Two distinct parsing stages in nonword reading aloud: Evidence from Russian

Xenia Schmalza, Alexander Porshnev and Eva Marinus

ABSTRACT

Word reading partly depends on the activation of sublexical letter clusters. Previous research has studied which types of letter clusters have psychological saliency, but less is known about cognitive mechanisms of letter string parsing. Here, we take advantage of the high degree of context-dependency of the Russian orthography to examine whether consonant–vowel (CV) clusters are treated as units in two stages of sublexical processing. In two experiments using a nonword reading task, we use two orthogonal manipulations: (a) insertion of a visual disruptor (#) to assess whether CV clusters are kept intact during the early visual parsing stage, and (b) presence of context-dependent grapheme–phoneme correspondences (GPCs; e.g., п[a] → /l/; п[x] → /lj/), to assess whether CV clusters remain intact or are split during the print-to-speech conversion stage. The results suggest that although CV clusters are initially processed as perceptual units in the early visual parsing stage, letters and not CV clusters drive print-to-speech conversion.

A major component of reading is the computation of phonology from orthography. This conversion requires numerous preceding steps. First, the reader needs to ascertain the identity and location of letters. Second, these need to be clustered in such a way that they map onto meaningful segments: For example, the word “touch” needs to be clustered into the letter groups t, ou, and ch, because these map directly onto phonemes. After this clustering has occurred, the orthographic segments can be matched onto their phonological equivalents and bound into whole words, and pronunciation and access to meaning can occur.

The current paper is concerned with the cognitive processes underlying parsing letter strings into sublexical clusters. Although parsing has been researched throughout the last decades (e.g., Bowey, 1990; Marinus & de Jong, 2011; Perry, 2013; Prinzmetal, Treiman, & Rho, 1986; Rey, Jacobs, Schmidt-Weigand, & Ziegler, 1998; Rey, Ziegler, & Jacobs, 2000), open questions remain about the exact mechanisms that underlie the parsing procedure. The question of the nature of the units that are important for the sublexical system has received a lot of attention (e.g., Kim, Taft, & Davis, 2004; Perry, 2013; Schmalz et al., 2014). Several studies have shown the importance of graphemes, where a grapheme is defined as the letter or letter cluster that corresponds to a single phoneme (Marinus & de Jong, 2011; Rey et al., 1998, 2000). This is intuitive, because graphemes map onto phonemes and are thus an important guide to a word’s pronunciation: As shown with the “touch” example above, the system needs to know which letters form a grapheme, because a failure to do so would lead to an incorrect pronunciation such as /tɔʊkh/.

It is likely that in addition to units that serve to predict the pronunciation, other orthographic units have psychological reality. This is especially important for polysyllabic words, as these contain a greater number of linguistically plausible units than shorter
words. Empirical evidence shows that polysyllabic words are parsed into orthographic-syllabic units (e.g., Prinzmetal et al., 1986; Taft, 1992). More recent evidence suggests that the consonant–vowel (CV) structure of a letter string is an important factor affecting how it is parsed (Chetail, Balota, Treiman, & Content, 2015; Chetail & Content, 2012; Chetail, Drabs, & Content, 2014). For example, Chetail and Content (2014) showed that the physical length of briefly presented words is underestimated for hiatus words: Hiatus words contain a vowel cluster that falls between a syllable boundary (e.g., oasis, where the oa cluster maps onto separate syllables). Since the orthographic structure, through the double vowel, signals the presence of fewer clusters than are present in reality, participants tend to rate these words as taking less space on the screen than control words with an equivalent amount of orthographic and phonological units (e.g., opera). Given the brief presentation duration used in this experiment, the findings suggest that information about the co-occurrence of two vowels, signalling separate speech units, is processed at an early visual stage.

A CV-based parsing mechanism is implemented in the connectionist dual processes (CDP)+/++ models (Perry, Ziegler, & Zorzi, 2007, 2010). Here, a prelexical stage is implemented as an attentional window that places the letters into a template of onsets (initial consonant of a syllable), vowels, and codas (final consonants of a syllable). The importance of such a visual parsing mechanism based on syllabic structure requires that the visual system identifies letters as consonants and vowels at an early processing stage. Thus, the parsing mechanism can explain the results of previous studies that found that CV structure is used to parse words into their underlying linguistic structure.

The next question is whether early visual parsing and graphemic parsing represent the same mechanism, or whether these processes occur at different stages of the reading process. In the CDP+/++ model, letter strings are mapped onto a graphemic template, based on the graphemic structure, such that multi-letter graphemes (e.g., th, sh) are represented as a single unit. Thus, the parsing of letter strings into graphemes precedes the mapping of the letter string onto a CV-template (Perry et al., 2010). Providing evidence conflicting with this view, a recent experiment has used priming to elucidate the timing of grapheme parsing (Lupker, Acha, Davis, & Perea, 2012): The briefly presented prime disrupted either a two-letter grapheme (e.g., th), or a control two-letter consonant cluster (e.g., tr). There were no differences as a function of this manipulation in a lexical decision task, suggesting that an early visual stage does not distinguish whether two consonants map onto one phoneme or two.

From the literature, we can draw a tentative distinction: On the one hand, parsing can occur during early processing stages. This early process does not appear to be affected by the later print-to-speech conversion stage (Marinus & de Jong, 2008; Martensen, Maris, & Dijkstra, 2003), but by the CV structure of the letter string (Chetail & Content, 2014). On the other hand, words are also parsed into sublexical clusters, which facilitate the retrieval of a phonological word form (Rey et al., 2000; Schmalz et al., 2014; Taft, 1991; Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995). By definition, this process parses a letter string into units that map onto speech sounds; however, it does not seem to have an influence during the early stages of visual word recognition (Lupker et al., 2012).

In order to further investigate the processes underlying parsing, we use the Russian orthography, which has a high degree of context-dependency in deriving a letter string’s pronunciation (Ulicheva, Coltheart, Saunders, & Perry, 2016). Of interest to the current paper, the pronunciation of most consonant letters in Russian can be either non-palatalized or hard (e.g., /n/), or palatalized or soft (/nj/). By default, a consonant is hard if it occurs in isolation, if it precedes other consonants or the vowels a, o, y, b, v, or if it occurs in the final position of a word. Palatalization is signalled by the succeeding vowel: Consonants followed by the vowels e, ō, u, o, or я are soft (hereafter: soft, e.g., не → /n’e/). Critically, these rules involve the pronunciation of the consonant. Each “softening” vowel letter has a non-softening counterpart (а–я; о–e; у–o, etc.); the pronunciations of these vowel pairs within the context of a CV cluster are considered allophonic (e.g., in the syllables no, /ne/, and не, /n’e/).

While such context-dependent rules constitute the majority of Russian grapheme-to-phoneme correspondences, they are transparent, in the sense that they always apply: There are few exceptions (isolated loanwords) that cannot be pronounced correctly based on a set of grapheme–phoneme rules (Ulicheva et al., 2016). To disambiguate the pronunciation of a Russian consonant, information about the subsequent letter is almost always needed. In other words, CV clusters in Russian are highly predictive of the correct
pronunciation, while isolated consonants are not. Therefore, at the level of print-to-speech conversion, CV sequences are linguistically important orthographic units.

The predominance of these context-dependencies in the Russian orthography allows us to create a well-controlled and diverse item set to study the psychological salience of CV units. This is in contrast to other alphabetic orthographies, which have been used to address similar questions. For example, Martensen et al. (2003) studied the effects of visual segmentation and print-to-sound mapping on reading processes in Dutch. Dutch contains quite a few multi-letter correspondences (e.g., $\text{oe} \rightarrow /u:/$), but context-sensitive correspondences ($\text{c}s\text{a} \rightarrow /k/ ; \text{c}i \rightarrow /s/ $) are sparse, restricting the items that can be used in such an experiment. As a result, it is difficult to draw strong conclusions based on the outcomes of their study, as null results can reflect either the narrow item set or theoretical factors.

We use two methods to assess the psychological salience of Russian CV units at different processing stages. First, we use a visual disruption paradigm to study the early stages of parsing (Marinus & de Jong, 2008; Martensen et al., 2003; Perry, 2013; Taft, 2001). Here, a visual disruptor such as a space or hash sign (#) is inserted between letters that are proposed to form a unit, and in a control location. Research to date has found that inserting a visual disruptor slows down lexical decision and reading aloud latencies. Importantly, the size of the disruption effect depends on the position of the disruptor. It is assumed that a larger processing cost incurs when the disruptor breaks a salient orthographic cluster. Thus, the size of the disruption effect across positions can inform us about the orthographic units that have psychological reality. It is likely that this manipulation disrupts an early, pre-lexical stage of the parsing process, which may be independent of the print-to-speech conversion process (Marinus & de Jong, 2008; Martensen et al., 2003; Perry, 2013).

In the visual disruption manipulation, we are interested whether a visual disruption within a CV cluster ($\text{C#V}$) would affect reading latencies to a greater extent than a visual disruption in a control location, between adjacent orthographic syllables (CV#CV). This would support a model where a polysyllabic nonword would be parsed onto a graphemic template based on its syllabic structure. A CV cluster would form a salient unit, because it forms an orthographic syllable, while a VC cluster within a word is likely to be mapped into different slots if it crosses the orthographic syllable boundary. Alternatively, it is possible that phonological factors exert an influence on this early parsing process. If this is the case, interrupting CV clusters should be especially damaging when the vowel changes the consonant’s pronunciation (e.g., disrupting the CV cluster $\text{h}s\text{x}a /n\text{e}/$ should be less detrimental than disrupting the cluster $\text{h}s\text{x}a /n\text{e}/$), because the visual disruption would delay the context needed for the correct pronunciation of a soft consonant.

As a second manipulation, to test the saliency of CV clusters at a later, print-to-speech conversion stage, we manipulate the succeeding letter of a consonant: namely, whether it is succeeded by a vowel that softens its pronunciation, or by a vowel that allows it to retain the default hard pronunciation (i.e., $\text{a} &, \text{a}, \text{y}$, $\text{i}, \text{u}$, or $\text{?}$). This can be thought of as an analogue to whammies in English (Rastle & Coltheart, 1998). The dual route cascaded (DRC) model of reading aloud (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) assumes that the sublexical decoding mechanism operates in a serial, letter-by-letter fashion. In Russian, perceiving a consonant should activate its default, hard pronunciation. If the subsequent vowel softens the consonant, the system needs to cancel the hard pronunciation in favour of the soft pronunciation once the vowel is activated. This is proposed to cause interference, thus slowing down pronunciation latencies compared to an item where all letters receive the dominant pronunciation (Rastle & Coltheart, 1998). In the DRC, it is possible to implement context-sensitive rules (for the description of a Russian version of the DRC, see Ulicheva et al., 2016). Note that context-sensitive rules are also implemented in the English DRC. For example, the grapheme $\text{ch}$ is pronounced as $/\text{t}j/ $ (“church”) by default, but its pronunciation is changed to $/k/$ when it is succeeded by a consonant (“chrome”). Context-sensitive rules are still considered to be grapheme–phoneme correspondence rules, as they describe the pronunciation of a single grapheme (in the above examples, the grapheme $\text{ch}$).

The DRC predicts that we should find a whammy effect in Russian, where nonwords with soft consonants are harder to read aloud than nonwords where the consonants retain their default pronunciation. This would imply that, despite the linguistic–statistical saliency of CV clusters in Russian, serial sublexical decoding rests on a letter-by-letter conversion process. This would not suggest that larger units
are not important: We know from research in various orthographies that units larger than letters do affect reading aloud processes (e.g., Prinzmetal et al., 1986; Rey et al., 2000; Schmalz et al., 2014; Taft & Radeau, 1995). Rather, the critical question relates to the role of letters versus CV clusters during the stage of sublexical print-to-speech assignment. While the DRC would predict that a whammy effect would occur due to the role of letters as an intermediary step of sublexical decoding, the CDP+/++ would predict that letters are mapped onto the graphemic template before print-to-speech conversion is initiated, and that the pronunciation is retrieved for the entire cluster (Perry et al., 2007). In this latter view, we should find no whammy effect, because CV clusters such as /na/ and /ne/ would be mapped directly onto their phonological translation, without the intermediate stage of letter-by-letter processing.

In summary, we employ two manipulations: a visual disruption to tap an early visual parsing process, and the whammy manipulation, to tap the serial letter-by-letter decoding process, during which each grapheme is assigned a phoneme. Due to the linguistic saliency of CV clusters, we expect an inhibitory effect on reading aloud latencies if the CV cluster is disrupted (C#V) compared to when the disruptor is in a control location (V#C). If phonology is involved in this early parsing process, the visual disruption effect should be stronger when it occurs within a CV cluster in which the vowel changes the consonant’s pronunciation (C#Vs, or /i#e/, compared to C#Vh, or /i#a/). For the second manipulation, we expect to find a whammy effect, or overall slower response latencies when the nonword contains soft consonants, compared to when all consonants retain the default pronunciation. This would show that even in Russian, despite the importance of CV units in predicting pronunciation, letters are treated as an intermediary in sublexical processing. Finding no whammy effect would suggest that not all alphabetic orthographies require the use of letters as a gateway to sublexical decoding.

**Experiment 1**

**Method**

**Items**

We created 112 six-letter nonwords with three syllables each. Half had a CVCVCV structure, and the other half a VCVCVC structure. Out of the 56 CVCVCV-nonwords and VCVCVC-nonwords, half contained vowels that modified the consonants’ pronunciations (CV.CV.CV; VCV.CV.CV), and the other half only vowels that allowed the consonants to retain their default pronunciation (CVh.CVh.CVh; VCVh.CVh.CVh). Across the soft and hard conditions, each item pair contained the same consonants, such that the consonantal skeleton was matched pairwise across the hard/soft conditions. To study the effects of visual disruption, we also created two lists (A and B). Each list contained the same 112 items, but a different half included the visual disruptor, a hash sign (#) between the third and fourth letter (CVC#CV and VCV#C).

To reduce the potential of participants adopting unexpected non-linguistic strategies to determine consonant palatalization, the item set also included 20 filler nonwords, which had both softening and hardening vowels. To avoid confounds from word-level activation, we constructed nonwords that did not resemble any Russian words. Twenty graduate students from Nizhniy Novgorod, who were all native speakers of Russian and did not participate in the experiment, ensured that the item set did not contain nonwords that sounded similar to real words or phrases in Russian (see Protopapas, Gerakaki, & Alexandri, 2006, for a similar approach to minimize lexical similarity). The items are shown in Appendix A; a spreadsheet with the item characteristics can be downloaded at [https://osf.io/2fk5r/](https://osf.io/2fk5r/).

**Participants**

We tested 17 participants who were Russian native speakers and undergraduate students from Nizhniy Novgorod. Four students read both lists A and B in separate sessions more than one week apart. An additional six participants read only List A, and an additional 7 read only List B; the data thereby consisted of 10 readings of List A and 11 readings of List B. Note that both linear mixed effect (LME) and Bayes Factor analyses can deal with such mixed designs by using participants and items as random effects. The participants provided informed consent and received a small monetary compensation for each session.

**Procedure**

The items were presented in DMDX (Forster & Forster, 2003), in lower case on the screen for 2.5 seconds, or until the onset of a vocal response. The order of items was randomized, and nonwords with and
without hash signs appeared in mixed blocks. The participants were instructed to read aloud each nonword as quickly and accurately as possible, while ignoring the hash sign.

Results

The vocal responses for the experimental items were scored with the software CheckVocal (Protopapas, 2007). The full scored data, as well as the R script with the analyses, can be downloaded at https://osf.io/2fk5r/. We discarded non-responses or incorrect responses (2.4%). As the number of errors was small and equally distributed across conditions (mean accuracy rate in each condition ranging from 96.6% to 98.9%), we analysed only reaction times (RTs). The descriptive statistics from each cell are shown in Table 1. The trimming and transformation were decided based on the visual inspection of qq-plot. Log RTs showed an approximately normal distribution, and no outlier responses. Therefore, untrimmed log RTs are used as the dependent variable for all analyses.

We used both LME models (Baayen, Davidson, & Bates, 2008) and Bayes factors (BFs; Morey & Rouder, 2014; Rouder, Speckman, Sun, Morey, & Iverson, 2009) to analyse the data; p-values are provided for descriptive purposes only, as all of our conclusions are based on the Bayesian analyses. BFs present an alternative approach of data analysis to conventional significance testing. Here, one model (e.g., one containing the effect or interaction of interest) is compared to another model (e.g., one excluding the effect or interaction of interest). The BF quantifies the degree to which the data are compatible with one model over the other. Thus, it is particularly useful for interpreting null effects—in which case the data will be more in favour of a model excluding the effect of interest. A BF value greater than 10 is considered to provide strong evidence for the model that is tested compared to the model that it is tested against; a value between 3 and 10 provides some evidence, a value between one third and 3 means that the data cannot be used to adjudicate the two models, and values of smaller than one third provide some and strong evidence, respectively, against the model that is tested (Rouder et al., 2009). For the BF analyses, we used the package BayesFactor and its default parameters. We report BFs for all theoretically meaningful comparisons (see next paragraph).

A priori, we decided on analyses that would address the theoretically meaningful questions outlined in the introduction. First, it is of interest whether the disruption effect is greater for CV clusters than for a condition where the disruptor falls between two CV clusters (Analysis 1). This is addressed by a comparison of the size of the disruption effect for CVCVCV items (RTCVCVCV − RTCVCVCV) compared to VCVVCV items (RTVCVCVC − RTCVCVCVC; i.e., a CV structure by visual disruption interaction).

Second, we aim to determine whether visual disruption has a differential effect for CVh compared to CV clusters (Analysis 2). Here, we need to compare the size of the disruption effect, RTCVCVCV − RTCVCVCV, for CVhCVhCVh versus CVhCVhCVhCVh items. If there is no influence of print-to-speech conversion at the early visual parsing stage, we expect to find evidence for an additive model, with no interaction between vowel type and visual disruption.

Third, we aim to determine whether there is a whammy effect for nonwords containing softening vowels compared to nonwords containing hardening vowels (Analysis 3). To this end, we exclude items that contain a hash sign, and compare items with softening vowels (CVhCVhCVh, VCVhCVhC) to items with hardening vowels (CVhCVhCVh, VCVhCVhC).

Analysis 1: Disrupting CV clusters versus control items

The first question was whether the presence of the visual disruption affects reading latencies to a stronger extent when it disrupts a CV cluster, regardless of the type of vowel, compared to when it falls between two CV clusters. To this end, we conducted a 2 × 2 analysis, with CV structure (CVCVCV versus VCVVCV) and presence or absence of visual disruption as fixed factors, with log RTs as the dependent variable. In this and all subsequent analyses, the slopes of vowel type were allowed to vary across participants (Barr, Levy, Scheepers, & Tily, 2013), and the dichotomous predictor variables were recoded as 0.5 and −0.5. All models also included the fixed factor of previous RT, centred and transformed into z-scores for each individual participant (Baayen, 2008).

The LME showed a significant effect of visual disruption, \(t = 3.71, p = .0008\), reflecting longer latencies when the visual disruptor was present, and no main effect of CV structure, \(p = .85\). The interaction, however, was significant, \(t = 2.40, p = .0177\). This indicates a stronger disruption effect for CVC\#VCV than for VCV\#CVC items (see Table 1).
BF, the presence of the interaction was more than two times more likely than its absence given the data, though the BF did not reach the conventional threshold of 3, BF = 2.32 (±1.47).

**Analysis 2: Visual disruption of CVs compared to CVh clusters**

The next question concerns the effect of visual disruption across different types of CV units. If the context-dependency affects early visual parsing, we should find an interaction between the presence of a visual disruption and vowel type: The effect of disruption should be stronger for CV.CV.CV nonwords than for CVh.CVh.CVh nonwords. If all CV clusters are equally salient in Russian at the early parsing stage, regardless of palatalization, we should find a main effect of visual disruption, but no interaction between visual disruption and vowel type.

Here, we included only nonwords with a CVCVCV structure. A 2 × 2 LME analysis (Vowel Type × Visual Disruption) showed that the main effects of visual disruption and vowel type were significant: \( t = 5.27, p < .0001 \), and \( t = 4.47, p < .0001 \), respectively. This reflected faster response times for default-consonant items, and for items without visual disruption. The critical interaction between them did not reach significance, \( p = .55 \), and the Bayes factor analysis provided evidence against an interactive model compared to an additive one, BF = 0.20 (±1.79%).

**Analysis 3: Whammy effects**

The last question was whether we would find a whammy effect, or slower RTs for nonwords containing hardening vowels than for nonwords with softening vowels. Similarly, the Bayes factor provided strong evidence for an effect of vowel type, when we compared a model including vowel type as a predictor to a base-line model that contained only the random effects of subjects and items and the fixed effect of previous RT, BF = 41.12 (±2.38%).

**Discussion**

In the analyses of Experiment 1, we found that (a) visually disrupting orthographic syllables affects reading aloud latencies to a greater extent than inserting the disruptor between syllables, (b) visual disruption is identical when it disrupts a context-sensitive rule (i.e., the vowel changes the consonant’s pronunciation) and when it disrupts a cluster of the same orthographic structure, where no context-sensitive rule is present, and (c) in line with previous work on the whammy effect, nonwords with context-sensitive rules take longer to read aloud than nonwords without context-sensitive rules.

However, before we move on to the general discussion, there are several issues regarding the sample and items of Experiment 1. The sample size is rather small, and the use of a mixed design is somewhat unconventional. To ensure that our findings are not spurious effects due to noise, we provide a replication of Experiment 1.

Furthermore, there are potential confounds in the item set: The manipulation of hardening versus softening vowels could systematically affect the average bigram frequency across conditions by facilitating reading aloud latencies in one condition over the other. It is also possible that there are differences in frequency between hard and soft consonants. We therefore calculated the average bigram frequency of each item, and the consonant phoneme’s

---

Table 1. Descriptives of Experiment 1.

| Stimulus structure | Hardening vowels | Softening vowels | Softening vowels |
|-------------------|------------------|------------------|------------------|
|                   | RT (ms)          | SD                | SD                | RT (ms)          | SD                | SD                |
|                   | by participantsa | by itemsb         | by participantsa  | by itemsb         | by participantsa  | by itemsb         |
| CVCVCV            | 782.5            | 132.3             | 89.7              | 872.8            | 140.1             | 85.8              |
| CVC#VCV           | 866.7            | 153.5             | 120.4             | 951.5            | 161.3             | 111.0             |
| VCVVCVC           | 845.3            | 128.8             | 86.4              | 867.2            | 153.6             | 100.6             |
| VCVhCVh           | 858.8            | 122.9             | 108.6             | 884.1            | 155.3             | 105.9             |

Note: Average reaction times (RTs, in ms) as a function of vowel type, consonant structure, and visual disruptor condition. C = consonant; V = vowel.

*aStandard deviation calculated from the averages by participants. bStandard deviation calculated from averages by items.*
frequency. These varied across condition: Both bigram frequency and consonant phoneme frequency were lower for the “soft” (means: 18,304 and 9335, respectively) than for the “hard” conditions (means = 27,947 and 14,055, respectively). Thus, it is possible that bigram frequency and/or consonant phoneme frequency drives the apparent whammy effect reported in Experiment 1. We therefore designed a new experiment, which allowed us to replicate the results of Experiment 1, while matching for bigram frequency and consonant phoneme frequency across conditions.

Experiment 2

Method

Items
The items were 116 nonwords. As in Experiment 1, they had either a CVCVCV structure or a VCVCVC structure; half of each contained only softening vowels, and the other half only hardening vowels. Again, there were two lists, in each of which a different half of the items contained a visual disruptor (#) between the third and the fourth letter. Unlike Experiment 1, we could not use the same consonantal skeleton across “hard” and “soft” conditions, as this would have greatly restricted the number of possible items. However, the items were matched across conditions, both on the average bigram frequency and on the average consonant phoneme frequency. The items are shown in Appendix B; a spreadsheet with the item characteristics can be downloaded at https://osf.io/2fk5r/.

Participants
The participants were 30 undergraduate students from Nizhniy Novgorod, who participated in exchange for course credit. All participants read both lists; 14 started with List A, and 16 started with List B. Between sessions, there was a break of at least 2 days.

Procedure
The procedure was identical to that of Experiment 1.

Results
The participants’ responses were scored with the software CheckVocal as correct, incorrect, or no response (Protopapas, 2007). Fifteen responses were scored as non-responses (0.2% of all responses) and were excluded for all analyses. Inspection of individual participants’ accuracy rates indicated that one participant had a very low accuracy rate (44%) for items with a visual disruption. Specifically, the participant substituted the visual disruption symbol with letters that were not present in the nonword. This suggests that the participant did not understand the instructions. Therefore, we exclude this participant from all subsequent analyses. The accuracy and (trimmed) RTs for each condition are provided in Table 2.

For the remaining 29 participants, the overall error rate was 2.38%. Across conditions, the accuracy was high and approximately equally distributed (ranging from 96.4% to 98.5%). We therefore did not analyse accuracy further.

For the RT analyses, we excluded all incorrect responses. As for Experiment 1, further exclusion and transformation procedures were based on visual inspection of qq-plots. This resulted in the exclusion of four datapoints with RT < 333 ms. An inverse transformation of the trimmed RTs yielded an approximately normal distribution and is used as the dependent variable in the subsequent analyses. Note that the trimming and transformation decisions were taken prior to analysing the data and are therefore not conditional on the outcome of the analyses.

The questions of interest and analyses were identical to those in Experiment 1.

Analysis 1
Again, here we were interested in the effect of visual disruption for words with a CVCVCV structure, compared to a VCVCVC structure. We found a main effect of visual disruption, t = 5.18, p < .0001, with faster RTs for items with than for those without disruption, and a main effect of syllable structure, t = 3.18, p = .0017, with longer latencies for VCVCVC items than for CVCVCV items. The interaction was also significant, t = 2.66, p = .0085, reflecting a stronger disruption effect for CVCVCV words than VCVCVC words. As in Experiment 1, the Bayes factor provided anecdotal evidence for the presence of this interaction, BF = 2.65 (±0.73%).

Analysis 2
Here, we were interested in the potentially differential effect of visual disruption of CV_s versus CV_h clusters. For CVCVCV-items only, we found a main effect of visual disruption, t = 6.15, p < .0001, reflecting longer RTs for disrupted than for undisturbed items, and a
main effect of consonant pronunciation (hard versus soft), $t = 5.12$, $p < .0001$, reflecting longer RTs for soft than for hard consonants. The interaction was not significant, $t = 0.66$, $p = .511$. The Bayes factor provided evidence against this interaction, BF$ = 0.20$ ($±1.52$).

**Analysis 3**

For the items with no visual disruption, we assessed whether or not items with soft consonants (whammies) were pronounced more slowly than items with hard consonants. Here, we found a main effect of consonant type, $t = 5.34$, $p < .0001$. The Bayes factor provided very strong evidence for this effect, BF$ = 75,106$ ($±1.09\%$).

**Discussion**

The theoretically important results of Experiment 2 are identical to those of Experiment 1: We found an interaction between the location (between versus within syllable) and presence of the disruptor (Analysis 1), a visual disruption effect of similar magnitude for CV$_s$ and CV$_h$ clusters (Analysis 2), and a whammy effect, or slower RTs for items with softening vowels than for items where the consonants retain their default pronunciation (Analysis 3).

The only difference between the two experiments is the significance of the main effect of structure (CVCVCV versus VCVCVC; Analysis 1) in Experiment 2. This relates to overall faster reaction times for the CVCVCV than the VCVCVC items (see Table 2). As there are no systematic differences across the two experiments, it is not clear whether this discrepancy across experiments is driven by random noise or a hidden moderator. As this comparison is not of interest to our hypotheses, we do not discuss this further.

In both Experiments 1 and 2, Analysis 1 showed an interaction between the presence and location of the visual disruptor in the predicted direction (i.e., stronger disruption in the C#V than in the V#C conditions); however, in neither did the degree of evidence exceed the conventional cut-off of BF$ = 3$. As the design was identical across experiments, we seek to address the possibility that taken together, the two datasets provide stronger evidence. To this end, we collapsed the datasets of Experiments 1 and 2 and performed an identical Bayes factor analysis to Analysis 1 on this collapsed dataset. This combined dataset provided strong evidence for the presence of an interaction, BF$ = 44.6$ ($±1.88\%$).

**General discussion**

We conducted two Russian nonword reading aloud experiments with two manipulations: (a) A visual disruptor was inserted either within a CV cluster (CVC#VCV) or between CV clusters (VCV#CVC), and (b) the pronunciation of the consonant was either palatalized or not, as signalled by the subsequent vowels. In Experiment 1, the items were matched across the soft and hard conditions on the identity of the consonant graphemes and vowel phonemes, but differed in terms of bigram frequency and consonant phoneme frequency. In Experiment 2, the items were matched across all conditions on bigram frequency and consonant phoneme frequency. Across both experiments, the key analyses provided converging results, suggesting that the effects are stable.

The experiments aimed to address three questions. Firstly, we found that a visual disruption within CV clusters (i.e., disrupting an orthographic syllable) affects reading latencies to a greater degree than a disruption between CV clusters (i.e., when the orthographic syllable is not disrupted). This speaks to the saliency of orthographic syllables (CV clusters) in Russian, at an early, visual stage of processing. Second, the impact of a visual disruption on reading speed did not differ for CV clusters containing

| Stimulus structure | Hardening vowels | Softening vowels |
|--------------------|------------------|------------------|
|                    | RT (ms)          | SD by participants$^a$ | SD by items$^b$ | RT (ms) | SD by participants$^a$ | SD by items$^b$ |
| CVCVCV             | 677.9            | 157.0            | 37.8            | 737.0    | 163.4            | 41.0            |
| CVC#VCV            | 734.2            | 181.4            | 54.5            | 781.3    | 177.0            | 49.3            |
| VCVCVC             | 714.0            | 159.2            | 42.4            | 766.8    | 182.5            | 45.9            |
| VCV#CVC            | 742.0            | 174.3            | 35.2            | 773.7    | 174.2            | 33.3            |

Note: Average reaction times (RTs, in ms) as a function of vowel type, consonant structure, and visual disruptor condition. C = consonant; V = vowel.

$^a$Standard deviation calculated from the averages by participants. $^b$Standard deviation calculated from averages by items.
context-sensitive rules (CV_s) or those consisting of only context-insensitive rules (CV_i). This suggests that early visual parsing is not influenced by the nature of print-to-speech mappings. Third, we found a whammy effect, suggesting that in Russian, as in English (Rastle & Coltheart, 1998), letters act as a gateway between visual processing and print-to-speech conversion.

Our findings suggest that parsing occurs in two different stages: During an early, pre-lexical stage, Russian words are parsed into CV clusters. The context-dependency of specific conversion rules emerges at a later stage. Previous studies have suggested that visual disruptions may have an effect on a stage that precedes grapheme–phoneme conversion (Marinus & de Jong, 2008; Martensen et al., 2003; Perry, 2013). Here, we explicitly test and confirm this claim, as we find no influence of phonological characteristics on visual parsing based on orthographic structure. Our results also converge with those of Chetail and colleagues (Chetail et al., 2015; Chetail & Content, 2012; Chetail et al., 2014), who have found evidence for parsing at early visual processing stages. The presence of an early visual parsing stage, which is driven by orthographic CV structure, could be implemented by the CDP+/++ models (Perry et al., 2007, 2010), as these use the syllabic structure to map a word onto the graphemic template. In its current form, however, the input scheme of the CDP++ maps graphemes, rather than letters, onto the CV-template, such that a graphemic parsing stage necessarily precedes this CV-based parsing process (Perry et al., 2010).

Future research could assess what factors drive the development of the salience of CV structure as a cue in early visual parsing. In Russian, CV clusters are linguistically important, because they map consistently onto phonology. In many alphabetic orthographies, such as French and English, the CV structure serves as a cue to syllabic parsing. Determining the syllabic boundaries at the beginning of the visual word recognition process is essential for print-to-speech mapping, articulatory planning, and lexical stress assignment. In the course of reading acquisition, children may learn to treat syllables as salient psycholinguistic units, precisely because they serve as informative cues to a word’s pronunciation (Ziegler & Goswami, 2005). However, in the time course of the visual word recognition process, it is unlikely that feedback from phonological processes directly influences syllabic parsing: In line with our results and interpretation, previous work suggests a pre-phonological locus of parsing based on CV structure (Chetail & Content, 2014; Prinzmetal et al., 1986).

During the print-to-speech translation process, we show slower processing of items with context-sensitive rules than of items without such context-sensitive rules. This is in line with the DRC model (Coltheart et al., 2001), as it suggests that even in an orthographic system with a high level of context-dependency, where surrounding letters need to be taken into account to compute a pronunciation, sublexical decoding is driven by letters. As letters are processed serially, a whammy effect occurs when the system encounters a context-sensitive rule: Initially, an incorrect phoneme is activated (e.g., in Russian, a hard consonant), which needs to be de-activated when the subsequent letter is processed, if it changes the pronunciation of the preceding letter. The finding of a whammy effect is in conflict with the view that the mapping of CV clusters onto syllable-based templates are directly involved in print-to-speech translation, as implemented in the CDP+/++ (Perry et al., 2007).

In the DRC, the presence of whammies should increase response latencies (Rastle & Coltheart, 1998). Our results are consistent with this explanation. However, we cannot fully rule out additional or alternative processes that may have contributed to the RT difference between the hard and soft conditions. While the items of Experiment 2 were equated for consonant phoneme frequency and bigram frequency, it is possible that there are uncontrolled factors, such as letter frequency, the frequency of vowel phonemes (which cannot be calculated due to reduction of non-stressed vowels in Russian), or ease of articulation, which may slow down the RTs for CV clusters with soft compared to hard consonants.

Our explanation rests on the assumption that the whammy effect operates on the behavioural level as described above and by Rastle and Coltheart (1998). Nevertheless, even if one prefers a different explanation for the whammy effect, our finding of no interaction between the presence of a whammy and the presence of a visual disruption across two experiments suggests two independent processes. Our proposed explanation of slower RTs for the soft than for the hard consonants is based on the DRC model and is supported by previous empirical studies (Coltheart & Rastle, 1994; Roberts, Rastle, Coltheart, & Besner, 2003); however, future studies are needed to corroborate or refine the exact cognitive mechanisms that are responsible for the observed whammy effect in Russian and in other orthographies.
In addition to an effect of the presence of whammies, the number of whammies present in an item should have a cumulative effect on response latencies (Rastle & Coltheart, 1998). In our experiments, the CV, CV, CV, CV nonwords contained three whammies, while the VC, CV, CV, CV nonwords contained only two (as the first vowel does not have a preceding consonant). Indeed, we found a stronger whammy effect for the former than the latter type of items in Experiment 1: Taking into account only items without visual disruption, the whammy effect was 90.3 ms for the CV, CV, CV, CV condition, and 21.9 ms in the VC, CV, CV, CV condition. In Experiment 2, however, there was no difference, as the whammy effect was 55.7 ms in the CV, CV, CV, CV condition, and 54.6 ms in the VC, CV, CV, CV condition. It is not clear why we would expect a different result across experiments. One potential explanation is a confound with bigram frequency in Experiment 1: In general, items in the hard condition had higher bigram frequency than those in the low condition. This difference was larger for CV, CV, CV, CV than for VC, CV, CV, CV items (a difference between hard and soft conditions of 11,540 and 7745, respectively). As the Bayes factor for a three-way interaction between structure (CV, CV, CV, CV, VC, CV, CV), consonant type (hard, soft) and experiment yielded equivocal evidence, BF = 1.61 (±1.58%), however, it is also possible that this discrepancy arose due to random noise. We consider this latter possibility more likely, because research on the effect of bigram frequency on reading processes has shown, if anything, a small inhibitory effect (e.g., Chetail et al., 2015; Gernsbacher, 1984).

The effect of a single whammy could be relatively small (Schmalz, Beyersmann, Cavalli, & Marinus, 2016), which would mean that the difference between a double- and a triple-whammy can only be reliably detected with higher statistical power. As the items in the current study had either double- or triple-whammies, we maximized the size of main effect of consonant type, which accounts for the consistent evidence for a whammy-effect across both experiments.

In summary, the current study provides insights into parsing mechanisms underlying reading aloud. We report evidence for two separable parsing procedures: The first depends on an item’s orthographic CV structure, and probably reflects pre-lexical visual processes. The second appears to operate by processing each letter sequentially while retrieving a phonological equivalent. The role of letters in sublexical conversion in Russian, with its heavily context-dependent print-to-speech conversion principles, emphasizes the universality of the importance of letters in alphabetic orthographies. The presence of two distinct parsing mechanisms opens new questions: While previous literature has been concerned with isolating psychologically salient units (e.g., Bowey, 1990; Prinzmetal et al., 1986; Schmalz et al., 2014; Taft, 1992), future research can aim to determine which units are important for which parsing system, what factors determine the saliency of different units, and, more broadly, how the two parsing systems interact with other parts of the reading system.

Note

1. We thank Anastasia Ulicheva for providing the bigram frequency and consonant frequency counts. Note that calculating the frequency of vowel phonemes for nonwords is unfeasible, because the pronunciation of a vowel depends on where the participant places stress; unstressed vowels are reduced in Russian. As stress assignment varies across participants, there is no single correct pronunciation for each vowel.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59(4), 390–412. doi:10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278. doi:10.1016/J.Jml.2012.11.001

Bowey, J. A. (1990). Orthographic onsets and rimes as functional units of Reading. Memory & Cognition, 18(4), 419–427.

Chetail, F., Balota, D., Treiman, R., & Content, A. (2015). What can megastudies tell us about the orthographic structure of English words? The Quarterly Journal of Experimental Psychology, 68(8), 1519–1540. doi:10.1080/17470218.2014.963628

Chetail, F., & Content, A. (2012). The internal structure of chaos: Letter category determines visual word perceptual units. Journal of Memory and Language, 67(3), 371–388.

Chetail, F., & Content, A. (2014). What is the difference between OASIS and OPERA? Roughly five pixels: Orthographic structure biases the perceived length of letter strings. Psychological Science, 25(1), 243–249. doi:10.1177/0956797613500508
Chetail, F., Drabs, V., & Content, A. (2014). The role of consonant/vowel organization in perceptual discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(4), 938–961.

Coltheart, M., & Rastle, K. (1994). Serial processing in Reading aloud: Evidence for dual-route models of Reading. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1197–1211.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). Drc: A dual route cascaded model of visual word recognition and Reading aloud. *Psychological Review, 108*(1), 204–256. doi:10.1037/0033-295X.108.1.204

Forster, K. I., & Forster, J. (2003). Dmdx: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers, 35*, 116–124.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysynomy. *Journal of Experimental Psychology: General, 113*(2), 256–281.

Kim, J., Taft, M., & Davis, C. (2004). Orthographic-phonological links in the lexicon: When lexical and sublexical information conflict. *Reading and Writing: An Interdisciplinary Journal, 17*, 187–218.

Lupker, S. J., Acha, J., Davis, C. J., & Perea, M. (2012). An investigation of the role of grapheme units in word units in Reading. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1491–1516.

Marinus, E., & de Jong, P. F. (2008). The use of sublexical clusters in normal and dyslexic readers. *Scientific Studies of Reading, 12*(3), 253–280. doi:10.1080/10888430802132246

Marinus, E., & de Jong, P. F. (2011). Dyslexic and typical-Reading children use vowel digraphs as perceptual units in Reading. *The Quarterly Journal of Experimental Psychology, 64*(3), 504–516. doi:10.1080/17470218.2010.509803

Martensen, H., Maris, E., & Dijkstra, T. (2003). Phonological ambiguity and context sensitivity: On sublexical clustering in visual word recognition. *Journal of Memory and Language, 49*(3), 375–395. doi:10.1016/S0749-596X(03)00668-8

Morey, R. D., & Rouder, J. N. (2014). Package “BayesFactor”. Retrieved from http://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf

Perry, C. (2013). Graphemic parsing and the basic orthographic syllable structure. *Language and Cognitive Processes, 28*(3), 355–376. doi:10.1080/01690965.2011.641386

Perry, C., Ziegler, J., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP++ model of Reading aloud. *Psychological Review, 114*(2), 273–315. doi:10.1037/0033-295X.114.2.273

Perry, C., Ziegler, J., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of Reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology, 61*(2), 106–151.

Prinzmetal, W., Treiman, R., & Rho, S. H. (1986). How to see a Reading unit. *Journal of Memory and Language, 25*(4), 461–475. doi:10.1016/0749-596X(86)90038-0

Protopapas, A. (2007). Checkvocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*(4), 859–862. doi:10.3758/bf03192979

Protopapas, A., Gerakaki, S., & Alexandri, S. (2006). Lexical and default stress assignment in Reading Greek. *Journal of Research in Reading, 29*(4), 418–432. doi:10.1111/1467-9817.2006.00316.X

Rastle, K., & Coltheart, M. (1998). Whammys and double whammys: The effect of length on nonword Reading. *Psychonomic Bulletin & Review, 5*(2), 277–282.

Rey, A., Jacobs, A. M., Schmidt-Weigand, F., & Ziegler, J. C. (1998). A phoneme effect in visual word recognition. *Cognition, 68*(3), B71–B80. doi:10.1016/S0010-0277(98)00051-1

Rey, A., Ziegler, J., & Jacobs, A. (2000). Graphemes are perceptual Reading units. *Cognition, 75*, B1–B12.

Roberts, M., Rastle, K., Coltheart, M., & Besner, D. (2003). When parallel processing in visual word recognition is not enough: New evidence from naming. *Psychonomic Bulletin & Review, 10*(2), 405–414. doi:10.3758/Bf03196499

Rouder, J. N., Speckman, P. L., Sun, D. C., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237. doi:10.3758/PBr.16.2.225

Schmalz, X., Beyermann, L., Cavalli, E., & Marinus, E. (2016). Unpredictability and complexity of print-to-speech correspondences increase reliance on lexical processes: More evidence for the orthographic depth hypothesis. *Journal of Cognitive Psychology, 28*(6), 658–672. doi:10.1080/20445911.2016.1182172

Schmalz, X., Marinus, E., Robidoux, S., Palethorpe, S., Castles, A., & Coltheart, M. (2014). Quantifying the reliance on different sublexical correspondences in German and English. *Journal of Cognitive Psychology, 26*(8), 831–852.

Taft, M. (1991). *Reading and the mental lexicon*. Hove: Lawrence Erlbaum.

Taft, M. (1992). The body of the BOSS: Subsyllabic units in the lexical processing of polysyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1004–1014.

Taft, M. (2001). Processing of orthographic structure by adults of different Reading ability. *Language and Speech, 44*, 351–376.

Taft, M., & Radeau, M. (1995). The influence of the phonological characteristics of a language on the functional units of Reading: A study in French. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 49*(3), 330–348. doi:10.1037/1196-1961.49.3.330

Treiman, R., Mullenix, J., Bijeljac-Babic, R., & Richmond-Weltz, E. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124*(2), 107–136.

Ulicheva, A., Coltheart, M., Saunders, S., & Perry, C. (2016). Phonotactic constraints: Implications for models of oral Reading in Russian. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(4), 636–656. doi:10.1037/xlm0000203

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled Reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3–29. doi:10.1037/0033-2909.131.1.3
### Appendix A

| CV, hard | CV, soft | VC, hard | VC, soft |
|---------|----------|----------|----------|
| батыно (batyino) | батине (batine) | атывоб (atynob) | ятунеб (yatuneb) |
| бутьды (butydy) | бутьди (butydi) | утывоб (utymob) | ятунеб (yatuneb) |
| гонаану (gonanu) | генану (genanu) | онанку (onanok) | яннанеб (ynanenb) |
| горуы (goru) | герооби (geroubi) | орубы (oruby) | ериуби (eryubi) |
| дахьыно (daxygo) | дагыгой (dagygo) | ахьогод (ahyogod) | яхьенеб (yahyeneb) |
| доку (doku) | докуне (dokune) | окунд (okund) | екунед (ekyned) |
| кохасы (koxasy) | кохасы (koxasi) | охасык (oxasik) | ехасик (exhasik) |
| ласуру (lasuru) | ласуру (lasuri) | асурол (asurol) | ясворол (yasvorol) |
| мокосо (mokoso) | мокесе (mekese) | окосом (okosom) | екессел (ekessel) |
| мосаку (mosak) | мосаку (mosak) | осакум (osasum) | есаком (esakom) |
| мульху (mulyxu) | муйлыху (muylyhu) | улыхум (ulyhom) | юлихом (yulihom) |
| нутохи (nutox) | нутокейх (nutoki) | утокейх (utokeik) | ютейхин (yuteykin) |
| ратому (ratomu) | ратому (ratomy) | атомур (atomur) | ютамур (yutamur) |
| ронго (ronugo) | роноге (ronoge) | оногор (onoger) | енотер (enoter) |
| русогу (rusogu) | русогу (rusogu) | усогор (usogor) | юсегор (yusygor) |
| русоху (rusoxu) | русоху (rusoxu) | усухор (usuhor) | юсехор (yusychor) |
| салуху (salуху) | салуху (saluxu) | алюху (aluhu) | ялюхес (yaluhes) |
| согата (soqata) | согата (soqata) | огата (ogata) | егатяс (egatyas) |
| суганы (sugany) | суганы (sugany) | уганис (uganis) | еъянис (ejanis) |
| тасуло (tasu) | тасуле (tasule) | асулот (asulot) | ясулот (yasulot) |
| торо (torda) | тороде (torded) | оротод (orotod) | еротед (eroted) |
| хокубы (xokubu) | хокуби (xokubi) | окубы (okubys) | екюбис (ekyubis) |
| хорька (xoryka) | хорьки (xorki) | орьках (orkaks) | ериях (eryiaks) |
| хохлы (xoxilo) | хехиле (xehile) | ашылол (ashylol) | ехилол (exhilol) |
| хулула (xumala) | хоумаха (xumahax) | умаахах (umatxas) | юмаахах (ymamahax) |

### Appendix B

| CV, hard | CV, soft | VC, hard | VC, soft |
|---------|----------|----------|----------|
| багоозо (bagooz) | верясик (veryasik) | аланумг (alamug) | ёнисек (yonelex) |
| бадахо (badoz) | ветери (vyoteri) | алумумг (alamug) | ениёд (eniyod) |
| бобадо (bobado) | ветюсо (vetysyo) | азомол (azomol) | ениред (eniyared) |
| бозодо (bozo) | виляюр (vilyuryo) | улемаб (ulemab) | ерскос (erksos) |
| болуму (bolumu) | ваяятие (vyanatyef) | алоозол (alozol) | ерятил (eryatil) |
| бузабо (buzabo) | лёвея (lyoveya) | уломез (ulomes) | еселееб (esileeb) |
| бахого (baxogo) | лемани (lemani) | омадус (omadus) | ескевём (eskevm) |
| гулдубо (gudubo) | ливёси (livosy) | амозоз (amozoz) | етлюоз (etyuluz) |
| дагзааху (dagzaahu) | люсёря (lyusorya) | аходу (ahod) | етсёём (etyseem) |
| дазамо (dazamo) | льемэс (lyamezes) | элугут (elugut) | ивёрэд (iyveroed) |
| добулухо (doboluch) | месёля (meselya) | удаул (udal) | илнеб (ilynub) |
| дозыга (dogyga) | мириони (mirioni) | одолы (odoly) | илётол (iletol) |
| дываха (dya) | мяхнёл (myhnel) | улудоз (uludoz) | илдзог (ilyd) |
| дыважу (dyvyaj) | могуриэ (muguriye) | улудоз (uludoz) | имсил (imesil) |
| дадуха (daduka) | мянэл (myanele) | улудоз (uludoz) | имндих (imndib) |
| додула (dodula) | рёвере (ryvorer) | удульб (udulub) | иренж (irenj) |
| лузого (luzogo) | ротяся (rotyasya) | агозаз (agozaz) | иресёл (iresol) |
| любода (lyuboda) | римьес (rimyes) | оздыг (ozdyg) | искёрёд (yskored) |
| лухоха (lukoxa) | ритёмис (ritemis) | удун (udun) | ифарял (ifaryal) |
| льюма (lyumaa) | риммыла (rimmiyala) | олыхах (olykah) | ирювиом (yruviom) |
| мобаза (mobaza) | сёясюэ (syeasue) | азозух (azozux) | ямелиал (yameliy) |
| мындэб (myndeb) | сёврям (syaovyr) | убдаг (ubdag) | янорёг (yanorog) |
| мадабо (madabo) | сёсите (syosite) | удылаб (udylab) | ярилюм (yariyum) |
| мадабо (madabo) | сёмве (symbve) | оазбэд (ozaybd) | ясётюб (yasotyub) |
| магахо (maga) | сюрюта (syuryuta) | узадаб (uzadab) | ясклза (yaskylaz) |
| халумо (halum) | теноня (tenoniy) | обожоб (obogob) | ятёнёб (yatyon) |
| хоболо (kobolo) | тютиме (tyutime) | обулаг (obulag) | ятесаб (yatesab) |