

ПОДГОТОВКА УСЛОВНЫХ МЕТОК ДЛЯ ОЦЕНКИ УДОБОЧИТАЕМОСТИ ШРИФТОВ

In clause feature of a technique of an estimation legibility fonts is analyzed: preparation of conditional labels in a test material for the control of reading of the continuous text.

Для исследования удобочитаемости шрифта основной является методика ее оценки, связанная с определением скорости чтения. Одним из вариантов данной методики является установление времени чтения текста в шрифтовом оформлении с учетом контроля чтения, при этом особенностью способа контроля является учет условных меток в тестовом материале, способствующих: 1) определению более точного времени чтения текста, 2) оценке различимости шрифта на основании определения сходства букв русского алфавита, в т. ч. и сходства букв при наборе их конкретным шрифтом.

Способ контроля чтения заключается в том, что в тестовый материал в произвольно отобранные слова, длиной больше шести символов, умышленно вводятся условные метки, которые можно найти только при внимательном прочтении слова. По количеству правильно найденных меток и по времени их нахождения можно судить о внимательности прочтения текста. Метки, т. е. замена букв в слове, вводятся в текст по определенному принципу: заменяются не любые буквы, а только те, конструкции которых при чтении не всегда могут быть правильно идентифицированы, например, из-за сходства букв русского алфавита, из-за особенностей рисунка буквы данного шрифта или дефектов печати, из-за низкой различимости тестового шрифта.

Для уточнения условных меток тестового материала необходимо определить группы буквенных знаков русского алфавита. Алфавит разбивают на группы, объединив в каждой из них буквы со сходными признаками, чтобы легче проследить закономерности, возникающие в шрифте, согласно традиции в теории его построения. Существуют разные подходы к делению алфавита на группы, при этом группы могут быть установлены на основании анализа графических признаков символа, по геометрическим параметрам знаков шрифта и по субъективному их сходству. Однако в любую из полученных группировок всегда допустимо вносить изменения на основании анализа конкретного рисунка шрифта.

По форме и конструкции прописные буквы русского алфавита можно условно разделить на четыре основные группы [1]:

1. Прямые знаки — конструктивно состоящие из вертикальных и горизонтальных штрихов, пересекающихся под прямым углом (или близким к прямому): Г, Е, Н, П, Т, Ц, Ч, Ш, Щ.

2. Круглые или полукруглые знаки — полностью или частично состоящие из штрихов, которые образуют крупные (в рост всего знака) округлые элементы, чаще всего овальной или полуовальной формы: С, О, Э, Ю.

3. Знаки с диагональными штрихами — состоящие из диагональных штрихов или имеющие их в своей конструкции: А, Д, Ж, И, К, Л, М, У, Х. Знаки Д, Ж, К, Л отнесены в эту группу условно, так как в зависимости от индивидуального решения они могут вовсе не иметь диагональных штрихов. М, И могут также быть интерпретированы как прямые знаки.

4. Знаки с округлыми элементами — имеющие округлые элементы, по размеру не достигающие роста всего знака: Б, В, З, Р, Ф, Ъ, Ы, Ь, Я.

Аналогично могут быть разделены на группы и строчные буквы алфавита [2]. Существует следующее деление строчных букв русского алфавита на группы: 1) н, п, ц, и, ш, щ, ч; 2) о, ю, б, е, с, э, а; 3) т, г; 4) в, ь, ы, ь, з, я; 5) у, м, х; 6) р, ф; 7) к, ж, л, д.

Рисунок шрифта может вносить свои изменения в подобную группировку, например, знаки «к» и «н» или знаки «л» и «п» могут оказаться схожими, мало различимыми для глаза при нахождении их в сплошном тексте. Однако разбиение по геометрическим характеристикам (пропорциональность, контрастность шрифта и др.) не всегда совпадает с классификациями знаков по графическим признакам. Сложность сравнения заключается и в том, что в деление знаков алфавита по их графическим признакам является достаточно общим и не охватывает полностью особенности рисунков букв. Разделение букв на группы по геометрическим признакам проводится во взаимосвязи с особенностями конкретного шрифта, в данной статье полученные группы приведены для шрифта Times New Roman: 1) а, д, ц, и, н, ч, к, п, х; 2) в, л, ь, я, ь; 3) г, т, м; 4) ж, ы, щ, ш, ю; 5) з, э; 6) б, е; 7) о; 8) р, у, ф; 9) с. При этом группы 1–5 настолько близко расположены, что есть основания считать их одной общей группой.

Разбиение знаков на группы по геометрическим параметрам было осуществлено при помощи кластерного анализа значений и показало, что оно зависит не только от контрастности, пропорциональности шрифта и прочих характеристик, но также и от конкретного начертания знака шрифта. При этом разбиение по геометрическим характеристикам не всегда совпадает с классификациями знаков

по графическим признакам (как для шрифта Times New Roman, так и для других шрифтов).

Хотя графическая особенность знака при данном подходе к группировке знаков и не учитывается, однако в ряде случаев объединение по группам осуществляется с учетом не только геометрических параметров шрифта. Можно наблюдать тенденцию к объединению букв согласно особенностям конкретного шрифта, например, для шрифта Times New Roman к букве «б» наиболее близка буква «е», в то время как для шрифта Arial «б» объединяется с «р» и «ф», для шрифта Courier New — «б» с «ф».

На основе анализа субъективных суждений о схожести букв русского алфавита [3] была получена следующая картина относительно распределения букв в группы: 1) Ъ, Б, Р, Ы, Ь, В, Ф, О, С, З, Я, Ю; 2) Е; 3) Г, Ч, Т, Д, Н, К, Ж; 4) Ш, Щ, Ц, П; 5) А, Л, М, У, Х; 6) И, Й.

Группировка осуществлялась при наборе букв рубленным шрифтом в прописном начертании. Согласно полученным расчетам, наиболее похожими с субъективной позиции следует считать сочетания АЛ, НЧ, ГТ, ЪВ, БР, ОС.

Сравнение полученных группировок не может быть однозначным, так как в основном графические особенности знаков определяются согласно построению прописных букв, тогда как процент встречаемости прописных знаков в текстах изданий гораздо ниже встречаемости строчных знаков.

На основании вышеназванных подходов и полученных в результате их применения данных, а также с учетом объема тестового материала, используемого для определения удобочитаемости по методике скорости чтения, были отобраны наиболее характерные в данном случае буквенные знаки: «и» и «н», «н» и «п», «г» и «т» и некоторые другие. Всего для эксперимента выбрано 12 пар знаков (24 пары с учетом перемены букв местами). Сочетания знаков в соответствующем шрифтовом оформлении представлено на рис. 1–5.

Пара букв «к» и «н» была выбрана для проведения эксперимента, поскольку при двух подходах к группировкам букв эти знаки оказывались рядом, в одной группе. Что касается графических признаков, то при наборе текста мелким кеглем в шрифтах некоторых рисунков графическая особенность «к» (наличие наклонной линии) оказывается менее значимой, чем визуальная схожесть или геометрические параметры, которые оказывают влияние на различимость буквы. Аналогичный подход обусловил и выбор в качестве условных меток пары букв «л» и «п».

Пары букв «и»/«н» и «н»/«п» входят в одну группу при оценке их графических признаков и геометрических параметров, в то время как по субъективному сходству эти буквы оконча-

тельно не были выделены в одну группу, хотя больше 50 % испытуемых и назвали их в качестве схожих букв.

Пара букв «г» и «т» при всех трех подходах находится в одной группе. Аналогичные данные получены и для пары букв «ш»/«щ» и «ч»/«н».

Основываясь на субъективной оценке, пара букв «ю» и «о» была выделена для использования в качестве условных меток, поскольку 67 % испытуемых назвали эти знаки похожими.

Пара «о» и «с» по графическим признакам и по субъективной оценке находится в одной группе, причем по субъективной схожести она имеет весьма высокий процент (94 %), что и обусловило значимость этой пары для анализа удобочитаемости шрифтов. Аналогичные данные получены и для пары «в» и «з» (70 %).

Пары букв «о»/«р» и «б»/«о» были отобраны с целью оценки влияния на распознаваемость выносного элемента при общей графической схожести знака.

КН ИН НП ГТ ЮО

СО ОР БО ЩШ ЧН ВЗ ЛП

Рис. 1. Сочетания буквенных знаков Times New Roman

КН ИН НП ГТ ЮО СО

ОР БО ЩШ ЧН ВЗ ЛП

Рис. 2. Сочетания буквенных знаков ZapfElliptical711CTT BT

КН ИН НП ГТ ЮО СО

ОР БО ЩШ ЧН ВЗ ЛП

Рис. 3. Сочетания буквенных знаков AdonisCTT

КН ИН НП ГТ ЮО СО

ОР БО ЩШ ЧН ВЗ ЛП

Рис. 4. Сочетания буквенных знаков BookmanCTT

КН ИН НП ГТ ЮО СО

ОР БО ЩШ ЧН ВЗ ЛП

Рис. 5. Сочетания буквенных знаков NewBaskervilleCTT

Таблица 2

Ненайденные метки в тестовом материале для NewBaskervilleCTT / Times New Roman

| Знак исходный / измененный | Сумма найденных меток | | Группа различимости |
|----------------------------|-----------------------|------------|---------------------|
| | исходный | измененный | |
| к/н | 3/6 | 5/8 | I/II |
| н/к | 2/2 | | |
| и/н | 4/3 | 7/13 | II/III |
| н/и | 3/10 | | |
| н/п | 6/4 | 9/12 | II/III |
| п/н | 3/8 | | |
| г/т | 8/2 | 10/4 | II/I |
| т/г | 2/2 | | |
| ю/о | 14/9 | 16/15 | III/III |
| о/ю | 2/6 | | |
| с/о | 4/5 | 10/9 | II/II |
| о/с | 6/4 | | |
| о/р | 1/3 | 1/3 | I/I |
| р/о | 0/0 | | |
| б/о | 1/0 | 2/2 | I/I |
| о/б | 1/2 | | |
| ш/ш | 14/3 | 28/9 | III/II |
| ш/щ | 14/3 | | |
| ч/н | 8/4 | 14/11 | III/II |
| н/ч | 6/7 | | |
| в/з | 7/5 | 12/12 | III/III |
| з/в | 5/7 | | |
| л/п | 4/3 | 6/11 | II/II |
| п/л | 2/8 | | |

Согласно полученным данным, в шрифте NewBaskervilleCTT к числу плохо распознаваемых можно отнести следующие сочетания знаков «ю»/«о», «щ»/«ш», «ч»/«н», «в»/«з». Для шрифта Times New Roman: «и»/«н», «н»/«п», «ю»/«о», «з»/«в». Не вызывает затруднений при распознавании рисунок букв «о»/«р», «б»/«о» для обоих шрифтов, видимо, благодаря выносным элементам. Для NewBaskervilleCTT легкой для распознавания является пара «к»/«н», в то время как для Times New Roman эти буквы вызвали большую трудность при идентификации. Однако в последнем шрифте легче распознавались буквы «г» и «т».

Суммарное количество найденных меток в тесте для шрифтов Zapf Elliptical 711CTT BT и AdonisCTT представлено в табл. 3.

Для шрифта ZapfElliptical711CTT BT к числу плохо распознаваемых рисунков буквенных знаков можно отнести «ю»/«о» и «ш»/«щ». Однако такие знаки, как «г»/«т» и «ч»/«н», имеют малое суммарное значение найденных меток (или равное нулю) и относятся к первой группе различимости, хотя при трех подходах к группировке букв они оказывались в одной группе, т. е. при чтении именно они с наибольшей вероятностью могут быть неверно идентифицированы.

Согласно методике оценки удобочитаемости при нахождении испытуемым всех меток, допущенных в тексте, за выходное время чтения принимается время, зафиксированное при прочтении текста. Если количество найденных меток меньше максимального, тогда к полученному для данного испытуемого времени прибавляется время, приходящееся на каждую найденную метку. Подобный метод позволяет получить достаточно однородные экспериментальные данные и охарактеризовать различимость шрифтов. Полученные экспериментальные данные для пяти шрифтов представлены в табл. 1.

Доверительный интервал для шрифта Times New Roman составляет примерно 0,36, т. е. все шрифты, время чтения которых лежит за пределами $3,56 \pm 0,36$, имеют статистически значимую разницу. Поэтому шрифты ZapfElliptical 711CTT BT и AdonisCTT относятся к той же группе удобочитаемости, что и шрифт Times New Roman, в то время как шрифты BookmanCTT и NewBaskervilleCTT составляют следующую группу, и можно говорить о том, что их читаемость хуже, чем у шрифта Times New Roman. По данной методике удобочитаемость шрифтов ZapfElliptical 711CTT BT, AdonisCTT и Times New Roman следует признать примерно равной.

Поскольку исследование удобочитаемости связано с оценкой различимости отдельных знаков, то из анализа найденных меток, т. е. опечаток в тексте, можно сделать выводы относительно различимости шрифта, особенностей рисунка схожих букв в различных шрифтах. Суммарное количество найденных меток в тестовом материале для шрифтов NewBaskervilleCTT и Times New Roman представлено в табл. 2.

Под исходным знаком подразумевается буква, которая должна была присутствовать в тексте, исходя из смысла и правил написания слов, а измененный знак является условной меткой, которую и должен был найти испытуемый при чтении текста. Сумма найденных меток, т. е. нераспознанных знаков, указана как для каждой пары отдельно, так и для пар с учетом перемены местами. Группа различимости является условной категорией, позволяющей отнести букву в конкретном шрифтовом оформлении к числу способствующих или не способствующих удобочитаемости рисунка шрифта.

Таблица 1

Удобочитаемость шрифта по времени чтения

| Начертание шрифта | Время чтения, мин |
|--------------------------|-------------------|
| Times New Roman | 3,56 |
| ZapfElliptical 711CTT BT | 4,19 |
| AdonisCTT | 4,29 |
| BookmanCTT | 5,35 |
| NewBaskervilleCTT | 5,37 |

Таблица 3
Ненайденные метки в тестовом материале для ZapfElliptical 711 CTT BT / AdonisCTT

| Знак исходный / измененный | Сумма найденных меток | | Группа различимости |
|----------------------------|-----------------------|-------|---------------------|
| | 4/3 | 7/4 | |
| к/н | 4/3 | 7/4 | II/I |
| н/к | 3/1 | | |
| и/н | 7/6 | 10/14 | II/III |
| н/и | 3/8 | | |
| н/п | 0/1 | 3/2 | I/I |
| п/н | 3/1 | | |
| г/т | 0/12 | 4/20 | I/III |
| т/г | 4/8 | | |
| ю/о | 7/6 | 13/8 | III/II |
| о/ю | 6/2 | | |
| с/о | 3/1 | 9/3 | II/I |
| о/с | 6/2 | | |
| о/р | 2/3 | 3/7 | I/II |
| р/о | 1/4 | | |
| б/о | 1/0 | 2/0 | I/I |
| о/б | 1/0 | | |
| щ/ш | 4/5 | 16/8 | III/II |
| ш/щ | 12/3 | | |
| ч/н | 0/6 | 0/7 | I/II |
| н/ч | 0/1 | | |
| в/з | 1/0 | 3/2 | I/I |
| з/в | 2/2 | | |
| л/п | 5/3 | 8/16 | II/III |
| п/л | 3/13 | | |

При наборе текста мелким кеглем, например, в 10 пунктов, свисающий элемент буквы «щ» в шрифте может плохо распознаваться, что и объясняет наличие пары «ш»/«щ» в группе различимости.

Буквы «ю»/«о» часто смешиваются, поскольку «о» является составным элементом «ю», и наличие последней в словах рядом с буквами с ярко выраженной вертикалью создает эффект частотола, который не способствует различимости и удобочитаемости шрифта. В случае же, когда в качестве измененного знака в тексте присутствовала буква «о», то ее нераспознавание также можно объяснить наличием рядом с «о» прямых знаков, например «н» или «п», благодаря которым буква кажется читающему похожей на «ю». Данная пара является плохо распознаваемой для большей части из рассмотренных нами шрифтов.

Для шрифта AdonisCTT в третью группу различимости попали иные пары: «и»/«н», «г»/«т», «л»/«п». Последнее сочетание для других шрифтов находится в группе со средней различимостью, а пара «и»/«н» относится к третьей группе для шрифта Times New Roman.

Данные для шрифта BookmanCTT (табл. 4) показывают, что наименее различимы пары букв с ярко выраженной вертикальной линией: «и»/«н», «н»/«п», «г»/«т».

Таблица 4
Ненайденные метки в тестовом материале для шрифта BookmanCTT

| Знак исходный / измененный | Сумма найденных меток | | Группа различимости |
|----------------------------|-----------------------|----|---------------------|
| | 2 | 5 | |
| к/н | 2 | 5 | I |
| н/к | 3 | | |
| и/н | 5 | 13 | III |
| н/и | 8 | | |
| н/п | 2 | 16 | III |
| п/н | 14 | | |
| г/т | 8 | 19 | III |
| т/г | 11 | | |
| ю/о | 0 | 8 | II |
| о/ю | 8 | | |
| с/о | 1 | 6 | II |
| о/с | 5 | | |
| о/р | 1 | 2 | I |
| р/о | 1 | | |
| б/о | 1 | 2 | I |
| о/б | 1 | | |
| ш/ш | 7 | 10 | II |
| ш/щ | 3 | | |
| ч/н | 10 | 11 | II |
| н/ч | 1 | | |
| в/з | 3 | 3 | I |
| з/в | 0 | | |
| л/п | 0 | 2 | I |
| п/л | 2 | | |

Таким образом, кроме времени прочтения текста, оформленного тем или иным шрифтом, на удобочитаемость влияет различимость буквенных знаков шрифта, а распознаваемость и схожесть знаков, кроме исторически сложившегося их начертания, во многом зависит от конкретного рисунка шрифта.

Применение условных меток позволит учесть при оценке удобочитаемости шрифта как различимость знака, так и время чтения текста, набранного этим шрифтом. Работа выполняется при поддержке БРФФИ.

Литература

1. Дубина Н. Особенности построения шрифтов // Компьюарт. — 2001. — № 6. — С. 83–89.
2. Большаков М. В., Гречиго Г. В., Шицгал А. Г. Книжный шрифт. — М.: Книга, 1964. — 312 с.
3. Фрумкина Р. М. Прогноз в речевой деятельности. — М.: Наука, 1974. — 240 с.