Russian orthography and learning to read

Eugenia Kerek and Pekka Niemi
University of Turku
Finland

Abstract

The unique structure of Russian orthography may influence the organization and acquisition of reading skills in Russian. The present review examines phonemic-graphemic correspondences in Russian orthography and discusses its grain-size units and possible difficulties for beginning readers and writers. Russian orthography is governed by a hierarchical, relatively regular 3-tier system of rules, complicated by numerous exceptions. Many theorists find that the key to this regularised complexity lies in Russian morphology. This review presents the perspectives of prominent Russian linguists on what linguistic units Russian orthography represents, and it evaluates and analyses their relevance for contemporary reading research.

Keywords: reading acquisition, Russian, grapheme-to-phoneme regularity, grain-size unit

The bulk of reading research has been conducted on the English language. However, understanding is growing that theoretical models of reading and reading acquisition cannot be based on properties of only one language, even the most commonly used one (for a detailed review, see Share, 2008). Cross-linguistic comparisons have become popular (Bruck, Genesee, & Caravolas, 1997; Cossu, Shankweiler, Liberman, Katz, & Tola, 1988; Wimmer & Goswami, 1994), and in multilingual communities, several studies have investigated bilinguals’ reading strategies in diverse language combinations (Comeau, Cormier, Grandmaison, & Lacroix, 1999; Durgunoglu, 1997; Wade-Wooley & Geva, 1999). Linguistic diversity serves as a natural laboratory for identifying the language-specific properties of the reading process and its core components common to all languages and orthographies.

Reading instruction in shallow orthographies usually relies on basic phoneme-grapheme correspondences, and the acquisition of decoding skills in these orthographies is completed within fairly short periods of time. Deeper orthographies constitute more complexly organised systems, the mastery of which demands more time and effort (Seymour, Aro, & Erskine, 2003). Constraints placed by the degree of regularity in phoneme-grapheme correspondences influence processing strategies in different orthographies, as has recently been described by the psycholinguistic grain-size theory (Ziegler & Goswami, 2005, 2006). This theoretical framework suggests that learning to read in regular orthographies requires reliance on “small” psycholinguistic grain-size units, whereas the reduced reliability of small grain-size units in more irregular orthographies may encourage beginning readers to develop multiple grain-size reading

http://nflrc.hawaii.edu/rfl
units. Consequently, reading acquisition in irregular orthographies demands more time and
greater effort to develop several recoding strategies and a high degree of flexibility in reliance on
grain units of varying sizes. The inventory of such units is different for specific orthographies
and is the result of the interplay between the orthography’s inherent structural features, its
correspondences with the phonology in question, and the dominant methods of reading
instruction, which are in their turn often motivated by the nature of the orthography.

A detailed explanation of how orthographic complexity may slow down reading acquisition was
presented by Frost (1998). Frost’s strong phonological model suggests that during the initial
phase of word recognition, a fast prelexical computation occurs, when letters or letter clusters are
converted into phonemes or syllables. This procedure results in underspecified phonological
representations, based on unambiguous letters. Frost suggested that the phonological code used
for lexical access is more detailed in shallow orthographies and relatively impoverished in deep
ones. A reader of deeper orthographies must develop grain units of the optimal size for fast
conversion into preliminary phonological representations. A beginning reader then starts out with
phonological recoding of single letters into phonemes and, as the process of reading acquisition
continues, learns to convert larger units of letters into phonemic clusters. According to Frost, a
skilled reader is defined by the speed of the assembly process, the size of the computed units,
and the efficiency of lexical access using underspecified phonological information.

Analyses of specific features of orthographies may be used to predict grain-size units used in
reading. For example, based on findings from Danish, an orthography with multiple complexities,
Elbro (2006) presented a reading acquisition model based on the principle of economy where the
most reliable grapheme-phoneme correspondences are learned first, and more complex and less
productive associations are learned later. Elbro suggested that all deviations from the alphabetic
principle, even those linked to predictable pronunciations, have a disruptive effect on beginning
reading. He suggested the following sequence of literacy acquisition in Danish: (a) learning of
single letter-single sound correspondences, (b) learning of letter-sound patterns with conditional
pronunciations, (c) learning of spelling based on morphemic orthographic knowledge, and (d)
learning of word-specific orthographic patterns. According to Elbro, conditional pronunciations
are learned primarily in the smallest possible units (vowel-consonant combinations) rather than
inside bigger units (rimes). Morphemic and word-specific pronunciations take a long time to
learn, even though Elbro allows for the possibility that some high-frequency morphemes or
words may be recognised as whole patterns.

Russian is one of the world’s most widely used languages, one of the six official languages of the
United Nations. However, despite the current interest in reading strategies promoted by different
orthographies, the number of publications based on Russian data is strikingly small. The
attention of Russian linguists and psychologists has traditionally been directed at the difficulty of
spelling in Russian, and not reading, which was considered a relatively easy skill to acquire
(Inshakova, 2004). While theories dealing with the structure of Russian writing and the methods
of teaching it at school have been a focal point of discussions inside the Russian linguistic and
pedagogical research community, difficulties in reading acquisition for a long time remained the
domain of speech therapy and special education. Despite considerable effort in studying
developmental dyslexia in Russian (Kornev, 1997; Lalayeva & Venediktova, 2001), little is
known about the strategies a beginning Russian reader relies on.

*Reading in a Foreign Language* 21(1)
The scarcity of research is especially unfortunate because Russian phonology and orthography possess a number of unique features, making Russian an interesting test case for current theories of word recognition and reading acquisition. Russian has a rather elegant orthographic system: on the one hand, quite complex and hierarchical, and on the other hand, organised around a dominant principle and therefore, sufficiently regular and predictable, even though the number of exceptions is high. This combination of complexity and regularity is what makes Russian orthography interesting for comparative reading research. The present review aims to contribute to this research by undertaking an in-depth analysis of Russian phonological and orthographic structure and introducing the English-speaking reader to works by Russian scholars on this subject. A detailed description of Russian orthography and its possible difficulties for beginning readers have been presented in English by Grigorenko (2003, 2006). A short but informative article by Liberman (1980) is devoted to the same topic. The difference between the present review and previous similar accounts is that it puts the description of the special features of Russian orthography into the context of existing theories about the structural organization of this complex system and about reading acquisition in different types of orthographies. A discussion of the implications of the special features of Russian orthography for defining the dominant strategies used by beginning and skilled readers in Russian is the main goal of the present review.

**Major Features of Russian Morphology and Phonology**

Unlike major European languages, where grammatical meanings are often expressed analytically, that is, by grammatical constructions rather than morphology, Russian is a synthetic language with a vast variety of affixes and endings. In contrast to other morphologically rich languages, like Finnish or Hungarian, where morphemes are glued to each other, morphemic fusion is a widespread phenomenon in Russian. Morphemic and syllabic boundaries often do not coincide, and analyzing the morphological structure of a Russian word may be a complicated task. A word in Russian may have several prefixes, suffixes, and an ending, for example, пред|на|знач|ен|ный (“pred|na|znach|en|nyj,” having the purpose of).

The Russian phonological system contains 42 phonemes (differences of opinion concerning the phonemic inventory of Russian are outside the scope of this article, but see Bondarko, 1998). The syllabic structure of Russian allows closed syllables and consonant clusters in both the onset and coda positions. A consonant cluster in Russian may contain up to four consonants, for example, встреч|а (“vstrecha,” meeting), and each grapheme in such clusters usually corresponds to a separate phoneme (but see exceptions in Table 1).

Languages with a synthetic type of grammar often have phonological systems based on paradigmatic connections. As a result of the modifications phonemes undergo in speech, Russian phonology is characterised by a vast variety of positional vowel and consonant alternations. These paradigmatic relations between phonemes determine the system of phoneme-grapheme relations in Russian.

In Russian orthography, stress is marked only in dictionaries and books for beginning readers or nonnative learners of Russian, while skilled readers are expected to assign stress on the basis of...
their linguistic competence. However, stress constitutes a central feature of the phonetic structure of a Russian word, and stress assignment is a vital factor in reconstructing the holistic shape of a Russian word. Erroneous stress assignment hampers comprehension or may even change word meaning, for example, мук’а (“muk’a,” flour) and м’ука (“m’uka,” suffering).

Table 1. Some common consonant clusters and their pronunciations

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Pronunciation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Здн</td>
<td>[зн]</td>
<td>праздник (holiday)</td>
</tr>
<tr>
<td>Рдц</td>
<td>[рц]</td>
<td>сердце (heart)</td>
</tr>
<tr>
<td>Лнц</td>
<td>[лц]</td>
<td>солнце (sun)</td>
</tr>
<tr>
<td>Стн</td>
<td>[ст]</td>
<td>лестница (stairs)</td>
</tr>
<tr>
<td>Вств</td>
<td>[ст]</td>
<td>чувство (feeling)</td>
</tr>
</tbody>
</table>

Other clusters

| Жч       | [шц]          | мужчина (man) |
| Зч       | [щц]         | извозчик (cabman) |
| Сч       | [цц]          | счастье (happiness) |
| Чт       | [штц]         | что (what) |
| Чн       | [шнц]         | конечно (of course) |
| Тц       | [цц]          | вкратце (in brief) |
| Дц       | [цц]          | двадцать (twenty) |
| Тч       |               | лётчик (pilot, flyer) |
| Дч       | [чц]          | докладчик (reporter) |
| Тся      | [шц]          | учиться ((he) studies) |
| Ться     | [щц]          | учиться (to study) |

Stress in Russian is free (can fall on any syllable in a word) and mobile (can move to another syllable). The mobility of Russian stress has a morphological motivation: it is often shifted in word formation or inflexion, for example, город (“gorod,” city) and города (“goroda,” cities). Due to its unpredictability, stress assignment in Russian demands a well-developed visual anticipation ability, which is often problematic for beginning readers whose cognitive resources are entirely focused on the task of sequential phonological recoding (see Rayner, 1986).

Stress in Russian is strongly centred. The stressed syllable is pronounced with much more prominence than the remaining syllables. A Russian word is a complex unity of syllables with varying degrees of vowel intensity and length, organised around the stressed syllable. Stressed vowels increase in length and intensity, while unstressed ones undergo different types of reduction depending on their positions relative to the stressed one. Especially affected are the vowel phonemes o and e, which respectively sound like a and i when unstressed, for example, сова (“sova,” owl, pronounced as “sava”), and река (“reka,” river, pronounced as “rika”). Russian orthography does not represent vowel reduction in unstressed positions, which is one of the sources of its numerous irregularities.

The Russian consonant system is much more complex than the vowel system (Bondarko, 1998), but paradigmatic consonant alternations are more systematic than the vowel alternations. One of the most fundamental characteristics of the Russian consonant system is the distinction between
soft (palatalised) and hard consonants. Sometimes the hardness-softness dichotomy is the only
meaning-distinguishing feature, for example, вес (“ves,” weight) and весь (“vesj,” whole), брат
(“brat,” brother), and брать (“bratj,” take). The most typical consonant alternations in Russian
are progressive assimilations of voiced or unvoiced obstruents (“d” in “vodka” is pronounced as
“t” because of the following “k”), and devoicing of voiced obstruents in word-final position, like
in German. As a result, Russian contains numerous homophones, and this is known to cause
difficulties for beginning writers, for example, кот (“kot,” cat) and код (“kod,” code).

Language-specific features of the Russian morphological and phonological systems, as described
above, influence the process of language acquisition (Ceytlin, 2000). Grigorenko (2006) pointed
out that the whole Russian phonological system is characterised by extraordinary fluidity, where
the quantity and quality of both vowel and consonant phonemes are highly dependent on their
positions. Bondarko (1998) suggested that high variability in the phonetic form of a Russian
word might require complex psycholinguistic mechanisms of speech perception and production.
The existence of strongly centred stress in Russian, where the stressed syllable gains prominence
at the expense of unstressed ones, results in the division of disyllabic and multisyllabic Russian
words into two perceptual parts, the prestressed-stressed, and the poststressed, where sounds are
sometimes totally blurred. The diminished distinctness of phonological representations of lexical
items in long-term memory may cause difficulties in the development of phonological awareness
and in reading acquisition (Elbro, Borstrøm, & Petersen, 1998; Elbro & Pallesen, 2002). This
idea is supported by the results of experiments with prereaders and beginning readers of Russian,
reported by Lepskaya (1987) and Bogomazov (2001), who showed that the indistinctness of the
poststressed parts of multisyllabic words in Russian delays the development of phonological
awareness for the phonemes used in those parts of the words.

Irregularities of Grapheme-phoneme Correspondences in Russian

Not all language-specific features of Russian morphology and phonology are reflected in writing.
As a result, the spelling of Russian words may be difficult for beginning readers and writers.
This part of the review classifies the irregularities of Russian orthography from the standpoint of
their predictability.

Theorists of Russian writing largely accept Baudouin de Courtenay’s (1963) division of the
Russian writing system into two broad levels: graphics and orthography. These terms, widely
used in the Russian scientific and pedagogical literature, may be confusing for Western
specialists. Instead of being used to denote a system of writing as a whole, in Russian, the term
orthography is reserved for describing cases where the choice of grapheme in writing is not clear
(weak positions). To make the choice, the writer must try to find a related word in which he or
she can identify the phoneme in question in a strong position. For vowels, the strong position, in
which the phonemic quality is at its best, is under stress; for consonants, the strong position is
before a vowel. The term alphabetic rules is used to describe one-to-one phoneme-grapheme
correspondences (Baudouin de Courtenay, 1963; Zinder, 1987). The combinatorial possibilities of
encoding phonemes within syllables are determined by the rules of graphics.
Predictable Irregularities (I): Alphabetic Rules and Graphics

Russian uses the Cyrillic alphabet of 33 letters, including 2 special markers, soft sign ь and hard sign ъ, which have no phonemic values (see Table 2).

<table>
<thead>
<tr>
<th>Letter</th>
<th>Closest English letter</th>
<th>Approximation of pronunciation in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Аа</td>
<td>a</td>
<td>father</td>
</tr>
<tr>
<td>Бб</td>
<td>b</td>
<td>box</td>
</tr>
<tr>
<td>Вв</td>
<td>v</td>
<td>vest</td>
</tr>
<tr>
<td>Гг</td>
<td>g</td>
<td>get</td>
</tr>
<tr>
<td>Дд</td>
<td>d</td>
<td>do</td>
</tr>
<tr>
<td>Ее</td>
<td>e</td>
<td>yet</td>
</tr>
<tr>
<td>Её</td>
<td>e</td>
<td>yolk</td>
</tr>
<tr>
<td>Жж</td>
<td>zh</td>
<td>vision</td>
</tr>
<tr>
<td>Зз</td>
<td>z</td>
<td>zebra</td>
</tr>
<tr>
<td>Ии</td>
<td>i</td>
<td>yield</td>
</tr>
<tr>
<td>Йй</td>
<td>i</td>
<td>short “у́й”</td>
</tr>
<tr>
<td>Кк</td>
<td>k</td>
<td>kinky</td>
</tr>
<tr>
<td>Лл</td>
<td>l</td>
<td>lip</td>
</tr>
<tr>
<td>Мм</td>
<td>m</td>
<td>mother</td>
</tr>
<tr>
<td>Нн</td>
<td>n</td>
<td>nice</td>
</tr>
<tr>
<td>Оо</td>
<td>o</td>
<td>song</td>
</tr>
<tr>
<td>Пп</td>
<td>p</td>
<td>put</td>
</tr>
<tr>
<td>Рр</td>
<td>r</td>
<td>rock</td>
</tr>
<tr>
<td>Сс</td>
<td>s</td>
<td>soup</td>
</tr>
<tr>
<td>Тт</td>
<td>t</td>
<td>top</td>
</tr>
<tr>
<td>Уу</td>
<td>u</td>
<td>book</td>
</tr>
<tr>
<td>Фф</td>
<td>f</td>
<td>fat</td>
</tr>
<tr>
<td>Хх</td>
<td>kh</td>
<td>home</td>
</tr>
<tr>
<td>Цц</td>
<td>ts</td>
<td>mats</td>
</tr>
<tr>
<td>Чч</td>
<td>ch</td>
<td>chicken</td>
</tr>
<tr>
<td>Шш</td>
<td>sh</td>
<td>shop</td>
</tr>
<tr>
<td>Щщ</td>
<td>shch</td>
<td>shch</td>
</tr>
<tr>
<td>Ъъ</td>
<td>“</td>
<td>silent</td>
</tr>
<tr>
<td>Ыь</td>
<td>i, y</td>
<td>busy, bill</td>
</tr>
<tr>
<td>Ьь</td>
<td>ь</td>
<td>onion</td>
</tr>
<tr>
<td>Ээ</td>
<td>e</td>
<td>band</td>
</tr>
<tr>
<td>Йю</td>
<td>iu</td>
<td>new</td>
</tr>
<tr>
<td>Яя</td>
<td>ia</td>
<td>yuppy</td>
</tr>
</tbody>
</table>

Phoneme-grapheme correspondences are not always straightforward in Russian because many letters in the Russian alphabet are not bound to representing only one phoneme each. Attention must be paid to the so-called syllabic (Ivanova, 1966) or positional (Moiseev, 1987) principle of
Russian graphics. The essence of this is that phonemic values cannot be assigned to the consonants and vowels in a syllable without evaluating either or both the left or right context of a given grapheme. One of the most common examples is the representation of palatalised consonant phonemes: 15 consonant letters out of the 21 each correspond to two phonemes, where one letter signifies both hard and soft consonant phonemes. Russian has no diphthongs, and vowel phonemes are not divided into long and short. Six vowel phonemes are represented in writing by 10 letters (see Table 3). This includes 4 letters that each represent a combination of the original phoneme with the phoneme ы (as in yes) or with a preceding soft consonant. For this purpose, the so-called combined graphemes (Kuzmina, 2005) or syllabemes (Pavlova, 2000) я, ё, ю, е, и are used (see Table 3).

Table 3. Russian vowel letters, corresponding phonemes, and English approximations of their pronunciation

<table>
<thead>
<tr>
<th>Russian vowel letter</th>
<th>Phoneme</th>
<th>English approximation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Аа</td>
<td>A</td>
<td>as in father</td>
</tr>
<tr>
<td>Ее</td>
<td>y+е</td>
<td>as in yes</td>
</tr>
<tr>
<td>Её</td>
<td>y+о</td>
<td>as in yoghurt</td>
</tr>
<tr>
<td>Ии</td>
<td>Ие</td>
<td>as in see</td>
</tr>
<tr>
<td>Оо</td>
<td>O</td>
<td>as in object</td>
</tr>
<tr>
<td>Ўу</td>
<td>Oо</td>
<td>as in look</td>
</tr>
<tr>
<td>Бьы</td>
<td>И</td>
<td>as in bill</td>
</tr>
<tr>
<td>Ээ</td>
<td>E</td>
<td>as in get</td>
</tr>
<tr>
<td>Юю</td>
<td>y+oo</td>
<td>as in you</td>
</tr>
<tr>
<td>Яя</td>
<td>y+a</td>
<td>as in yard</td>
</tr>
</tbody>
</table>

In positions where a consonant is not followed by a vowel, for example, at the end of a word or before another consonant, softness is indicated by the soft sign ь. The soft sign has no phonemic value and is sometimes treated as part of a complex grapheme or a diacritic (Scherba, 1983; Zinder, 1987). The pronunciation of the consonants, depending on their hardness or softness, is illustrated in Table 4.

Table 4. Rules for reading hard-soft consonants

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Hard (before а, о, у, ы, э, й)</th>
<th>Soft (before е, ё, и, ю, я, ь)</th>
</tr>
</thead>
<tbody>
<tr>
<td>б</td>
<td>был (was)</td>
<td>бил ((he) beat)</td>
</tr>
<tr>
<td>в</td>
<td>высота (height)</td>
<td>весна (spring)</td>
</tr>
<tr>
<td>г</td>
<td>гусь (goose)</td>
<td>деньги (money)</td>
</tr>
<tr>
<td>д</td>
<td>дар (gift)</td>
<td>дядя (uncle)</td>
</tr>
<tr>
<td>з</td>
<td>роза (rose)</td>
<td>зима (winter)</td>
</tr>
<tr>
<td>к</td>
<td>конь (horse)</td>
<td>кино (cinema)</td>
</tr>
<tr>
<td>л</td>
<td>лук (onion)</td>
<td>люк (hatch)</td>
</tr>
<tr>
<td>м</td>
<td>мать (mother)</td>
<td>мел (chalk)</td>
</tr>
<tr>
<td>н</td>
<td>новый (new)</td>
<td>небо (sky)</td>
</tr>
<tr>
<td>п</td>
<td>папка (file)</td>
<td>песня (song)</td>
</tr>
<tr>
<td>р</td>
<td>рама (frame)</td>
<td>говорю ((I) speak)</td>
</tr>
<tr>
<td>с</td>
<td>сумка (bag)</td>
<td>сестра (sister)</td>
</tr>
<tr>
<td>т</td>
<td>там (there)</td>
<td>тема (topic)</td>
</tr>
<tr>
<td>ф</td>
<td>шкафы (wardrobes)</td>
<td>физика (physics)</td>
</tr>
<tr>
<td>х</td>
<td>холод (cold)</td>
<td>стихи (verses)</td>
</tr>
</tbody>
</table>
The application of the syllabic principle in this case is very functional because it allows encoding of the phonological opposition *hardness-softness* in a rational and economical way by using 4 extra vowel letters instead of 15 extra consonant letters. Some linguists see the relationship between soft and hard consonants in Russian phonology as hierarchical, where the soft consonants are the marked members of the opposition, and therefore, the vowel letters that do not signal softness are not explicitly perceived as signalling hardness (Bondarko, 1981; Kuzmina, 1981). Hardness is seen as the default attribute of a consonant letter. However, others view the use of *a, y, o, э*, a consonant and the absence of the soft sign, or word-final position as markers of the preceding consonant’s hardness (Skoblikova, 1974).

Another case of irregular phoneme-grapheme correspondence in Russian involves the *soft-set* vowel letters in the opposite version of the syllabic principle, when a single letter stands for two phonemes, the phoneme *ü* and a vowel. This function is also contextual, and the reader has to analyse the context in front of the soft-set letter because it is used when *ü* is combined with a vowel in word-initial position (e.g., *юбка*, “jubka,” skirt), after a vowel (e.g., *моя*, “moja,” my), or after *ь/*ъ (e.g., *бельё*, “beljo,” linen). This irregularity is often difficult for beginning writers, who tend to use the combination *ü* + vowel instead of the syllabemes *я, ё, ю*, and *е* (Kuzmina, 2005).

These irregularities in Russian graphics have serious implications for the practice of reading instruction. The written syllable in Russian is an inseparable unit of mutually dependent parts, which according to Ivanova (1966), is a basic reading and writing unit. The consonant-vowel (CV) segments inside different types of syllables are often used as the central reading unit by primer authors. Determining whether the consonant should be palatalised is only possible by looking at the next (vowel) letter, and therefore, approaches to reading CV syllables, especially sound blending inside a syllable, is one of the central issues in the history of reading instruction in Russia. Elkonin (1988) considered developing anticipatory orientation to the vowel the central element of the reading process in Russian and suggested that initial reading instruction should focus on the development of this orientation.

Previously, Russian primers separated palatalised and non-palatalised consonants, introducing in the beginning only words that contained “hard” consonants. Today, consonant letters are presented at once as signifying both the palatalised and non-palatalised phonemes. The mechanism of decoding a CV syllable is considered to be the same for both types of consonant phonemes; it demands analysing the context to the right of the consonant letter. Most modern primers address the irregularities of Russian graphics by presenting letters in blocks to prepare young learners for the application of the syllabic principle. Children are trained to detect a CV syllable inside a word and read it as a unit, while other members of intrasyllabic consonant clusters are added on to the blended CV unit (Goretsky, Kiriushkin, & Shanko, 1988). Kornev (1997) suggested that syllables become the operative grain-size units in reading acquisition in Russian after only 3–6 months of reading instruction, whereas letters remain the main grain-size units in writing acquisition, at least during the 1st year of instruction.

To sum up, although Russian graphics can often represent one-to-one phoneme-grapheme correspondences, they also contain numerous irregularities. Their rules are highly positional and
Predictable Irregularities (II): Orthography

The complexity of phoneme-grapheme correspondences in Russian significantly increases when, on top of the rules of Russian graphics, responsible for reflecting phonemes in their strong positions, the rules of orthography are applied to the phonemes in weak positions and many letters acquire phonemic values that are not defined by the rules of graphics. Rusakova and Ceytlin (1999) found that in a standard Russian text, 70% of graphemes are determined by graphic rules, and 30% by both graphic and orthographic rules.

Russian linguists’ decades-long heated debate on the underlying principle of Russian orthography was aimed at deciding which linguistic unit is central for it, phoneme or morpheme. This debate was prompted by differing definitions of what constitutes a phoneme. Phonemically indefinite segments (weak positions) are interpreted in different ways by the Moscow and St. Petersburg phonetic schools, which has a direct impact on their choices of unit of analysis for Russian orthography. Linguists of the Moscow school (Avanesov, 1956; Panov, 1979; Reformatskij, 1955) see positionally alternating phonemic values within the same morpheme as a single unit consisting of a strong phoneme present in the strong position and other weak members of the paradigm, which appear in the positions of neutralization (weak positions). This phonemic row is, according to the Moscow school, reflected in writing through its strong-position representative. While the Moscow school emphasised paradigmatic connections between sounds, the St. Petersburg school adopted a linear approach. St. Petersburg linguists (Bondarko, 1998; Ivanova, 1971; Scherba, 1983) do not accept the idea of a phoneme as a complex of strong and weak phonemes and view cases of phonemic indistinctness in weak positions as phonemic alternations. The prevailing mechanism of writing in Russian is supposed to be morphological analogy, aiming at the visual unity of the morpheme.

Both theoretical approaches describe the same phenomenon with the help of different linguistic units and levels of abstraction. The basic idea of the phonemic principle (Moscow school) is that positional alternations of sounds belong to one and the same phoneme, always reproduced in writing by the same letter, whereas the morphemic principle (St. Petersburg school) sees different phonemes alternating within the same morpheme and represented by the same letter to preserve morphological continuity.

Russian linguists disagree on whether cases where graphemes represent phonemes in weak positions are part of regular phoneme-grapheme correspondences and can be seen as some kind of secondary alphabetic values (Gvozdev, 1954; Ivanova, 1971; Osipov, 1970; Scherba, 1983; Selezniova, 1981; Skoblikova, 1974). The primary source of these debates is the regularity of Russian orthography’s deviations from transparency. Grigorenko (2003) described writing in Russian as largely morphophonemic and noted a logical, regular structure underlying word alterations that may seem irregular at the surface level. Ivanova (1966) wrote that writing in Russian dissociates from pronunciation, but only at certain points such as morphemic boundaries and the absolute end of the word for consonants and inside the morphemes for vowels.
Unpredictable Irregularities: The Traditional Principle

Russian orthography, like many others, is not based exclusively on one principle, but rather on a combination of several. According to Liberman (1980, p. 54), “Russian spelling is phonemic to a point, very morphological and largely traditional.” Russian linguists disagree on exactly how many principles are at work. Most definitions take into account the regularity of the rules upon which a principle is based (Selezniova, 2004; Zinder, 1987). Predictable irregularities are verifiable; a phoneme in the weak position can be checked by finding a related word with a different stress pattern where the same phoneme is in the strong position, for example, вода ("vod’a," water) and водка (“v’odka,” vodka). If a phoneme in the weak position cannot be related to a phoneme in the strong position, the spelling of the word simply has to be memorised, for example, собака (“sobaka,” dog), where the phonemic quality of о is blurred, and no other word or word form has it in the strong position (under stress). Often, these are words of unclear etymology or loan words from other languages. Sometimes the written form of such a word will adequately represent its phonemes, and sometimes not. These spellings are often grouped together as the traditional principle. Ivanova (1977) described the traditional principle as one where phonemes in weak positions are represented by one of the phonologically possible letters. Though the choice of a grapheme is arbitrary, the number of available graphemes is, in fact, limited. Besides, some common consonant clusters in Russian have a pronunciation that traditionally deviates from each consonant's respective representation in writing. Some examples are presented in Table 1.

Words spelled according to the traditional principle are often problematic for beginning writers. However, it is not clear whether the unpredictable nature of orthographic irregularities in the spelling of these words affects their reading in any way.

Educational Practice

Varying definitions of the underlying principles of Russian orthography, even though these definitions are not always stated explicitly, determine to a large extent the structure and the choice of reading instruction materials in Russian schools. Several types of reading and writing programs with supplementary materials, based on different principles and approaches to reading acquisition, are available for use in Russian schools. Most programs are phonics-based, but they differ in their presentation of irregularities of Russian orthography.

Traditionally, most of the reading and writing instruction programs used in Russian schools have been based on the ideas of the St. Petersburg school. Reading instruction in Russia does not deal explicitly with the grapheme-phoneme irregularities on the orthographic level after the basic grapheme-phoneme correspondences and the rules of graphics have been learned. Instead, the rules of Russian orthography are extensively covered during writing lessons and supplemented by training in stress assignment, the blending of phonemes into syllables, morphological analysis, and lexicology (the study of lexical meanings and derivational patterns of words).

Since the 1970s, reading and writing instruction programs of a new type have been created by groups of researchers interested in the ideas of the Moscow school and the possibility of guiding
students in their discovery of the underlying phonological principles of Russian. Such programs (Elkonin, Tsukerman, & Bugrimenko, 1995; Repkin, Levin, Timchenko, & Zhedek, 1994; Soloveichik, 2008) draw the attention of students to positional changes in phonemes inside words rather than to morphological connections between words. These programs are of an experimental nature and aim at helping students to see the phonology and orthography of Russian as complex but logically organised systems. The spelling exercise shown in Figure 1 demands that the learner make a choice between two possible letters based on the position of the sound that they signify.

![Figure 1](image)

Figure 1. A spelling exercise from Soloveichik’s (2008) textbook *K tainam nashego jazyka* [Towards the mysteries of our language].

According to Share (2008), the predominant emphasis on reading accuracy is the result of the anglocentric focus in reading acquisition research. Goswami (2006) pointed out that for children learning to read in most European languages, the main challenge is to achieve fluency rather than accuracy (see also Wimmer & Mayringer, 2002). Reading instruction in Russian schools is speed-focused and dominated by reading aloud from Grade 1; teachers traditionally measure students’ success in reading acquisition twice a year by counting the number of words each child can read per minute. Reading instruction in Russian schools focuses mainly on introducing children to basic phoneme-grapheme correspondences and to blending sounds inside CV syllables, whereby special attention is paid to vowels as the main actors in contextual effects inside CV syllables (Kostromina & Nagayeva, 1999; Omorokova, 1997; Starzhinskaya, 1988). Beginning readers in Russian schools are expected to reach the stage of accurate syllabic reading by the end of the first grade. After a child has achieved the stage of syllabic reading, he or she gets little help from the teacher in reaching complete fluency. At the same time, the demands placed on the beginning readers’ fluency at school are quite high, often prompting parents to prepare for the pressures of the school program by teaching children to read prior to school entry. Recently, the educational standards have been changed to include not only the quantity of words read per minute but also reading comprehension.
Learning to Read Russian: Possible Grain-Size Units

Share (2008) noted that any orthography in effect can be seen as an attempt to compromise between providing beginners with an efficient self-teaching mechanism and at the same time providing distinctive word- or morpheme-specific visual orthographic configurations needed for automatised skilled word recognition. Orthographies differ in the emphasis they put on different parts of this dual function. Grigorenko (2006, p. 319) wrote that the Russian language is “heaven for experts and hell for novice writers and readers.” Indeed, many features of Russian orthography, such as letter-sound patterns with conditional pronunciations, morphemic spellings, and word-specific orthographic patterns, are the same as those described by Elbro (2006) as disruptive for the process of reading acquisition in deeper orthographies.

According to the psycholinguistic grain-size theory (Ziegler & Goswami, 2005, 2006), beginning readers in deeper orthographies develop multiple reading strategies to cope with the irregularities presented by such writing systems. The structure of Russian orthography, which is at the same time both sufficiently complex and quite consistent in its complexity, presupposes the existence of several grain-size units. On the basis of our analysis of Russian orthography as a system and of Russian linguists’ ideas about its internal structure, we discuss possible grain-size units for reading in Russian and research questions, which in our opinion are relevant for international reading research and could be explored using Russian-language material. Obviously, this discussion is of a speculative nature, and the suggestions made must be verified by experimental results.

As shown above, Russian linguists from different schools of thought hold conflicting views on the nature of Russian orthography. The most disputed problem in the linguistic theory of Russian orthography, whether its regularised complexities are motivated by morphological or by phonological factors, has relevance for choosing possible grain-size units of reading in Russian. If the description of Russian writing as phonemic-morphological (Selezniova, 1997) is accepted, reading acquisition in Russian should logically at a certain point involve morphemes as grain-size units. If, on the other hand, Russian readers rely on the type of phonemes suggested by the Moscow school, which views the sounds in weak positions as positional variants of the original phoneme in the strong position, the focus might shift from morphemes to Moscow-school type phonemes (i.e., abstract phoneme-sized units, which can be read in different ways depending on their positions in words). Decoding in this case proceeds in a relatively linear way as the reader copes with cases of phonemic indistinctness (weak positions) without needing to access the morpheme inside of which this phoneme is read. At the same time, different types of literacy-related processes and tasks may require different types of grain-size units.

Bogomazov (2005) came up with a hypothesis, supported by experimental data, which can be considered an elegant compromise between the conflicting views on the nature of Russian orthography and the phonological units it encodes. Bogomazov suggested that Russian children gradually develop a two-level phonological system, starting with phonemes of the St. Petersburg school, which mainly serve the processes of speech perception (and reading), and then adding a more abstract level of Moscow-school phonemes, which are used in speech production (and writing). Bogomazov noted that the development of Moscow-school phonemes occurs primarily
inside the consonant system, responsible for expressing lexical meanings, while vowels mainly serve to produce grammatical forms. If a morpheme is located in the prestressed part of the word, as is often the case with morphemes carrying lexical meaning in Russian, the formation of Moscow-school-type phonemes in this part of the word occurs more rapidly. In terms of reading acquisition, the implication of this theory is that beginning readers of Russian rely on grapheme-phoneme correspondences, whereas skilled readers often use morphemes as grain-size reading units. Bogomazov noted that adult foreigners learning to read and write in Russian might in some cases develop a two-level phonological system in reverse because they are primarily influenced by the written variant of Russian, so they tend to pronounce Russian words the way they are written, without applying the required phonetic modifications.

There is yet another candidate for the role of a larger grain-size unit in Russian. Besides basic phonemic recoding, reading in Russian requires constant attention to the intrasyllabic structure because the quality of the consonant phoneme depends on the following vowel letter. The CV syllable in Russian orthography is an especially tightly cohesive unit, and most reading instruction programs include special exercises aimed at promoting syllabic reading. While positionality is a structural feature in Russian graphics, and the syllabic principle is entirely based on syntagmatic relations, the morphological principle emphasises the vertical, paradigmatic connections. How the skill of using and coordinating these two operational units is acquired and developed has not been investigated experimentally. This dual task might require considerable attentional resources and complicate the process of reading skill automatisation. The processing costs of attending in parallel to different grain-size units (phoneme, syllable, and morpheme) can be expected at different points of reading acquisition in Russian.

Still, some Russian linguists disagree with the traditional division of the Russian system of writing into graphics and orthography. Kuzmina (1981) of the Moscow school argued that writing in Russian does not presuppose consecutive stages where the rules of graphics are applied first, and then the initial draft is corrected according to the rules of orthography. The process has one stage, taking into account two types of conditions for the choice of grapheme: phonological (graphic) and morphological (orthographic). Linguists of the St. Petersburg school do not accept that the division between graphics and orthography presupposes dual coding. Selezniova (1981) noted that a distinction should be made between the acquisition of rules, which is indeed done in two stages, and their application, which is a unitary process. Kuzmina (2005) studied beginning writers’ errors and concluded that spontaneous acquisition of rules of graphics and orthography does not occur in a linear sequence. After the initial stage, when children only use one-to-one phoneme-grapheme correspondences, graphic and orthographic rules seem to be acquired in parallel. A presentation of all of the principles applicable to written Russian can be found in Table 5.

Lexical stress is another important aspect of Russian phonology that might influence the development of reading strategies in Russian. Stress patterns in Russian words vary greatly, and stress assignment is vital for reconstructing the spoken form of the written word. In particular, the place of stress determines the type of position (strong or weak) of the vowel phonemes in a word. On the other hand, stress dependencies in multisyllabic words span over several syllables, and therefore, the place of stress in a word cannot be determined by looking at a part. This is a bit like the chicken-and-egg problem: you cannot read the word accurately unless you know the
stress pattern, but you cannot know the stress pattern unless you read the word. That is why in Russian primers and other books for beginning readers, stress is marked by a special diacritic (’’) above the stressed vowel. However, because stress in Russian is often located far from a word’s beginning, finding the stressed syllable is not effortless for beginning readers.

<table>
<thead>
<tr>
<th>Level</th>
<th>Regularity</th>
<th>Irregularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonemic</td>
<td>1:1 phoneme-grapheme correspondences</td>
<td>No</td>
</tr>
<tr>
<td>Syllabic</td>
<td>Vowels or softness-hardness markers determine the softness-hardness of the preceding consonant; some letters are syllabemes</td>
<td>Exceptions: ж, ш, ч, ш, ц, й</td>
</tr>
<tr>
<td>Morphological-phonemic</td>
<td>Preserves the unity of morphemes and represents a phoneme in a weak position</td>
<td>Morphemes and phonemes in positions that have no other verifiable counterparts</td>
</tr>
<tr>
<td>Traditional</td>
<td>Words are spelled in a way that does not follow any of the above principles</td>
<td>Item-based irregularity</td>
</tr>
</tbody>
</table>

Note. a St. Petersburg school phonemes or Moscow school phonemes in strong positions. b Moscow school phonemes in weak positions for which there are counterparts in strong positions.

Skilled Reading in Russian: Possible Grain-Size Units

Ziegler and Goswami (2005) argued that reading development is a continuum stretching well into adulthood, with skilled reading bearing developmental footprints of the constraints imposed by the orthographic system. From this point of view, an analysis of language-specific features of the Russian orthography may help in defining processing strategies that skilled readers in Russian might use. According to the minimality principle postulated by Frost (1998), in word recognition, skilled readers initially rely on minimal phonological representation, where prosodic information and segments containing phonological ambiguity are not available prior to lexical access and can be activated later. What kind of minimal representation, sufficient for lexical access, can be produced by skilled readers without knowing the place of stress in a word in Russian? Because knowing the place of stress is affected primarily by the reader’s knowledge of the quality of vowel phonemes, their representation in the initial phase of word recognition cannot be complete. At the same time, vowel letters in Russian are involved together with soft and hard signs in contextual effects, signalling the softness-hardness of the preceding consonant. Therefore, they cannot be completely absent from the initial phonological representation, unless we suppose that information about consonant softness-hardness is not important at this stage. In Russian linguistic theory, soft and hard consonants are considered to be separate phonemes; consequently, some information about the vowel letter or the soft or hard sign following the consonant is plausibly present in this preliminary phonological representation. This might be very general, such as letters signalling softness (or hardness). Voicedness is one more characteristic of consonant phonemes, which depends on the quality of the preceding or following phoneme. The reader needs to know if the following phoneme is a vowel and if the
preceding or the following consonant is voiced or voiceless. This information should probably also be present in the unspecified representation.

Initial representations for skilled readers in Russian can thus take the form of syllables or morphemes with unspecified phonemic values for those consonant and vowel letters that do not represent the phonemes in the spoken form of the word. These values can be assigned after lexical access has occurred and information about lexical stress has become available. Skilled readers in Russian, just like skilled writers, might then apply the Moscow type of phonemes (phonemic rows) in the prelexical access stage and specify the phonemic values for phonemically indistinct word segments at a later stage of reading recognition on the basis of the information about strong and weak positions derive from stress assignment. This concerns primarily the vowel phonemes, because, as shown above, the identity of the consonantal skeleton of the word can be more or less defined already in the initial representations. Frost (1998) described his own experiments in Hebrew and of his colleagues in English (Berent & Perfetti, 1995), which showed that initial unspecified representations at the prelexical access stage in these languages have a consonantal basis, where the vowel information is supplied in the second cycle of processing. The same might be true for skilled reading in Russian. In Russian, consonants are more informative than vowels (Bogomazov, 2001; Silchenkova, 2002); for example, vowels are the first to be omitted when words are shortened. However, the strong degree of positionality inherent in Russian orthography makes extracting at least some information about the vowel letters indispensable even for skilled readers.

Note that word recognition processes in silent reading and reading aloud are not identical (Frost, 1998; Share, 2008). In particular, while for silent reading, minimal preliminary phonological information might suffice, reading aloud demands well-specified phonological representations. Consequently, grain-size units for reading aloud and silent reading might also be different. Reading aloud demands much more detailed representations for vowels and precise and well-timed assignment of stress.

Conclusion

In the present review, the Russian writing system has been described as a complex yet sufficiently consistent system that presents a beginning reader with several levels of complexity and consistency to cope with. Other researchers who have written on the topic share this viewpoint (Grigorenko, 2006; Liberman, 1980). In this article, the writing system has been examined from the point of view of possible grain-size units involved in both beginning and skilled reading in Russian. This approach has been prompted by recent suggestions of Ziegler and Goswami (2005, 2006) that differences in orthographic systems result in differences in reading strategies developed by beginning readers. The discussion of principles of writing in Russian and the controversies related to their definitions highlight the possibility of multiple grain-size units not only in writing, which is widely acknowledged, but also in reading.

A detailed analysis of the phonological and morphological structure of Russian words serves as a basis for a prognosis of the possible complexities of learning to read in Russian. Such specific traits of Russian phonology as differentiation between soft and hard consonants, which is
signalled in writing by the following letter, demand special attention and long-term strategic training in blending syllable components by beginning readers. The fact that Russian phonotactics allows complex onsets does not make this task any easier. Morphemic (phonemic, according to the Moscow school) continuity in written Russian demands well-developed morphological awareness. However, analyzing a Russian word into its constituent morphemes is often difficult because of wide-spread morphemic fusion. The development of morphological awareness is also hampered by suprasegmental features such as the extreme variability of Russian stress patterns and the fact that distinctiveness in the pronunciation of different parts of Russian words is dependent on their positions in relation to the stressed syllable. Phonetic modifications in the quality of Russian consonants also prevent a Russian speaker from retaining an image of a morpheme as a distinct perceptive unit.

In many ways the difficulties that beginning readers in Russian encounter might be similar to the ones outlined for Danish learners by Elbro (2006), for example, letter-sound patterns with conditional pronunciations (syllables), spellings based on morphemic orthographic knowledge, and word-specific orthographic patterns (traditional spellings). Danish is one of the deeper European orthographies, and its complexities slow down the process of reading acquisition, as shown by Seymour, Aro, and Erskine (2003). However, the comparable degree of complexity in Russian does not necessarily mean that it will bring about the same results as in Danish. The reason for this optimism is the predictability of many irregularities in written Russian words. Coping with complexities, of course, demands more training, but because many of them obey certain rules, eventually mastering them could bring about a qualitative change in reading strategies and an emergence of multiple grain-size units, which beginning readers learn to rely on. These hypothetical suggestions should, of course, be verified experimentally.

Some topical research questions concerning reading acquisition in Russian are the following. How does the high variability in the phonological and morphological structure of Russian lexical items influence learning to read in this language? How do the irregularities of the Russian writing system influence reading acquisition? How are reading accuracy and fluency in Russian related to the complexities of Russian orthography? What is the developmental progression in terms of grain-size units in reading aloud? How is the lack of congruency between syllables and morphemes in the structure of Russian words resolved at different stages of reading acquisition? Future studies could explore these issues.

An investigation of the process of reading in Russian should include lexical items of different complexity levels, for example, the ones containing and not containing different types of irregularities. A longitudinal study could reveal the dynamics of the development of reading skills and the changes in the influence of the item complexity on the rate of progress in reading acquisition.

The possible grain-size units in skilled reading in Russian, discussed above, might serve as a basis for silent reading—the most common reading activity for advanced readers. Researchers concerned with skilled reading might be interested in answering the following questions. What is the prevailing mechanism of skilled reading aloud in Russian? Is it a process occurring in several stages, involving the build-up of well-specified phonological representations on top of the initial minimal ones? If so, what do these minimal phonological representations look like? Does skilled
reading involve dissecting syllables or morphemes as larger grain-size units? What are the developmental footprints of the language-specific traits of Russian orthography in the reading strategies of skilled readers of Russian? The answers to these complex questions could inform reading instruction in Russian.

Finding out what grain-size units are operative at different stages of reading acquisition and for different types of reading activity in Russian will help to develop effective instruction methods for achieving reading accuracy and fluency and setting balanced reading proficiency norms for Russian primary school students. Moreover, investigating the development of reading skills in Russian will contribute to cross-linguistic reading research in answering questions about the relationship between the regularity of an orthography and predictability of its complexities and progress in reading acquisition.

Acknowledgments

This research was supported by grant No. 53432 from the Research Council for Culture and Society, Academy of Finland, to Pekka Niemi. We thank the anonymous reviewers for their helpful comments and suggestions.

Note

1. In all examples from Russian, the word within quotation marks is not a phonetic transcription of the actual pronunciation of the Russian word, but the representation of the word spelled with Latin letters, unless specifically mentioned otherwise.

References


Kerek & Niemi: Russian orthography and learning to read


Lalayeva, R. I., & Venediktova, L. V. (2001). Diagnostika i korrekcija narushenij chtenija i pis’ma u mladshih shkolnikov [Diagnosis and treatment of younger school-aged children’s reading and writing difficulties]. St. Petersburg, Russia: Sojuz.


Reformatskij, A. A. (1955). Vvedenie v jazykoznanije [Introduction to linguistics]. Moscow:
Uchпедгиз.
About the Authors

Eugenia Kerek is a doctoral student at the University of Turku, Finland. She is a native Russian speaker, and her PhD research deals with reading acquisition in Russian and preliteracy skills of bilingually reared children.

Pekka Niemi is a professor of psychology at the University of Turku, Finland. His main research interest is development of cognitive processes, particularly reading acquisition and difficulties associated with it. Address for correspondence: Department of Psychology, University of Turku, Turku, FIN-20014, Finland. E-mail: peknie@utu.fi