

Carrying across Language and Code

*Nick Montfort and
Natalia Fedorova*

July 2012

TROPE-12-04

Abstract

With reference to electronic literature translation projects in which we have been involved as translators or as authors of the source work, we argue that the process of translation can expose how language and computation interrelate in electronic literature.

(Prepared for the Translating E-Literature Conference, Université Paris 8 Vincennes-Saint-Denis, June 12-14, 2012)

A technical report from

The Trope Tank
Massachusetts Institute of Technology
77 Massachusetts Ave, 14N-233
Cambridge, MA 02139 USA
<http://trope-tank.mit.edu>

© 2012 Nick Montfort and Natalia Fedorova
This work is licensed under the Creative Commons
Attribution-ShareAlike 3.0 Unported License.
To view a copy of this license, visit:
<http://creativecommons.org/licenses/by-sa/3.0/>
or send a letter to Creative Commons, 444 Castro Street,
Suite 900, Mountain View, California, 94041, USA.

We discuss several electronic literature projects in which we have been involved as translators or as authors of the source work. Our emphasis will be on computational works rather than ones involving complex surfaces of text or multimedia elements.

- The transliteration of "Concrete Perl," a set of concrete poems realized as Perl programs, from the Latin alphabet to Cyrillic.
- The translation of "The Two," a simple poetry generator with narrativity and certain design ambiguities, from English to French, Spanish and Russian.
- The translation of "Through the Park," a simple narrative generator, from English to Russian, with some discussion of translations of a remix of this generator.
- The translation of "Epigraphic Clock," a cybertext poem, from English to Russian.
- The translation of two interactive fictions, focusing on one that has been completed: "Olvido Mortal," which was translated from Spanish to English.

The first three are poetry generators that operate in the level of letter, word, and sentence; "Epigraphic Clock" works mainly on words, but also at the letter level. Finally, the two interactive fiction pieces incorporate elaborate models of simulated worlds and of language.

"Concrete Perl" by Nick Montfort

http://nickm.com/poems/concrete_perl.html

Transliterated to Cyrillic by Nick Montfort & Natalia Fedorova

Difficulties arise in translation, or in this case in *transliteration*, even when the engagement with language is completely at the lexical, alphabetic level. The four programs that constitute "Concrete Perl" programs are 32 character-long machines (32-character Perl programs) for generating concrete poems. We found it impossible to convert some of these programs to a non-Latin alphabet (one not represented in ASCII) and keep the code as short as it was originally.

Because the constraint used in the original program is defined over both code and text (represented in strings), a change in either one — and certainly a change in both — brings the program up against, or out of, the constraint in new and different ways.

Alphabet Expanding

```
{print$,$"x($+=.01),a..z;redo}
```

Расширение алфавита

```
{binmode STDOUT,utf8;print"x($+=.01),chr for 1072..1103;redo}
```

All the Names of God

```
{print"a"x++$..."x$,$,=,_,redo}
```

Все имена бога

```
sub n{for(1072..1103){if($_[1]){push@_,n($_[0].chr,$_[1]-1)}else{print$_[0].chr._}}return @_}{binmode STDOUT,utf8;n("","$.++");redo}
```

Letterformed Terrain

```
{print$,$_=(a..z)[rand$=];redo}
```

Буквенный ландшафт

```
{binmode STDOUT,utf8;print$,chr((1072..1103)[rand$=]);redo}
```

The only case in which the length of transliterated program is the same is in “Gegemonia ASCII” (“ASCII Hegemony”). This poem portrays the dogged, insistent, inescapable march of every printable symbol in the classic ASCII character set, an American encoding that came to be used globally. The new version of the poem is simply the same as the original one, as if denying the possibility of other alphabets. The translated title of this program, in submission to the hegemony of ASCII, has been transliterated to the Latin alphabet instead of being written in Cyrillic. It is also consistent with one of the first works of Russian electronic literature, “ROMAN” (“The Novel”), a collective hypertext by Roman Leibov and others. In this work, the entire Russian language text was represented in its ASCII transliteration.

ASCII hegemony

```
{print" ".chr for 32..126;redo}
```

Gegemonia ASCII

```
{print" ".chr for 32..126;redo}
```

In this case, our seemingly straightforward and very minor translation project, which was really a transliteration project, nevertheless revealed that conventional and popular programming languages are not neutral on the matter of human languages or even alphabets. To switch out of the Roman Alphabet, the plain ASCII encoding for “Global English,” requires extra overhead and a longer program, even if one does not mind using the obligatory English keywords.

“The Two” by Nick Montfort

http://nickm.com/poems/the_two.html

Translated to French by Serge Bouchardon

Translated to Spanish by Carlos León

Translated to Russian by Natalia Fedorova

“The Two” is a poetry generator that produces stanzas that are also stories. Each “stanzory” has a first line describing people by profession or role and a second line with generated pronouns. The reader is left to resolve this reference. In cases where one masculine and one feminine pronoun is produced, the reader must rely on cultural stereotypes or some other means of reasoning. In translating the textons of “The Two” to French, Spanish, and Russian, it was necessary to somehow work around the way

that these languages almost always specify gender when introducing a person by profession or role.

To translate this work into French, a devious technique is needed — one which Serge Bourchardon discovered and used. First, use only those few nouns designating professions and roles that are not themselves specific to a gender: “dentiste” is an example. But if this word were to appear in the poem, it would require a definite article in front of it, “le” or “la,” and this article would reveal the gender of the dentist in question.

So, a further constraint is employed: only those professions or roles that begin in vowels are used. If the noun begins in a vowel, the article used is “l’” and does not reveal whether the noun is masculine or feminine. The dentist does not appear, but the automobile driver, the economist, the entrepreneur, the anesthesiologist, and other characters whose names begin in vowels in French are there on the first line of the stanzas. In a few cases, nouns that are grammatically masculine but which are used to refer to both men and women are used.

In Spanish, the task of translating such fragments of text without referring to gender is even harder. There are words for professions and roles that are not inherently masculine or feminine, such as *agente*, but in Spanish, even though *agente* begins with a vowel, it must appear as “el agente” or “la agente” and so the gender of the agent must be revealed. Carlos León used “la persona” now and then and in other cases simply evokes the situation in the first line, capturing the original situations without naming the two participants with specific nouns.

police officer — perpetrator
shopper — cashier
babysitter — child
indigent — librarian
driver — pedestrian
rescuer — survivalist

l'agent de police — l'automobiliste
le procurer — l'agresseur
l'éclargist — l'équilibriste
le medicine — l'anesthesiste
l'entraîneur — l'haltérophile
l'élève — professeur
l'économiste — l'entrepreneur

Desde su silla de juez se acerca a la silla de los acusados
Al aproximarse al cajero, mira a la persona encargada de cobrar
Deja de lavar la ropa y se le acerca mientras juega con sus juguetes
Al salir de clase, llama a la puerta del despacho
Aunque vive en la calle, se gira hacia la persona que está en el mostrador de la librería
Desde su coche llama a la persona que cruza la calle
Tras separarse del resto del equipo de rescate, encuentra a la persona extraviada

Translating the names of professions into Russian requires making stylistical choices. In Russian, there are professions indicated with masculine and feminine nouns as well as “non-feminine” and “non-masculine” professions, either of which can be paired (meaning that there is a term for the other gender) or not. In official and business writing, “non-feminine” professions such as *служитель правопорядка* (police officer) and *спасатель* (rescuer) are grammatically masculine. So-called unpaired professions such as *водитель* (driver) do not even have synthetic feminine forms at all. On the other hand, “non-masculine” professions such as *няня* (nurse) do not have a synthetic masculine form. In semi-formal or colloquial style feminine versions of the paired professions can be used: *покупатель – покупательница* (client), *кассир – кассирша* (cashier), *студент – студентка* (student), *учитель – учительница* (teacher), *библиотекарь – библиотекарьша* (librarian). And the word *жертва* (victim) is, predictably, always feminine. The feminine forms of the paired professions do designate women performing these professional roles, but have a connotation of familiarity. The neutral and official way of describing people by their professions would still be masculine.

So by choosing to use the nouns in masculine gender Fedorova made the first two lines sound like a neutral report of court testimony. However, the third line, where the resolution is given, can hardly be imagined to be in official or business discourse. It introduces the third voice of an external and a more informed observer, making the text polyphonic.

In Russian — as in the other two languages into which “The Two” was translated — the differences in syntax require some changes in code. In Russian, for instance, the agreement of transitive and intransitive verbs demands different forms of pronoun or a pronoun and a preposition. The language data are therefore represented differently in the original and in the Russian translation, and there is different code to process and display this data.

berates
begs
confesses to
asks for advice from
seeks help from
expresses sympathy to
smacks
defies
surrenders to
hugs

('упрекает', 'его.', 'ее.')
('умоляет', 'его.', 'ее.')
('ударяет', 'его.', 'ее.')
('обнимает', 'его.', 'ее.')
('просит', 'его о помощи.', 'ее о помощи.')
('открывается', 'ему.', 'ей.')
('сочувствует', 'ему.', 'ей.')
'

('сдается','ему.','ей.')

('бросает','ему вызов.','ей вызов.')

('спрашивает','у него совета.','у нее совета.')

In a nutshell, the French, Spanish and Russian translations, given the constraints and new connotations presented in the target language, suggest that word-level text generating machines involve significant engagement between the language and computation, engagement which is revealed in the process of translation. Preserving the meaning involves both changes in code and sentence construction.

“Through the Park” by Nick Montfort
http://nickm.com/poems/through_the_park.html
Translated to Russian by Natalia Fedorova

“Excerpts from the Chronicles of Pookie & JR” by J.R.Carpenter
<http://luckysoup.com/pookieandjr/>
Translated to Spanish by Laura Borràs Castanyer
Translated to Catalan by Laura Borràs Castanyer
Translated to Russian by Natalia Fedorova

“Through the Park,” a story generator, works on language at the sentence level, applying a simple and general narrative technique; code modification is not necessary neither in translation of “Through the Park,” nor in translation of its remixes.

“Through the Park” contains 25 sentences; nine of these are removed each time the system runs and the remaining sixteen are printed in their original order. The sentences are:

The girl grins and grabs a granola bar.
 The girl puts on a slutty dress.
 The girl sets off through the park.
 A wolf whistle sounds.
 The girl turns to smile and wink.
 The muscular man paces the girl.
 Chatter and compliments cajole.
 The man makes a fist behind his back.
 A wildflower nods, tightly gripped.
 A snatch of song reminds the girl of her grandmother.
 The man and girl exchange a knowing glance.
 The two circle.
 Laughter booms.
 A giggle weaves through the air.
 The man's breathing quickens.
 A lamp above fails to come on.
 The man dashes, leaving pretense behind.
 Pigeons scatter.
 The girl runs.
 The man's there first.

Things are forgotten in carelessness.
 The girl's bag lies open.
 Pairs of people relax after journeys and work.
 The park's green is gray.
 A patrol car's siren chirps.

The sentences were consciously written to suggest (although not directly assert) that the two characters might be in a friendlier or more antagonistic relationship, and that the situation is more playful or sinister. In this simple system, there is no representation of the underlying fabula or story levels that is separate from a potential text, which may or may not be included in the final, realized discourse. However, there is a very simple model of a real narrative technique, ellipsis. This system has a model of how to omit some of the story — by performing a very simple procedure of dropping certain sentences.

Linguistically, and specifically in terms of discourse linguistics, it is problematic to include pronouns or other words that refer to other sentences; if such words are used, “she” or “he” might appear before “the girl” or “the man” are introduced. The more cohesive a text is — the more it uses reference to tie sentences to earlier sentences — the harder it is to elide a sentence from it without adjusting the other sentences. This is why the original sentences use almost no explicit reference of this sort, although there is a mention of “the two” midway through the list of sentences.

This issue of reference does not present a problem for an attentive, precise translator. The sentences would probably be translated in a usable way by such a translator if they were simply presented without any additional instruction. However, it is true that a less careful translator who decided to work more loosely and introduce pronouns and reference would encounter difficulties. The resulting text would not work properly in a translated system.

The same system was re-used by writer and artist J.R. Carpenter to create “Excerpts from Chronicles of Pookie & JR”. Similarly, pronouns were not used in the English, Spanish and Catalan versions of the “Excerpts”. Keeping the same syntax structure did not present any translation challenge. The need to maintain an ambiguity of tone or emotion did complicate the translation process to some extent, further highlighting the particular way in which the original sentences were constructed. However, each of the sentences could be translated, resulting in a Russian system that produced ellipses with the same sorts of effects as the original English system.

In sum, the way that computation and language engaged with one another in “Through the Park” and in the remix of it was clearly not as low-level and as profound as in the previous two systems. The relative ease of translation, and the ability to translate while leaving the code untouched, help to show this.

“Epigraphic Clock” by John Cayley
<http://programmatalogy.shadoof.net/applets/epigraphicclock/>
Translation to Russian by Natalia Fedorova (in progress)

“Epigraphic Clock”, a work in the literary clock series by John Cayley, is a poem that

consists of 365 words. The length of each stanza depends on the number of days in the month to which it corresponds. The piece “spells” time in two-word phrases of year, month, day, hours, minutes, and seconds. Translation of the phrases is hindered by the need for the two “time spelling” words to agree in case, gender and number in Russian. For instance, prepositions *о, во, на, и* require prepositional case, and a verb in the past tense has to agree in gender with a noun or a preposition. Meanwhile two words in phrase are highlighted in the text indicating the seconds and change in accordance with the clock. Thus when the phrase is rooted within the main text of the poem, the lack of agreement is excused by the context of the words and does not stand out. Grammatical adaptations on this level are unnecessary. The current incomplete state of “Точное время” reflects the complexity of word-level generative poetry translation, which demands, as “The Two” translation demonstrated, changes both in code and sentence structure. Although the initial translation of the poem text took a week to complete, a full translation that implements the generation mechanism in the new language requires more effort.

“Olvido Mortal” by Andrés Viedma Peláez

http://www.terra.es/personal/a_viedma/olvido.htm

Translated to English by Nick Montfort

Finally, we turn to the most computationally elaborate literary systems, interactive fiction programs that accept short phrases of typed input. The need to modify code (or to use different code libraries for a different language) is a very obvious in the case of interactive fiction, which involves not only assembling textons to generate output but also parsing and recognizing natural-language commands.

Almost a decade ago, in 2003, one of us, (Nick Montfort), was asked to translate *Olvido Mortal*, literally “mortal forgetting,” into English. The author and others in the Spanish interactive fiction community had attempted a translation into English which they called “Shattered Memory.” Although they knew English well enough to correspond with Montfort, this translation did not work out well and was not fully intelligible. The difficulty probably stemmed, in part, from the challenge of translating interactive fiction, which requires the full comprehension of the interactor, much as a riddle does. Prominent IF author and critic Emily Short wrote, “I find that playing an IF game that has been imperfectly translated is much more difficult than reading imperfectly-translated prose.”

This preliminary translation also didn’t give the impression of an English literary work, nor did it function in the same way that a literary translation would have. Finally, “Shattered Memory” fell victim to one of the few rules in the Interactive Fiction Competition, the main way of circulating new short games in the English-language IF community. There is a rule that prohibits works based on previously released works, and this was interpreted to prohibit translations.

The translation done by Montfort, “Dead Reckoning,” could not address this ban on translations in the IF Competition directly, but it could give this work another chance to be experienced by the English-language interactive fiction community – and it did accomplish that. The translation was more successful than the preliminary attempt, at

least, as it was reviewed positively.

“Olvido Mortal” was written using InformATE (“Inform Ahora Totalmente en Español” or “Inform, Now Totally in Spanish”), a library that adapts Graham Nelson’s Inform 6 system to Spanish. This library for Inform 6 was developed originally by José Luis Díaz, who uses the pseudonym “Zak McKraken.” InformATE not only supplies a Spanish parser; it also replaces the English-like Inform 6 keywords of the programming language itself with Spanish ones. To translate it, a conversion into the original Inform 6 was undertaken, so that code that was in Spanish-like InformATE was converted to English-like Inform 6. The variable names, however, were left in Spanish. Strings of Spanish were commented out and the new English strings were added beneath them. In this translation, there was no attempt to “clean up” the original code, something other e-lit translators have done, even though it is often easier, with the benefits of hindsight and with more time and a fresh perspective, to see how things might have been written more efficiently.

“Whom the Telling Changed” by Aaron Reed

<http://www.aaronareed.net/telling.html>

Translation to Russian by Natalia Fedorova (in progress)

“Whom the Telling Changed” (the Russian translation will be called “Кому не быть прежним, услышав это”) by Aaron Reed is an IF based on the Epic of Gilgamesh; a translation to Russian is now being undertaken. The interactor has a choice to be either a healer or a warrior. The translation project is enabled by Rinform, the Russian version of Graham Nelson’s Inform 6, which was developed by Denis Gaev in 2005. The system can generate case forms for nouns and recognize these forms in commands. It can also process most of the verb forms with prefixes.

However, one issue with the standard operation of the Rinform system is that the neutral tone of the universal modern English “you,” conventionally used in interactive fiction and used in “Whom the Telling Changed” specifically, is difficult to maintain in Russian. The stylistic complication is due to the importance of the distinction between the second person plural **Вы** and singular **ты**: the first is too formal and official and the second is too informal and at certain cases impolite in modern Russian. Still, “ты” is most widely used in the original Russian IF (“Книга судьбы” by Vsevolod Zubarev and “déjà vu” by Setke) and in the Russian translations of “All Roads” by John Ingold, “Spider and Web” by Andrew Plotkin, “Vespers” by Jason Devlin, and “Blue Chairs” by Chris Climas. However, the second person plural is used in Denis Gaev’s translation of the classic “Adventure” by Crowther and Woods. Gaev’s translation of “Through the Looking Glass” by Gareth Rees does not address the story to Alice, as Rees’s game does, and so does not use the second-person pronoun. The tone of Gilgamesh would require either the singular **ты** or third person narration by analogy with *byliny*, the Russian epic songs.

Conclusion

Gregory Rabassa has stated that “translation is essentially the closest reading one can

give a text,” [1] which suggests that the translation of a computational system to produce linguistic or narrative creativity would involve a very deep analysis and understanding of the system. John Zuern points out that paying attention to what happens “when we translate (or don’t translate) electronic texts will lead to finger-grained insights into the relationship between ‘electronic’ as a category and ‘literature’ as a category.” [2] We have tried to show that translation not only helps to relate categories, but also aids in understanding the interplay of computation and language in particular works. Just as literary translation allows for an extremely close reading and for new insights about a text, the translation of electronic literature can allow for a better understanding of how they are literary and also of the specific ways in which computing and language come together in them.

References

- [1] Rabassa, Gregory. “No Two Snowflakes are Alike: Translation as Metaphor.” p. 6. *The Craft of Translation*. Eds. Biguenet, John and Shulte, Rainer. The University of Chicago Press: Chicago, 1989.
- [2] Zuern, John. “Generally Electronic, Particularly Literary—or the Other Way Around?: How the Electronic Literature Collection Pushes Comparativism.” *Society for Literature, Science, and the Arts Conference*. Kitchener, ON. September 22-25, 2011.