Getting to grips with "linguistic complexity" from a cross-linguistic perspective

Peter Arkadiev
Freiburg Institute for Advanced Studies
(formerly) Johannes Gutenberg-Universität Mainz

FRIAS
FREIBURG INSTITUTE
FOR ADVANCED STUDIES
ALBERT-LUDWIGS-
UNIVERSITÄT FREIBURG
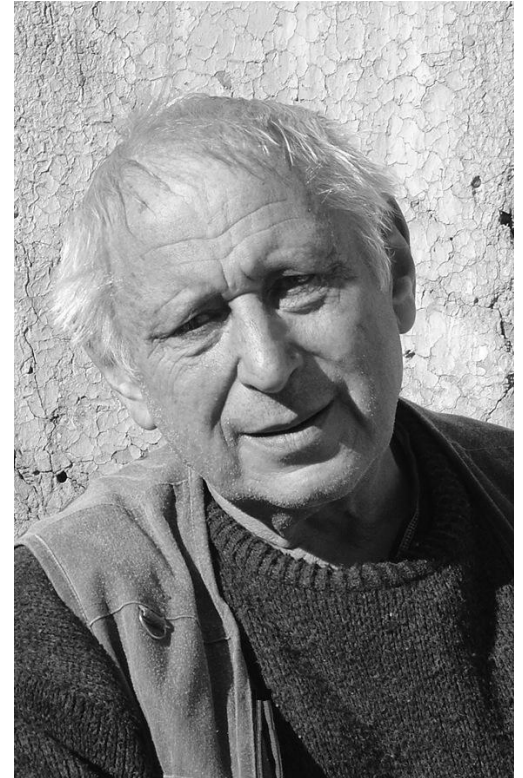
JOHANNES GUTENBERG
UNIVERSITÄT MAINZ
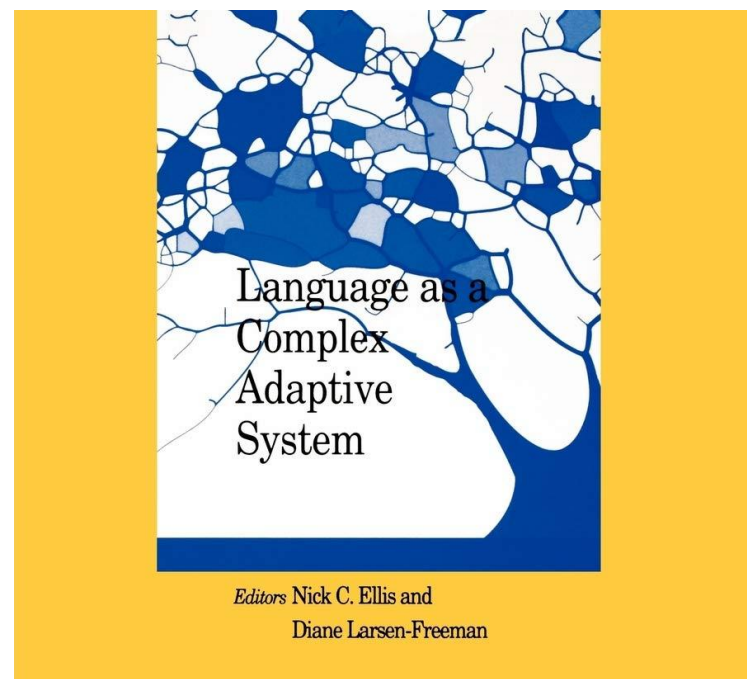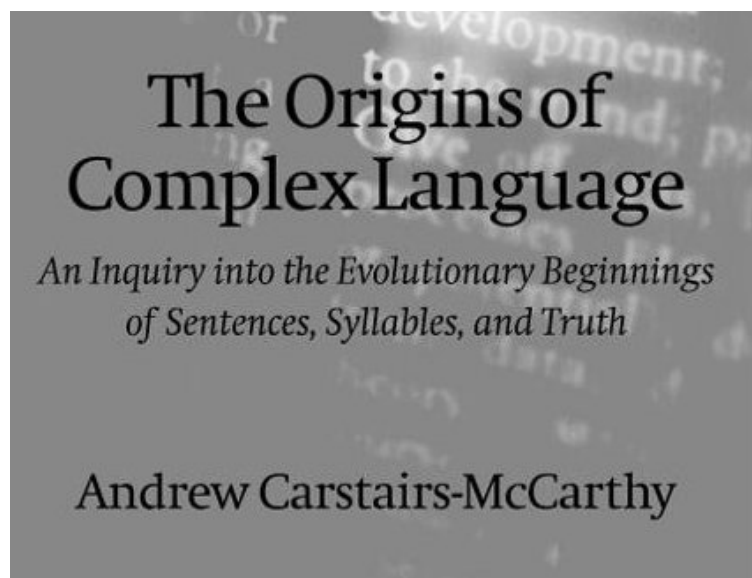
JG|U

Epigraph

Alexander Kibrik (1939-2012):

- "Language itself is simple,
it is the linguists' conceptions
thereof which are complex,
due to their inadequacy."

("Linguistic postulates", 1983/1992,
originally in Russian, my translation)

Why "linguistic complexity"?

Linguists tend to speak about human language being "complex":

Why "linguistic complexity"?

Nowadays linguists tend to speak about languages being more or less "complex":

- "Languages seem amazingly complex, in particular when one tries to learn foreign languages after childhood and youth." (Forker 2021: 1)

- "The idea of simpler and more complex language structure has both logic and common experience on its side." (Joseph 2021: 3)

- "The languages evidencing the least complexity of all of the world's languages are creoles." (McWhorter 2011: 3)

Why "linguistic complexity"?

Nowadays linguists tend to speak about languages being more or less "complex":

- "Languages seem amazingly complex, in particular when one tries to learn foreign languages after childhood and youth." (Forker 2021: 1)

- "The idea of simpler and more complex language structure has both logic and common experience on its side." (Joseph 2021: 3)

- "The languages evidencing the least complexity of all of the world's languages are creoles." (McWhorter 2011: 3)

Why "linguistic complexity"?

Nowadays linguists tend to speak about languages being more or less "complex":

- "Languages seem amazingly complex, in particular when one tries to learn foreign languages after childhood and youth." (Forker 2021: 1)

- "The idea of simpler and more complex language structure has both logic and common experience on its side." (Joseph 2021: 3)

- "The languages evidencing the least complexity of all of the world's languages are creoles." (McWhorter 2011: 3)

Why "linguistic complexity"?

Nowadays linguists tend to speak about languages being more or less "complex":

- "Languages seem amazingly complex, in particular when one tries to learn foreign languages after childhood and youth." (Forker 2021: 1)

- "The idea of simpler and more complex language structure has both logic and common experience on its side." (Joseph 2021: 3)

- "The languages evidencing the least complexity of all of the world's languages are creoles." (McWhorter 2011: 3)

Why "linguistic complexity"?

This has not always been the case:

- for much of the 20th century, linguists, implicitly or explicitly, tended to assume that all languages are "equally complex".

- A useful overview of the relevant *Ideengeschichte* can be found in Joseph & Newmeyer 2012.

- See also a bibliography in Kortmann & Schröter 2020.

Why "linguistic complexity"?

This has not always been the case:

- for much of the 20th century, linguists, implicitly or explicitly, tended to assume that all languages are "equally complex".

- A useful overview of the relevant *Ideengeschichte* can be found in Joseph & Newmeyer (2012).

- See also a bibliography in Kortmann & Schröter (2020).

Why "linguistic complexity"?

The "equi-complexity" thesis:

- "Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology [word structure] and syntax [sentence structure], is about the same as any other." (Hockett 1958: 180)

- "A central finding of linguistics has been that all languages, both ancient and modern, spoken by both 'primitive' and 'advanced' societies, are equally complex in their structure." (Fortson 1997: 4)

Why "linguistic complexity"?

The "equi-complexity" thesis:

- "Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology [word structure] and syntax [sentence structure], is about the same as any other." (Hockett 1958: 180)

- "A central finding of linguistics has been that all languages, both ancient and modern, spoken by both 'primitive' and 'advanced' societies, are equally complex in their structure." (Fortson 1997: 4)

Why "linguistic complexity"?

The "equi-complexity" thesis:

- "Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology [word structure] and syntax [sentence structure], is about the same as any other." (Hockett 1958: 180)

- "A central finding of linguistics has been that all languages, both ancient and modern, spoken by both 'primitive' and 'advanced' societies, are equally complex in their structure." (Fortson 1997: 4)

Why "linguistic complexity"?

The "equi-complexity" thesis:

- "Objective measurement is difficult, but impressionistically it would seem that the total grammatical complexity of any language, counting both morphology [word structure] and syntax [sentence structure], is about the same as any other." (Hockett 1958: 180)

- "A central finding of linguistics has been that all languages, both ancient and modern, spoken by both 'primitive' and 'advanced' societies, are equally complex in their structure." (Fortson 1997: 4)

Why "linguistic complexity"?

Not really:

- "Although the argument of a priori equi-complexity cannot be excluded, it is, however, not falsifiable: for every change in complexity it can be argued that there is another component, in another domain of language structure, pragmatics, or even culture, where the amount of complexity would be levelled out." (Kusters 2003: 12)

- "There are languages […] which have a simple(r) or less complex grammar and others with a (considerably more) complex grammar." (Kortmann & Szmrecsanyi 2009: 266)

Why "linguistic complexity"?

Not really:

- "Although the argument of a priori equi-complexity cannot be excluded, it is, however, not falsifiable: for every change in complexity it can be argued that there is another component, in another domain of language structure, pragmatics, or even culture, where the amount of complexity would be levelled out." (Kusters 2003: 12)

- "There are languages […] which have a simple(r) or less complex grammar and others with a (considerably more) complex grammar." (Kortmann & Szmrecsanyi 2009: 266)

Why "linguistic complexity"?

Not really:

- "Although the argument of a priori equi-complexity cannot be excluded, it is, however, not falsifiable: for every change in complexity it can be argued that there is another component, in another domain of language structure, pragmatics, or even culture, where the amount of complexity would be levelled out." (Kusters 2003: 12)

- "There are languages […] which have a simple(r) or less complex grammar and others with a (considerably more) complex grammar." (Kortmann & Szmrecsanyi 2009: 266)

Why "linguistic complexity"?

Not just a matter of linguistics:

- "Language complexity is usually not measured for its own sake. The purpose of measuring complexity is usually to learn something about language, society, the brain or other realworld phenomena, in other words, to use language complexity as an explanatory variable for addressing fundamental research questions." (Ehret et al. 2021: 8)

Why "linguistic complexity"?

Methodological concerns:

- "In linguistics, there is little consensus on how to define, measure, and compare complexity across languages." (Bentz et al. 2023: 9)

- "[T]here is no goldstandard or real-world benchmark against which complexity measurements could be evaluated" (Ehret et al. 2021: 2)

Why "linguistic complexity"?

Methodological concerns:

- "In linguistics, there is little consensus on how to define, measure, and compare complexity across languages." (Bentz et al. 2023: 9)

- "[T]here is no goldstandard or real-world benchmark against which complexity measurements could be evaluated." (Ehret et al. 2021: 2)
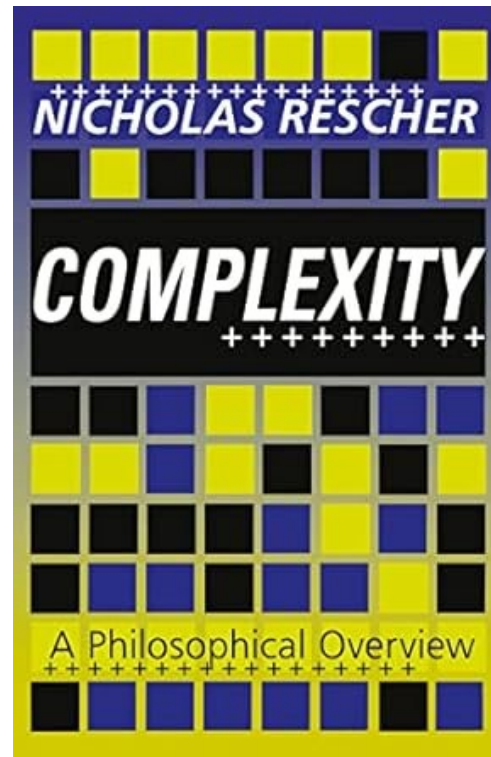
Understanding "linguistic complexity"

Nicholas Rescher (1928-2024)

Understanding "linguistic complexity"

Rescher (1998: 1):

- "Complexity is first and foremost a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure, be it organizational or operational."

- "Any sort of system or process […] will be to some extent complex."

- "Accordingly, all manner of things can be more or less complex: […] mind-engendered processes (languages or instructions) […]"

Understanding "linguistic complexity"

Rescher (1998: 1):

- "Complexity is first and foremost a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure, be it organizational or operational."

- "Any sort of system or process [...] will be to some extent complex."

- "Accordingly, all manner of things can be more or less complex: [...] mind-engendered processes (languages or instructions) [...]"

Understanding "linguistic complexity"

Rescher (1998: 1):

- "Complexity is first and foremost a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure, be it organizational or operational."

- "Any sort of system or process […] will be to some extent complex."

- "Accordingly, all manner of things can be more or less complex: […] mind-engendered processes (languages or instructions) […]"

Understanding "linguistic complexity"

Rescher (1998: 1):

- "Complexity is first and foremost a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure, be it organizational or operational."

- "Any sort of system or process [...] will be to some extent complex."

- "Accordingly, all manner of things can be more or less complex: [...] mind-engendered processes (languages or instructions) [...]"

Understanding "linguistic complexity"

Rescher (1998: 1):

- "Complexity is first and foremost a matter of the number and variety of an item's constituent elements and of the elaborateness of their interrelational structure, be it organizational or operational."

- "Any sort of system or process [...] will be to some extent complex."

- "Accordingly, all manner of things can be more or less complex: [...] mind-engendered processes (languages or instructions) [...]"

Understanding "linguistic complexity"

Modes of complexity (based on Rescher 1998: 9):

- epistemic mode: related to linguists' knowledge about languages;

- ontological mode: related to the properties of languages;

- functional: related to the use of languages by speakers.

Much harder to distinguish in practice than to postulate in theory.

Understanding "linguistic complexity"

Modes of complexity (based on Rescher 1998: 9):

- epistemic mode: related to linguists' knowledge about languages;

- ontological mode: related to the properties of languages;

- functional: related to the use of languages by speakers.

Much harder to distinguish in practice than to postulate in theory.

Understanding "linguistic complexity"

Modes of complexity (based on Rescher 1998: 9):

- epistemic mode: related to linguists' knowledge about languages;

- ontological mode: related to the properties of languages;

- functional: related to the use of languages by speakers.

Much harder to distinguish in practice than to postulate in theory.

Understanding "linguistic complexity"

Modes of complexity (based on Rescher 1998: 9):

- epistemic mode: related to linguists' knowledge about languages;

- ontological mode: related to the properties of languages;

- functional mode: related to the use of languages by humans.

Much harder to distinguish in practice than to postulate in theory.

Understanding "linguistic complexity"

Modes of complexity (based on Rescher 1998: 9):

- epistemic mode: related to linguists' knowledge about languages;

- ontological mode: related to the properties of languages;

- functional mode: related to the use of languages by humans.

Much harder to distinguish in practice than to postulate in theory.

Understanding "linguistic complexity"

Two distinct but related notions:

absolute complexity	relative complexity
system-based	user-based
"objective" properties of a linguistic system	features of the linguistic system that make it difficult for learners
complexity per se	difficulty

Understanding "linguistic complexity"

- For speakers of both English and Russian, learning German as a second language is difficult, but the reasons for this can be quite distinct.

- Kabardian is extremely difficult to learn for speakers of any European language, but for the speakers of the related Abaza it is apparently relatively easy.

- "*complex* should not be confused with *complicated*" (Mufwene et al. 2017: 3)

Understanding "linguistic complexity"

- For speakers of both English and Russian, learning German as a second language is difficult, but the reasons for this can be quite distinct.

- Kabardian is extremely difficult to learn for speakers of any European language, but for the speakers of the related Abaza it is apparently relatively easy.

- "*complex* should not be confused with *complicated*" (Mufwene et al. 2017: 3)

Understanding "linguistic complexity"

- For speakers of both English and Russian, learning German as a second language is difficult, but the reasons for this can be quite distinct.

- Kabardian is extremely difficult to learn for speakers of any European language, but for the speakers of the related Abaza it is apparently relatively easy.

- "*complex* should not be confused with *complicated*" (Mufwene et al. 2017: 3)

Understanding "linguistic complexity"

Another dichotomy:

global complexity	local complexity
language "as a whole"	particular subsystem or phenomenon

Understanding "linguistic complexity"

phenomenon	English	German
grammatical gender	none	three
grammatical case in nouns	none	four
declensions of nouns	none	several
irregular morphology in general	less	more
word order in main and dependent clauses	same	different
tenses	more	less

Understanding "linguistic complexity"

phenomenon	English	German
grammatical gender	none	three
grammatical case in nouns	none	
declensions of nouns	none	
irregular morphology in general	less	
word order in main and dependent clauses	same	different
tenses	more	less

English is
"less complex"

Understanding "linguistic complexity"

phenomenon	English	German
grammatical gender	none	three
grammatical case in nouns	none	
declensions of nouns	none	
irregular morphology in general	less	
word order in main and dependent clauses	same	different
tenses	more	

English is "less complex"

English is "more complex"

Understanding "linguistic complexity"

- Against "global complexity":

- "[N]o comprehensive proposal exists to date for measuring the degree of complexity of an entire language, nor is there even agreement on precisely what should be measured" (Joseph & Newmeyer 2012: 360)

- "[I]t is in fact impossible to define the notion of overall complexity in an objective meaningful way. At best, the "overall complexity" of a language can be understood as a vector (one-dimensional matrix) of separate values." (Deutscher 2009: 247)

Understanding "linguistic complexity"

- Against "global complexity":

- "[N]o comprehensive proposal exists to date for measuring the degree of complexity of an entire language, nor is there even agreement on precisely what should be measured" (Joseph & Newmeyer 2012: 360)

- "[I]t is in fact impossible to define the notion of overall complexity in an objective meaningful way. At best, the ''overall complexity'' of a language can be understood as a vector (one-dimensional matrix) of separate values." (Deutscher 2009: 247)

Understanding "linguistic complexity"

Local complexity:

- manifests itself differently in each local subsystem
- finding appropriate quantitative measures is itself a difficult task.

Understanding "linguistic complexity"

The perfect tenses in English and German:

Parameter	English	German
No. of components	2 (auxiliary + participle)	2 (auxiliary + participle)
No. of auxiliaries	1 (*have*)	2 (*haben* vs. *sein*)
No. of regular types of participle formation	1 (*work ~ worked*)	2 (*machen ~ gemacht; sprach ~ gesprochen*)
Interactions with verb morphology	no	yes (prefixed verbs)
Forms of the auxiliary	few	many
Contiguity of components	usually contiguous	often non-contiguous
Constraints on use	very difficult to measure quantitatively	

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as elaboration:
  - number of distinctions encoded (e.g. tenses, genders or grammatical categories absent in other languages);
  - number of forms encoding a category (e.g. distinct declensions or conjugations, or markers within a word or sentence);
  - number of grammatical rules (e.g. different ordering or phonological adjustment).

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as elaboration:
    - number of distinctions encoded (e.g. tenses, genders or grammatical categories absent in other languages);
    - number of forms encoding a category (e.g. distinct declensions or conjugations, or markers within a word or sentence);
    - number of grammatical rules (e.g. different ordering or phonological adjustment).

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as elaboration:
  - number of distinctions encoded (e.g. tenses, genders or grammatical categories absent in other languages);
  - number of forms encoding a category (e.g. distinct declensions or conjugations, or markers within a word or sentence);
  - number of grammatical rules (e.g. different ordering or phonological adjustment).

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as elaboration:
  - number of distinctions encoded (e.g. tenses, genders or grammatical categories absent in other languages);
  - number of forms encoding a category (e.g. distinct declensions or conjugations, or markers within a word or sentence);
  - number of grammatical rules (e.g. different ordering or phonological adjustment).

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as elaboration:
    - number of distinctions encoded (e.g. tenses, genders or grammatical categories absent in other languages);
    - number of forms encoding a category (e.g. distinct declensions or conjugations, or markers within a word or sentence);
    - number of grammatical rules (e.g. different ordering or phonological adjustment).

Types of "linguistic complexity"

Elaboration: English tense-aspect system

- present vs. past + different types of future

- simple vs. progressive vs. perfect

- combinations of the above

(1)   *An otherwise working day is a day that an employee would have been working had the day not been a public holiday.* (Internet)

Types of "linguistic complexity"

Elaboration: English tense-aspect system

- present vs. past + different types of future

- simple vs. progressive vs. perfect

- combinations of the above

(1) *An otherwise working day is a day that an employee would have been working had the day not been a public holiday.* (Internet)

Types of "linguistic complexity"

Elaboration: English tense-aspect system

- present vs. past + different types of future

- simple vs. progressive vs. perfect

- combinations of the above

(1)   An otherwise working day is a day that an employee would have been working had the day not been a public holiday. (Internet)

Types of "linguistic complexity"

Elaboration: English tense-aspect system

- present vs. past + different types of future

- simple vs. progressive vs. perfect

- combinations of the above

(1) *An otherwise working day is a day that an employee would have been working had the day not been a public holiday.* (Internet)

Types of "linguistic complexity"

Elaboration: English tense-aspect system

- present vs. past + different types of future
- simple vs. progressive vs. perfect
- combinations of the above

(1)   *An otherwise working day is a day that an employee would have been working had the day not been a public holiday.* (Internet)

Types of "linguistic complexity"

Grammatical elaboration is related to redundancy:

- "[L]anguages can be far more complex than merely the minimum required for effective communication. A great deal of complexity is redundant historical baggage." (Deutscher 2009: 245-6)

- Redundancy is necessary for successful communication (hearer-based perspective), but is costly for production (speaker-based perspective) (cf. Leufkens 2023: 98-9).

Types of "linguistic complexity"

Grammatical elaboration is related to redundancy:

- "[L]anguages can be far more complex than merely the minimum required for effective communication. A great deal of complexity is redundant historical baggage." (Deutscher 2009: 245-6)

- Redundancy is necessary for successful communication (hearer-based perspective), but is costly for production (speaker-based perspective) (cf. Leufkens 2023: 98-9).

Types of "linguistic complexity"

Grammatical elaboration is related to redundancy:

- "[L]anguages can be far more complex than merely the minimum required for effective communication. A great deal of complexity is redundant historical baggage." (Deutscher 2009: 245-6)

- Redundancy is necessary for successful communication (hearer-based perspective), but is costly for production (speaker-based perspective) (cf. Leufkens 2023: 98-9).

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as opacity (lack of "transparency"):
  - deviations from the "one-meaning-one-form" principle;
  - deviations from "iconic" ordering of elements;
  - unpredictability and irregularity of forms;
  - phonological fusion.

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as opacity (lack of "transparency"):
  - deviations from the "one-meaning-one-form" principle;
  - deviations from "iconic" ordering of elements;
  - unpredictability and irregularity of forms;
  - phonological fusion.

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as opacity (lack of "transparency"):
  - deviations from the "one-meaning-one-form" principle;
  - deviations from "iconic" ordering of elements;
  - unpredictability and irregularity of forms;
  - phonological fusion.

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as opacity (lack of "transparency"):
  - deviations from the "one-meaning-one-form" principle;
  - deviations from "iconic" ordering of elements;
  - unpredictability and irregularity of forms;
  - phonological fusion.

Types of "linguistic complexity"

Distinct facets of complexity:

- complexity as opacity (lack of "transparency"):
    - deviations from the "one-meaning-one-form" principle;
    - deviations from "iconic" ordering of elements;
    - unpredictability and irregularity of forms;
    - phonological fusion obliterating boundaries between elements;
    - etc.

Types of "linguistic complexity"

Opacity: person marking in Abaza (Northwest Caucasian, Russia)

(2) *də-b-b-əj-ṭ*
3SG.H-2SG.F-see-PRS-DCL
'You see him/her.'

(3) *də-b-pš-əj-ṭ*
3SG.H-2SG.F-look-PRS-DCL
'S/he looks at you.'

DCL – declarative mood
F – feminine gender, H – human
PRS – present tense
SG – singular number

© Yury Koryakov

Types of "linguistic complexity"

Opacity: person marking in Abaza (Northwest Caucasian, Russia)

(2) *də-b-b-əj-ṭ*
3SG.H-2SG.F-see-PRS-DCL
'You see him/her.'   subject

(3) *də-b-pš-əj-ṭ*
3SG.H-2SG.F-look-PRS-DCL
'S/he looks at you.'   object

DCL – declarative mood
F – feminine gender, H – human
PRS – present tense
SG – singular number

Types of "linguistic complexity"

Opacity: person marking in Abaza (Northwest Caucasian, Russia)

(2) *də-b-b-əj-ṭ*
     3SG.H-2SG.F-see-PRS-DCL
     'You see him/her.'   object

(3) *də-b-pš-əj-ṭ*
     3SG.H-2SG.F-look-PRS-DCL
     'S/he looks at you.'   subject

DCL – declarative mood
F – feminine gender, H – human
PRS – present tense
SG – singular number

Types of "linguistic complexity"

Opacity: person marking in Abaza (Northwest Caucasian, Russia)

(2) *də-b-b-əj-ṭ*
    3SG.H-2SG.F-see-PRS-DCL
    'You see him/her.'

(3) *də-b-pš-əj-ṭ*
    3SG.H-2SG.F-look-PRS-DCL
    'S/he looks at you.'

Same affixes in the same order, but their interpretation depends on the lexical class of the verb they co-occur with

DCL – declarative mood
F – feminine gender, H – human
PRS – present tense
SG – singular number

Types of "linguistic complexity"

Different types of opacity are pervasive in natural languages.

- "A striking fact about languages is that it is exceptional for them to display a systematic one-to-one relation between meaning and form, i.e. languages are never completely transparent." (Hengeveld & Leufkens 2018: 139)

- The grammars of some constructed languages, e.g. Esperanto, were designed with a specific aim to achieve maximal transparency.

Types of "linguistic complexity"

Different types of opacity are pervasive in natural languages:

- "A striking fact about languages is that it is exceptional for them to display a systematic one-to-one relation between meaning and form, i.e. languages are never completely transparent." (Hengeveld & Leufkens 2018: 139)

- The grammars of some constructed languages, e.g. Esperanto, were designed with a specific aim to achieve maximal transparency.

Types of "linguistic complexity"

Different types of opacity are pervasive in natural languages:

- "A striking fact about languages is that it is exceptional for them to display a systematic one-to-one relation between meaning and form, i.e. languages are never completely transparent." (Hengeveld & Leufkens 2018: 139)

- The grammars of some constructed languages, e.g. Esperanto, were designed with a specific aim to achieve maximal transparency.

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- elaboration in a given grammatical domain can be (more or less) straightforwardly measured and compared across languages;

- measuring (relative) opacity of a grammatical domain is much more difficult;

- recently, a number of sophisticated information-theory based measures have been developed for some types of opacity.

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- elaboration in a given grammatical domain can be (more or less) straightforwardly measured and compared across languages;

- measuring (relative) opacity of a grammatical domain is much more difficult;

- recently, a number of sophisticated information-theory based measures have been developed for some types of opacity.

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- elaboration in a given grammatical domain can be (more or less) straightforwardly measured and compared across languages;

- measuring (relative) opacity of a grammatical domain is much more difficult;

- recently, a number of sophisticated information-theory based measures have been developed for some types of opacity.

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- elaboration in a given grammatical domain can be (more or less) straightforwardly measured and compared across languages;

- measuring (relative) opacity of a grammatical domain is much more difficult;

- recently, a number of sophisticated information-theory based measures have been developed for some types of opacity.

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- Are these two facets of complexity comparable resp. commensurable?

- Which of them is "more important"?

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- "Measures of both ''explicitness/sophistication'' and ''irregularity'' firmly belong in our intuition of complexity, but there is no non-arbitrary way of weighing up their relative importance." (Deutscher 2009: 249)

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- "The LOW ENTROPY CONJECTURE is the hypothesis that enumerative morphological complexity is effectively unrestricted, as long as the average conditional entropy, a measure of integrative complexity, is low."
(Ackerman & Malouf 2013: 436)

Types of "linguistic complexity"

Complexity as elaboration vs. complexity as opacity:

- "The LOW ENTROPY CONJECTURE is the hypothesis that enumerative morphological complexity is effectively unrestricted, as long as the average conditional entropy, a measure of integrative complexity, is low."
(Ackerman & Malouf 2013: 436)

"Complexity" and flagging/indexing

- My FRIAS project "Argument flagging in head-marking languages: Types of interaction between argument-coding systems".

- A cross-linguistic investigation of the possible interactions between the two most important formal systems of encoding predicate-argument relations.

- Not a study of "complexity" *per se.*

- However, the phenomena in question can be assessed in terms of "elaboration", "redundancy" and "opacity".

"Complexity" and flagging/indexing

- My FRIAS project "Argument flagging in head-marking languages: Types of interaction between argument-coding systems".

- A cross-linguistic investigation of the possible interactions between the two most important formal systems of encoding predicate-argument relations.

- Not a study of "complexity" *per se.*

- However, the phenomena in question can be assessed in terms of "elaboration", "redundancy" and "opacity".

"Complexity" and flagging/indexing

- My FRIAS project "Argument flagging in head-marking languages: Types of interaction between argument-coding systems".

- A cross-linguistic investigation of the possible interactions between the two most important formal systems of encoding predicate-argument relations.

- Not a study of "complexity" *per se.*

- However, the phenomena in question can be assessed in terms of "elaboration", "redundancy" and "opacity".

"Complexity" and flagging/indexing

- My FRIAS project "Argument flagging in head-marking languages: Types of interaction between argument-coding systems".

- A cross-linguistic investigation of the possible interactions between the two most important formal systems of encoding predicate-argument relations.

- Not a study of "complexity" *per se.*

- However, the phenomena in question can be assessed in terms of "elaboration", "redundancy" and "opacity".

"Complexity" and flagging/indexing

Basic notions:

- predicates (usually, verbs) encode situations, and arguments (usually, nouns and noun phrases, as well as pronouns) express their participants;

- flagging ("dependent-marking"): formal marking of arguments for their grammatical or semantic relation to the predicate;

- indexing ("head-marking"): formal marking of predicates for such properties of (some of) their arguments as person, number, gender and grammatical or semantic relation.

"Complexity" and flagging/indexing

Basic notions:

- predicates (usually, verbs) encode situations, and arguments (usually, nouns and noun phrases, as well as pronouns) express their participants;

- flagging ("dependent-marking"): formal marking of arguments for their grammatical or semantic relation to the predicate;

- indexing ("head-marking"): formal marking of predicates for such properties of (some of) their arguments as person, number, gender and grammatical or semantic relation.

"Complexity" and flagging/indexing

Basic notions:

- predicates (usually, verbs) encode situations, and arguments (usually, nouns and noun phrases, as well as pronouns) express their participants;

- flagging ("dependent-marking"): formal marking of arguments for their grammatical or semantic relation to the predicate;

- indexing ("head-marking"): formal marking of predicates for such properties of (some of) their arguments as person, number, gender and grammatical or semantic relation.

"Complexity" and flagging/indexing

Basic notions:

- predicates (usually, verbs) encode situations, and arguments (usually, nouns and noun phrases, as well as pronouns) express their participants;

- flagging ("dependent-marking"): formal marking of arguments for their grammatical or semantic relation to the predicate;

- indexing ("head-marking"): formal marking of predicates for such properties of (some of) their arguments as person, number, gender and grammatical or semantic relation.

"Complexity" and flagging/indexing

German

(4) *Der Dichter hilft dem Komponisten.*
     'The poet helps the composer.'

"Complexity" and flagging/indexing

German

(4) *Der Dichter hilft dem Komponisten.*
     'The poet helps the composer.'

     flagging by morphological case

"Complexity" and flagging/indexing

German

(4)   *Der Dichter* hilf*t* *dem Komponisten.*
       'The poet helps the composer.'

(4') *Du* hilf*st* *dem Komponisten.*
       'You help the composer.'

indexing of the subject by person-number affixes in the verb

"Complexity" and flagging/indexing

Cross-linguistic variation (simplified version):

flagging	indexing	example
no	no	Lao
yes	no	Japanese
no	yes	Navajo
yes	yes	German

"Complexity" and flagging/indexing

No flagging, no indexing:
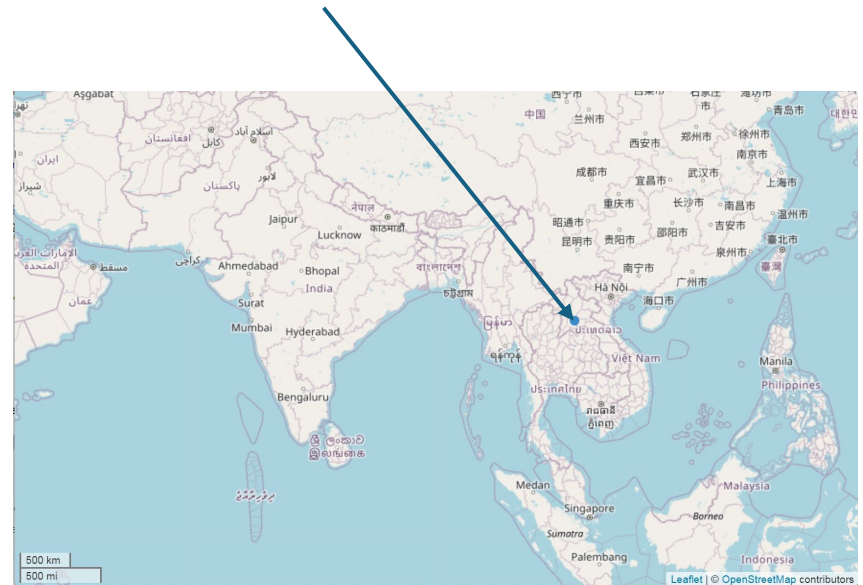
Lao (Tai-Kadai, Laos; Enfield 2007: 365):

(5) *caw⁴*   *haj⁵*   *khòòj⁵*   *haa⁵-lòò⁴*   *kiip⁵*
     2SG   give   1SG   five-hundred   kip
     'You gave me 500 kip.'

1 – 1st person
2 – 2nd person
SG – singular number

Map created with Lingtypology (Moroz 2017)

"Complexity" and flagging/indexing

Flagging, no indexing:

Japanese (Transeurasian, Japan; constructed):

(6) *shōjo=ga*   *shōnen=o*   *mi-ta*
     girl=NOM   boy=ACC   see-PST
     'The girl saw the boy.'

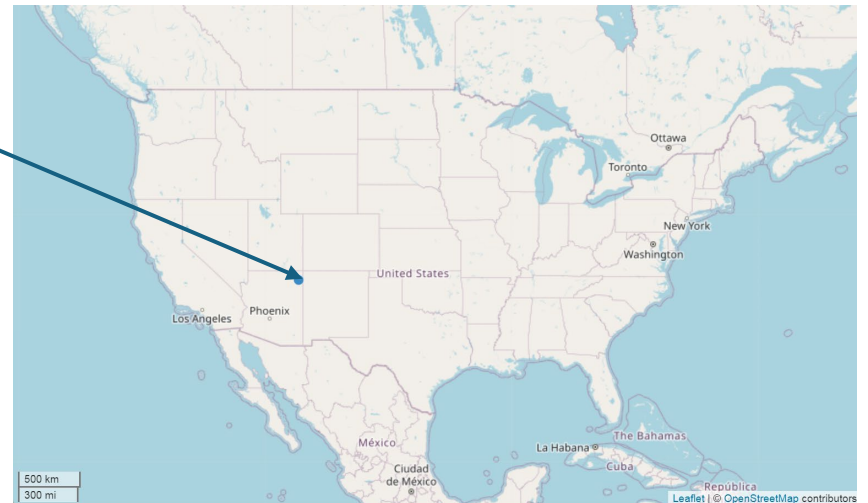ACC – accusative case
NOM – nominative case
PST – past tense

"Complexity" and flagging/indexing

Indexing, no flagging:

Navajo (Athabaskan, USA; Kibrik 2012: 229):

(7) *ˀashkii*   *ˀatʼééd*   *łįįˀ*   *y-e-i-ø-ní-lóóz*
      boy   girl   horse   her-to-it-he-PFV-lead
     'The boy brought the horse to the girl.'

PFV – perfective aspect

Map created with Lingtypology (Moroz 2017)

"Complexity" and flagging/indexing

Cross-linguistic variation (simplified version):

flagging	indexing	example
no	no	Lao
yes	no	Japanese
no	yes	Navajo
yes	yes	German

"Complexity" and flagging/indexing

Cross-linguistic variation (simplified version):

flagging	indexing	example
no	no	Lao
yes	no	Japanese
no	yes	Navajo
yes	yes	German

"Complexity" and flagging/indexing

- My project is about languages with both indexing and flagging.

- Moreover, it is about languages that have more indexing than German:
  - not just "subject-verb agreement";
  - indexing of at least both subject and object;
  - such languages are not infrequent.

"Complexity" and flagging/indexing

- My project is about languages with both indexing and flagging.

- Moreover, it is about languages that have more indexing than German:

  - not just "subject-verb agreement";
  - indexing of at least both subject and object;
  - such languages are not infrequent.

"Complexity" and flagging/indexing

- My project is about languages with both indexing and flagging.

- Moreover, it is about languages that have more indexing than German:
  - not just "subject-verb agreement";
  - indexing of at least both subject and object;
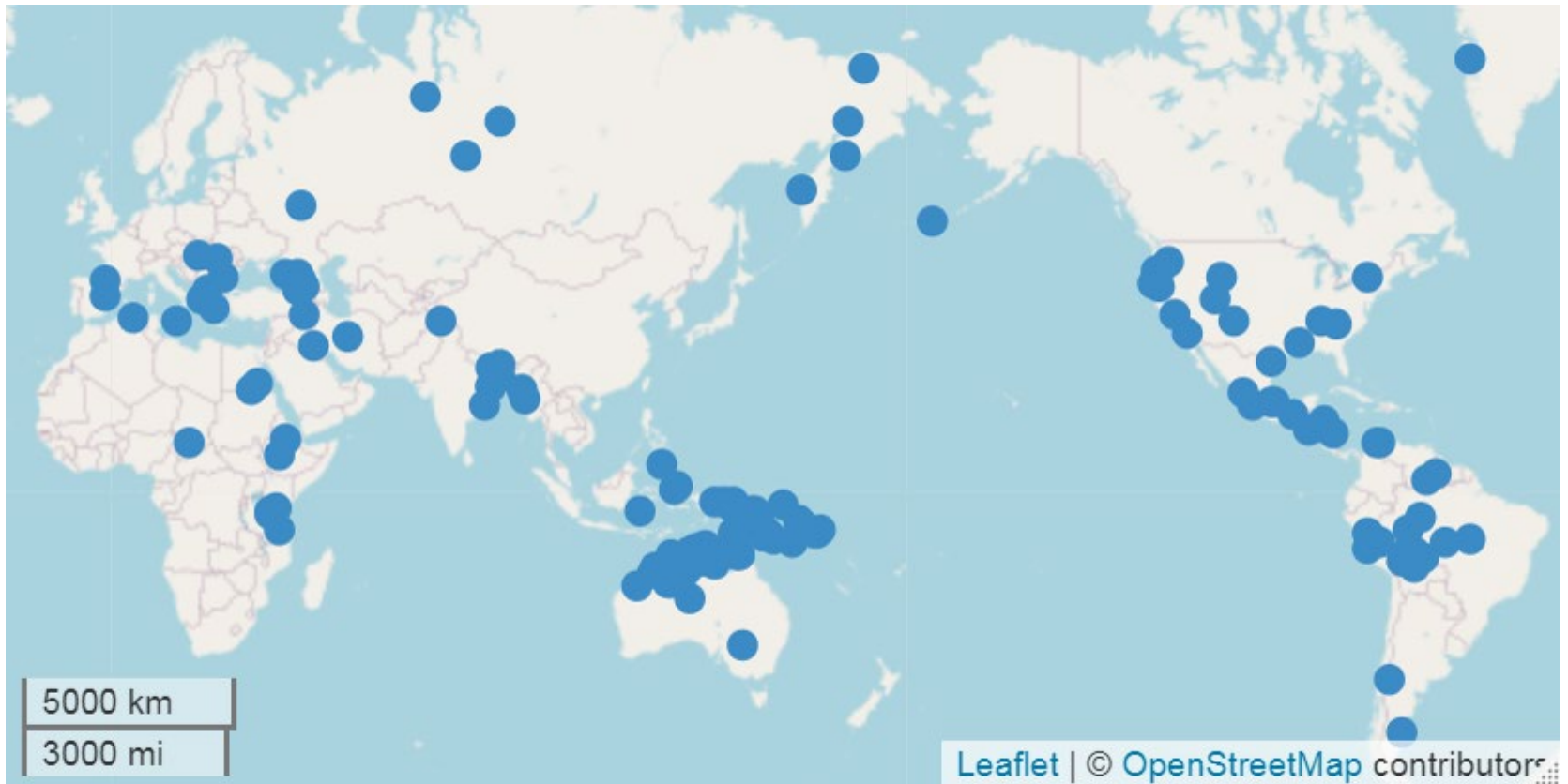 - such languages are not infrequent.

"Complexity" and flagging/indexing

- My project is about languages with both indexing and flagging.

- Moreover, it is about languages that have more indexing than German:

  - not just "subject-verb agreement";
  - indexing of at least both subject and object;
  - such languages are not infrequent.

"Complexity" and flagging/indexing

- My project is about languages with both indexing and flagging.

- Moreover, it is about languages that have more indexing than German:
  - not just "subject-verb agreement";
  - indexing of at least both subject and object;
  - such languages are not infrequent.

"Complexity" and flagging/indexing

Map created with Lingtypology (Moroz 2017)

"Complexity" and flagging/indexing

- Languages with both indexing and flagging fall into two broad types based on how the two systems interact:
  1. Languages where the domains of use of flagging and indexing do not overlap (complementary distribution).
  2. Languages where the domains of use of flagging and indexing overlap giving rise to double-marking.

"Complexity" and flagging/indexing

- Languages with both indexing and flagging fall into two broad types based on how the two systems interact:

    1. Languages where the domains of use of flagging and indexing do not overlap (complementary distribution).

    2. Languages where the domains of use of flagging and indexing overlap giving rise to double-marking.

"Complexity" and flagging/indexing

- Languages with both indexing and flagging fall into two broad types based on how the two systems interact:

  1. Languages where the domains of use of flagging and indexing do not overlap (complementary distribution).

  2. Languages where the domains of use of flagging and indexing overlap giving rise to double-marking.

"Complexity" and flagging/indexing

- Complementarity of flagging and indexing

Yimas (Lower Sepik-Ramu, Papua New Guinea)

(8) *namat   uraŋk   narmaŋ kɨ-n-ŋa-r-umpun*
     man.PL   coconut   woman   3SG.DO-3SG.SBJ-give-PRF-3PL.IO
     'The woman gave the coconut to the men.' (Foley 1986: 94)

(9) *tnumut-ɲan   ama-na-irm-n*
     sago_palms-OBL   1SG.SBJ-ASP-stand-PRS
     'I am standing at the two sago palms.' (Foley 1991: 165)

ASP – aspect marker
DO – direct object, IO – indirect object
OBL – oblique case
PL – plural, SG – singular
PRF – perfect   SBJ – subject
PRS – present tense

Map created with Lingtypology (Moroz 2017)

"Complexity" and flagging/indexing

- Complementarity of flagging and indexing

Yimas (Lower Sepik-Ramu, Papua New Guinea)

(8) *namat   uraŋk   narmaŋ ki-n-ŋa-r-umpun*
    man.PL   coconut   woman   3SG.DO-3SG.SBJ-give-PRF-3PL.IO
    'The woman gave the coconut to the men.' (Foley 1986: 94)

(9) *tnumut-ŋan   ama-na-irm-n*
    sago_palms-OBL   1SG.SBJ-ASP-stand-PRS
    'I am standing at the two sago palms.' (Foley 1991: 165)

ASP – aspect marker
DO – direct object, IO – indirect object
OBL – oblique case
PL – plural, SG – singular
PRF – perfect   SBJ – subject
PRS – present tense

Map created with Lingtypology (Moroz 2017)

"Complexity" and flagging/indexing

- Complementarity of flagging and indexing
  - indexing for the most prominent and frequently occurring "core" participants (agent, patient, recipient);
  - flagging for all other, less prominent, so-called "peripheral" participants (locations, instruments, causes and others);
  - an apparently logical and economical system;
  - yet most languages are not like Yimas.
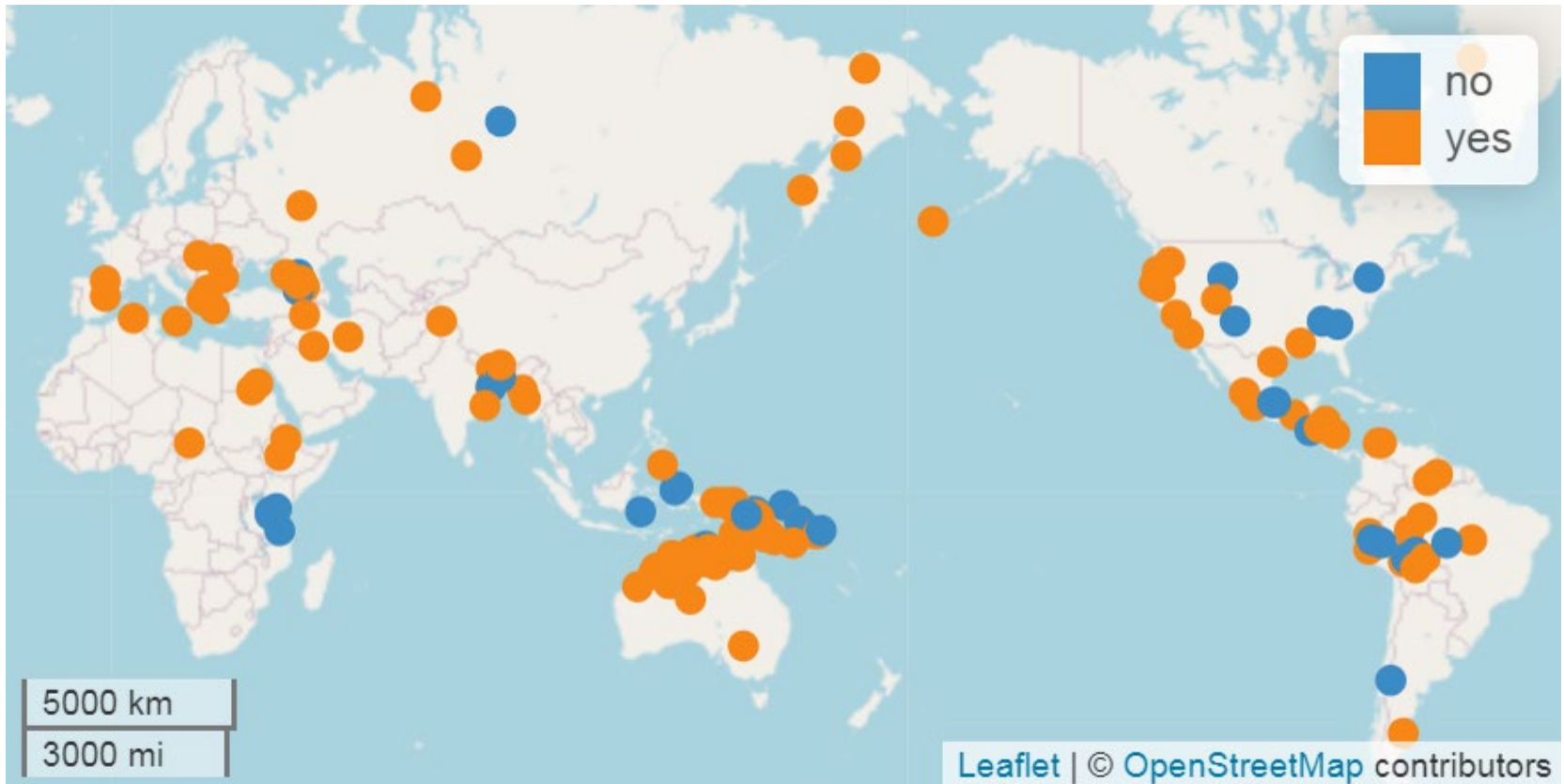
"Complexity" and flagging/indexing

- Complementarity of flagging and indexing
  - indexing for the most prominent and frequently occurring "core" participants (agent, patient, recipient);
  - flagging for all other, less prominent, so-called "peripheral" participants (locations, instruments, causes and others);
  - an apparently logical and economical system;
  - yet most languages are not like Yimas.

"Complexity" and flagging/indexing

- Complementarity of flagging and indexing
  - indexing for the most prominent and frequently occurring "core" participants (agent, patient, recipient);
  - flagging for all other, less prominent, so-called "peripheral" participants (locations, instruments, causes and others);
  - an apparently logical and economical system;
  - yet most languages are not like Yimas.

"Complexity" and flagging/indexing

- Complementarity of flagging and indexing
    - indexing for the most prominent and frequently occurring "core" participants (agent, patient, recipient);
    - flagging for all other, less prominent, so-called "peripheral" participants (locations, instruments, causes and others);
    - an apparently logical and economical system;
    - yet most languages are not like Yimas.

"Complexity" and flagging/indexing

Double-marking

"Complexity" and flagging/indexing
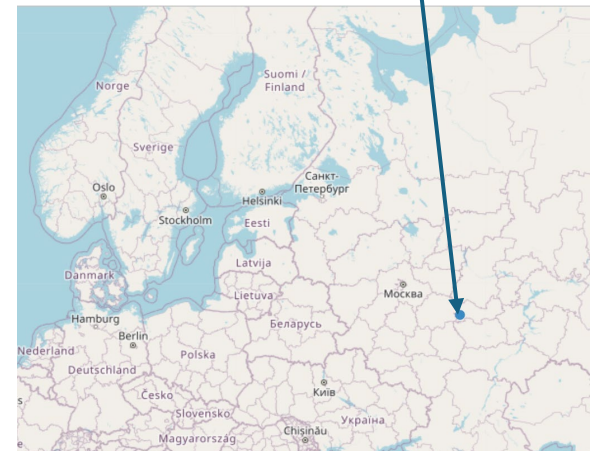
- Double-marking

Moksha (Uralic; Russia; Toldova et al. 2018: 575):

(10) a. *vas'ɛ*   *ker'-s'*   *šuftə*
      Vasya   cut-PST.3SG.SBJ   tree
      'Vasya cut a tree.'

      b. *vas'ɛ*   *ker'-əz'ə*   *šuft-t*
      Vasya   cut-PST.3SG.SBJ.3SG.OBJ   tree-DEF.SG.GEN
      'Vasya cut the tree.'

DEF – definite   OBJ – object
GEN – genitive case   SBJ – subject
PST – past tense   SG – singular number

"Complexity" and flagging/indexing
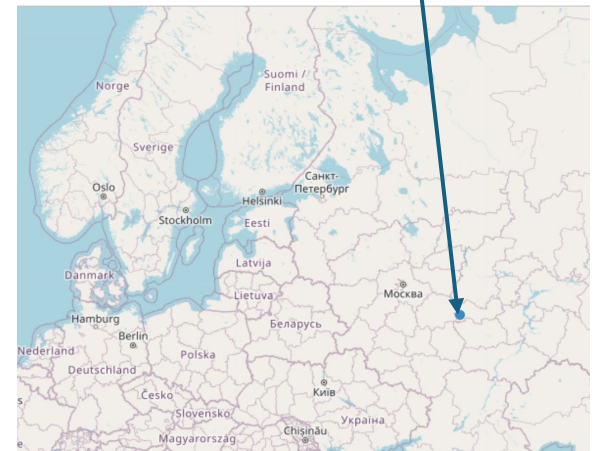
- Double-marking

Moksha (Uralic; Russia; Toldova et al. 2018: 575):

(10) a. *vas'ɛ*   *ker'-s'*   *šuftə*
           Vasya   cut-PST.3SG.SBJ tree
           'Vasya cut a tree.'

      b. *vas'ɛ*   *ker'-əz'ə*   *šuft-t*
           Vasya   cut-PST.3SG.SBJ.3SG.OBJ   tree-DEF.SG.GEN
           'Vasya cut the tree.'

DEF – definite	OBJ – object
GEN – genitive case	SBJ – subject
PST – past tense	SG – singular number

"Complexity" and flagging/indexing

- Even more double-marking

Basque (isolate, Spain, France; Saltarelli 1988: 242)

(10) *ni-k*   *zu-ri*   *gezurr-a esan*   *n-i-zu-n*

      1SG-ERG 2SG-DAT lie-DET   tell   1SG-AUX-2SG-PST

      'I told you a lie.'

\- elaboration and redundancy

AUX – auxiliary verb	PST – past tense
DAT – dative case	SG – singular number
DET – determiner	
ERG – ergative case	

"Complexity" and flagging/indexing

- Even more double-marking

Basque (isolate, Spain, France; Saltarelli 1988: 242)

(10) *ni-k*   *zu-ri*   *gezurr-a esan*   *n-i-zu-n*
    1SG-ERG 2SG-DAT lie-DET   tell   1SG-AUX-2SG-PST
    'I told you a lie.'

- elaboration and redundancy

AUX – auxiliary verb   PST – past tense
DAT – dative case   SG – singular number
DET – determiner
ERG – ergative case

"Complexity" and flagging/indexing

- Even more double-marking

Basque (isolate, Spain, France; Saltarelli 1988: 242)

(10) *ni-k*   *zu-ri*   *gezurr-a*   *esan*   *n-i-zu-n*
     1SG-ERG   2SG-DAT   lie-DET   tell   1SG-AUX-2SG-PST
     'I told you a lie.'

(10') *ni-k*   *zu-ri*   *gezurr-a*   *esan*   *di-zu-t*
     1SG-ERG   2SG-DAT   lie-DET   tell   AUX-2SG-1SG
     'I have told you a lie.'

- opacity:
  - one-to-many correspondences between flagging and indexing;
  - different order of person-number affixes in different tenses.

"Complexity" and flagging/indexing

- Even more double-marking

Basque (isolate, Spain, France; Saltarelli 1988: 242)

(10) *ni-k*   *zu-ri*   *gezurr-a*   *esan*   *n-i-zu-n*
     1SG-ERG   2SG-DAT   lie-DET   tell   1SG-AUX-2SG-PST
     'I told you a lie.'

(10') *ni-k*   *zu-ri*   *gezurr-a*   *esan*   *di-zu-t*
     1SG-ERG   2SG-DAT   lie-DET   tell   AUX-2SG-1SG
     'I have told you a lie.'

- opacity:
  - one-to-many correspondences between flagging and indexing;
  - different order of person-number affixes in different tenses.

"Complexity" and flagging/indexing

- Flagging and indexing often operate independently of each other and are subject to distinct, even if partially overlapping, preferences and constraints.

- "[A] whole with fewer parts engaged is several multilateral interactions can generate more interactive complexity" (Mufwene et el. 2017: 4)

"Complexity" and flagging/indexing

- Flagging and indexing often operate independently of each other and are subject to distinct, even if partially overlapping, preferences and constraints.

- Elements "engaged is several multilateral interactions can generate more interactive complexity" (Mufwene et el. 2017: 4)

"Complexity" and flagging/indexing

- My cross-linguistic study of flagging and indexing highlights the following "complexity"-related aspects of grammars:

  - elaboration across a variety of dimensions;

  - aspects of potential opacity in either of the domains;

  - most importantly, the dimension of interactions of flagging and indexing with each other (and with other domains of grammar) – an aspect often neglected in works on "linguistic complexity".

"Complexity" and flagging/indexing

- My cross-linguistic study of flagging and indexing highlights the following "complexity"-related aspects of grammars:
  - elaboration across a variety of dimensions;
  - aspects of potential opacity in either of the domains;
  - most importantly, the dimension of interactions of flagging and indexing with each other (and with other domains of grammar) – an aspect often neglected in works on "linguistic complexity".

"Complexity" and flagging/indexing

- My cross-linguistic study of flagging and indexing highlights the following "complexity"-related aspects of grammars:
  - elaboration across a variety of dimensions;
  - aspects of potential opacity in both domains;
  - most importantly, the dimension of interactions of flagging and indexing with each other (and with other domains of grammar) – an aspect often neglected in works on "linguistic complexity".

"Complexity" and flagging/indexing

- My cross-linguistic study of flagging and indexing highlights the following "complexity"-related aspects of grammars:
  - elaboration across a variety of dimensions;
  - aspects of potential opacity in both domains;
  - most importantly, the dimension of interactions of flagging and indexing with each other (and with other domains of grammar) – an aspect often neglected in works on "linguistic complexity".

Envoy

- Complexity "is multilayered, multifaceted and multidimensional in nature." (Kuiken 2023: 84)

- There is no meaningful way of assessing the "global complexity" of a language.

- "Modular" and "local" approaches to complexity are more promising.

Envoy

- Complexity "is multilayered, multifaceted and multidimensional in nature." (Kuiken 2023: 84)

- There is no meaningful way of assessing the "global complexity" of a language.

- "Modular" and "local" approaches to complexity are more promising.

Envoy

- Complexity "is multilayered, multifaceted and multidimensional in nature." (Kuiken 2023: 84)

- There is no meaningful way of assessing the "global complexity" of a language.

- "Modular" and "local" approaches to complexity are more promising.

Envoy

- Complexity "is multilayered, multifaceted and multidimensional in nature." (Kuiken 2023: 84)

- There is no meaningful way of assessing the "global complexity" of a language.

- "Modular" and "local" approaches to complexity are more promising.

Envoy

- Quantitative measures of "linguistic complexity" have limited application and have to rely on in-depth qualitative analysis and take into account many (not always independent) variables.

- Most generalisations about language complexity in the literature seem to be premature given both conceptual shortcomings and the quality of empirical coverage.

Envoy

- Quantitative measures of "linguistic complexity" have limited application and have to rely on in-depth qualitative analysis and take into account many (not always independent) variables.

- Most generalisations about "linguistic complexity" in the literature seem to be premature given both conceptual shortcomings and the quality of empirical coverage.

Envoy

- Approaches to "linguistic complexity" reveal epistemological and methodological problems that linguistics shares with other disciplines.

Envoy

- Approaches to "linguistic complexity" reveal epistemological and methodological problems that linguistics shares with other disciplines.

- An area of possible mutually-enriching interactions?

Thank you for your attention!
Danke für Ihre Aufmerksamkeit!

References

Ackerman F. & R. Malouf. 2013. Morphological organization: The low conditional entropy conjecture. *Language* 89.3, 429–464. https://doi.org/10.1353/lan.2013.0054

Bentz Chr., X. Gutierrez-Vasques, O. Sozinova & T. Samardžić. 2023. Complexity trade-offs and equi-complexity in natural languages: a meta-analysis. *Linguistics Vanguard special issue on Linguistic Complexity* 9, 9–25. https://doi.org/10.1515/lingvan-2021-0054

Deutscher G. 2009. "Overall complexity": a wild goose chase? In: G. Sampson, D. Gil & P. Trudgill (eds.) *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 243–251.

Ehret K., A. Blumenthal-Dramé, Chr. Bentz & A. Berdicevskis. 2021. Meaning and measures: Interpreting and evaluating complexity metrics. *Frontiers in Communication special issue on language complexity* 6, art. 640510. https://doi.org/10.3389/fcomm.2021.640510

Enfield N. J. 2007. *A Grammar of Lao*. Berlin, New York: Mouton de Gruyter.

Foley W. A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.

Foley W. A. 1991. *The Yimas Language of New Guinea*. Stanford, CA: Stanford University Press.

References

Forker D. 2021. Complexity and its relation to variation. *Frontiers in Communication special issue on language complexity* 6, art. 632468. https://doi.org/10.3389/fcomm.2021.632468

Fortson B. W., IV. 2010. *Indo-European Language and Culture. An Introduction*. 2nd ed. Chichester: Wiley Blackwell.

Hengeveld K. & S. Leufkens. 2018. Transparent and non-transparent languages. *Folia Linguistica* 52.1, 139–175. https://doi.org/10.1515/flin-2018-0003

Hockett Ch. F. 1958. *A Course in Modern Linguistics*. New York: MacMillan.

Joseph J. E. 2021. Why does language complexity resist measurement? *Frontiers in Communication special issue on language complexity* 6, art. 624855. https://doi.org/10.3389/fcomm.2021.624855

Joseph J. E. & Fr. J. Newmeyer. 2012. 'All languages are equally complex'. The rise and fall of a consensus. *Historiographia Linguistica* 39.3, 341–368. https://doi.org/10.1075/hl.39.2-3.08jos

Kibrik A. A. 2012. What's in the head of head-marking languages? In: P. Suihkonen, B. Comrie & V. Solovyev (eds.), *Argument Structure and Grammatical Relations: A Cross-Linguistic Typology*. Amsterdam, Philadelphia: John Benjamins, 211–240.

References

Kortmann B. & V. Schröter. 2020. "Linguistic Complexity." In: M. Aronoff (ed.), *Oxford Bibliographies in Linguistics*. New York: Oxford University Press. Last modified: 15 January 2020; DOI: 10.1093/obo/9780199772810-0254

Kortmann B. & B. Szmrecsanyi. 2009. World Englishes between simplification and complexification. In: Th. Hoffmann & L. Siebers (eds.), *World Englishes – Problems, Properties and Prospects: Selected papers from the 13th IAWE conference*. Amsterdam, Philadelphia: John Benjamins, 265–285.

Kuiken F. 2023. Linguistic complexity in second language acquisition. *Linguistics Vanguard special issue on Linguistic Complexity* 9, 83–93. https://doi.org/10.1515/lingvan-2021-0112

Kusters W. 2003. *Linguistic Complexity. The Influence of Social Change on Verbal Inflection*. Utrecht: LOT Publications.

Leufkens S. 2023. Measuring redundancy: the relation between concord and complexity. *Linguistics Vanguard special issue on Linguistic Complexity* 9, 95–106. https://doi.org/10.1515/lingvan-2020-0143

McWhorter J. *Linguistic Simplicity and Complexity. Why Do Languages Undress?* Berlin, Boston: De Gruyter Mouton.

References

Moroz G. 2017. lingtypology: easy mapping for Linguistic Typology. https://CRAN.R-project.org/package=lingtypology

Mufwene S., Fr. Pellegrino & Chr. Coupé. 2017. Complexity in language: A multifaceted phenomenon. In: S. Mufwene, Fr. Pellegrino & Chr. Coupé (eds.), *Complexity in Language. Developmental and Evolutionary Perspectives.* Cambridge: Cambridge University Press, 1–29.

Rescher N. 1998. *Complexity. A Philosophical Overview*. New Brunwick, London: Transaction Publishers.

Saltarelli M. 1988. *Basque*. London etc.: Croom Helm.

Toldova S. et al. 2018. *Èlementy mokšanskogo jazyka v tipologičeskom osveščenii* [*Elements of Moksha in a Typological Perspective*]. Moscow: Buki-Vedi.