

УДК 811.161.1'232

doi 10.17072/2073-6681-2022-4-46-55

Трудное для чтения слово: критерии, типология, корпусная статистика (на материале русской речи носителей китайского языка)

Кун Чунься

к. филол. н.

Санкт-Петербургский государственный университет

199034, Россия, г. Санкт-Петербург, Университетская наб., 7–9. kong.chunxia@mail.ru

SPIN-код: 7661-9044

ORCID: <https://orcid.org/0000-0001-9812-1457>*Статья поступила в редакцию 02.08.2022**Одобрена после рецензирования 21.09.2022**Принята к публикации 28.09.2022***Информация для цитирования**

Кун Чунься. Трудное для чтения слово: критерии выделения, типология, корпусная статистика (на материале русской речи носителей китайского языка) // Вестник Пермского университета. Российская и зарубежная филология. 2022. Т. 14, вып. 4. С. 46–55. doi 10.17072/2073-6681-2022-4-46-55

Аннотация. Статья посвящена описанию и систематизации слов, трудных для неподготовленного чтения на неродном языке. Материалом для анализа стали 40 спонтанных монологов чтения, записанных от 20 информантов, носителей китайского языка, поровну юношей и девушек, с разным уровнем владения русским языком (B2 или C1) и с разным психотипом (экстраверты, амбиверты, интроверты). Информанты читали два текста: сюжетный рассказ М. Зощенко «Рубашка фантази» и несюжетный отрывок из повести В. Короленко «Слепой музыкант». В записанных монологах были выделены слова с разными приметами спонтанности и ошибками, которые затем рассматривались в лексико-фонетическом, лексико-грамматическом, собственно лексическом и лексико-синтаксическом аспектах. Анализ материала показал, что критериями трудности для чтения китайцев стали такие признаки, как длина словоформы, частотность слова в русском языке, степень сложности грамматической формы, отсутствие слова в лексическом минимуме для соответствующего уровня РКИ, наличие у слова стилистических помет, особое отношение китайцев к слову и идиоматичность словосочетания. На основе этих критериев была разработана рабочая типология трудных для чтения слов и на примерах показано, как именно иностранцы преодолевают эти трудности (коммуникативные помехи). Оказалось, что наибольшие трудности у китайцев вызывают слова со сложной грамматической формой. Преодолевая такую помеху, информанты часто хезитируют, скандируют слово (произносят его по слогам), обрывают и повторяют слова, иногда даже вербализируют свои трудности и сомнения с помощью метакоммуникативных вставок. Случались в ходе чтения заминки и речевые сбои и на знакомых и нетрудных словах, что связано уже с общей неподготовленностью такого типа речевого задания. Полученная типология может быть полезна не только при общем анализе устной речи в рамках коллоквиалистики, но и при обучении чтению на неродном языке в практике преподавания русского языка в иностранной, в том числе китайской, аудитории.

Ключевые слова: спонтанный монолог; чтение; трудное слово; речевой сбой; речевой корпус; примета спонтанности.

Введение

Чтение является видом многоуровневой речевой деятельности, которая состоит в данном случае из техники чтения и понимания при чте-

нии (см.: [Сапунова 2009: 6; Звуковой корпус... 2013: 150]).

Техника чтения формируется на основе изучения фонетических, лексических, грамматиче-

ских особенностей языка, и это формирование требует постоянной тренировки. Владение техникой и умением чтения обеспечивает «перцептивную переработку письменного текста, перекодирование зрительных сигналов в смысловые единицы – восприятие графических знаков и соотношение их с определенными значениями. Выделяют несколько уровней формирования навыка чтения, которые поочередно осваиваются человеком в процессе обучения чтению» [Безруких 2007] (цит. по: [Звуковой корпус... 2013: 154]).

Одним из ключевых звеньев формирования техники чтения является узнавание слова, на которое влияют частотность, длина слова и умение понять контекст. Считается, что чем чаще встречается слово в речи, тем легче читающему найти его в своем *ментальном лексиконе*¹ и тем легче его опознать. Кроме всего прочего, на скорость распознавания слова влияет его длина. Как правило, чем длиннее слово, тем больше времени занимает его опознание. Вся полученная информация создает определенный контекст, который часто помогает понять значение слова [Сапунова 2009: 23].

Об особом звене при чтении – узнавании слова – пишет также В. А. Артемов, который предлагает важное понятие «механизма вероятностного прогнозирования», осуществляемого на трех этапах. По мнению исследователя, читающие прогнозируют слова по начальным буквам, угадывают предложения по начальным словам, затем они рефлексуют дальнейшее развитие событий по полученной информации [Артемов 1999: 15]. Такого же мнения придерживается и Т. А. Козлякова, которая считает, что «важными психологическими компонентами процесса чтения являются смысловое и вербальное прогнозирование» [Козлякова 2009].

Понимание же текста зависит от его характеристик (структура, логика, эмоциональность, лексика, грамматика, стиль и т. д.) и от индивидуально-психологических особенностей читающего (см.: [Азимов, Шукин 2009]). Понимание достигается с помощью восприятия и осмысления прочитанного. При восприятии читающий, с одной стороны, на основе уже полученных знаний (в том числе об иностранном языке), прогнозирует языковые единицы и описываемые смыслы прочитанного текста, с другой – воспринимает общий смысл текста, вкладываемый в него автором, с помощью своего багажа знаний.

При чтении и понимании важную роль играет также объем кратковременной оперативной памяти [Daneman 1988]: читающий с хорошей памятью лучше и дольше помнит прочитанное и реже совершает движение глазами назад.

Все вышеупомянутые мнения свидетельствуют о том, что чтение действительно является сложным видом рецептивной речевой деятельности. Поэтому при анализе чтения такие факторы, как техника чтения и понимание читаемого, должны изучаться взаимосвязанно. По мнению Ф. де Соссюра, «в языке нельзя отделить ни мысль от звука, ни звук от мысли; этого можно достигнуть лишь путем абстракции, что неизбежно приведет либо к чистой психологии, либо к чистой фонологии» [Соссюр 2003: 18]. При этом «очевидно, что техника чтения разных людей будет различаться и зависеть от многих факторов. Поэтому и продуцируемые во время чтения спонтанные монологи разных людей будут иметь ряд отличительных особенностей, зависящих от разных социальных и психологических характеристик информантов» [Сапунова 2009: 7].

Следовательно, восприятие осложняется еще более, когда говорящие читают текст на неродном языке, поскольку их уровень владения иностранным языком чаще всего отстает от коммуникативных требований, и, соответственно, эти два процесса (чтение и понимание) занимают больше времени и требуют больше усилий. Как справедливо отмечает И. В. Архипова, чтение на родном языке (зрелое чтение) приводит сразу к восприятию и пониманию содержания прочитанного текста, а при чтении на неродном языке восприятие языкового материала проходит через два этапа: первый – идентификация слова, второй – переработка смысла языковой единицы [Архипова 2007]. Особенно когда такую задачу выполняют читающие спонтанно, не готовясь к ней заранее, что, безусловно, вызывает целый ряд трудностей, порождающих разного рода сбои в речевом продукте (прочитанном тексте). Именно такое неподготовленное чтение на неродном языке и стало объектом настоящего исследования.

Цель

В ходе чтения на неродном языке наибольшие трудности порождает лексический состав текста, поэтому в ходе исследования была поставлена цель: выявить, какие именно слова в исходных текстах являются трудными для информантов-китайцев и как они с этими трудностями справлялись в ходе неподготовленного (спонтанного) чтения.

Материал и методы

Исследование выполнено на основе корпусного подхода. Источником материала стал блок русской интерферированной речи носителей китайского языка (RIK) в составе корпуса русской монологической речи, созданного в СПбГУ и известного как «Сбалансированная аннотиро-

ванная текстотека» (САТ) (см. о нем подробнее: [Звуковой корпус... 2013; Богданова-Бегларян, Шерстинова, Зайдес 2017; Богданова-Бегларян и др. 2019]). Конкретным материалом для анализа стали 40 монологов неподготовленного чтения на русском языке, записанных от 20 носителей китайского языка, поровну сюжетных (рассказ М. Зоценко «Рубашка фантази») и несюжетных (отрывок из повести В. Короленко «Слепой музыкант») текстов.

Информантами были 20 носителей китайского языка, обучающихся (на момент записи) в вузах Санкт-Петербурга. Состав информантов был строго сбалансирован социологически: по гендеру (10 юношей и 10 девушек) и уровню владения русским языком: 10 человек с уровнем В2 (ТРКИ-2) и 10 – с уровнем С1 (ТРКИ-3) согласно Российской государственной системе тестирования иностранных граждан [Центр... 2022]. По возрасту все информанты составляют достаточно однородную группу: 23–28 лет. Кроме того, все информанты прошли психологическое тестирование с использованием теста Г. Айзенка (см.: [Eysenk H. J., Eysenk S. B. G 1964; Личностный опросник... 1995]) на китайском языке [http://types.yuzeli.com/survey/epq85]. В результате теста в группе информантов было выявлено 5 экстравертов, 6 амбивертов и 9 интровертов. Все эти метаданные информантов учитывались в ходе исследования.

Классификация и статистика

Качество прочтения иностранцами русского слова зависит от многих факторов, которые и стали критериями выделения в читаемых текстах группы трудных слов (ТС), а также позволили создать их рабочую типологию. Среди таких критериев – наличие единицы в лексическом минимуме для соответствующего уровня ТРКИ, ее длина, частотность в русском языке и лексико-фонетические характеристики, соответствие китайскому менталитету, стилистические, грамматические характеристики и ряд других. На основе этих критериев и была построена типология ТС, полезная для анализа рассматриваемых монологов чтения. Были выявлены следующие группы трудных для китайцев слов.

1. Слова и словоформы, содержащие фонетически трудные фрагменты (как правило, сочетания согласных²): *к прачке, вздохнуть, хамство*:

- 1) *кэ-э* [*прачку прачке* (И12, жен., А, С1, чтение сюж.)³;
- 2) *без* [*хамУ-Уства* (И17, муж., И, В2, чтение сюж.)⁴.

2. Многосложные (многоморфемные, длинные) слова: *чистоплотный, необъяснимый*. При неподготовленном чтении большие усилия ин-

формант затрачивает именно на распознавание слова, на осознанный или неосознанный контроль непрерывности чтения и качества речи, а длина слова во многом влияет на скорость такого распознавания и в целом на скорость чтения, ср.:

- 3) *какой чисто...* [*плотный* (И14, жен., А, В2, чтение сюж.);
- 4) *нео... необъяс...* [*ни-мА-я тьма* (И17, муж., И, В2, чтение несюж.).

В процессе чтения подобных слов информанту приходится преодолевать возникающие трудности. Так, И14 в контексте (3) прочитал слово *чистоплотный* с разрывом на две части (*чисто... [плотный]*): такая хезитация в середине слова дала ему время для перевода взгляда на следующую часть слова, и таким образом он прочитал слово полно и правильно. В контексте (4) И15 дважды оборвал себя на слове *необъяснимая* (*нео... необъяс...*) и только после продолжительной хезитационной паузы (//) дочитал его со скандированием, замедляя тем самым темп чтения. Но даже при такой степени «осторожности» он все же поставил неправильное ударение в этом слове.

3. Разговорные формы сравнительной степени (компаратива) наречий и прилагательных: *поприличней, покрасивей, поскорей*; в учебниках по РКИ даются прежде всего литературные варианты этих форм – с суффиксом *-ее-* (*поприличнее, покрасивее, поскорее*), – поэтому разговорные варианты вызывают некоторую заминку у читающего иностранца в ходе восприятия их письменного облика. Во многих случаях информанты просто заменили оригинальные варианты их литературными эквивалентами:

- 5) *поприличнее* [*одеться* (И20, муж., А, С1, чтение сюж.);
- 6) *хотелось какую-нибудь рубашку / покрас...* *покрасивЕе купить* (И19, муж., Э, В2, чтение сюж.).

В контексте (5) информант И20 после прочтения словоформы *поприличнее* сделал остановку (паузу хезитации – /): возможно, он заметил свою ошибку, но все же не стал возвращаться и не исправил эту неточность. В следующем контексте (6) И19, предчувствуя трудность в завершении слова (взгляд снова опередил язык), оборвал это слово, затем повторил его, но при этом был, видимо, слишком напряжен, акцентировал больше внимания на суффиксе и в итоге неуместно перенес на него ударение (*покрасивЕе*).

4. Сложные грамматические формы (прежде всего причастия и деепричастия): *разопревшие, опрокинутым, реющею, колеблющеюся, зарисовывающуюся, притягиваясь*⁵; одновременно эти формы оказываются, как правило, и длинными,

многосложными, что повышает степень их трудности для читателя, ср.:

7) *опро...* [*опро...*] *кнУтыми туным э-э* [*о-про-к-нУтым* (И19, муж., Э, В2, чтение сюж.);

8) *колеблЮ...* [*колеблЮщ...*] *колеблЮщеюсь лЮ...* *колеблЮщеюси* (И20, муж., А, С1, чтение сюж.).

В контексте (7) И19 явно испытывает большие затруднения при прочтении длинного и грамматически сложного слова, о чем свидетельствуют многократные сбои. Сначала он оборвал слово, потом повторил его с разрывом и акцентологической ошибкой, похезитировал с помощью вставки неречевого звука (э-э) и длинной физической паузы, затем проскандировал это слово, но опять ошибся в ударении. В следующем контексте (8) также наблюдаются разрывы, повторы и акцентологическая ошибка в прочтении слова *колеблющеюся*. При этом И20 два раза реализовал квазислово (квазиформу) *колеблющеюсь*. Согласно мнению И. В. Королевой, при чтении квазислова у читающего активизируется ментальный лексикон, запускается механизм поиска фонетического или семантического сходства с другими словами [Королева 2006: 43]. Поэтому такая ошибка не случайна: вполне возможно, что информант реализовал в этом случае более знакомый для себя (иностранца) слог *-юсь* в конце возвратного глагола в форме 1 л., ед. ч.

5. Имена собственные. В текстах было одно женское имя (*Лукерья Петровна*) и две фамилии писателей – *Зоценко* и *Короленко*⁶. Русские имена, в отличие от привычных для информантов китайских, часто провоцируют появление в их речи разных хезитационных явлений и ошибок, ср.:

9) *ЗоцЁнко* (И17, муж., А, В2, чтение сюж.);

10) *н-н Лу-керья* [*Де...*] *тровна* (И15, жен., И, В2, чтение сюж.).

6. Стилистически маркированные слова, не входящие в лексические минимумы для иностранцев [Лексический минимум... 2014, 2018]. Такие слова сложны для прочтения именно в связи с наличием у них стилистических помет, поскольку иностранные учащиеся осваивают прежде всего кодифицированный литературный язык, представленный в словарях и учебных материалах. Стилистически маркированные единицы редко становятся предметом специального изучения в иностранной аудитории. В анализируемых текстах обнаружился целый ряд таких слов⁷: *заскочил* (разг.), *этакую* (прост., мест.), *барышни* (устар.), *двугривенный* (разг.), *голубушка* (разг.), *расстарайся* (разг.), *теперича* (прост.), *давеча* (прост.), *завсегда* (прост.), *в аккурат* (прост.), *простирнуть* (прост.), *напялил* (разг.),

загляденье (разг.), *мать честная!* (прост.), *небось* (прост.), *стираная* (разг.), *морда* (прост.), *хамство* (разг.), *подымалось* (книж., устар.) (см. соответствующие пометы в разных томах словаря: [МАС 1999]):

11) *голу...* / *голубушка голубУшка* (И11, жен., И, С1, чтение сюж.);

12) *небо...* *небось* (И2, муж., И, С1, чтение сюж.).

В обоих контекстах можно видеть обрывы и повторы, которые помогают говорящему выйти из «точки сбоя» и продолжить чтение.

7. Нечастотные слова, также не входящие в лексические минимумы для иностранцев [Лексический минимум... 2014, 2018] и потому также незнакомые или малознакомые информантам. В материале исследования встретилось довольно много таких слов, ср.: *воротник* (17,8 ipm⁸), *хватался* (11,3 ipm), *воротничок* (5,69 ipm), *льдина* (6,18 ipm), *лиман* (0,93 ipm), *сверкать* (24,6 ipm), *лачуга* (0,95 ipm), *солома* (9,9 ipm), *фимиам* (0,63 ipm), *раскинуться* (5,69 ipm), *шевелиться* (20,79 ipm), *рокотать* (1,45 ipm), *неизведанный* (2,38 ipm) (см.: [Ляшевская, Шаров 2009]). Обнаружились в читаемых текстах и слова, которых вообще нет в частотном словаре: *пристежными*, *к завтраму*, *фантази*, *глаженая*, *манжетки*:

13) *неизведами* [*неизведанными* (И1, жен., Э, В2, чтение несюж.);

14) *фи...* [*миам...*] *фи...* [*фимиама* (И3, жен., И, В2, чтение несюж.).

В контексте (13) информант И1 реализовал квазислово, затем, после паузы хезитации, повторил слово корректно. В следующем контексте (14) слово также прочитано информантом И3 многочисленными сбоями.

8. Этически неприемлемые слова. Так, слово *чёрт*, хорошо знакомое всем читающим, вызывает у китайцев отчетливо негативную ассоциацию, поэтому его прочтение сопровождалось многочисленными хезитационными явлениями (физическая пауза хезитации и вставка неречевых звуков). По результатам опроса, проведенного в группе из 322 китайцев разного пола, возраста и уровня владения русским языком, почти все респонденты знают слово *чёрт* и знают много выражений, содержащих его, но при этом треть из них выразили негативное отношение к данному слову (см. об этом подробнее: [Кун Чунья 2021a]):

15) *чёрт* [*их знает*] *думаю* (И11, жен., И, С1, чтение сюж.);

16) *э-э чёрт э* [*их э-э значит знает* (И16, муж., А, В2, чтение сюж.).

9. Идиоматический оборот: *чёрт их знает, что за чёрт! ах, чёрт! Мать честная!* Возможно, те особенности, которые выявлены при прочтении

слова *чёрт*, связаны не только с особым отношением китайцев к этому слову, но и с фактом его употребления в тексте М. Зощенко исключительно в составе идиом (незнание совокупности смысла компонентов и условий употребления идиомы). При «пословной» стратегии чтения сверхсловные единства являются для иностранцев своего рода «камнем преткновения», поскольку идиомы могут состоять из компонентов, семантически не связанных с окружающими словами, и выглядят текстовой аномалией, ср.:

- 17) *ах чёрт / э-э чего думаю / делать* (И15, жен., И, В2, чтение сюж.);
 18) *что за мать / честная* (И13, муж., А, С1, чтение сюж.).

Анализ материала показал, что чтение выделенных трудных слов действительно сопровождается разнообразными приметами спонтанности (см. об этом подробнее: [Кун Чунься 2021б]). Статистика возникших у говорящего затруднений (запинок, речевых сбоев) при прочтении ТС представлена в таблице. Доля затруднений высчитывалась отдельно в рамках каждой выделенной группы трудных слов. Скажем, в группе фонетически трудных слов затруднения встретились в 61,7 % случаев употребления таких слов. Значит, еще 38,3 % фонетически трудных слов были прочитаны без затруднения. Иными словами, 100 % – это количество единиц внутри каждой группы.

Доля трудных слов разного типа, прочитанных с затруднениями (в %)
 The Proportion of Difficult Words of Different Types Read with Difficulties (in %)

Группа трудных слов	Подгруппа	Доля ТС, прочитанных с затруднениями	Доля ТС, прочитанных с затруднениями, в целом по группе
Лексико-фонетическая	Фонетически трудные слова	61,7	54,0
	Длинные слова	42,5	
Лексико-грамматическая	Разговорные грамматические формы	56,7	64,5
	Сложные грамматические формы	67,9	
Лексическая	Имена собственные	41,9	61,3
	Стилистически окрашенные слова	58,5	
	Нечастотные слова	62,4	
	Этически неприемлемые слова	18,3	
Лексико-синтаксическая	Идиоматический оборот	27,5	27,5

Из таблицы видно, что чаще всего информанты испытывали какие-либо затруднения при чтении слов из лексико-грамматической группы (64,5 %), особенно форм причастий и деепричастий от возвратных глаголов (*колеблющуюся, зарисовавшуюся, притягиваясь*) (67,9 %).

Затруднений при прочтении слова *чёрт* в материале исследования встретилось не слишком много (всего 18,3 % от всех его реализаций), но показалось целесообразным и это слово отнести к числу трудных, сформировав группу этически неприемлемых слов, потому что это расширяет наши представления о возможных причинах затруднений китайцев при чтении на русском языке.

Помимо слова *чёрт*, трудности разного типа и разной степени вызывали и некоторые другие, хорошо знакомые информантам, слова (см. некоторые наблюдения такого рода в работе: [Богданова-Бегларян, Кун Чунься 2021]). Например,

многие затруднились в постановке ударения в словах *лезет, надеяться, примерить* и под.:

- 19) *завтра вечеринка надо к завтрашнем / могу ли надеЯться / надеяться говорит можно* (И1, жен., Э, В2, чтение сюж.);
 20) *на горло / не лезёт / не лезет / не лезёт* 忘丁 <‘забыла’ – перевод мой. – Ч. К.> (И18, жен., Э, С1, чтение сюж.);
 21) *вот / примерИл / эту рубаху / и-и / как-то / и / как-то не / по себе стало* (И18, жен., Э, С1, чтение сюж.).

По примеру (19) видно, что слово *надеяться* хорошо знакомо информанту: совершив спонтанную акцентологическую ошибку, он тут же, после небольшой заминки, практически в режиме *онлайн*⁹, исправился и повторил слово с правильным ударением.

В втором (20) примере информант И18 долго колеблется между правильным и неправильным вариантами прочтения слова *лезет*, сомневается

в ударении, но останавливается все-таки на неправильном варианте и при этом, с помощью метакоммуникативной вставки на родном языке (忘了 <‘забыла’>), оправдывает себя: он (в данном случае – она), дескать, знает это слово, просто забыл(а) его ударение.

Заключение

Проведенный анализ показал, таким образом, что самыми трудными для прочтения носителями китайского языка оказались сложные грамматические формы (причастия и деепричастия). Кроме того, ключевыми характеристиками, способными повлиять на качество чтения, стали длина, частотность и стилистические характеристики слова. Столкнувшись с ТС, китайские информанты невольно хезитируют, делают ошибки, заменяют слова и звуки, обрывают себя, перестраивают фразу на ходу и допускают прочие речевые сбои, которые можно отнести к так называемым *приметам спонтанности* и которые свидетельствуют, в числе прочего, о недостаточной сформированности у иностранных испытуемых навыка чтения и о незнании ими некоторых аспектов русского языка.

Предложенная типология ТС может быть полезна в методическом плане: не только для анализа материала чтения на неродном языке, но и для совершенствования практики преподавания РКИ, в частности в китайской аудитории.

Примечания

¹ Под *ментальным лексиконом* в лингвистике понимается как минимум словарный запас конкретной языковой личности, а как максимум – «сложная многоярусная система пересекающихся полей, представляющих собой упорядоченную по разным основаниям информацию как о явлениях действительности, так и о связанных с ними языковых единицах, сложную сеть взаимосвязей, увязывающую огромное количество знаний в памяти человека» [Залевская 1990: 87–88], или, иными словами, «“словарь в голове” индивида, функционирующий в соответствии с закономерностями психического развития человека – носителя языка и культуры» [Золотова 2005: 3]. Нет сомнения, таким образом, что «ментальный лексикон определяется не только словарем родного языка говорящего, но и всем его жизненным опытом, всей совокупностью его знаний об окружающей действительности. А значит, этот лексикон принципиально различен у разных людей и в целом не совпадает со словарем языка» [Богданова 2011: 57]. В последнее время изучение ментального лексикона,

его устройства и функционирования, выдвигается на первый план, оказываясь на пересечении интересов лингвистики и психологии. Ср. еще одно подходящее определение: ментальный лексикон – это «сеть, в которой каждое слово связано со всеми другими» [Караулов 1987: 85]. Исследования ментальных лексиконов посвящены, прежде всего, «выяснению принципиальных вопросов: как извлекаются из памяти слова, как они в ней репрезентированы и организованы; как взаимодействуют лингвистические и психологические факторы при формировании репрезентации лексического знания и обеспечении лексической обработки в реальном времени» [Глазанова 2001: 7–8]. В практике преподавания РКИ все эти моменты нужно непременно учитывать.

² Для китайского языка не характерно стечение согласных, которое и разрезается гласной вставкой (см. литературу об этом: [Чэн Чэнь 2021]).

³ В атрибуции ко всем примерам в статье указан номер информанта (И1, И2, ...), его пол, психотип (Э – экстраверт, А – амбиверт, И – интроверт), уровень ТРКИ (В2 или С1), а также тип исходного текста для чтения – сюжетный или несюжетный. Знак (l) означает физическую паузу хезитации, которая может быть разной длительности: (l) – краткая, фактически – запинка говорящего, (ll) – более долгая. О других особенностях орфографического представления материалов блока RIK см.: [Чэн Чэнь 2021: 156–157; Кун Чунься 2022: 143–144].

⁴ Здесь и далее прописная буква в слове означает неправильное ударение. Удвоение буквы означает растяжку соответствующего звука, что также относится к приметам спонтанности [там же].

⁵ Стоит отметить, что здесь не только форма сложна, но и лексика во многом устаревшая и трудная для осмысления (например: *зарисовавшаяся*).

⁶ Далеко не все информанты прочитывали фамилии авторов текстов, но большинство их реализаций были некорректными.

⁷ Стоит отметить, что трудность при чтении подобных незнакомых или малознакомых слов возникает далеко не всегда, поскольку короткие слова в любом случае дают возможность читающему для быстрого их распознавания.

⁸ IPM (англ. *instances per million*) – стандартное представление частоты токена или леммы, вычисляемое относительно условного корпуса в миллион единиц, независимо от объема реального корпуса.

⁹ Вслед за В. И. Подлесской и А. А. Кибриком [Подлесская, Кибрик 2007], в работе выделяются две основные стратегии, которые использует говорящий для того, чтобы исправить ошибку или преодолеть речевое затруднение / колебание: сразу после ошибочного произнесения фрагмента – *онлайн коррекция* – или дистантно, после какой-то новой порции текста, в течение которой говорящий продолжает помнить об ошибке и возвращается к ней, чтобы исправить, – *офлайн коррекция*.

Источники

Азимов, Шукин 1999 – Азимов Э. Г., Шукин А. И. *Словарь методических терминов (теория и практика преподавания языков)*. СПб.: ИКАР, 1999. 472 с.

Лексический минимум... 2014 – Лексический минимум по русскому языку как иностранному. Второй сертификационный уровень. Общее владение. Андришина Н. П. (ред.). 5-е изд. СПб.: Златоуст, 2014. 164 с.

Лексический минимум... 2018 – Лексический минимум по русскому языку как иностранному. Третий сертификационный уровень. Общее владение. Андришина Н. П. (ред.). СПб.: Златоуст, 2018. 200 с.

Личностный опросник... 1995 – Личностный опросник ЕРІ (методика Г. Айзенка). Альманах психологических тестов. Римский Р. Р., Римская С. А. (ред.). М.: КСП, 1995. С. 217–224.

Ляшевская, Шаров 2009 – Ляшевская О. Н., Шаров С. А. *Частотный словарь современного русского языка* (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009. 1112 с.

МАС 1999 – *Словарь русского языка: в 4 т.* А. П. Евгеньева (ред.). М.: Русский язык: Полиграфресурсы, 1999. Т. 1: А–Й, 702 с.; Т. 2: К–О, 736 с.; Т. 3: П–Р, 750 с.; Т. 4: С–Я, 797 с.

Центр... 2022 – Центр языкового тестирования [Электронный ресурс]. URL: <https://testingcenter.spbu.ru/ekzameny/russia/trki.html> (дата обращения: 20.01.2022).

艾森克人格问卷 (EPQ) 成人版 [Личностный опросник Г. Айзенка (EPQ) для взрослых]. URL: <https://types.yuzeli.com/survey/epq85> (дата обращения: 14.12.2020). (На китайском)

Список литературы

Артемов В. А. *Психология обучения иностранным языкам*. М.: Просвещение, 1999. 228 с.

Архипова И. В. *Некоторые особенности понимания текста при чтении на иностранном языке. Социально-экономические и технические систе-*

мы: исследование, проектирование, оптимизация. 2007. № 6. С. 300–302.

Безруких М. М. *Почему современные подростки плохо читают? Школьная библиотека*. 2007. № 9–10. С. 92–95.

Богданова Н. В. *Русское слово в трех режимах фиксации – словарь, ментальный лексикон и реальное употребление (лексикографический и лингвометодический аспекты)*. Русский язык за рубежом. 2011. № 6 (229). С. 56–64.

Богданова-Бегларян Н. В., Шерстинова Т. Ю., Зайдес К. Д. *Корпус «Сбалансированная Аннотированная Текстотека»: методика многоуровневого анализа русской монологической речи. Анализ разговорной русской речи (АР³-2017): труды седьмого междисциплинарного семинара*. Кочаров Д. А., Скрелин П. А. (ред.). СПб.: Политехника-принт, 2017. С. 8–13.

Богданова-Бегларян Н. В. и др. *Корпус «Сбалансированная аннотированная текстотека» (САТ): изучение специфики русской монологической речи / Н. В. Богданова-Бегларян, О. В. Блинова, К. Д. Зайдес, Т. Ю. Шерстинова // Труды Института русского языка им. В. В. Виноградова. Вып. 21. Национальный корпус русского языка: исследования и разработки / ред. А. М. Молдован, отв. ред. выпуска В. А. Плуныян. М., 2019. С. 111–126.*

Глазанова Е. В. *Типы связей в ментальном лексиконе и экспериментальные методы их исследования: дис. ... канд. филол. наук*. СПб., 2001. 237 с.

Залевская А. А. *Слово в лексиконе человека: Психолингвистическое исследование*. Воронеж: Изд-во Воронеж. ун-та, 1990. 205 с.

Звуковой корпус как материал для анализа русской речи: колл. монография. Часть 1. Чтение. Пересказ. Описание / отв. ред. Н. В. Богданова-Бегларян. СПб.: Филол. ф-т СПбГУ, 2013. 532 с.

Золотова Н. О. *Ядро ментального лексикона человека как естественный метаязык: автореф. дис. ... д-ра филол. наук*. Тверь, 2005. 44 с.

Караулов Ю. Н. *Русский язык и языковая личность*. М.: Наука, 1987. 264 с.

Козлякова Т. А. *Роль чтения в обучении иностранных учащихся на начальном этапе // Труды Белорусского государственного технологического университета*. 2009. № 5. С. 211–213.

Королёва И. В. *Роль лингвистических факторов в развитии процессов чтения (экспериментальное исследование на материале русского языка: дис. ... канд. филол. наук*. СПб., 2006. 227 с.

Кун Чунься. *Неподготовленное чтение как разновидность спонтанной речи (о приметах*

спонтанности в чтении на неродном языке) // Вестник Пермского университета. Российская и зарубежная филология. 2021а. Т. 13, № 2. С. 36–46. doi 10.17072/2073-6681-2021-2-36-46

Кун Чунья. Слово ЧЁРТ в русской речи китайцев, или о причинах запинок в ходе неподготовленного чтения на неродном языке // Коммуникативные исследования. 2021б. Т. 8(2). С. 310–322.

Кун Чунья. Специфика неподготовленного чтения на неродном языке (комплексное исследование на материале русской речи носителей китайского языка): дис. ... канд. филол. наук. СПб., 2022. 162 с.

Подлеская В. И., Кибрик А. А. Самоисправление говорящего и другие типы речевых сбоев как объект аннотирования в корпусах устной речи // Научно-техническая информация. 2007. Серия 2(23). С. 1–39.

Сапунова Е. М. Неподготовленное чтение как вид речевой деятельности и тип устного спонтанного монолога (на материале русского языка): дис. ... канд. филол. наук. СПб., 2009. 237 с.

Соссюр Ф. де. Курс общей лингвистики (Извлечения). Психоллингвистика в очерках и извлечениях. Хрестоматия. Радзиховская В. К. (ред.). М.: Академия, 2003. С. 184–195.

Чэн Чэнь. Хезитации в русской устной речи носителей китайского языка. СПб.: Нестор-История, 2021. 231 с.

Daneman M. Individual Differences in Reading Skills. Handbook of Reading Research. R. Barr, P. D. Pearson, M. L. Kamil, P. Mosenthal (eds.). Vol. II. 1996. P. 512–539.

Eysenk H. J., Eysenk S. B. Manual of the Eysenck Personality Inventory. London: Univ. of London press, 1964. 24 p.

References

Artemov V. A. *Psikhologiya obucheniya inostrannym yazykam* [Psychology of Teaching Foreign Languages]. Moscow, Prosveshcheniye Publ., 1999. 228 p. (In Russ.)

Arkhipova I. V. Nekotorye osobennosti ponimaniya teksta pri chtenii na inostrannom yazyke [Some features of understanding texts when reading in a foreign language]. *Sotsial'no-ekonomicheskie i tekhnicheskie sistemy: issledovanie, proektirovanie, optimizatsiya* [Socio-Economic and Technical Systems: Research, Design, Optimization], 2007, issue 6, pp. 300–302. (In Russ.)

Bezrukikh M. M. Pochemu sovremennye podrostki plokhо chitayut? [Why are today's teenagers bad at reading?]. *Shkol'naya biblioteka* [School Library], 2007, issue 9–10, pp. 92–95. (In Russ.)

Bogdanova N. V. Russkoe slovo v trekh rezhimakh fiksatsii – slovar', mental'nyy leksikon i real'noe upotreblenie (leksikograficheskiy i lingvometodicheskiy aspekty) [Russian word in three modes of fixation – a dictionary, mental lexicon, and real usage (lexicographical and linguistic-methodological aspects)]. *Russkiy yazyk za rubezhom* [Russian Language Abroad], 2011, issue 6 (229), pp. 56–64. (In Russ.)

Bogdanova-Beglarian N. V., Sherstinova T. Yu., Zaydes K. D. Korpus 'Sbalansirovannaya annotirovannaya tekstoteka': metodika mnogourovnevnogo analiza russkoy monologicheskoy rechi [The corpus 'Balanced annotated text collection': A methodology for multilevel analysis of Russian monological speech]. *Analiz razgovornoy russkoy rechi (AR3-2017): Trudy sed'mogo mezhdistsiplinarnogo seminara* [Analysis of Colloquial Russian Speech (AR3-2017): Papers of the 7th interdisciplinary seminar]. Ed. by D. A. Kocharov, P. A. Skrelin. St. Petersburg, Politekhniko-print Publ., 2017, pp. 8–13. (In Russ.)

Bogdanova-Beglarian N. V., Blinova O. V., Zaydes K. D., Sherstinova, T. Yu. Korpus 'Sbalansirovannaya annotirovannaya tekstoteka' (SAT): izuchenie spetsifiki russkoy monologicheskoy rechi [Corpus 'Balanced annotated text collection (textotec)' (SAT): Studying the specificity of Russian monological speech]. *Trudy Instituta russkogo yazyka imeni V. V. Vinogradova. Natsional'nyy korpus russkogo yazyka: issledovaniya i razrabotki* [Proceedings of the V. V. Vinogradov Russian Language Institute. Issue 21. The Russian National Corpus: Research and Developments]. Ed. by A. M. Moldovan, issue editor V. A. Plungyan. Moscow, 2019, pp. 111–126. (In Russ.)

Glazanova E. V. *Tipy svyazey v mental'nom leksikone i eksperimental'nye metody ikh issledovaniya*. Diss. kand. filol. nauk [Types of connections in the mental lexicon and experimental methods of their research. Cand. philol. sci. diss.]. St. Petersburg, 2001. 237 p. (In Russ.)

Zalevskaya A. A. *Slovo v leksikone cheloveka: Psikholingvisticheskoe issledovanie* [Word in the Human Lexicon: A Psycholinguistic Study]. Voronezh, Voronezh State University, 1990. 205 p. (In Russ.)

Zvukovoy korpus kak material dlya analiza russkoy rechi. Kollektivnaya monografiya. Chast' I. Chtenie. Pereskaz. Opisanie [A Speech Corpus as Material for the Analysis of Russian speech. Collective Monograph. Pt. 1. Reading. Retelling. Description]. Ed. by N. V. Bogdanova-Beglarian. St. Petersburg, Faculty of Philology of Saint Petersburg State University Press, 2013. 532 p. (In Russ.)

Zolotova N. O. *Yadro mental'nogo leksikona cheloveka kak estestvennyy metayazyk*. Avtoref diss. dokt. filol. nauk [The core of the human mental lexicon as a natural meta-language. Abstract of Dr. philol. sci. diss.]. Tver, 2005. 44 p. (In Russ.)

Karaulov Yu. N. *Russkiy yazyk i yazykovaya lichnost'* [The Russian Language and Language Personality]. Moscow, Nauka Publ., 1987. 264 p. (In Russ.)

Kozlyakova T. A. Rol' chteniya v obuchenii inostrannykh uchashchikhsya na nachal'nom etape [The role of reading at the early stage of teaching foreign students]. *Trudy Belorusskogo gosudarstvennogo tekhnologicheskogo universiteta* [Proceedings of the Belarusian State Technological University], 2009, issue 5, pp. 211–213. (In Russ.)

Koroleva I. V. *Rol' lingvisticheskikh faktorov v razviti protsessov chteniya (eksperimental'noe issledovanie na materiale russkogo yazyka)*. Diss. kand. filol. nauk [The role of linguistic factors in the development of reading processes (an experimental study on the material of the Russian language. Cand. philol. sci. diss.]. St. Petersburg, 2006. 227 p. (In Russ.)

Kong Chunxia. *Nepodgotovlennoe chtenie kak raznovidnost' spontanoy rechi (o primetakh spontannosti v chtenii na nerodnom yazyke)* [Unprepared reading as a kind of spontaneous speech (about signs of spontaneity in reading in a non-native language)]. *Vestnik Permskogo universiteta. Rossiyskaya i zarubezhnaya filologiya* [Perm University Herald. Russian and Foreign Philology], 2021a, issue 13 (2), pp. 36–46. doi 10.17072/2073-6681-2021-2-36-46. (In Russ.)

Kong Chunxia. *Slovo CHERT v russkoy rechi kitaytsev, ili o prichinakh zapinok v khode nepodgotovlennogo chteniya na nerodnom yazyke* [The word chort (devil) in the Russian speech of the Chinese, or on the reasons for stumbling during unprepared reading in a non-native language]. *Kommunikativnye issledovaniya* [Communication Studies], 2021b, issue 8(2), pp. 310–322. (In Russ.)

Kong Chunxia. *Spetsifika nepodgotovlennogo chteniya na nerodnom yazyke (kompleksnoe issledovanie na materiale russkoy rechi nositeley kitayskogo yazyka)*. Diss. kand. filol. nauk [Features of unprepared reading in a non-native language (a comprehensive study on the material of the Russian speech of native Chinese speakers). Cand. philol. sci. diss.]. St. Petersburg, 2022. 162 p. (In Russ.)

Podlesskaya V. I., Kibrik A. A. *Samoispravleniya govoryashchego i drugie tipy rechevykh sboev kak ob'ekt annotirovaniya v korpusakh ustnoy rechi* [Speaker's self-corrections and other types of speech failures as the object of annotation in speech corpora]. *Nauchno-tekhnicheskaya informatsiya* [Scientific and Technical Information], 2007, issue 2(23), pp. 1–39. (In Russ.)

Sapunova E. M. *Nepodgotovlennoe chtenie kak vid rechevoy deyatel'nosti i tip ustnogo spontanogo monologa (na materiale russkogo yazyka)*. Diss. kand. filol. nauk [Unprepared reading as a type of speech activity and a type of oral spontaneous monologue (a case study of the Russian language). Cand. philol. sci. diss.]. St. Petersburg, 2009. 237 p. (In Russ.)

Saussure F. de. *Kurs obshchey lingvistiki (Izvlacheniya)* [Course in general linguistics (extracts)]. *Psikholingvistika v ocherkakh i izvlecheniyakh* [Psycholinguistics in Essays and Extracts]. Moscow, Academia Publ., 2003, pp. 184–195. (In Russ.)

Cheng Chen. *Khezitatsii v russkoy ustnoy rechi nositeley kitayskogo yazyka* [Hesitations in Russian Speech of Chinese Speakers]. St. Petersburg, Nestor-Istoria Publ., 2021. 231 p. (In Russ.)

Daneman M. Individual differences in reading skills. *Handbook of Reading Research*. Ed. by R. Barr, P. D. Pearson, M. L. Kamil, P. Mosenthal. 1996, vol. 2, pp. 512–539. (In Eng.)

Eysenk H. J., Eysenk S. B. *Manual of the Eysenck Personality Inventory*. London, London University Press, 1964. 24 p. (In Eng.)

Words Difficult to Read: Criteria, Typology, Corpus Statistics (based on the Russian speech of native Chinese speakers)

Kong Chunxia

PhD Graduate

Saint Petersburg State University

7–9, Universitetskaya Embankment, St. Petersburg, 199034, Russian Federation. kong.chunxia@mail.ru

SPIN-code: 7661-9044

ORCID: <https://orcid.org/0000-0001-9812-1457>

Submitted 2 Aug 2022

Revised 21 Sep 2022

Accepted 28 Sep 2022

For citation

Kong Chunxia. Trudnoe dlya chteniya slovo: kriterii, tipologiya, korpusnaya statistika (na materiale russkoy rechi nositeley kitayskogo yazyka) [Words Difficult to Read: Criteria, Typology, Corpus Statistics (based on the Russian speech of native Chinese speakers)]. *Vestnik Permskogo universiteta. Rossiyskaya i zarubezhnaya filologiya* [Perm University Herald. Russian and Foreign Philology], 2022, vol. 14, issue 4, pp. 46–55. doi 10.17072/2073-6681-2022-4-46-55 (In Russ.)

Abstract. The article is devoted to the description and systematization of words that are difficult for unprepared reading in a non-native language. The material for the analysis included 40 spontaneous reading monologues recorded from 20 informants (10 male and 10 female), native Chinese speakers, with different levels of proficiency in Russian (B2 or C1) and with different psychotypes (extroverts, ambiverts, introverts). Informants were reading two texts: M. Zoshchenko's narrative story *The Fantasy Shirt* and a non-narrative excerpt from V. Korolenko's story *The Blind Musician*. In the recorded monologues, words with different signs of spontaneity and errors were singled out, and they were then examined in the lexico-phonetic, lexico-grammatical, lexical, and lexico-syntactical aspects. The analysis of the material showed the following features to be the difficulty criteria for the Chinese readers: the length of the word form, the word frequency in Russian, the degree of complexity of the grammatical form, the absence of the word in the basic vocabulary list for the corresponding level of Russian as a foreign language, the presence of stylistic usage labels in dictionary entries for the word, the special attitude of the Chinese to the word, the idiomaticity of the word combination. Based on these criteria, a typology of difficult-to-read words was developed. The paper provides examples showing how foreigners overcome these difficulties (communicative obstacles). It has been found that the most difficult for Chinese are words with a complex grammatical form. When faced with such words, informants often make hesitation pauses, syllabize, cut off and repeat words, sometimes even verbalize their difficulties and doubts with the help of metacommunicative inserts. There were also noted hiccups and speech failures when the informants were reading familiar and easy words, which was due to the general unpreparedness of this type of speech task. The typology obtained can be useful not only in the general analysis of oral speech by specialists studying colloquial language but also in the teaching of reading in a non-native language as part of teaching the Russian language in a foreign, including Chinese-speaking, audience.

Key words: spontaneous monologue; reading; difficult word; speech failure; speech corpus; signs of spontaneity.