

For Comment: Church Slavonic Plural Rules

Ponomar White Paper¹

Plural rules are descriptions of how languages handle plurals of nouns or unit expressions. For example, English has two forms: *one hour* and *two (three, 1.5, 0) hours*. Russian has several forms: a singular (*один час*), a paucal (*два, три, четыре часа*), and a plural (*пять часов*).

Definitions of plural rules are required to create the initial data for a language-based locale definition in the Unicode Common Locale Data Repository (CLDR). The CLDR provides locale-specific information to software applications, including parsing data, calendar information, collation rules, and more. The availability of a CLDR locale for a given language is necessary in order for this language to be properly supported in a wide variety of computer systems, including those manufactured by Apple, Google, and IBM – to name just the main vendors relying on CLDR for their internationalization implementations.

The Ponomar Project is involved in the development of CLDR locale data for Church Slavonic (ISO 639-2 code *cu*). Presently vendors either do not provide a Church Slavonic locale or erroneously use Russian locale data. This leads to various faulty software implementations in the support of Church Slavonic text on the computer. Therefore, the presence of CLDR data for Church Slavonic is crucial.

Your comment is requested on the following information about Church Slavonic plural rules. These rules were derived on the basis of the corpus of modern (post-Nikonian) Church Slavonic texts. However, within this corpus, different rules exist in the liturgical texts vs. the Elizabethan Bible, in particular, with regard to the treatment of the numeral 12. It is proposed that the usage of the liturgical texts be taken to be standard, meaning that the number 12 will be used with the dual, not the plural, form. This is perhaps the only controversial point of this document.

These data were compiled on the basis of the discussion in Gamanovitch (1964, pp. 89ff) and extensive search of the Church Slavonic corpus via the *Eureka!* search engine available here: <http://www.ponomar.net/cgi-bin/search.cgi>. Additional information on plural rules in the CLDR is available here: <http://cldr.unicode.org/index/cldr-spec/plural-rules>. Your comments are appreciated.

Table 1: Usage of Plurals in Church Slavonic

	First Declension	Second Declension	Third Declension	Fourth Declension
	рѣѡ	женѧ	зѧповѣдѧ	имѧ
one	1 рѧѣѡ	женѧ	зѧповѣдѧ	имѧ
other	1/2 рѧѣѧ	женѧѧ	зѧповѣдѧѧ	имѧѧѧ
two	2 рѧѣѧ	женѣѧ	зѧповѣдѧѧ	имѧѧѧ
few	3 рѧѣѧѧ	женѧѧѧ	зѧповѣдѧѧѧ	имѧѧѧѧ

¹ This paper is part of the Slavonic Computing Initiative at the Ponomar Project. Please direct comments to Aleksandr Andreev (aleksandr.andreev@gmail.com)

many 5	παιτῶν ραββί	παιτῆν ραββῶν	пѣтъ рабѣ
other fraction	τέτταρτην ραβὰ	τέτταρτη ραβὰ	чѣтвѣрτѣ ραβὰ

On the bases of the rules elucidated in Tables 1 and 2, the following CLDR rules are proposed for standardization:

Category	Examples	Rules
one	1, 21, 31, ...	$n \bmod 10$ is 1 and $n \bmod 100$ is not 11
two	2, 12, 22,	$n \bmod 10$ is 2
few	3, 4, 23, 24	$n \bmod 10$ in 3..4 and $n \bmod 100$ not in 13..14
many	0, 5-11, 13-20	$n \bmod 10$ is 0 or $n \bmod 10$ in 5..9 or $n \bmod 100$ in 13..14 or $n \bmod 100$ is 11
other	0.5, 1.5, etc.	<i>everything else</i>