Restrictions on the ontological category of indefinite pronoun series in the languages of Europe

ConSOLe XXVI, UCL

Daria Bikina
daria.bikina@gmail.com

HSE School of Linguistics (Moscow)

February 14-16, 2018
Outline

Introduction
Outline

Introduction

Data and methods
Outline

Introduction

Data and methods

Characteristics of pronominal series

Characteristics of pronominal systems

Discussion
Outline

Introduction

Data and methods

Characteristics of pronominal series

Characteristics of pronominal systems
Outline

Introduction

Data and methods

Characteristics of pronominal series

Characteristics of pronominal systems

Discussion
Forming indefinite pronouns

[Haspelmath 1997]: 22

Indefinite pronouns normally occur in **series** which have one member for each of the major **ontological categories** such as person, thing, property, place, time, manner, amount, plus a few others.

... In the most common case, indefinite pronouns consist of (i) a stem indicating ontological category, plus (ii) a formal element shared by all members of indefinite pronoun series, such as *some-* and *any-* in English. ... This element will be called **indefiniteness marker**.
Forming indefinite pronouns

However, an indefinite pronoun system may not be a simple multiplication of possible stems and indefiniteness markers.
Forming indefinite pronouns
Not just a multiplication: three Russian series

<table>
<thead>
<tr>
<th>Category</th>
<th>koe-series</th>
<th>to-series</th>
<th>libo-series</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON</td>
<td>koe-kto</td>
<td>kto-to</td>
<td>kto-libo</td>
</tr>
<tr>
<td>THING</td>
<td>koe-čto</td>
<td>čto-to</td>
<td>čto-libo</td>
</tr>
<tr>
<td>PLACE.IN</td>
<td>koe-gde</td>
<td>gde-to</td>
<td>gde-libo</td>
</tr>
<tr>
<td>PLACE.ILL</td>
<td>koe-kuda</td>
<td>kuda-to</td>
<td>kuda-libo</td>
</tr>
<tr>
<td>PLACE.EL</td>
<td>koe-otkuda</td>
<td>otkuda-to</td>
<td>otkuda-libo</td>
</tr>
<tr>
<td>TIME</td>
<td>*koe-kogda</td>
<td>kogda-to</td>
<td>kogda-libo</td>
</tr>
<tr>
<td>MANNER</td>
<td>#koe-kak</td>
<td>kak-to</td>
<td>kak-libo</td>
</tr>
<tr>
<td>REASON</td>
<td>*koe-počemu</td>
<td>počemu-to</td>
<td>počemu-libo</td>
</tr>
<tr>
<td>AMOUNT</td>
<td>?koe-skol’ko</td>
<td>skol’ko-to</td>
<td>?skol’ko-libo</td>
</tr>
<tr>
<td>PROPERTY</td>
<td>#koe-kakoj</td>
<td>kakoj-to</td>
<td>kakoj-libo</td>
</tr>
<tr>
<td>CHOICE FROM A SET</td>
<td>koe-kotoryj</td>
<td>*kotoryj-to</td>
<td>?kotoryj-libo</td>
</tr>
<tr>
<td>POSSESSION</td>
<td>koe-čej</td>
<td>čej-to</td>
<td>čej-libo</td>
</tr>
</tbody>
</table>
## Forming indefinite pronouns

The problem of different stems: two English series

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Interrogative stem</th>
<th>Nominal stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON</td>
<td><em>anywho</em></td>
<td>anybody</td>
</tr>
<tr>
<td></td>
<td><em>somewho</em></td>
<td>somebody</td>
</tr>
<tr>
<td>THING</td>
<td><em>anywhat</em></td>
<td>anything</td>
</tr>
<tr>
<td></td>
<td>#somewhat</td>
<td>something</td>
</tr>
<tr>
<td>PLACE</td>
<td>anywhere</td>
<td>—</td>
</tr>
<tr>
<td>TIME</td>
<td><em>anywhen</em></td>
<td>anytime</td>
</tr>
<tr>
<td></td>
<td><em>somewhen</em></td>
<td>sometime</td>
</tr>
<tr>
<td>MANNER</td>
<td>anyhow</td>
<td>—</td>
</tr>
<tr>
<td>REASON</td>
<td><em>anywhy</em></td>
<td>—</td>
</tr>
<tr>
<td></td>
<td><em>somewhy</em></td>
<td>—</td>
</tr>
</tbody>
</table>

Daria Bikina  
daria.bikina@gmail.com  
HSE School of Linguistics (Moscow)  
Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
The current research

- Do the gaps in indefinite pronoun systems occur accidentally?
- If not, what triggers the absence of a specific form in the system?
  - type of a stem?
  - properties of the ontological category?
  - semantics of indefinite pronoun series?
  - language properties?
  - areal influence?
- What theoretical implications do the results provide?
Data

A sample of 21 European languages

- Partially based on a Haspelmath’ sample ([Haspelmath 1997]: 244—317)
- This data has been extended and re-checked with native speakers to establish gaps

- English
- German
- Dutch
- Swedish
- French
- Italian
- Modern Greek
- Lithuanian
- Russian
- Ukrainian
- Polish
- Czesh
- Serbian
- Bulgarian
- Finnish
- Hungarian
- Moksha Mordvin
- Western Mari
- Basque
- Turkish
- Georgian

Daria Bikina  dariabikina@gmail.com

Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
Language sample
Ontological categories

- The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008]).

I use the include-all-you-can approach: extended set

- A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.

- place: location, place: direction, and place: source are distinguished in many Slavic languages, including Russian: gde 'where', kuda 'to where', otkuda 'where from'.

- A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.

- property vs. choice from a set: English what kind vs. which.

- possessor: Polish kto 'who' vs. czyj 'whose'.

- verb for a pronominal verb: Moksha Mordvin mej-t'@ms (what-vbz-inf) 'to do something'.

Daria Bikina daria.bikina@gmail.com
HSE School of Linguistics (Moscow)
Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
Ontological categories

- The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008])
- I use the include-all-you-can approach: extended set
Ontological categories

- The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008])
- I use the include-all-you-can approach: extended set
  - A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.
Ontological categories

The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008])

I use the include-all-you-can approach: extended set

- A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.

- **PLACE:** LOCATION, **PLACE:** DIRECTION and **PLACE:** SOURCE are distinguished in many Slavic languages, including Russian: *gde* 'where', *kuda* 'to where', *otkuda* 'where from'
Ontological categories

- The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008])
- I use the include-all-you-can approach: extended set
  - A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.
  - PLACE: LOCATION, PLACE: DIRECTION and PLACE: SOURCE are distinguished in many Slavic languages, including Russian: *gde* 'where', *kuda* 'to where', *otkuda* 'where from'
- PROPERTY vs. CHOICE FROM A SET: English *what kind* vs. *which*
Ontological categories

- The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008])

- I use the include-all-you-can approach: extended set
  - A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.
  - **PLACE**: LOCATION, PLACE: DIRECTION and PLACE: SOURCE are distinguished in many Slavic languages, including Russian: *gde* 'where', *kuda* 'to where', *otkuda* 'where from'
  - **PROPERTY vs. CHOICE FROM A SET**: English *what kind vs. which*
  - **POSSESSOR**: Polish *kto* 'who' vs. *czyj* 'whose'
The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008])

I use the include-all-you-can approach: extended set

- A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.
- PLACE: LOCATION, PLACE: DIRECTION and PLACE: SOURCE are distinguished in many Slavic languages, including Russian: *gde* 'where', *kuda* 'to where', *otkuda* 'where from'
- PROPERTY vs. CHOICE FROM A SET: English *what kind* vs. *which*
- POSSESSOR: Polish *kto* 'who' vs. *czyj*'whose'
- VERB for a pronominal verb: Moksha Mordvin *mej-t’-əms* (what-VBZ-INF) 'to do something'
Ontological categories

- The specific set of ontological categories is under discussion ([Jackendoff 1990], [Haspelmath 1997], [Award 2001], [Hengeveld, Mackenzie 2008], [Hengeveld, Mackenzie 2008])
- I use the include-all-you-can approach: extended set
  - A category is distinguished if there is a language in the sample which has a specific interrogative form of this category.
  - PLACE: LOCATION, PLACE: DIRECTION and PLACE: SOURCE are distinguished in many Slavic languages, including Russian: *gde* 'where', *kuda* 'to where', *otkuda* 'where from'
  - PROPERTY vs. CHOICE FROM A SET: English *what kind vs. which*
  - POSSESSOR: Polish *kto* 'who' vs. *czyj* 'whose'
  - VERB for a pronominal verb: Moksha Mordvin *mej-t’-əms* (what-VBZ-INF) 'to do something')
## Ontological categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERSON</td>
<td>who</td>
</tr>
<tr>
<td>THING</td>
<td>what</td>
</tr>
<tr>
<td>PLACE: LOCATION</td>
<td>where</td>
</tr>
<tr>
<td>PLACE: DIRECTION</td>
<td>to where</td>
</tr>
<tr>
<td>PLACE: SOURCE</td>
<td>where from</td>
</tr>
<tr>
<td>TIME</td>
<td>when</td>
</tr>
<tr>
<td>MANNER</td>
<td>how</td>
</tr>
<tr>
<td>AMOUNT</td>
<td>how many</td>
</tr>
<tr>
<td>REASON</td>
<td>why</td>
</tr>
<tr>
<td>PROPERTY</td>
<td>what (kind)</td>
</tr>
<tr>
<td>CHOICE FROM A SET</td>
<td>which</td>
</tr>
<tr>
<td>POSSESSOR</td>
<td>whose</td>
</tr>
<tr>
<td>(VERB)</td>
<td>to do what</td>
</tr>
</tbody>
</table>
Factors discussed

- Characteristics of pronoun series
Factors discussed

- Characteristics of pronoun series
Factors discussed

- Characteristics of pronoun series
  - Type of the stem
Factors discussed

- Characteristics of pronoun series
  - Type of the stem
  - Semantics of the series
Factors discussed

- Characteristics of pronoun series
  - Type of the stem
  - Semantics of the series
- Characteristics of pronominal systems of the sample
Factors discussed

- Characteristics of **pronoun series**
  - Type of the stem
  - Semantics of the series
- Characteristics of **pronominal systems** of the sample
  - Absence of a non-derived (simple) interrogative of a particular category
Factors discussed

- Characteristics of **pronoun series**
  - Type of the stem
  - Semantics of the series
- Characteristics of **pronominal systems** of the sample
  - Absence of a non-derived (simple) interrogative of a particular category
  - Number of different stem types
Factors discussed

- Characteristics of **pronoun series**
  - Type of the stem
  - Semantics of the series

- Characteristics of **pronominal systems** of the sample
  - Absence of a non-derived (simple) interrogative of a particular category
  - Number of different stem types
Type of the stem

▶ Interrogative / relative pronoun
e.g., German irgendwer, irgendwas
Type of the stem

- Interrogative / relative pronoun
e.g., German irgendwer, irgendwas

- Another indefinite pronoun
e.g., German jemand → irgendjemand;
Hill Mari ma-gən’ät (what-INDF1) → ta-ma-gən’ät (INDF2-what-INDF1)
Type of the stem

- **Interrogative / relative pronoun**
  e.g., German irgendwer, irgendwas

- **Another indefinite pronoun**
  e.g., German jemand → irgendjemand;
  Hill Mari ma-gən’ät (what-INDF1) → ta-ma-gən’ät
  (INDF2-what-INDF1)

- **Generic noun**
  e.g., English anybody, anytime
Type of the stem

- Interrogative / relative pronoun
e.g., German irgendwer, irgendwas

- Another indefinite pronoun
e.g., German jemand → irgendjemand;
Hill Mari ma-gən’ät (what-INDF1) → ta-ma-gən’ät (INDF2-what-INDF1)

- Generic noun
e.g., English anybody, anytime

- Numeral 'one'
e.g., English anyone, French quelqu’un
Type of the stem

- Interrogative / relative pronoun
e.g., German irgendwer, irgendwas

- Another indefinite pronoun
e.g., German jemand → irgendjemand;
Hill Mari ma-gən'ät (what-INDF1) → ta-ma-gən'ät (INDF2-what-INDF1)

- Generic noun
e.g., English anybody, anytime

- Numeral 'one'
e.g., English anyone, French quelqu’un

- Null stem (any, some as PROPERTY, CHOICE FROM A SET and AMOUNT in English)
Type of the stem

- Interrogative / relative pronoun
  e.g., German irgendwer, irgendwas

- Another indefinite pronoun
  e.g., German jemand → irgendjemand;
  Hill Mari ma-gən’ät (what-INDF1) → ta-ma-gən’āt (INDF2-what-INDF1)

- Generic noun
  e.g., English anybody, anytime

- Numeral 'one'
  e.g., English anyone, French quelqu’un

- Null stem (any, some as PROPERTY, CHOICE FROM A SET and AMOUNT in English)

- Special stem type (non-regular)
Type of the stem

![Graph showing the distribution of stem types]

- **Count**: yes
- **Count**: no

- **Indefinite**: few
- **Interrogative**: high
- **Noun**: moderate
- **Null**: low
- **Numerical**: low
- **Relative**: low
- **Special**: low

Daria Bikina  
daria.bikina@gmail.com  
HSE School of Linguistics (Moscow)  
Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
Type of the stem

Pronominal series based on numeral 'one' often are restricted to PERSON and CHOICE FROM A SET, which is probably associated with individuality.
The sample of series was very roughly divided into 4 groups:

- Specific indefinites
- Non-specific indefinites (used in non-veridical contexts)
- Free-choice indefinites
- Negative indefinites
Semantics of series

- The results showed that in general, semantics does not correlate with the absence of particular indefinite pronoun forms.
Semantics of series

- The results showed that in general, semantics does not correlate with the absence of particular indefinite pronoun forms.
- Free-choice indefinites show fewer gaps due to the fact that many of them are sluicing-based, and thereof less grammaticalized, which implies the lack of restrictions.
Semantics of series

- The results showed that in general, semantics does not correlate with the absence of particular indefinite pronoun forms.
- Free-choice indefinites show fewer gaps due to the fact that many of them are sluicing-based, and thereof less grammaticalized, which implies the lack of restrictions.

**Figure:** Desemanticization of indefinite pronouns ([Haspelmath 1997])
Figure: Proportion of absent and present indefinite pronoun forms by ontological category
Absence of a non-derived base interrogative

<table>
<thead>
<tr>
<th></th>
<th>Non-derived form</th>
<th>Derived form</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUANTITY</td>
<td>French: <em>combien</em></td>
<td>English: <em>how many</em></td>
</tr>
<tr>
<td>REASON</td>
<td>English: <em>why</em></td>
<td>Bulgarian: <em>za-štto</em> (for-what)</td>
</tr>
<tr>
<td>POSSESSOR</td>
<td>Russian: <em>čej</em></td>
<td>Moksha: <em>kin’</em> (who.Gen)</td>
</tr>
</tbody>
</table>

Daria Bikina  
daria.bikina@gmail.com
HSE School of Linguistics (Moscow)

Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
Absence of a non-derived base interrogative: POSSESSOR

Figure: Percentage of absent forms

Languages with special POSSESSOR pronouns; prepositional constructions; case forms of 'who' / 'which'.
Absence of a non-derived base interrogative: POSSESSOR → PERSON
Absence of a non-derived base interrogative: REASON

Figure: Percentage of absent forms

Languages with special REASON pronouns; prepositional constructions; dative/causal forms of 'what'.

Daria Bikina  daria.bikina@gmail.com
HSE School of Linguistics (Moscow)
Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
Absence of a non-derived base interrogative: REASON

- REASON indefinite pronouns show unexpected behaviour: not the absence, but the presence of a non-derived REASON interrogative triggers the absence of a corresponding indefinite form

Daria Bikina  
daria.bikina@gmail.com

Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
Absence of a non-derived base interrogative: REASON

- REASON indefinite pronouns show unexpected behaviour: not the absence, but the presence of a non-derived REASON interrogative triggers the absence of a corresponding indefinite form

- One more argument in favour of originality of why-pronouns ([Bromberger 1992], [de Villiers 1991], [de Villiers 1996], [Rizzi 2001], [Thornton 2008])
Absence of a non-derived base interrogative: REASON

- REASON indefinite pronouns show unexpected behaviour: not the absence, but the presence of a non-derived REASON interrogative triggers the absence of a corresponding indefinite form

- One more argument in favour of originality of why-pronouns ([Bromberger 1992], [de Villiers 1991], [de Villiers 1996], [Rizzi 2001], [Thornton 2008])
Number of possible stems

Figure: Percentage of absent forms by number of possible stems (1, 2, 3, 4) and by language: PERSON, TIME, REASON, MANNER
PCA analysis
Discussion

- These are the properties of languages and not the properties of particular indefinite pronoun series that influence the way a system of indefinite pronouns is organized.
Discussion

- These are the properties of languages and not the properties of particular indefinite pronoun series that influence the way a system of indefinite pronouns is organized.
- Our results make a contribution to the idea of how the ontological categories may be ranked.
Discussion

- These are the properties of languages and not the properties of particular indefinite pronoun series that influence the way a system of indefinite pronouns is organized.

- Our results make a contribution to the idea of how the ontological categories may be ranked. [Mackenzie 2009]: There is a hierarchy of semantic categories that reflects the level of their cognitive complexity:

\[
\text{individual} \subset \text{place} \subset \text{time} \subset \text{manner} \subset \text{quantity} \subset \text{reason}
\]

Daria Bikina  
daria.bikina@gmail.com  
HSE School of Linguistics (Moscow)

Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
Does the percentage of gaps correlate with the frequency of corresponding stems of different ontological categories?
Does the percentage of gaps correlate with the frequency of corresponding stems of different ontological categories?

What exactly in the semantics of reason pronouns conflicts with the ability to form indefinite pronouns?
Acknowledgements

I would like to thank:

▶ Maria Kholodilova for supervising me,
▶ George Moroz for helping me with the quantitative part of the research,
▶ the audience of TyLex’17 at HSE for useful comments,
▶ and all my language consultants for their inestimable help.

This research has been supported by Russian Foundation for Basic Research, project No. 16-06-00536.
Thank you for your attention!
References


Daria Bikina  dariabikina@gmail.com
HSE School of Linguistics (Moscow)
Restrictions on the ontological category of indefinite pronoun series in the languages of Europe
References


References
