A Collaborative Project for the Documentation of the Shughni Language

Amanda Barie, Darya Bukhtoyarova, Raphael Finkel, Andrew Hippley, Mark Richard Lauersdorf, Jeanmarie Rouhier-Willoughby, Gregory Stump, University of Kentucky

Guqblismo Alamshoov*, Shohnazar Mirzoev*, Gulnoro Mirzovafoeva, Shahlo Nekshoovae*, Khorgo State University, *Institute of Humanities of the Tajik Academy of Sciences

Project Participants

In Tajikistan

Muqblismo Alamshoov, Head of the Department of Pamirian Languages, Institut of Humanities of the Tajik Academic of Sciences, Khorgo State University
Shohnazar Mirzoev, Vice Director of the Institute of Humanities of the Tajik Academy of Sciences, Foreign Language Department, Khorgo State University
Gulnoro Mirzovafoeva, English Department, Khorgo State University
Shahlo Nekshoova, Department of Tajik Language and Literature, Khorgo State University; student at the Institute of Humanities of the Tajik Academy of Sciences.

In Kentucky

Stanley Brunn, Professor of Geography, University of Kentucky
Raphael Finkel, Professor of Computer Science, University of Kentucky
Andrew Hippley, Associate Professor of Linguistics, University of Kentucky
Diane King, Assistant Professor of Anthropology, University of Kentucky
Mark Richard Lauersdorf, Associate Professor of Languages and Linguistics, University of Kentucky
Jeanmarie Rouhier-Willoughby, Associate Professor of Russian and Linguistics, University of Kentucky
Igor V. Sopronenko, Documentary Film Producer/Director, Signatory Media Production LLC, Lexington, Kentucky
Gregory Stump, Professor of Linguistics, University of Kentucky

Project Resources

The Collaboratory for Research in Computing for Humanities (RCH), a research unit of the University of Kentucky, brings together faculty and students from engineering, computer science, and the humanities to develop tools to facilitate and support collaborative research projects in the Digital Humanities, and encourages and supports interdisciplinary projects among partners from UK and around the world.

The main lab (the "Digital Research Incubator") of the RCH provided an ideal environment and testbed for digitally-enhanced research collaboration during the Shughni Workshop in July 2008.

• all work sessions were recorded in digital audio and video and stored as lossless, archive-quality media files
• all elicited examples were keyboarded directly as digital text and projected onto the whiteboard for immediate group verification, correction, annotation, and discussion, and archived as written documentation of sessions
• verbal morphology data were entered directly into KATR (a set-based extension of DATS) for on-the-fly paradigm generation and verification, as well as immediate adjustment of KATR rules
• all digital materials were immediately uploaded onto a networked server for anytime/anywhere retrieval by team members

KATR

KATR is a language for lexical knowledge representation designed and implemented by Roger Evans and Gerald Gardner (Evans & Gardner 1996). KATR is an extension of DATS designed and implemented by Raphael Finkel and Gregory Stump: its main deviation from DATS is that it allows a word's morphosemantic properties to be represented as an unordered set. Finkel has used KATR to devise a formal definition of a fragment of Shughni verb morphology; this definition allows verb paradigms to be computed automatically. This facilitates the checking of hypothesised forms with Shughni and is an important step towards the development of an extremely precise model of Shughni morphology. Below is a paradigm of forms which Finkel's program generates for the Shughni verb 'mam':'

Outcomes and Student Research

Some preliminary outcomes of our research

Andrew Hippley & Gregory Stump, 'Periphrasis and Shughni verb inflection', paper presented at the 1st International Conference on Language Documentation and Conservation University of Hawai‘i at Mānoa, March 12-14, 2009

Student research

Since the inception of our work on Shughni, student participation has been an important component of the project. In the spring semester of 2007, Greg Stump had students in his Grammatical Analysis course investigate Shughni morphosyntax through the field: Language modeling for elicitation and documentation of Shughni, paper presented at the 1st International Conference on Language Documentation and Conservation University of Hawai‘i at Mānoa, March 12-14, 2009

Singular

<table>
<thead>
<tr>
<th>Stem</th>
<th>1sg</th>
<th>2sg</th>
<th>3sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
</tbody>
</table>

Plural

<table>
<thead>
<tr>
<th>Stem</th>
<th>1pl</th>
<th>2pl</th>
<th>3pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
</tbody>
</table>


date

The Collaboratory for Research in Computing for Humanities (RCH), a research unit of the University of Kentucky, brings together faculty and students from engineering, computer science, and the humanities to develop tools to facilitate and support collaborative research projects in the Digital Humanities, and encourages and supports interdisciplinary projects among partners from UK and around the world.

The main lab (the "Digital Research Incubator") of the RCH provided an ideal environment and testbed for digitally-enhanced research collaboration during the Shughni Workshop in July 2008.

• all work sessions were recorded in digital audio and video and stored as lossless, archive-quality media files
• all elicited examples were keyboarded directly as digital text and projected onto the whiteboard for immediate group verification, correction, annotation, and discussion, and archived as written documentation of sessions
• verbal morphology data were entered directly into KATR (a set-based extension of DATS) for on-the-fly paradigm generation and verification, as well as immediate adjustment of KATR rules
• all digital materials were immediately uploaded onto a networked server for anytime/anywhere retrieval by team members

KATR

KATR is a language for lexical knowledge representation designed and implemented by Roger Evans and Gerald Gardner (Evans & Gardner 1996). KATR is an extension of DATS designed and implemented by Raphael Finkel and Gregory Stump: its main deviation from DATS is that it allows a word's morphosemantic properties to be represented as an unordered set. Finkel has used KATR to devise a formal definition of a fragment of Shughni verb morphology; this definition allows verb paradigms to be computed automatically. This facilitates the checking of hypothesised forms with Shughni and is an important step towards the development of an extremely precise model of Shughni morphology. Below is a paradigm of forms which Finkel's program generates for the Shughni verb 'mam':'

Outcomes and Student Research

Some preliminary outcomes of our research

Andrew Hippley & Gregory Stump, 'Periphrasis and Shughni verb inflection', paper presented at the Southeast Morphology Meeting, University of Surrey, Guildford, United Kingdom, September 26, 2008.
Andrew Hippley & Gregory Stump, 'The wonders of morphology: The Shughni past tense', English Department Colloquium, University of Kentucky, November 14, 2008

Student research

Since the inception of our work on Shughni, student participation has been an important component of the project. In the spring semester of 2007, Greg Stump had students in his Grammatical Analysis course investigate Shughni morphosyntax through the field: Language modeling for elicitation and documentation of Shughni, paper presented at the 1st International Conference on Language Documentation and Conservation University of Hawai‘i at Mānoa, March 12-14, 2009

Singular

<table>
<thead>
<tr>
<th>Stem</th>
<th>1sg</th>
<th>2sg</th>
<th>3sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
</tbody>
</table>

Plural

<table>
<thead>
<tr>
<th>Stem</th>
<th>1pl</th>
<th>2pl</th>
<th>3pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
<tr>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
<td>wuz</td>
</tr>
</tbody>
</table>