

## When is Orthography Not Just Orthography? The Case of the Novgorod Birchbark Letters

ANDREW DOMBROWSKI  
*University of Chicago*

### Introduction

In the corpus of the Novgorod birchbark letters (BBLs), a tendency to use the Cyrillic letters {o, ъ<sup>1</sup>} and {e, ѣ<sup>2</sup>} completely interchangeably emerged during the 12<sup>th</sup>-13<sup>th</sup> centuries. Occurring in 31% of the documents dated to the first quarter of the 12<sup>th</sup> century, this practice became practically ubiquitous (up to 92% prevalence) by the middle of the 13<sup>th</sup> century (Zaliznjak 2004:25). This orthographic merger applies to all etymological \*o, \*e, \*ŭ, and \*ĭ, including, crucially, etymological \*ŭ and \*ĭ that are expected to have been lost during the jer shift (also known as *weak jers*).<sup>3</sup> For this reason, this merger has most frequently been analyzed as purely orthographic (Živov 1984/2006:93-96, Zaliznjak 2002, Zaliznjak 2004:23-24). Nuorluoto 2007, however, proposes that this merger (at least as applied to stressed vowels) reflects intermediate stages in the sound change \*ŭ > /o/. However, this proposal nevertheless partially relies on orthographic principles (Nuorluoto 2007:181).

In this paper, I argue instead that this merger (which can be termed the jer-letter merger for brevity) is entirely explicable in terms of the underlying phonology of the Old Novgorod dialect. In particular, I propose that the outcome of the jer shift in the Old Novgorod dialect was affected by a Finnic substratum. According to this proposal, language shift from Finnic to Slavic (roughly contemporaneously with the jer shift in Slavic) resulted in the imposition of Finnic phonotactic constraints onto Novgorod Slavic, which in turn resulted in the frequent preserva-

---

<sup>1</sup> This letter is known as the back jer, and originally denoted a back lax vowel [ʊ].

<sup>2</sup> This letter is known as the front jer, and originally denoted a front lax vowel [ɪ].

<sup>3</sup> Etymologies and transcriptions of Cyrillic orthography are given according to the orthography customary in Slavic linguistics. IPA correspondences are as follows: ć = [tʃ], c = [ts], š = [ʃ], ŭ = [ʊ], ĭ = [ɪ].

tion of etymological \*ŭ and \*ĭ in weak position. This analysis sheds new light on the role of language contact in the history of Russian as well as demonstrates the value of historical dialectology for resolving orthographic problems.

## 1 Orthographic Merger of {o, ъ} and {e, ъ}

### 1.1 Description

In a sizeable portion of the BBL corpus, the Cyrillic letters {o, ъ} are used interchangeably to denote etymological \*o and \*ŭ, while the Cyrillic letters {e, ъ} are used to denote etymological \*e and \*ĭ. In other varieties of Old Russian (OR), it is commonplace for the sound changes \*ŭ > /o/ and \*ĭ > /e/ (in strong position) to be reflected in the orthography, usually through the use of the letters {o, e} to reflect /o/ and /e/ from etymological jers. The orthographic merger found in the BBLs differs from the more common pattern in OR in two salient ways. First, it applies to all instances of \*ŭ and \*ĭ, not just to strong jers. Second, it is a complete merger, in which the four letters involved are used essentially interchangeably. This results in orthography strikingly distinct from that found elsewhere in Old Russian, as illustrated in the examples below.

The examples below include both orthographic and phonetic representations for the lexemes in the Novgorod BBLs and in OR before and after the jer shift. The forms given for OR are not drawn from specific sources, but are rather representations of how the given lexemes would be likely to be represented given preexisting knowledge about OR orthography and phonology (c.f. Borkovskij 1963, Živov 1984/2006, Vlasto 1986, Živov 2006). The phonetic representations given for the Novgorod dialect are in all instances approximations based on the proposals outlined in this paper. As such, they are provided primarily for the sake of expository clarity. The BBLs reflect approximately 400 years of development and multiple sociolects; as such, the phonetic representations here are necessarily simplified. They are primarily designed to reflect the underlying phonology of the documents from the 12<sup>th</sup> to early 14<sup>th</sup> centuries that merge {o, ъ} and {e, ъ}, and should not be interpreted as applying to all documents attested in the BBL corpus.

(1) \*čĭto 'what':

	Novgorod	OR before jer shift	OR after jer shift
Cyrillic	четъ	чѣто	чѣто, что
Transcription	četŭ	čĭto	čĭto, čto
Phonetic	[tʃɛ'tɔ] / [tʃɛ'tɔ]?	[tʃɪ'tɔ]	[tʃtɔ]

*When is Orthography Not Just Orthography*

It should be noted that this representation of 'what' is not the only variant attested in the BBLs; instead, it has been chosen to clearly represent the degree to which the representation of vowels in the BBLs can differ from that found elsewhere in OR. The BBLs attest all logical possibilities of vowel representation in this lexeme as well as variation due to the merger of \*č and \*c (a merger termed *cokan'e*) in north Russian, the jer shift, and other consonantal assimilations.<sup>4</sup> OR after the jer shift can optionally preserve the older orthography with the jer vowel intact; this is due both to the frequency with which older texts were copied and to the higher prestige associated with older orthographic practices (c.f. Živov 1984/2006:96). The use of a full vowel letter for the first vowel in \*čito is quite unexpected, insofar as jer vowels preceding a stressed syllable with a non-jer are the first to be lost in the jer shift (Borkovskij 1963:98). The two variants given for the phonetic representation of 'what' in the Novgorod dialect, with different initial consonants, reflect the above-mentioned merger of \*č and \*c. In this merger, [ts] is the most likely outcome, but the possibility of other reflexes, like [tɕ], cannot be ruled out.

Example (2) below gives a more representative illustration of this variation:

(2) \*poklonŭ 'bow'

	Novgorod	OR before jer shift	OR after jer shift
Cyrillic	ПОКЛОНО, ПОКЛОНЪ	ПОКЛОНЪ	ПОКЛОНЪ, ПОКЛОН
Transcription	poklono, poklonŭ	poklonŭ	poklonŭ, poklon
Phonetic	[pɔ'klɔnɔ]?	[pɔ'klɔnɔ]	[pɔ'klɔn]

The two forms cited above for Novgorod comprise the clear majority of attested forms of this common word (frequently used in greeting formulas), although some further variation is also attested.<sup>5</sup> The correspondence between word-final vowels in this case is particularly clear because this lexeme does not appear to have the nominative singular ending /-e/ that is uniquely characteristic of the Old Novgorod dialect.

Example (3) further demonstrates the extent of possible variation within a given lexeme:

<sup>4</sup> The full list of orthographic variants for the nominative / accusative case of 'what' is as follows: *čit[o], čito, ceto, čto, cto, sto, co, četo, čit-, čtŭ, što, čo, co, če, četŭ, c[t]o* (Zaliznjak 2004:818).

<sup>5</sup> The full list of variants is as follows: *poklonŭ, poklon, pokolono, poklo, pklonŭ, pklono, poklno, pokono, poklŭ{n}, {po}poklonŭ, {plo} poklono, ]po]k[lo]nŭ, poklonŭ*. Many of these variants are abbreviations and should not, therefore, be interpreted as indicative of phonological developments.

(3) \*vŭzĭmi 'take-2SG.IMPV'

	Novgorod	OR before jer shift	OR after jer shift
Cyrillic	вѣзьми, возьми, вѣзѣми, возѣми, возми, восми	вѣзьми	вѣзьми, возьми возми
Transcription	vŭzĭmi, vozĭmi, vŭzemi, vozemi, vozmi, vosmi	vŭzĭmi	vŭzĭmi, vozĭmi, vozmi
Phonetic	[vɔzɛ'ɫmi]?	[vuzɪ'ɫmi]	[vɔz'ɫmi]

The table in (3) includes all forms attested for this verb form in the BBL corpus, with the exception of some fragmentary attestations (Zaliznjak 2004:718-719). Notable in (3) is the presence of forms where the weak jer is not represented (*vozmi*, *vosmi*). These forms basically parallel post-jer shift OR forms. The presence of examples like this illustrates the important point that at no point in time does the BBL corpus reflect a linguistically monolithic population (Vermeer 1997). Furthermore, at no point does the prevalence of the {o, ъ} ~ {e, ъ} merger exceed 92%. In other words, it is quite possible that the attested tokens of a given lexeme reflect multiple underlying phonological systems.

## 1.2 Chronology

The merger of the {o, ъ} and {e, ъ} letters occurs with varying frequency throughout the time span of the BBLs. In the earliest chronological layer, comprising documents from the eleventh century, it occurs in only 8% of documents. The frequency of this merger rises steadily throughout the twelfth century, reaching a prevalence of 68% in documents from the last quarter of the twelfth century. It is ubiquitous, occurring with approximately 90% frequency, during the thirteenth century before dropping off in the fourteenth century. Only 14% of documents from the latest, post-1400 layer of the BBL corpus display the merger (Zaliznjak 2004:25).

The chronological trajectory of the {o, ъ} ~ {e, ъ} merger closely tracks three

other “orthographic” practices, the most notable of which is the interchangeable use of the Cyrillic letters {e, ѣ} for etymological \*e and \*ě.<sup>6</sup> The chronologies of these two practices correlate closely ( $r = .8$ ). Another close match is found between the jer-letter merger and the use of the Cyrillic letter {ѣ} for /f/ ( $r = .87$ ). Yet another strong correlation is found between the jer-letter merger and the practice of writing /u/ after a consonant as Cyrillic {y} ( $r = .86$ ), although the absolute levels of prevalence differ considerably. Zaliznjak 2004 also provides information pertaining to the chronological evolution of other orthographic practices in the BBLs (Zaliznjak 2004:25-33). No strong correlation can be observed between the jer-letter merger and any of the other eight evolving orthographic practices that Zaliznjak charts in the BBLs.<sup>7</sup>

Two points are especially worth noting with regard to the above chronology. Firstly, out of the three phenomena that correlate closely with the jer-letter merger, only two are obviously orthographic in nature. The interchangeable use of the Cyrillic letters {e, ѣ} could quite possibly reflect underlying phonological developments. While a full discussion of this problem goes beyond the scope of this paper, it may be noted that although \*ě > /i/ is frequent in north Russian dialects, the shift of \*ě > /e/ is also well-attested both in East Slavic and in Slavic as a whole. This suggests that the thirteenth century may have been a period of intense linguistic change in the Old Novgorod dialect, rather than a period of intense change in orthographic practice. Furthermore, the twelfth century is not only the period in which the jer-letter merger increases to the point of ubiquity, but is also the period in which the jer shift has been independently hypothesized to have been carried out in Novgorod (Zaliznjak 2004:59-60). In order to fully evaluate the significance of this overlap, some background information about the jer shift and early Slavic orthography is necessary.

## **2 The Jer Shift and Orthography**

### **2.1 The Jer Shift**

The phonological inventory of Common Slavic included two phonemic high lax vowels, [ʊ] and [ɪ], known as the back and the front jers respectively in traditional terminology. After the establishment of Slavic orthography (and the major geographical dispersion of Slavic), a sound change took place that has been frequently termed the jer shift.

Broadly speaking, the jer shift involved the elimination of [ʊ] and [ɪ] when

---

<sup>6</sup> The Common Slavic phoneme denoted as /ě/ was phonetically realized as [æ] in Common Slavic, but was raised to [ɛ], [e], or even [i] throughout East Slavic.

<sup>7</sup> The strongest observable correlation is a slight negative correlation with the use of the Cyrillic letter {ѣ} to represent /ja/ ( $r = -.48$ ). Besides that, correlations are negligible ( $r$  ranging from  $-.03$  to  $-.29$ ).

phonotactically permissible. Jer vowels that were eliminated are traditionally known as weak jers. When phonotactic restrictions prevented deletion of the jer vowels, they shifted to other vowels (Bethin 1998:104-5). The phonotactic details are to a certain extent both language-dependent and context-dependent, but an influential approximation is known as Havlik's Law<sup>8</sup>, which can be stated as follows: a jer is weak in word-final position and before a non-jer vowel. A jer before a weak jer is strong. A jer before a strong jer is weak. This results in vowel-zero alternations within paradigms, e.g. Common Slavic \**šivici* 'tailor-NOM.SG' & \**šivica* 'tailor-GEN.SG' > Ukrainian *švec* & *ševca* 'ibid' (Bethin 1998:105).

The jer shift spread from the southwest northwards within Slavic, beginning in the 9th century in southwest Slavic. However, the jers were retained into the 13th century on the NE periphery (i.e., Novgorod) and possibly later on the NW periphery (Bethin 1998:104-5, Zaliznjak 2004:58). As already mentioned, the jer shift in Novgorod seems to have overlapped with the emergence of the jer-letter merger. Since the evidence provided by written sources is crucial for determining the chronology and progression of the jer shift, an understanding of early Slavic orthographic practice is crucial for any evaluation of the BBL evidence.

## 2.2 Early Slavic Orthography

Slavic orthography was developed after at least the early stages of differentiation of Slavic into East, West, and South Slavic. The use of Cyrillic to write Slavic spread along with Orthodox Christianity from the south to the north. However, the orthographic system designed for ninth century South Slavic was only an approximate fit for later East Slavic due to a range of systematic phonological isoglosses separating these branches of Slavic. Simultaneously, South Slavic orthographic rules (as well as morphology and syntax) were imbued with the prestige of the liturgical language itself.<sup>8</sup>

In this context, a complex norm arose governing the application of Cyrillic orthographic principles by speakers of East Slavic. "Norm" in this context should, of course, not be understood as the equivalent of a modern literary language. Although full standardization never occurred, the prevalence of orthographic corrections in manuscripts suggests a broadly normative approach to orthography (Živov 2006:65-69). The learning process involved the repetition of CV sequences reflecting liturgical pronunciation. On the basis of this, writers learned to make orthographic generalizations ranging from fairly simple and inductive to more elaborate, depending on register, education level, and the specific orthographic issue at hand (Živov 1984/2006:93-95, Živov 2006:48-55).

---

<sup>8</sup> The relationship between the liturgical Church Slavonic language (based on East South Slavic) and Old Russian has even been analyzed in terms of diglossia (Uspenskij 1994). This analysis is far from universally accepted (cf. Kamčatnov 2005:49-56), but the very existence of this debate is nonetheless indicative of the underlying complexity of the sociolinguistic situation.

## *When is Orthography Not Just Orthography*

As a result of this, both orthographic and linguistic information is relevant for the analysis of any unusual or interesting features observable in early Slavic sources. On one extreme, some phenomena are clearly explicable only in terms of orthography – for instance, variation in the style of writing given letters. On another extreme, some phonological features (like the devoicing of word-final consonants), are reflected directly in the written text, and are best explained in terms of phonology rather than orthography. The jer-letter merger potentially involves both criteria. While it has generally been analyzed in terms of orthography, as outlined above, I will argue below that the jer-letter merger actually reflects the underlying phonology of the Old Novgorod dialect.

### **3 Finnic Substrate and the Jer Shift**

Independent evidence exists that a Finnic substrate played a significant role in the development of the northern dialects of Russian, including but not necessarily limited to the Old Novgorod dialect. Language shift from Finnic to Slavic has been hypothesized as an explanatory factor for the development of several phenomena observed in north Russian, ranging from phonological phenomena like *cokan'e* (the non-differentiation of the affricates [ts] and [tʃ]) to the emergence of nominative objects of infinitives (Timberlake 1974, Vermeer 2000, Čekmonas 2001).

In addition to a range of other evidence indicating a widespread Finnic-speaking population in early medieval north Russia, the BBL corpus contains the first written attestation of any variety of Finnic. BBL #292 is dated to the beginning of the 13<sup>th</sup> century and is written in an early form of Karelian (Haavio 1964). This suggests that a period of active bilingualism culminating in language shift from Finnic into Slavic was roughly contemporaneous with the emergence of the jer-letter merger.

While it is impossible to reconstruct the precise phonotactics of the variety of Finnic spoken in north Russia at that time, in general Finnic phonotactics do not tolerate complex consonant clusters of the type that the jer shift produced in Slavic (Laakso 2001:186). Since grammatical interference (or imposition, in the terminology proposed by van Coetsem 1988 and 2000) often results from language shift, it is logical to suppose that the influence of Finnic phonotactics may have affected the process of the jer shift in the Old Novgorod dialect. This resulted in the much more frequent preservation of weak jers, which is directly reflected in the jer-letter merger.

### **4 Conclusions**

In the above sections, it has been proposed that the interchangeable use of the Cyrillic jer letters and {o, e} in the Novgorod BBLs reflects a different outcome of

the jer shift triggered by the imposition of phonotactic constraints from a Finnic substrate. This proposal requires no new assumptions. Instead it relies only on the direct evidence of the BBL corpus, preexisting evidence suggesting widespread language shift from Finnic, and the generalization that language shift may often result in structural changes. This analysis does suggest a revision of the history of the jer shift in East Slavic, but this is not inherently problematic. The jer shift reached the northern periphery of East Slavic well after the emergence of other dialectal traits differentiating the Novgorod region from other varieties of East Slavic, and therefore variation in the outcome of the jer shift is not overly surprising. Furthermore, this proposal has some specific advantages, which are briefly sketched below.

First, the above analysis would result in a much more transparent and simple relationship between orthography and phonology in the BBL corpus. This is particularly advantageous because it is generally hypothesized that the writers of the BBLs were much less trained in elaborate orthographic principles than, for instance, scribes in monasteries. Other proposals have been made that link the jer-letter merger to phonology, but these proposals still rely on the hypothesis that the jer shift in Novgorod basically paralleled that found elsewhere in East Slavic, albeit somewhat later (Nuorluoto 2007). In such an approach, instances of the jer-letter merger applying to weak jers must still be explained through orthographic practices.

Furthermore, this proposal permits a principled and coherent chronology of the jer shift in Novgorod (and its reflection in the jer-letter merger observable in the BBLs). In the first stage, roughly corresponding to the 11<sup>th</sup> century, the orthographic merger is infrequent, which corresponds to the predictions that can be made from pre-existing knowledge about phonology and orthographic practices before the jer shift. In the emergence stage, corresponding to the 12<sup>th</sup> century, the jer shift takes place in Novgorod and the increasing prevalence of the jer-letter merger basically tracks the progress of the jer shift. By the 13<sup>th</sup> century, a stable situation emerges in which the jer shift has taken place but with much more frequent preservation of weak jers due to the influence of substrate Finnic phonotactics. The prevalence of the orthographic jer-letter merger declines in the 14<sup>th</sup> century and onward. This likely is due to an increasingly monolingual population in closer contact with the rest of Russian, as a result of which the jer reflexes found elsewhere in Russia prevailed.

## References

- Bethin, Christina. 1998. *Slavic Prosody: Language Change and Phonological Theory*. Cambridge, New York: Cambridge University Press.
- Borkovskij, V. I. and P. S. Kuznecov. 1963. *Istoričeskaja grammatika russkogo*

*When is Orthography Not Just Orthography*

*jazyka*. Moscow: ANSSSR.

Čekmonas, Valerij. 2001. On Some Circum-Baltic Features of the Pskov- Novgorod (Northwestern Central Russian) Dialect. In Östen Dahl and Maria Koptjevskaja-Tamm, eds. *The Circum-Baltic Languages: Typology and Contact*, 339-359, Amsterdam: John Benjamins..

Haavio, Martti. 1964. The Oldest Source of Finnish Mythology: Birchbark Letter No. 292. *Journal of the Folklore Institute* 1(2): 45-66.

Kamčatnov, A. M. 2005. *Istorija ruskogo literaturnogo jazyka*. Moscow: Academia.

Laakso, Johanna. 2001. The Finnic Languages. In Östen Dahl and Maria Koptjevskaja-Tamm, eds. *The Circum-Baltic Languages: Typology and Contact*. Amsterdam: John Benjamins: 179- 212.

Nuurluoto, Johani. 2007. The Interchangeability of the Graphemes <o> and <ъ> in Old Russian Birchbark Documents: A Graphical Effect or a Reflection of Sound Change in Progress? In Johani Nurluoto, ed., *Topics on the Ethnic, Linguistic and Cultural Making of the Russian North*. *Slavica Helsingiensa* 32: 176-195.

Timberlake, Alan. 1974. *The nominative object in Slavic, Baltic, and West Finnic*. Munchen: Sagner.

Uspenskij, B. A. 1994. *Kratkij očerk istorii ruskogo jazyka*. Moscow: Gnozis.

Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Universitätsverlag, C. Winter.

Vermeer, Willem. 1997. Notes on Medieval Novgorod Sociolinguistics. *Russian Linguistics* 21(1):23-47.

Vermeer, W. 2000. On the status of the earliest Russian isogloss: four untenable and three questionable reasons for separating the progressive and the second regressive palatalization of Common Slavic. *Russian Linguistics* 24(1):5 – 29.

Vlasto, A. P. 1986. *A Linguistic History of Russia to the End of the Eighteenth Century*. Oxford: Clarenton Press.

Andrew Dombrowski

- Zaliznjak, A. A. 2002. Drevnerusskaja grafika so smešeniem  $\upsilon - o$  i  $\upsilon - e$ . in Zaliznjak, A. A. *Russkoe imennoe sklonenie (s priloženiem izbrannyx rabot po sovremennomu russkomu jazyku i obščemu jazykoznaniju)*. Moscow: Jazyki slavjanskoj kul'tury.
- Zaliznjak, A. A. 2004. *Drevnenovgorodskij dialekt*. Moscow: Jazyki slavjanskoj kul'tury.
- Živov, V. M. 1984/2006. Pravila i proiznošenie v russkom cerkovnoslavjanskom pravopisanii XI-XIII veka. In Živov, V. M. *Vostočnoslavjanskoe pravopisanie XI-XIII veka*. Moscow: Jazyki slavjanskoj kul'tury: 76-131.
- Živov, V. M. 2006. Norma, variativnost' i orfografičeskie pravila v vostočno-slavjanskom pravopisanii XI-XIII veka. In Živov, V. M. *Vostočnoslavjanskoe pravopisanie XI-XIII veka*. Moscow: Jazyki slavjanskoj kul'tury: 9-76.

adombrow@uchicago.edu