pragmatically marked and neutral (Ponariadov 2010, Asztalos & Tánczos 2014, Asztalos 2016)

→ An ongoing SOV > SVO change in contemporary Udmurt?

Sociolinguistic and areal background

- Udmurts: mostly bilingual speakers (Udmurt–Russian):
  - *old speakers*: balanced or Udmurt-dominant bilingualism
  - *young speakers*: balanced or **Russian-dominant** bilingualism (Salánki 2007: 59)
    - → influence of Russian on the Udmurt spoken by young speakers (**Russian: (S)VO + head-initial**)

- **Southern** areas: **Tatar** is also spoken (a Turkic language which is (**S)OV + head-final**) → influence of SOV Tatar on Udmurt in these areas

The Udmurt Republic

Research questions and methods

- Can we assume that Udmurt is shifting **from (S)OV to (S)VO (and from the head-final to the head-initial) type under the influence of Russian?**
  - ➢ **How frequent** are SOV and SVO orders and the head-final and the head-initial variants of the same phrase **in the same context**?

    (→ *quantitative* method)
  - ➢ Can (S)VO and head-initial phrases occur in **neutral** sentences?
    (→ *qualitative* aspect)
  - ➢ Is there a difference in terms of word order preferences of
    - **young** vs. **old** speakers of Udmurt
    - Udmurts living in **Udmurtia** vs. Udmurts living in **Tatarstan**?

- The examined constructions: Dryer's (1992) *correlation pairs*

Data 1: The questionnaire and the informants

- fieldwork (2014–2015)
- 90 respondents (all native speakers of Udmurt), from all main dialectal groups

  → grouped into: 1. respondents living **in Udmurtia**

  2. **in Tatarstan**

  → age groups: 1. born between 1935–1965

  2. 1970 and 2002

- **neutral** sentences → context: '*What happened?*'
- non-representative survey

Question types

- **closed-ended questions** (all questions were given in Udmurt)**:**

1. completing sentences by ordering and conjugating/declining given words:
   - *What's new?*
   - *Nothing interesting. Yesterday _____ (soup, Mary, to cook).*

2. grammaticality judgement about head-initial phrases:

   - *What's new?*

   - *Georgy fell off a ladder.*   *a) correct*   *b) not really correct*   *c) incorrect*

3. grammaticality judgement about both the **head-initial** and the **head-final** variants of the same sentence (**comparison**)

- **open-ended question:** writing a few sentences about a picture

Data 2: Textual analysis

- % of SOV and SVO sentences, and the head-initial vs. head-final variants in
    - old folklore texts from the end of the **19th cent.** vs. blog posts
      from **2012-15**
    - old newspaper articles from **1924 vs. contemporary**
      newspaper articles

- information structural analysis of SVO sentences and of sentences containing head-initial phrases
    → only pragmatically marked, or also neutral?

Results

Textual analysis

- SVO order and head-initial variants of certain phrases: higher frequency in contemporary texts than in old ones

Textual analysis

- VO sentences can also be neutral:

Odig   džyte   öťi   kollegaosme
one   evening   I.invited   my colleagues
doram   kunoje.
to my place   to guest
'One evening I invited my colleagues to my place.'

Questionnaire: Head-initial vs. head-final

- Although SOV and **head-final** orders turned out to be still **more frequent** and more acceptable than SVO and head-initial ones,

- **SVO** and **head-initial** variants were also produced and judged grammatical by a part of the informants, and

- **some speakers** sometimes **preferred the head-initial** variants over the head-final ones

Generational and areal differences 1.

- **younger** speakers produced **more frequently** and judged more favourably SVO and the head-initial variants than old speakers

  → *apparent time-hypothesis*: an age-stratified variation

    can be the sign of a linguistic change in progress (cf.

    Labov 1963, Trudgill 1992)

- speakers **from Udmurtia** produced **more frequently** and judged more favourably the **head-initial** variants than speakers of Udmurt from Tatarstan

Generational and areal differences 2.

→ *old speakers from Tatarstan*: almost exceptionless preference for SOV and head-initial variants

→ *younger ones from Udmurtia*: the highest % of production and acceptance of the head-initial variants

Verb + postpositional phrase – areal + generational difference

Mon   **śulmaśkiśko**   **D'ima**   **śaryś.**

I   worry   Dima   about

'I am worried about Dima.'

Verb + Subject (existential sentences) – areal difference

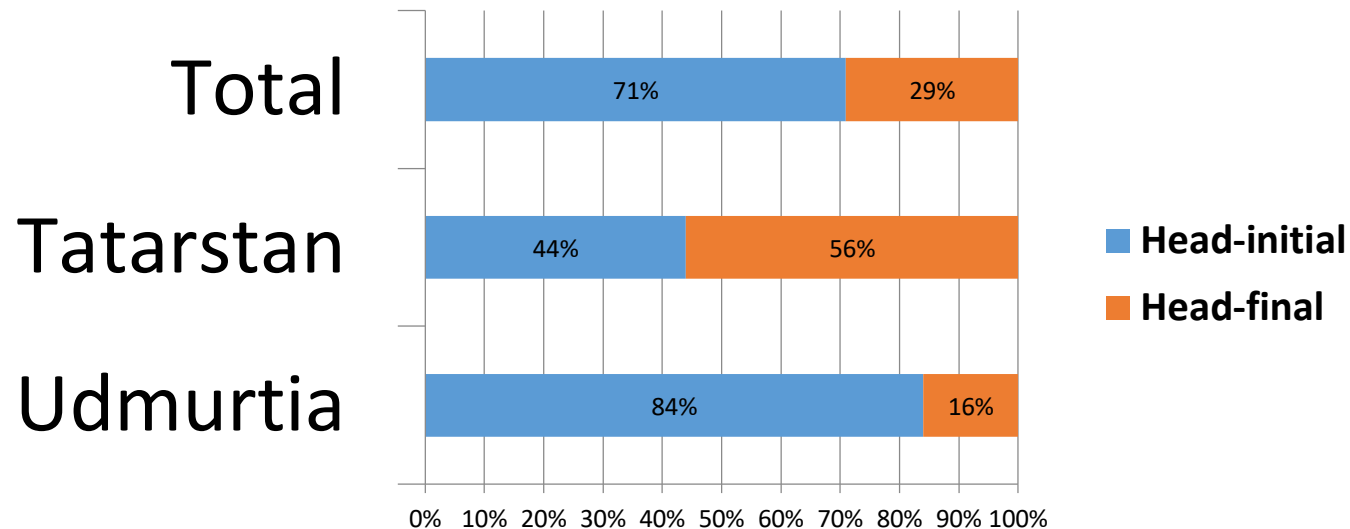Context: 'Our village is big and nice.'

(1) a. ***Otyn*** *vań* *kinoťeatr* *no* *klub*.

       there   is   cinema   and   clubhouse

   b. ***Otyn*** *kinoťeatr* *no* *klub* *vań*.

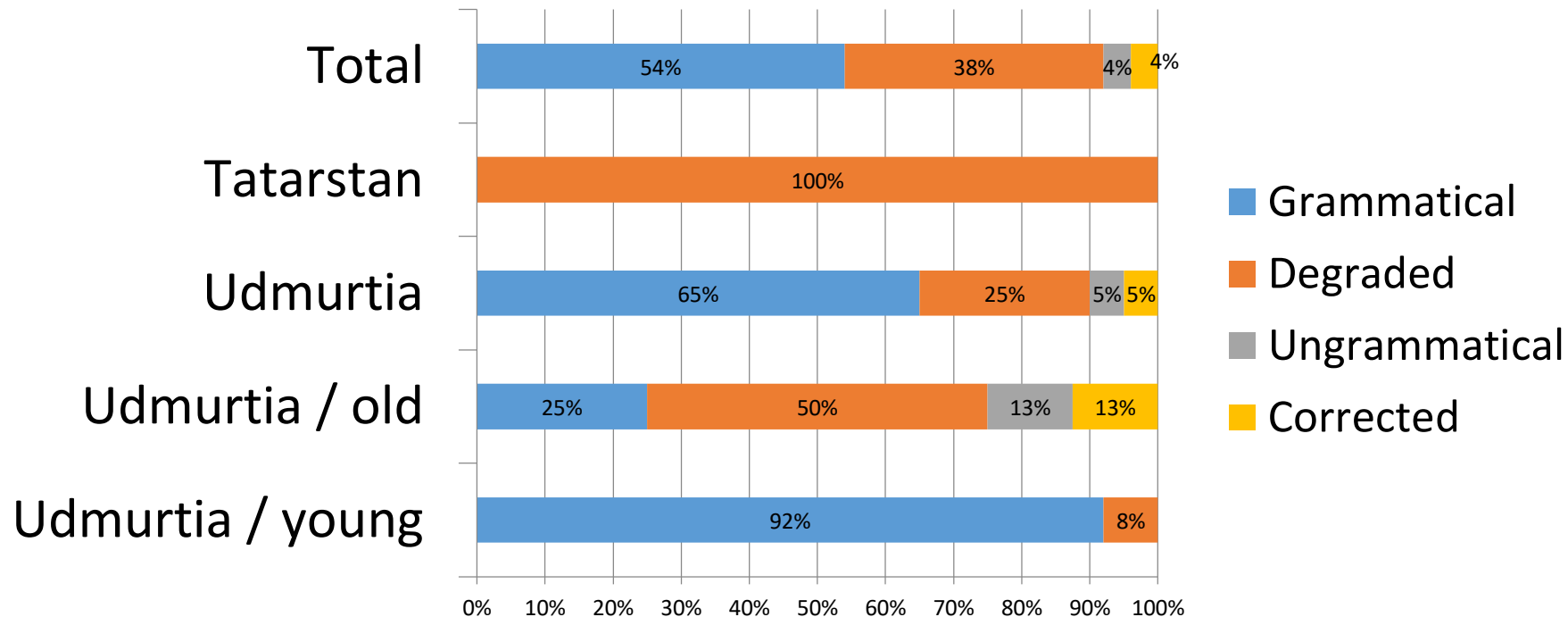   'There is a cinema and a clubhouse in there.'

	Head-initial	Head-final
Total	71%	29%
Tatarstan	44%	56%
Udmurtia	84%	16%

Auxiliary + Verb phrase – areal + generational difference

Jegitjoslen **potiz** **šuldyrjaśkemzy** *(…)*

of_youngs wanted having_fun

'Young people wanted to have fun (…)' (and they went to the disco.)

Conclusions

- Although (S)OV and head-final orders are still more frequent in Udmurt, but
- **(S)VO** and **head-initial** orders can also be used in **neutral** contexts
- **generational differences** → **ongoing typological shift** from the (S)OV to the (S)VO (and from head-final to head-initial) type
- due to the **influence of Russian** (which is **(S)VO + head-initial**) – a part of the younger speakers are Russian-dominant bilinguals
- **influence of the Tatar** language (which is **(S)OV + head-final**) on the Udmurt spoken in Tatarstan → slows down the change in trilingual areas

Thank you for your attention!

The fun part: Udmurt pop music :)!

- Despacito in Udmurt:
  https://www.youtube.com/watch?v=AGu2vhfbTUE

- Udmurt rap:
  https://www.youtube.com/watch?v=Y3ILJJXA8XY

References

- Asztalos, Erika 2018. *Szórendi típusváltás az udmurt nyelvben*. Doctoral dissertation. Eötvös Loránd University.

- Asztalos, Erika & Tánczos, Orsolya 2014. *Competing Grammars in nowadays Udmurt.* Conference presentation. 7th Budapest Uralic Workshop. 2014. 02. 03–05.

- Asztalos, Erika 2016. A fejvégű grammatikától a fejkezdetű felé: generációs különbségek a mai udmurt beszélőközösségben a szórendhasználat és -megítélés terén. In É. Kiss Katalin – Hegedűs Attila – Pintér Lilla (ed.) *Nyelvelmélet és kontaktológia 3.* Szent István Társulat. Budapest–Piliscsaba. 126–156.

- Bereczki, Gábor 2003. A magyar nyelv finnugor alapjai. Universitas Könyvkiadó. Budapest.

- Brody, Jill 1984. Some problems with the concept of basic word order. *Linguistics* 22. 711–736. Croft, William 1990. *Typology and Universals*. Cambridge: Cambridge University Press.

- Bulychov, M. N. 1947. *Poriadok slov v udmurtskom predlozhenii*. Udmurtgosizdat. Izhevsk.

- Csúcs, Sándor 1990. *Chrestomathia Votiacica*. Tankönyvkiadó. Budapest.

- Dryer, Matthew S. 1992. The Greenbergian word order correlations. Language 68. 81–138.

- Dryer, Matthew S. 1995. Frequency and pragmatically unmarked word order. In Downing, Pamela A. –Noonan, Michael (szerk.), *Word Order in Discourse* (Typological Studies in Language 30), 105–35. Amsterdam/Philadelphia: John Benjamins Publishing Conmpany.

- Gavrilova, T. G. 1970. *Poriadok slov v udmurtskom prostom povestvovatelnom predlozhenii*. Zapiski Udmurtkogo NII istorii, ekonomiki, literatury i jazyka pri Sovete Ministrov Udmurtskoj ASSR. Izhevsk.

- Givón, Talmy 1979. *On Understanding Grammar*. New York: Academic Press.

- Greenberg, Joseph Harold 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In: Greenberg, Joseph Harold (ed.): Universals of Language. MIT Press. Cambridge, MA. 58–90.

References 2.

- Ponarjadov, V. V. 2010. Poriadok slov v permkikh jazykakh v sravnitelno-tipologicheskom osveshchenii (prostoe predlozhenie). Syktyvkar.

- Salánki, Zsuzsanna 2007. Az udmurt nyelv mai helyzete. Doktori disszertáció. Kézirat. Eötvös Loránd Tudományegyetem. Budapest.

- Siewierska, Anna 1988. *Word Order Rules*. London: Croom Helm.

- Song, Jae Jung 2001. *Linguistic Typology*. *Morphology and Syntax* (Longman Linguistics Library). Harlow and London: Pearson Education

- Suihkonen, Pirkko 1990. Korpustutkimus kielitypologiassa sovellettuna udmurttiin. Suomalais-ugrilaisen Seuran toimituksia 207. Suomalaisugrilainen Seura. Helsinki.

- Trudgill, Peter 1992. Introducing Language and Society. Penguin Group. London.

- Timerkhanova, N. N. 2011. Osobennost porjadka slov v prozaicheskikh proizvedenijakh G. E. Vereshchaginai v sovremennom udmurtkom jazyke. In: Tipologicheskie aspekty mnogojazychiia v sovremennom obrazovatelnom prostranstve. Izdatelstvo „Udmurtskii universitet". Izhevsk. 180–185.

- Vennemann, Theo 1974. Analogy in Generative Grammar: The Origin of Word Order. In L. Heilmann (szerk.), *Proceedings of the Eleventh International Congress of Linguists*, 79–83. Bologna: Il Mulino.

- Vilkuna, Maria 1998. Word Order in European Uralic. In: Siewierska, Anna (ed): Constituent Order in the Languages of Europe. Empirical approaches to language typology 20–1. Mouton de Gruyter. Berlin–New York. 173–233.

- Winkler, Eberhard 2011. Udmurtische Grammatik. Veröffentlichungen der Societas Uralo-Altaica 81. Harrassowitz. Wiesbaden.

- Zhuikov, S. P. 1937. Osnovy grammatiki udmurtkogo jazyka: tezisy k pervoi respublikanskoi iazykovoi konferencii. Udmurtgosizdat. Izhevsk