

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/20022> holds various files of this Leiden University dissertation.

Author: Chahine, Nadine

Title: Reading Arabic : legibility studies for the Arabic script

Date: 2012-10-25

The Experiment

Imagine you are a doctor, and a patient comes to you with an ailment. Now, it is not an illness that is terminal, but it is serious enough that it could affect this patient's quality of life. Imagine that you had several types of medicine that you could give, and your basis for judgment is your gut instinct as a doctor as well as the traditional way of prescription, which is tried but not tested. You could follow these traditions, but in your heart, would you not want to know that you are giving the very best solution? Would you not want a more scientific answer?

Design is no medicine, and we are not in the business of treating people. We do prescribe solutions though to a multitude of visual communication problems. We are in the business of presenting information to the general public. Let us go back to the world of make belief and try to imagine that you are a designer and a client wants you to design a schoolbook in Arabic. Which typeface would you choose?

Chapter 3 painted the landscape of available Arabic typefaces and highlighted the conventional usages of the different styles. These conventions are simply just that. Designers often disagree on the type style best suited for continuous reading. Technical limitations have morphed the Naskh style into a system of either two or four forms per letter. This solution is far from the complex dynamic forms of manuscript Naskh. At a time when these technical limitations no longer exist, which style of text face should we recommend? This is at the heart of this research. The Arabic typographic repertoire includes simplified styles that have garnered acceptance and have become part and parcel of everyday communication. Technical breakthroughs allow us to return to a more complex system. So how does one move forward?

Chapter 6 further illuminated the canvas. Reading Arabic is an activity that is seriously understudied. The few studies that have been done have shown that the process of reading Arabic is different than reading languages such as English or French. Different reading metrics, different parts of brain activation, and a host of unanswered questions lead to this research: What is the effect of the complexity of word formation in the Arabic Naskh style on reading speed? How do the various interpretations of Naskh affect reading speed and eye movement patterns? How do character clusters influence the eye's optimal viewing position? What role do the vocalization marks play in reading?

As discussed in chapter 5, a special typeface family was designed along the Naskh styles that are seen in the Arabic typographic repertoire today. Afandem has three levels of increasing complexity: Simplified with two forms per character, Traditional with four forms per character alongside a discreet amount of ligatures, and Dynamic with a multitude of different forms per character that attach in a more fluid and organic fashion.

A multitude of questions tempt this researcher. How do these styles affect reading speed? Knowing that the norm is to drop them out, what happens if we include vocalization marks? The ways to go about finding the answers were varied, but the choice

settled on the use of eye tracking for the simple reason that it is the most intimate of methods. It is able to give highly accurate reading measures, as well as being able to further the investigation of the eyes' movements while reading. This method opens insights into the reading of Arabic, and offers real-time viewing of readers' interaction with the text. Seventy-two subjects read sample Arabic texts on screen with an embedded camera that tracked their eye movement. The samples showed 3 different interpretations of the Naskh style with different levels of complexity, once with and once without vocalization. The eyes' movements were recorded every 8 milliseconds, thus giving an intimate view of the mechanics of reading Arabic.

To date, this is the first comprehensive study to investigate Arabic typeface legibility through the use of eye tracking equipment. It is also one that abandoned the confines of a lab environment and engaged with the readers in their daily context. The experiment design collected statistics that address several reading measures such as total reading time, fixation¹ duration, number of fixations, number of regressions², and the respective lengths of forward and regressive saccades³. These measures are all indicators of reading speed, and as such, the method used and the results collected are directly positioned to answer the main hypothesis regarding the effect of the complexity of the word shape on reading speed. The research results add clear and statistically reliable findings to the fields of Arabic linguistics and Arabic typeface design. Questions such as which typeface to use for faster reading, or when and where to use vocalized texts can now be answered.

Hypothesis

The complexity of the word form (as can be seen in the different interpretations of Naskh) decreases reading speed. This effect is expected to be more profound in younger readers. The presence of vocalization marks removes ambiguity and therefore decreases the number of regressions.

Method

Research method

This experiment dealt with issues relating to both design and linguistics. On the one hand, it exposed subjects to different typefaces that were varying in complexity (from Simplified to Traditional to Dynamic as discussed in Chapter 4). The aim here was to measure reading speed, but even that is open to interpretation: What is a good measure of reading speed? Is it the total amount of time needed to read a given passage? Is it the duration of a single fixation? What about other reading

¹ Fixation (visual, def.): "The orienting of the eyeball so that the projection of the viewed object falls on the fovea and is in focus. The object or location in space is called the fixation point and lies along the fixation line, which can be drawn from the fovea through the pupil to the object" (Reber, Allen, & Reber, 2009).

² Regression (def.): "In reading, any eye movement back over material already read. The frequency of such regressions is related to the difficulty of the material and the reading skills of the individual" (Reber et al., 2009).

³ Saccade (def.): "A quick eye movement, a jump of the eyes from one fixation point to another. Saccadic movements are seen most clearly during reading and the scanning of visual displays" (Reber et al., 2009).

measures, such as the total number of fixations needed to read a given passage, and the number of regressions made during that process?

One method that was certainly tempting to use was the naming test. Basically, one measures the amount of time needed for a participant to read a word or non-word. This was a very attractive option since it does away with the cognitive time needed for comprehension. If one were to go with the definition of legibility as the ease with which words are encoded, then it would have been perfect.

However, this experiment had a linguistic layer as well. Other than the issue of typeface legibility, it dealt with the role of vocalization and its effect on reading speed. Given that Arabic texts are usually un-vocalized, and that the presence of vowels would help to disambiguate a word's meaning, the pure measure of word naming was not enough. The presence of vowels was hypothesized to reduce the amount of regressions, and to measure that one had to look at the total process of reading passages of real texts rather than series of non-words. This does not negate that word naming would be a good method to use, but simply that it was not sufficient to answer all questions at this point.

What sort of information can one glean with eye tracking? To put it simply, there is the stimulus presented as well as a record of the position of both eyes every 8 milliseconds. Analysis of this data yields the duration of every fixation as well as its position with regards to the stimulus. Further analysis leads to the calculation of the various reading measures that are of interest to the research question. The data presented in the following pages will be the aggregate of the individual reading measures of the 72 participants tested.

Overall, the use of eye tracking as a research method is very well suited to answer the hypothesis, as well as provide a clearer picture of the mechanics of reading Arabic. The data collected will be adequate for answering all of the questions asked in this dissertation⁴.

Materials

The experiment was designed to test 3 different typefaces (Simplified, Traditional, and Dynamic Afandem) in 2 conditions (with and without vocalization). This meant that there needed to be 6 different texts to be read by every participant. A trial paragraph was added at the beginning of the experiment in order for the participants to learn the system of how the experiment was going to proceed and, more importantly, to set them at ease. That brought the total number of texts to 7, and these were all selected from the classical fable *Kalilah wa dimnah*. Each text was made up of 1 paragraph that told a short story about humans or animals and led to a moral lesson to be understood by the end of the paragraph (Fig. 6.1–6.4).

These 7 paragraphs (Appendix) were presented to the subjects in the same order every time. The first paragraph was set in a partially vocalized Ruq'aa typeface so as not to give a familiarity advantage to one of the styles being tested. The 6 conditions to be tested were rotated across the different texts in order to counterbalance⁵ the

⁴ There was also one more reason why eye tracking was selected as the method, and that was the effect of character clusters or ligatures on fixation positions. The Dynamic version of Afandem, like the manuscript Naskh, has words in which several characters cluster together to form quite complex ligatures. This was expected to affect both the fixation position, as well as the fixation duration. This kind of information is best looked at through eye tracking. This investigation will be developed in further research.

⁵ Counterbalancing (def.): "An experimental procedure for controlling irrelevant factors. A good example

مقتطفات من كتاب كليله ودمنة لابن المقفع (بالتصريف)

مثل الغراب والأسود

قال دمنة: زعموا أنّ غراباً كان له وكرّاً في شجرة على جبل، وكان قريباً منه جحر ثعبانٍ أسود. فلما كان الغراب إذا أفرغ عمد الثعبان الى فراخه فأكلها فاشتد هذا الأمر على الغراب وهزن. فشكا ذلك إلى صديقه له من بنات أوى، قال له: أريدُ مُشاورتك في أمرٍ قد عزمت عليه. قال: وما هو؟ قال الغراب: قد عزمْتُ أن أذهبَ إلى الثعبان الأسود إذا نام فأنقر عينيه فأفقاها لعلّي أستريح منه. قال له ابن أوى: بس الحيلة التي املت! فالتمسَ أمراً نصيب فيه بُغيتك من الثعبان من غير أن تعرّض نفسك للهلاك. وإياك أن تكون مثل الطائر الأبيض الذي أراد قتلَ السرطان فقتل نفسه.

Fig 6.1 Trial test showing partially vocalized text in Ruqaa style.

مَثَلُ الْحَمَّالِينَ وَالرَّجُلِ الَّذِي أَصَابَ كَنْزًا

وَمَنْ اسْتَكْتَرَ مِنْ جَمْعِ الْكُتُبِ وَقِرَاءَةِ الْعُلُومِ مِنْ غَيْرِ أَعْمَالِ الرَّوِيَّةِ فِيمَا يَقْرُؤُهُ، كَانَ خَلِيقًا أَنْ لَا يُصِيبُهُ، إِلَّا مَا أَصَابَ الرَّجُلَ الَّذِي زَعَمَتِ الْعُلَمَاءُ أَنَّهُ اجْتَازَ الْمَفَاوِزَ؛ فَظَهَرَ لَهُ مَوْضِعُ آثَارِ كَنْزٍ. فَجَعَلَ يَحْفِرُ وَيَطْلُبُ فَوْقَ عَلَى شَيْءٍ مِنْ عَيْنٍ وَوَرَقٍ، فَقَالَ فِي نَفْسِهِ: إِنْ أَنَا أَخَذْتُ فِي نَقْلِ هَذَا الْمَالِ قَلِيلًا قَلِيلًا، طَالَ عَلَيَّ، وَقَطَعَنِي الْأَسْتِغَالُ بِنَقْلِهِ وَإِحْرَازِهِ عَنِ اللَّذَّةِ بِمَا أَصَبْتُ مِنْهُ. وَلَكِنْ سَأَسْتَأْجِرُ أَقْوَامًا، يَحْمِلُونَهُ إِلَى مَنْزِلِي، وَأَكُونُ أَنَا آخِرَهُمْ. وَلَا يَكُونُ بَقِي وَرَائِي شَيْءٌ، يَشْغَلُ فِكْرِي بِنَقْلِهِ. وَأَكُونُ قَدْ اسْتَعْنْتُ لِنَفْسِي فِي إِرَاحَةِ بَدَنِي عَنِ الْكَدِّ بِبَيْسِيرِ أُجْرَةٍ، أُعْطِيهَا لَهُمْ. ثُمَّ جَاءَ كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ؛ فَجَعَلَ يُحْمَلُهُمْ مَا يُطِيقُ؛ فَيَنْطَلِقُ بِهِ إِلَى مَنْزِلِهِ هُوَ، فَيَفُوزُ بِهِ، حَتَّى إِذَا لَمْ يَبْقَى مِنَ الْكَنْزِ شَيْءٌ؛ انْطَلَقَ خَلْفَهُمْ إِلَى مَنْزِلِهِ، فَلَمْ يَجِدْ فِيهِ مِنْ الْمَالِ شَيْئًا. وَإِذَا كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ قَدْ فَازَ بِمَا حَمَلَهُ لِنَفْسِهِ. وَلَمْ يَكُنْ لِلرَّجُلِ مِنْ ذَلِكَ إِلَّا الْعَنَاءُ وَالتَّعَبُ، لِأَنَّهُ لَمْ يُفَكِّرْ فِي آخِرِ أَمْرِهِ.

Fig 6.2 Sample of text in the vocalized version of the Simplified style.

مَثَلُ الْحَمَّالِينَ وَالرَّجُلِ الَّذِي أَصَابَ كَنْزًا
 وَمَنْ اسْتَكْتَرَ مِنْ جَمْعِ الْكُتُبِ وَقِرَاءَةِ الْعُلُومِ مِنْ غَيْرِ إِعْمَالِ الرَّوِيَّةِ فِيمَا يَقْرُوهُ، كَانَ خَلِيقًا
 أَنْ لَا يُصِيبَهُ، إِلَّا مَا أَصَابَ الرَّجُلَ الَّذِي زَعَمَتِ الْعُلَمَاءُ أَنَّهُ اجْتَازَ الْمَفَاوِزَ؛ فَظَهَرَ لَهُ مَوْضِعُ
 آثَارِ كَنْزٍ. فَجَعَلَ يَجْفِرُ وَيَطْلُبُ فَوْقَ عَلَى شَيْءٍ مِنْ عَيْنٍ وَوَرَقٍ، فَقَالَ فِي نَفْسِهِ: إِنْ أَنَا أَخَذْتُ
 فِي نَقْلِ هَذَا الْمَالِ قَلِيلًا قَلِيلًا، طَالَ عَلَيَّ، وَقَطَعَنِي الْإِشْتِغَالُ بِنَقْلِهِ وَإِحْرَازِهِ عَنِ اللَّذَّةِ بِمَا
 أَصَبْتُ مِنْهُ. وَلَكِنْ سَأَسْتَأْجِرُ أَقْوَامًا، يَحْمِلُونَهُ إِلَى مَنْزِلِي، وَأَكُونُ أَنَا آخِرَهُمْ. وَلَا يَكُونُ بَقِي
 وَرَائِي شَيْءٌ، يَشْغَلُ فِكْرِي بِنَقْلِهِ. وَأَكُونُ قَدْ اسْتَعْنَتُ لِنَفْسِي فِي إِرَاحَةِ بَدَنِي عَنِ الْكَدِّ
 بِبَسِيرِ أُجْرَةٍ، أُعْطِيهَا لَهُمْ. ثُمَّ جَاءَ كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ؛ فَجَعَلَ يَحْمِلُهُمْ مَا يُطِيقُ؛ فَيَنْطَلِقُ
 بِهِ إِلَى مَنْزِلِهِ هُوَ، فَيَفُوزُ بِهِ، حَتَّى إِذَا لَمْ يَبْقَى مِنَ الْكَنْزِ شَيْءٌ؛ أَنْطَلِقُ خَلْفَهُمْ إِلَى مَنْزِلِهِ، فَلَمْ
 يَجِدْ فِيهِ مِنَ الْمَالِ شَيْئًا. وَإِذَا كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ قَدْ فَازَ بِمَا حَمَلَهُ لِنَفْسِهِ. وَلَمْ يَكُنْ
 لِلرَّجُلِ مِنْ ذَلِكَ إِلَّا الْعَنَاءُ وَالتَّعَبُ، لِأَنَّهُ لَمْ يُفَكِّرْ فِي آخِرِ أَمْرِهِ.

Fig 6.3 Sample of text in the vocalized version of the Traditional style.

مَثَلُ الْحَمَّالِينَ وَالرَّجُلِ الَّذِي أَصَابَ كَنْزًا
 وَمَنْ اسْتَكْتَرَ مِنْ جَمْعِ الْكُتُبِ وَقِرَاءَةِ الْعُلُومِ مِنْ غَيْرِ إِعْمَالِ الرَّوِيَّةِ فِيمَا يَقْرُوهُ، كَانَ خَلِيقًا
 أَنْ لَا يُصِيبَهُ، إِلَّا مَا أَصَابَ الرَّجُلَ الَّذِي زَعَمَتِ الْعُلَمَاءُ أَنَّهُ اجْتَازَ الْمَفَاوِزَ؛ فَظَهَرَ لَهُ مَوْضِعُ
 آثَارِ كَنْزٍ فَجَعَلَ يَجْفِرُ وَيَطْلُبُ فَوْقَ عَلَى شَيْءٍ مِنْ عَيْنٍ وَوَرَقٍ، فَقَالَ فِي نَفْسِهِ: إِنْ أَنَا أَخَذْتُ
 فِي نَقْلِ هَذَا الْمَالِ قَلِيلًا قَلِيلًا، طَالَ عَلَيَّ، وَقَطَعَنِي الْإِشْتِغَالُ بِنَقْلِهِ وَإِحْرَازِهِ عَنِ اللَّذَّةِ بِمَا
 أَصَبْتُ مِنْهُ. وَلَكِنْ سَأَسْتَأْجِرُ أَقْوَامًا، يَحْمِلُونَهُ إِلَى مَنْزِلِي، وَأَكُونُ أَنَا آخِرَهُمْ. وَلَا يَكُونُ بَقِي
 وَرَائِي شَيْءٌ، يَشْغَلُ فِكْرِي بِنَقْلِهِ. وَأَكُونُ قَدْ اسْتَعْنَتُ لِنَفْسِي فِي إِرَاحَةِ بَدَنِي عَنِ الْكَدِّ
 بِبَسِيرِ أُجْرَةٍ، أُعْطِيهَا لَهُمْ. ثُمَّ جَاءَ كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ؛ فَجَعَلَ يَحْمِلُهُمْ مَا يُطِيقُ؛ فَيَنْطَلِقُ
 بِهِ إِلَى مَنْزِلِهِ هُوَ، فَيَفُوزُ بِهِ، حَتَّى إِذَا لَمْ يَبْقَى مِنَ الْكَنْزِ شَيْءٌ؛ أَنْطَلِقُ خَلْفَهُمْ إِلَى مَنْزِلِهِ، فَلَمْ
 يَجِدْ فِيهِ مِنَ الْمَالِ شَيْئًا. وَإِذَا كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ قَدْ فَازَ بِمَا حَمَلَهُ لِنَفْسِهِ. وَلَمْ يَكُنْ
 لِلرَّجُلِ مِنْ ذَلِكَ إِلَّا الْعَنَاءُ وَالتَّعَبُ، لِأَنَّهُ لَمْ يُفَكِّرْ فِي آخِرِ أَمْرِهِ.

Fig 6.4 Sample of text in the vocalized version of the Dynamic style.

Table 6.1 The order in which the conditioned were counterbalanced.

Style	Order	P1	P2	P3	P4	P5	P6	P7
S T D (vocalized first)	1	R	Sv	Suv	Dv	Duv	Tv	Tuv
S T D (unvocalized first)	2	R	Suv	Sv	Duv	Dv	Tuv	Tv
S D T (vocalized first)	3	R	Sv	Suv	Tv	Tuv	Dv	Duv
S D T (unvocalized first)	4	R	Suv	Sv	Tuv	Tv	Duv	Dv
T S D (vocalized first)	5	R	Tv	Tuv	Sv	Suv	Dv	Dalt
T S D (unvocalized first)	6	R	Tuv	Tv	Suv	Sv	Duv	Daltv
T D S (vocalized first)	7	R	Tv	Tuv	Dv	Duv	Sv	Suv
T D S (unvocalized first)	8	R	Tuv	Tv	Duv	Dv	Suv	Sv
D S T (vocalized first)	9	R	Dv	Duv	Sv	Suv	Tv	Tuv
D S T (unvocalized first)	10	R	Duv	Dv	Suv	Sv	Tuv	Tv
D T S (vocalized first)	11	R	Dv	Duv	Tv	Tuv	Sv	Suv
D T S (unvocalized first)	12	R	Duv	Dv	Tuv	Tv	Suv	Sv

design of the experiment and avoid the introduction of confounding variables⁶. As such, by the time all participants had taken the tests, and with the exception of the first text, every paragraph was read in all of the 6 conditions. This avoided the trap of one paragraph being easier than the others and thus giving advantage to the condition in which it is set. Another reason for this rotation of conditions is that it neutralizes the order effects⁷. The behavior and response of participants can change over time due to fatigue or other factors, and this is why it is essential that the order in which the conditions are tested is varied. A counterbalanced design “reduces the chances of the order of treatment or other factors adversely influencing the results” (Shuttleworth, 2008).

The presence of 6 conditions made the balancing of the texts to be very tricky. In the counterbalancing of condition that a within-subjects study needs, one needs to present each participant with all the texts in all the conditions and the variability is controlled by a Greco-Latin square table⁸. Unfortunately, no Greco-Latin table exists for a six-condition experiment and so the solution was to take the 3 typeface styles as a base and use a three-factor Greco-Latin square table. This led to 6 different possible combinations. Since every typeface would be presented with and without

is fatigue: if a subject is to perform two tasks, x and y, the order of running would be counterbalanced so that fatigue factors were spread equally across both x and y. To wit: one half of the subjects would perform task x first and one half task y first. If the experiment were run so that all subjects performed x first, it would not be possible to tell whether y was more difficult than x or merely that the subjects were getting tired” (Reber et al., 2009).

6 Confound (def.): “In experimental work, to fail to separate two variables with the result that their effects cannot be independently ascertained. If in an experiment on memory and age all the older participants are female and all the young are male, then the sex and age are “confounded” and the memory data cannot be properly interpreted” (Reber et al., 2009).

7 Order effects (def.): “Effects that are attributed to the order in which treatments are presented in an experiment. Order effects can confound an experiment and typically counter-balancing procedures are employed as a control” (Reber et al., 2009).

8 Greco-Latin square or Orthogonal Latin squares (def.): “A balanced two-way classification scheme in which each condition occurs just once in each row and column... This balancing is often incorporated into experimental designs so that the order of administration of treatments is perfectly balanced across subjects. Two such Latin squares are orthogonal if, when combined, the same pair of symbols occurs only once in the combined square. This composite square is called a Greco-Latin square.” (Reber et al., 2009)

vocalization, the total number of possible combinations ended up at 12 (Table 6.1). In reality, the permutations would have been much more, but in order to achieve a practical design for the experiment, the decision was made to keep the order of vocalized to non-vocalized texts constant i.e. all the styles would be presented first with vowels then without, and vice versa. For example, one possible combination would be: Simplified (vocalized, non-vocalized), Traditional (vocalized, non-vocalized), and Dynamic (vocalized, non-vocalized). Another would be: Simplified (non-vocalized, vocalized), Traditional (non-vocalized, vocalized), Dynamic (non-vocalized, vocalized).

A slide with a focus cross was first presented. Each of the six paragraphs was presented on a slide, followed by another slide that had 2 questions related to the text just read (Appendix). This served as a quick check for comprehension to make sure that the participants were reading for comprehension rather than simply going through the mechanics of reading without the effort needed to understand the text.

The texts had a comparable number of characters per line. This was done so as to avoid the effect of line length on reading speed. This effect was seen in reading English and, in order to avoid confounding factors, the number of characters per line had to be maintained roughly constant in the experiment set up.

To put it all together, one sample combination of the 6 conditions looked like this:

1. Ruqaa text with partial vocalization (trial test)
2. Simplified Naskh without vocalization
- 2a. Comprehension check
3. Simplified Naskh with vocalization
- 3a. Comprehension check
4. Traditional Naskh without vocalization
- 4a. Comprehension check
5. Traditional Naskh with vocalization
- 5a. Comprehension check
6. Calligraphic Naskh without vocalization
- 6a. Comprehension check
7. Calligraphic Naskh with vocalization
- 7a. Comprehension check

The texts selected for the experiment were all from the same book as mentioned earlier. This was done on purpose so that they would have the same tone of voice, style of language, and level of vocabulary. The book from which they were taken, *Kalilah wa dimnah*, is a classic fable that used to be taught in schools up until a few years ago. These texts are no longer in the curriculum today and the subjects tested were unfamiliar with their content. This book was recommended by 2 linguists who were consulted in the selection of the material, and this choice was later approved by the director of Arabic language at the school network in which most of the testing was done. He looked over all the material, checked to make sure that all grammar and vocalization was correct, and approved the difficulty level as appropriate. The sample text was of average difficulty level for the 16-year-old group.

The test material was followed by a comprehension check of two questions. These were open-ended questions that had to be answered verbally. The questions were straightforward ones relating directly to incidents that were described in the text. They assured that the subject had actually understood what was going on in the story narrated, but they did not attempt to test that the moral of the story was understood as that was of a higher cognitive level than the comprehension check needed. In other words, it was possible that the subject read the text, understood the story, but failed to grasp the wisdom behind it. The questions avoided that scenario.

Moreover, the phrasing of the questions was done in a way where there was only one possible correct answer.

Since the typeface style was the key variable in question, serious efforts were made to ensure that the design of the three typefaces corresponded perfectly well with the hypothesis to be tested in the experiment. The different styles were checked with two type designers with extensive knowledge of the Arabic script and Arabic type design for confirmation regarding experiment requirements. The success of the experiment and the viability of the results rest on the validity of the experiment set up, as well as the appropriateness of the material used.

As explained in Chapter 4, the three versions of Afandem faithfully represent the three genres that are in question. Variables such as weight, contrast, optical size, axis, terminal treatment, and proportion were all controlled so as to stay constant throughout the different conditions. However, the change in the number of forms per character and the stacking order in which characters come together to form words resulted in a few unwanted results. One was the increased density of the words and reduced line width in the more complex styles. Another was the increased fluidity and dynamic energy, especially in the Dynamic version. Still, these effects were a natural byproduct of the increased complexity, and any attempt to keep them constant would have thrown into jeopardy the ability of each typeface to represent its respected genre. As such, these qualities had to be accepted as unavoidable. The only measure that could be taken were the use of extended swashes in the Dynamic that helped to increase the line width and bring it closer to what was seen in the Simplified and Traditional samples.

Another approach to this dilemma would have been to keep the same column width across all three typefaces by having a different word count per line. However, in terms of the confounding factors of different width in pixels vs. different word count per line, the first was seen to be the lesser of two evils. The experiment was complex to the extent that testing under both these scenarios would have made it impossibly long to test and analyze within the time frame of this research.

The paragraphs were presented in the same position on screen in relation to the starting point on the top right corner. The variability in line length across the different texts and typeface styles was reflected in a varying amount of white space in the left margin. The point size, tracking, word spacing, and leading remained constant across all samples. The presence of vocalization marks in no way changed the line breaks or any of the typeface metrics. The non-vocalized texts kept two forms of vocalization that are grammatically required even in non-vocalized texts: the Fathatein and the Shadda. Vocalized texts included all possible marks including the Sukoun, except in positions where the mark was irrelevant such as before a long vowel. This is the standard method of vocalizing Arabic texts and the director of Arabic language at the school network approved the samples presented.

The text was set in black on white and with ample line spacing and word spacing so as to avoid crowding of the visual source. This was to have enough space for the vowels to be included without any clashes. The text samples were created in Adobe Illustrator with vector outlines using un-hinted grey scale rendering technique.

Apparatus

The eye tracker used is the Tobii T120 model available during the year 2010. It is integrated in a stand-alone 17-inch computer monitor with a screen resolution of 1280×1024 pixels. The monitor connects to PC via several cables, and the recording is controlled via Tobii Studio software. The eye tracker records the movement of both

pupils, and the freedom of head movement is within a volume of 30x22x30 cm. The subject is expected to sit at a distance of 50-70 cm away from the screen though the best results can be found at around 60 cm away. The data rate of recording is 120Hz and accuracy is at 0.5 degrees. Drift is less than 0.3 degrees (Tobii, 2009).

The main advantage of the T120 Eye Tracker is that it is portable and allows head movement while still being accurate. As such, it removes the necessity of bringing subjects into a lab, as well as the discomfort of using chin-rests. The set-up of the tracker is similar to a natural computer environment, and there are great benefits from non-intrusive procedures that allow users to behave as naturally as possible.

Tobii Studio is the software used for the data collection as well as the calibration process. Other than gaze data, it also records, in real time: stimuli presented on screen, user camera and audio, screen content, timestamps, logged events and a few other indicators of user behavior. It allows for adding variables such as gender and age, and offers replay and visualization tools.

Participants

The number of participants in each group was calculated to be 36 per age group, bringing the total number of subjects to be 72. The 12 different variations of the experiment stimuli had to be read by an equal amount of participants and so the number of participants needed to be a multiple of 12. Psychologists recommend that the number of participants is around 30 so as to increase the likelihood of achieving statistically reliable results⁹ and this brought the number of subjects per group to 36.

The material and texts prepared were tailored towards a sample group of 16–17 year olds. However, the hypothesis expects that the effect of the complexity of the style is more pronounced in younger readers. This is why it was elected to also test with a younger group of ages 13–14. The reading material would be of a higher difficulty level for them, and this most likely affected some of the results as will be discussed later in the chapter.

The 72 participants (36 of each group) were randomly selected from a network of participating schools in Beirut. The younger group had 24 boys and 12 girls, and the older group had 19 boys and 17 girls. All students were native Lebanese, born and brought up in the Lebanon. Their first language is Arabic and their second language is English. The students of these schools come from average income families. The literacy rate of the area, and Lebanon in general, is very high.

Students who wear glasses or contact lenses were not selected. This was done as a precaution against eye tracking errors even though the manufacturer's marketing and product specifications claim that these would not have an effect. Still, given the length of time needed for the calibration and set-up, and the large number of hurdles presented by the eye tracker, it was prudent to go with the safe choice.

This network of schools is religious, and the students are expected to be familiar with vocalized manuscript Naskh (similar to the Dynamic Vocalized condition) as this is the way that the Quran is written. The Arabic reading material in these schools are usually partially vocalized so the students are also used to reading with

⁹ Statistical significance (def.): "The degree to which an obtained result is sufficiently unlikely to have occurred under the assumption that only chance factors were operating, and therefore the degree to which it may be attributed to systematic manipulations. The degree itself is typically specified and denoted as probability, e.g. $p < 0.05$ means that the results obtained (or more extreme results) could only have occurred by chance in fewer than 5 cases in 100. The smaller the p value, the more significant the results; that is, the less likely that they occurred by chance" (Reber et al., 2009).

the vowels present. The reason this is interesting to point out is the absence of Dynamic Naskh from everyday reading material. The usual text condition in everyday reading material in Lebanon is non-vocalized Simplified or Traditional Naskh. Selecting participants from a population that is also reading in Dynamic Naskh was the closest way to overcome the lack of familiarity with this complex style.

As this experiment dealt with underage students, written permission from parents was procured and organized with the schools' administrations. To their credit, the schools were supportive of the research and were willing to provide all the support that was asked of them. This included giving access to unused classrooms in which the testing can take place, providing the necessary furniture for the set-up, and coordinating the logistics of bringing the participants to the test. The school administrators were in charge of the selection of the participants and they were fully informed of the criteria for selection:

- . Random choice from the 3rd Intermediate (age 13–14) and 2nd Baccalaureate (age 16–17) classes.
- . Students are not to be selected based on their reading proficiency. This was respected as one could see from the variation in the reading times.
- . Both boys and girls can participate.
- . The only restriction is if the student wore contact lenses or glasses.

The students did not receive any financial compensation for their participation in the experiment, but they were duly thanked for their participation and effort at the end of the testing. All efforts were taken to make sure that the testing was undisturbed, and that the participants were not stressed by the exercise. The room doors had small windows where school supervisors could look in to check up on the progress of the testing without disturbing the recording and also to assist in case more students were to be sent in for more testing.

Design

The experiment aims to manipulate three distinct variables and so the independent variables¹⁰ to be tested are:

1. **Style:** The interpretation of the Arabic Naskh style in three increasing levels of complexity of word formation. These are Simplified, Traditional, and Dynamic Naskh, modeled to represent the repertoire of Naskh typefaces today.

2. **Vocalization:** This variable is of two levels. Texts are either fully vocalized or unvocalized. Every-day texts are usually unvocalized.

3. **Age:** This variable is a quasi-independent variable of two levels and the experiment is set up with two equal groups of 13- and 16-year-olds.

¹⁰ Independent variable (def.): "Any variable the values of which are, in principle, independent of the changes in the values of other variables. In an experiment, any variable that is specifically manipulated so that its effects upon the dependent variables may be observed. Also called the experimental variable, the controlled variable and the treatment variable" (Reber et al., 2009).

The manipulation of these independent variables is expected to have an effect on reading speed measures. Therefore, the dependent variables¹¹ are: Reading time, average fixation duration, number of fixations, number of regressions, and both forward and regressive saccade lengths. A detailed account of what these measures represent is to be expanded upon later in this chapter.

The experiment then is a counterbalanced 3x2x2 within-subjects design¹² (repeated measure design) for the style and vocalization variables and a between-subjects design¹³ for the Age variable.

A note here needs to be added regarding the external validity of legibility studies. As mentioned in chapter 5, one of the pitfalls that previous studies have fallen into was the testing of the relative legibility of specific typefaces. These typefaces bring with them a host of different aspects of the design, and while it is interesting to see how these typefaces perform, it is often hard to generalize the results as they are very much related to the typefaces at hand. This is why this study is designed around design characteristics (complexity of word formation) rather than around comparative performance of set typefaces.

In legibility studies, therefore, there are two options in how to set up the experiment: Either use existing typefaces, and this ends up measuring the relative legibility of these typefaces to each other; or, design a set of typefaces where design variables are controlled and the results are then related to specific design characteristics. The second option is more suited to answer the research question and has stronger external validity. In other words, such an experiment set-up is able to deliver answers to the relationship of the complexity of word formation and legibility while in the first option one can only deduce if one typeface is more legible than another. It is hoped that this focus on specific variables will support the external validity of the results.

Procedure

The experiment took place in Beirut, Lebanon, in March 2010. This was done for two reasons: the wish to test with native readers who are immersed in their script and reading it on a daily basis, and to avoid testing in a lab environment. Though there is nothing wrong with testing in a lab, the fact that the age group was relatively young, it was deemed easier to reach out to them in the context of their school environment. This gave a natural feel to the whole setup and the students showed no signs of anxiety. To the contrary, they were interested in what the experiment did, and were happy to participate.

The testing was done in unused classrooms on the same floors as classes were given. The set-up required a table onto which the screen and connected laptop PC were placed, and chairs for the researcher and participant. The 2 chairs were on opposite sides of the table and separated from view by the screen. This was done so that the

¹¹ Dependent variable (def.): "Any variable the values of which are, in principle, the result of changes in the values of one or more independent variables. In mathematics this notion of dependence is readily represented by an expression of the kind $y=f(x)$, where the values of y are dependent on the values of x . In psychology, the operative principle is that the behaviour of the subject under consideration is (like y) dependent upon the manipulation of some other factors (the analogue of x)" (Reber et al., 2009).

¹² Within-subjects design (def.): "An experimental design in which each subject is used in all conditions (Reber et al., 2009).

¹³ Between-subjects design (def.): "A research design in which different groups of subjects are run under different conditions" (Reber et al., 2009).

researcher would not inadvertently distract the participants. The laptop was placed behind the screen facing the researcher and served as the control center of the experiment. Tobii Studio offered the possibility to see, in real time, the eye movements across the stimuli, as well as a video recording of the subject throughout the test.

The lighting in the rooms used was relatively similar. The windows and natural light always came in sideways to the reader. Even if the lighting in one room was a bit brighter than another, there was no change in the lighting conditions during any one test, and so the subjects were reading all paragraphs under the same light, sound, and temperature conditions.

Participants were individually tested. All instructions were given in local dialect and were based on written notes so that the information given to each subject is consistent in both content and order of presentation. Participants read from screen the test material. The camera embedded within the screen made the eye movement recordings. The tests were conducted in a relatively quiet room in the school, with ample day light, with only the experimenter present. Tests took place on school days and specifically on the normal school hours between 8 am and 2pm to avoid fatigue.

Participants were first given a practice trial on how to initiate the reading and answer the comprehension questions. Participants answered the questions verbally. The trial paragraph and all questions were set in Ruqaa style and in a semi-vocalized state. This was to avoid giving more favor to any one of the conditions. Both groups read the same material. Participants will be told that this is a study about reading and eye movements.

The experiment scenario was as follows:

- . The subject was shown into the empty classroom and asked to sit at the assigned seat in front of the screen.
- . The subject was asked for his/her first name and age.
- . The researcher introduced herself and briefly explained her background and that the experiment is part of a PhD research into reading Arabic.
- . The researcher explained the purpose (to study the reading of Arabic on screen) and procedure of the experiment test (that each subject would read a few paragraphs and answer questions). The subject was assured that the experiment was not to be graded and the results are confidential.
- . The subject was asked to sit comfortably in the chair and to avoid large movements during the test. Small head movements are allowed.
- . Once the subject gave his/her ok to start the test, calibration started. Subjects looked at the dots presented on screen.
- . After the calibration process was completed, the actual test and recording began with the first trial paragraph. The subject was informed that the first slide was a trial in order to get used to the way the system works.
- . The subject was asked to verbally give the ok to move on to the next slide. They were discouraged from reading the paragraphs all over again as a few subjects were inclined to do.

- . The following step was to silently read the text on screen and to move on to the questions on the following slide. The subject was asked to answer the questions verbally and in the spoken dialect.
- . Once all texts were read and questions answered, the subject was thanked for his/her participation and was given a chance to ask questions (Some of the subjects showed interest in the topic). The subject then returned to class.

Analysis of Eye Tracking Data

Each sample text was presented on a full page but only a selection of the data was analyzed. This amounted to roughly to around 70-80% of the text on the page. Each selection was roughly 100 words, starting from the second or third line in the paragraph. The actual number of words in the 6 test paragraphs was, in that order: 101, 101, 102, 99, 103, and 105. Using a selection of the text was done so as to make sure that the readers are immersed in their reading. A study had shown that reading speed slowed down when the reader did not expect a change in the typeface (Sanocki, 1988; Walker, 2008), and for this reason, the first one or two sentences of each paragraph was not included in the data analysis.

Another reason for the selection was due to the nature of the texts used. The paragraphs started with an ambiguous moral statement that would be later clarified in the text. Though this decision was made prior to the testing, the actual eye tracking data confirmed its correctness as it showed that the opening statement often confused the subjects and many had had to reread it a few times. The portions of the texts selected were usually read more smoothly and with only an occasional necessity for rereading (Fig. 6.5).

Tobii Studio collected raw data that mainly presented the location of both pupils every 8 milliseconds and included an interface to visualize this (Tobii, 2009).

The data was checked manually, fixation by fixation, since the recordings contained ambiguous fixations that needed to be dealt with. For example, 2 fixations that were smaller than 125 ms and were on adjacent characters needed to be combined into one. Fixations that were less than 75 ms and were landing on white empty space were discarded. Fixations that were larger than 1 second were discounted from the calculation for fixation average as they were expected to be more than 1 fixation in reality. These adjustments were made after consulting with leading eye-tracking experts, and are typical of studies like this one¹⁴.

These were the basic principles used for analyzing the data:

- . Reading measures started with the first fixation in the target area.
- . Regressions into the text that came before the target area were not counted, except as 1 regression (statement not clear). This extra reading was not included in the reading time, fixation duration, saccade lengths, saccade lengths or number of fixations. Counting began as soon as the eye re-entered the target area, irrespective of where it landed.

Fig 6.5 (next 6 pages)
Visualization of eye movement in two subjects reading the 6 test paragraphs.

14 Deleted fixations, whether big or small, did not affect the reading time measure.

مثل السمكات الثلاث

زعموا أن غديراً كان فيه ثلاث سمكات: كيسة (حسنة الهائي)، وأكيس طئها، وعاجزة. وكان ذلك الغدير بنحوه من الأرض، لا يكاد يقويه أحد. ويقربه نهر حار. فاتفق أنه، اجتمار بذلك النهر صيادين، فأبصروا الغدير، فتواعدا أن يرجعا إليه بشياكهما، فيصيدا ما فيه من السمك. فسمعت السمكات قولهما: فأما أكيسهن، فارتابت بقولهما، فخرجت من المكان الذي يدخل فيه الماء من النهر إلى الغدير، فنجت بنفسها. وأما الكيسة الأخرى، فإتھا تهانت في الأمر حتى جاء الصيادان. فلما رأتهما، وعرفت ما يريدان، ذهبت لتخرج من حيث يدخل الماء؛ فإذا بهما قد صدرا ذلك المكان فحينئذ قالت: فرطت؛ وهذه عاقبة التفريط فكيف الخيلة على هذه الحال؟ ولما تنجح حيلة العجلة والإهراق. غير أن العليل لا يقطع الأمل. ثم إتھا تملوتت؛ فطفت على وجه الماء منقلبة على ظهرها تارة وتارة على بطنها. فظن الصيادان أنها ميتة، فوضعاها على الأرض بين النهر والغدير؛ فوثبت إلى النهر، فنجت. وأما العاجزة فلم تنزل التي يقبلان وإدبار حتى صيدت.

مثل طالبي العلم والصحيفة الصفراء

وكذلك من قرأ هذا الكتاب، ولم يفهم ما فيه، ولم يعلم غرضه ظهراً وباطناً، لم يتفح بما يبدو له من حظه ونقشه. كما لو أن رجلاً، قدّم له جوز صحيح، لم ينتفع به إلا أن يكسره، ويستخرج ما فيه. وكان أيضاً كالرجل الذي طلب علم الفصح من كلام الناس. فأتى صديقاً له من العلماء له علم بالفصاحة، فأعلمه حاجته إلى علم الفصح. فرسم له صديقه في صحيفة صفراء فصيح الكلام وتصريفه ووجهه. إنصرف بها إلى منزله، فجعل يكبر قراءتها، ولا يقف على معانيها، ولا يعلم تأويل ما فيها، حتى استأثرها كلها، فأعقد أنه قد أحاط بعلم ما فيها. ثم إنّه جلس، ذات يوم، في حفلة من أهل العلم والأدب، فأخذ في محاورتهم، فحرت له كلمة أخطأ فيها. فقال له بعض الجماعة: إنك قد أخطأت؛ والوجه غير ما تكلمت به. فقال: كيف أخطأت؟ وقد قرأت الصحيفة الصفراء وهي في منزلي؟ فكانت مفاثته هذه أوجه للحرية عليه وزافة ذلك قريباً من الجهل وبعيداً من الأدب.

Media: pt.01_Paper_1.png
Time: 00:00:00.000 - 00:01:49.753
Poker: post Word: All

مَثَلُ الرَّجُلِ الْهَارِبِ مِنَ الذَّنْبِ وَاللَّصُوصِ

قِيلَ إِنَّ رَجُلًا سَلَكَ مَفَاوِظَ فِيهَا خُوفٌ مِنَ السَّبَاعِ. وَكَانَ الرَّجُلُ خَيْرًا بَعِيثَ تِلْكَ الْأَرْضِ
وَخَوْفِهَا. فَبِمَا سَمِعَ غَيْرَ بَعِيدٍ، اعْتَرَضَ لَهُ ذَنْبٌ مِنْ أَجْدِ الذَّنَابِ وَأَضْرَاهَا. فَهَلَمَا رَأَى الرَّجُلُ
أَنَّ الذَّنْبَ قَدْ جَرَّدَ نَجْوَاهُ، خَافَ مِنْهُ، وَنَظَرَ حَوْلَهُ فَجَاءَ بِإِقْبَةِ الْإِقْبَةِ خَلْفَ وَجْهِهِ، فَذَهَبَ مُسْرِعًا
نَحْوَهَا. فَلَمَّا أَتَى الْوَادِيَّ، لَمْ يَرِ عَلَيْهِ قَبْلَةَ، وَرَأَى الذَّنْبَ قَدْ أَدْرَكَهُ، فَاتَّقَى نَفْسَهُ فِي الْمَاءِ،
وَهُوَ لَا يَجِدُ السَّبِيحَةَ، وَكَانَ يَفْرُقُ لَوْلَا أَنَّ نَهْرًا بِهِ قَوْمٌ مِنْ أَهْلِ الْقَرْيَةِ، فَتَوَالَّوْا لِإَخْرَاجِهِ،
فَأَخْرَجُوهُ وَقَدْ اشْرَفَ عَلَى الْهَلَاكِ. فَاتَّخَذَ الرَّجُلُ عِنْدَهُمْ، وَأَمِنَ عَلَى نَفْسِهِ مِنْ عَائِلَةِ
الذَّنْبِ، رَأَى عَلَى عِدْوَةِ الْوَادِي بَيْتًا، فَتَوَلَّى: أَدْخُلْ هَذَا الْبَيْتَ، فَاسْتَرَجِحْ بِهِ! فَهَلَمَا دَخَلَهُ وَجَدَ
جَمَاعَةً مِنَ اللَّصُوصِ، قَدْ قَطَعُوا الطَّرِيقَ عَلَى رَجُلٍ مِنَ التَّجَارِ، وَهُمْ يَقْتَسِمُونَ مَالَهُ، وَيَزِيدُونَ
قَتْلَهُ. فَهَلَمَا رَأَى الرَّجُلُ ذَلِكَ، خَافَ عَلَى نَفْسِهِ، وَمَضَى نَحْوَ الْقَرْيَةِ، فَأَسْنَدَ ظَهْرَهُ إِلَى حَائِطِ
مِنْ حِطَانِهَا، لِيَسْتَرِحَ مَا حَلَّ بِهِ مِنَ الْبَوْلِ وَالْإِعْيَاءِ، وَإِذْ سَقَطَ عَلَيْهِ الْحَائِطُ، فَاتَّخَذَ

Media: pt.01_Paper_1.png
Time: 00:00:00.000 - 00:01:39.273
Poker: post Word: All

مَثَلُ السُّلْحَفَاءِ وَالْبَطِّيْنِ

قَالَتِ الْأُمِّيُّ: زَعَمُوا أَنَّ قَدْرَهُ كَانَ عِنْدَهُ عُشْبٌ كَثِيرٌ، وَكَانَ فِيهِ بَطِّيَانٌ. وَكَانَ فِيهِ الْهَدِيرُ
سُلْحَفَاءَ بَيْنَهَا وَبَيْنَ الْبَطِّيْنِ مَوَدَّةٌ وَصِدَاقَةٌ قَدِيمَةٌ. فَاتَّفَقَ يَوْمًا أَنْ يُقْصَ الْمَاءُ فِي الْبَدِيرِ.
فَجَاءَتِ الْبَطِّيَانُ لِدَوَاعِ السُّلْحَفَاءِ، وَقَالَتَا: السَّلَامُ عَلَيْكَ أَيُّهَا السُّلْحَفَاءُ، فَإِنَّا ذَاهِمَتَانِ عَنْ
هَذَا الْمَكَانِ لِأَجْلِ تَقْصَانِ الْمَاءِ عَنْهُ. فَقَالَتِ السُّلْحَفَاءُ: إِنَّمَا بَيْنَ تَقْصَانِ الْمَاءِ عَلَى رِشْلِي؛
فَكَيْفِي، كَالسُّفِينَةِ، لَا أَقْدِرُ عَلَى الْعَيْشِ إِلَّا بِالْمَاءِ. فَأَمَّا أَنْتُمَا فَتَقْدِرَانِ عَلَى الْعَيْشِ حَيْثُ كُنْتُمَا.
فَاذْهَبَا بِي مَعَكُمْ. قَالَتَا: نَعَمْ. فَقَالَتْ: وَكَيْفَ السَّبِيلُ إِلَيَّ حَلِي؟ قَالَتَا: نَأْخُذُ بِطَرَفِي عَوْدٍ؛
وَنَقْبِضِينَ بِقَمِيكَ عَلَى وَسْطِيهِ، وَنَطْبِرُ بِكَ فِي الْجَوْءِ. وَإِيَّاكَ إِذَا تَمَيَّغَتِ النَّاسُ يَتَكَلَّمُونَ أَنَّ
تَنْطَلِقُ. لَمْ نَأْخُذْهَا، فَطَلَبْنَا فِي الْجَوْءِ، فَتَمَالَأَ النَّاسُ: عَجِبَ السُّلْحَفَاءُ بَيْنَ بَطِّيْنِ فَهَلَّا خَلَدَا!
فَلَمَّا تَمَيَّغَتْ ذَلِكَ، قَالَتَا: لَمْ نَأْخُذْهَا، فَتَمَالَأَ النَّاسُ: فَلَمَّا فَتَحَتْ فَاهَا بِالْثَّلِقِ، وَقَعَتْ
عَلَى الْأَرْضِ، فَتَمَلَّتْ.

مثل الطيب والجاهل

زعموا بأنه كان في بعض المدن طيب له رفيق وعليم. وكان ذا فطنة فيما يجري على يده من المعالجات. فكير ذلك الطيب، وضعف بصره. وكان ملك تلك المدينة ابن وحيد فأصابه المرض، فجيء بهذا الطيب. فلما سأل الفتى عن وجعه، فأخبره، فعرف داءه ودواءه، وقال: لو كنت أبصر لجمعت الأخلاط علي معرفتي بأجنابها، ولا أفتق في ذلك بأحد غيري. وكان في المدينة رجل جاهل، فبلغه الخبر؛ فأتاهم، وأدعى عثم القلب، وأعلمهم أنه خبير بمعرفة أخلاط الأدوية والعقاقير. فأموه الملك أن يدخل خزنة الأدوية، فيأخذ من أخلاط الدواء حاجته. فلما دخل الجاهل الخزنة، وعرضت عليه الأدوية؛ ولا يدري ما هي، أخذ في جملة ما أخذ منها صرة؛ فيما سب قاتل، وخاطبه بالأدوية؛ ولا علم له به ولا معرفة عنده بخبره. فلما تمت أخلاط الأدوية، سقى الفتى منه؛ فمات. فلما عرف الملك ذلك، دعا الجاهل، فسأله من ذلك الدوا؛ فمات من ساعته.

مثل الحملين والرجل الذي أصاب كثرًا

ومنى استكثر من جمع الكتب وقراءة العلوم من غير إعمال الروية فيما يقرؤه، كان خليقاً أن لا يصيبه، إلا ما أصاب الرجل الذي زعمت العلماء أنه اجتاز المفاوز؛ فظهر له موضع آتوا كثر. فجعل يخفر ويطلب فوق عيني شيء من عيني وورقي، فقال في نفسه: إن ما أخذت في نقل هذا المال قليلاً قليلاً، طال علي، وقطعتي الاشتغال بنقله وأخرله عن اللذة بما أصبت منه. ولكن سأسأجر أقواماً، يحملونه إلى منزلي، وأكون أنا آخرهم. ولا يكون بقي ورثي شيء، يشغل فكري بنقله. وأكون قد استعنت لنفسي في إراحة بقني عن الكد يسير أجره، أعطيها لهم. ثم جاء كل واحد من الحملين؛ فجعل يحملهم ما يطيق؛ فيسألون به إلى منزله، هور فيقولون: حتى إذا لم يبقى من الكثر شيء؛ أنطلق خلفهم إلى منزله فلم يجد فيهم من المال شيئاً. وإذا كل واحد من الحملين قد فاز بما حملته لنفسه. ولم يكن للرجل من ذلك إلا العناء والتعب، لأنه لم يفكر في آخر أمره.

مثل السمكات الثلاث

زعموا أن غديراً كان فيه ثلاث سمكات: كيسة (جبهة التائي)، وأكيس منها، وعاجزة. وكان ذلك الغدير بنجوة من الأرض، لا يكاد يقربه أحد. ويقربه نهر جار. فاتفق أنه اجتاز بذلك النهر صيادان، فأبصرا الغدير، فتولعا أن يرجعا إليه بشياكها، فيصيدا ما فيه من السمك. فسمعت السمكات قولها. فأملأكيسهن، فارتابت بقولها، فخرجت من المكان الذي يدخل فيه الماء من النهر إلى الغدير، فنجت بنفسها. وأما الكيسة الأخرى، فإنها تهافت في الأمر حتى جاء الصيادان. فلما رأتهما، وعرفت ما يريدان، ذهبت لتخرج من حيث يدخل الماء؛ فإذا بهما قد سدا ذلك المكان، فحينئذ قالت: فرطت؛ وهذا عاقبة التفريط فكيف الحيلة على هذا الخيال؟ ولما تنجح حملة العجلة والإرهاق. غم أن العاقل لا يقطع الأمل. ثم إنها تماوتت، فطفت على وجه الماء منقلباً على ظهرها تارة وتارة على بطنها. فظن الصيادان أنها ميتة، فوضعاها على الأرض بين النهر والغدير؛ فوثبت إلى النهر، وأمة العاجزة فلم تنزل في إقبال وإدبار حتى صيدت.

مثل طالب العلم والصحيفة الصفراء

وكذلك من قرأ هذا الكتاب، ولم يفهم ما فيه، ولم يعلم غرضه ظاهراً وباطناً، لم ينتفع بما يبدو له من خطه ونقشه. كما لو أن رجلاً، قدم له جوز صحيح، لم ينتفع به إلا أن يكسره، ويستخرج ما فيه. وكان أيضاً كالرجل الذي طلب علم الفصح من كلام الناس. فأتى صديقاً له من العلماء، له علم بالفصاحة، فأعلمه حاجته إلى علم الفصح. فرتسم له صديقه في صحيفة صفراء فصيح الكلام وتصاريفه ووجوهه. انصرف بها إلى منزله، فجعل يكثر قراءتها، ولا يقف على معانيها، ولا يعلم تأويل ما فيها، حتى استظهرها كلها، فأعتقد أنه قد أحاط بعلم ما فيها. ثم إنه جلس، ذات يوم، في مجلس من أهل العلم والأدب، فأخذ في محاورتهم، فجرت له كلمة أخطأ فيها. فقال له بعض الجماعة: إنك قد أخطأت؛ والوجه غير ما تكلمت به. فقال: كيف أخطئ؟ وقد قرأت الصحيفة الصفراء وهي في منزلي؟ فكأنت مقالته هذه أوجب للحجة عليه؛ وزاد ذلك قرأاً من الجهل ونقصاً عن الأدب.

مثل الرجل الهارب من الذئب واللصوص

قيل إن رجلاً سلك مفازة فيها خوف من السباع. وكان الرجل خيراً بوعيث تلك الأرض وخوفها. فلما سار غير بعيد، اعترض له ذئب من أحد الدئاب وأضرها. فلما رأى الرجل أن الذئب قاصد نحوه، خاف منه، ونظر حوله فلم يبق إلا قبة خلفه، فذهب مسرعاً نحوها. فلما أتى الوادي، لم ير عليه قنطرة، ورأى الذئب قد أدركه، فلقى نفسه في الماء، وهو لا يحسن السباحة، وكاد يغرق لولا أن يصر به قوم من أهل القرية، فتوافقوا لإخراجه فأخرجوه، وقد أشرف على الهلاك. فلما حصل الرجل عندهم، وأمن على نفسه من عاقبة الذئب، رأى على عذوة الوادي بيتاً، فقال: أدخل هذا البيت، فاستريح به. فلما دخله وجد جماعة من اللصوص، قد حفظوا الطريق على رجل من التجار، وهم يقتسمون ماله، ويريدون قتله. فلما رأى الرجل ذلك، خاف على نفسه، ومضى نحو القرية، فأسند ظهره إلى حائط من حيطانها، ليستريح ما حل به من الهول والإتياء، وإذا سقط عليه الحائط، مات.

مثل السلخانة والبطين

قالت الأثني: زعموا أن غديراً كان عنده غنم كثيرة، وكان فيه بطنان. وكان في الغدير سلخانة بينها وبين البطين مودة وصداقة قديمة. فاتفق يوماً أن تنقص الماء في الغدير. فجاءت البطنان لوداع السلخانة، وقالت: السلام عليك أيها السلخانة، فأنا ذاهبتان عن هذا المكان لأجل نقصان الماء عنه. فقالت السلخانة: إنما بيننا نصيبان الماء على مثلني؛ فكأنني، كالسيفينة، لا أقدر على العيش إلا بالماء. فلما أتتا فقديران على العيش حيث كنتم، فأذهبا بي معكما. قالتا: نعم. فقالت: وكف السبل إلى حلي؟ قالتا: تأخذ بطوننا، وتبضع بضمك علي وسطه، ونطير بك في الجو. وإياك إذا سمعت الناس يتكلمون أن تنطق. ثم أخذتاها، فطربتا في الجو. فقال الناس: عجب سلخانة بين بطنين قد حملتاها! فلما سمعت ذلك، قالت: يا الله! أعينكم أيها الكاس. فلما فهمت فاجها بالنطق، وقفلت على الأرض، فلبت.

Mediv: a7.3_Pagr_1.ppt
Time: 00:00:00.00 - 00:31:40.040
Font: post Modr All

مثل الطَّيِّبِ وَالْجَاهِلِ

زعموا أَنَّهُ كَانَ فِي بَعْضِ الْمَدِينِ طَبِيبٌ لَهُ رِفْقٌ وَعِلْمٌ. وَكَانَ ذَا فِطْنَةٍ فِيهَا يَجْرِي عَلَى يَدِهِ مِنَ الْمَعَالِجَاتِ. فَكَبُرَ ذَلِكَ الطَّيِّبُ، وَضَعُفَ بَصِيرُهُ. وَكَانَ لِمَلِكِ تِلْكَ الْمَدِينَةِ ابْنٌ وَجِيدٌ. فَأَصَابَهُ الْمَرَضُ، فَجِيءَ بِهِذَا الطَّيِّبِ. فَلَمَّا سَأَلَ الْفَتَى عَنْ وَجْعِهِ، فَأَخْبَرَهُ، فَعَرَفَ دَاءَهُ وَدَوَاءَهُ، وَقَالَ: لَوْ كُنْتُ أَبْصَرْتُ لَجَمَعْتُ الْأَخْلَاطَ عَلَى مَعْرِفَتِي بِأَجْناسِهَا، وَلَا أَتَقَى فِي ذَلِكَ بِأَحَدٍ غَيْرِي. وَكَانَ فِي الْمَدِينَةِ رَجُلٌ جَاهِلٌ، فَبَلَغَهُ الْخَيْرُ؛ فَأَتَاهُمْ، وَادَّعَى عِلْمَ الطَّبِّ، وَأَعْلَمَهُمْ أَنَّهُ خَيْرٌ بِمَعْرِفَةِ أَخْلَاطِ الْأَدْوِيَةِ وَالْعَقَاقِيرِ. فَأَمَرَ الْمَلِكُ أَنْ يَدْخُلَ خِزَانَةَ الْأَدْوِيَةِ، فَيَأْخُذَ مِنْ أَخْلَاطِ الدَّوَاءِ حَاجَتَهُ. فَلَمَّا دَخَلَ الْجَاهِلُ الْخِزَانَةَ، وَعَرَضَتْ عَلَيْهِ الْأَدْوِيَةُ؛ وَلَا يَدْرِي مَا هِيَ، أَخَذَ فِي جَمَلَةٍ مَا أَخَذَ مِنْهَا صُرَّةً؛ فِيهَا سِمْ قَاتِلٌ، وَخَلَطَهُ بِالْأَدْوِيَةِ؛ وَلَا عِلْمَ لَهُ بِهِ وَلَا مَعْرِفَةَ عِنْدَهُ بِجِنْسِهِ. فَلَمَّا تَمَّتْ أَخْلَاطُ الْأَدْوِيَةِ، سَقَمَ الْفَتَى مِنْهُ؛ فَمَاتَ. فَلَمَّا عَرَفَ الْمَلِكُ ذَلِكَ، دَعَا بِالْجَاهِلِ، فَسْتَقَاهُ مِنْ ذَلِكَ الدَّوَاءِ؛ فَمَاتَ مِنْ سَاعَتِهِ.

Mediv: a7.3_Pagr_1.ppt
Time: 00:00:00.00 - 00:32:17.140
Font: post Modr All

مَثَلُ الْحَمَّالِينَ وَالرَّجُلِ الَّذِي أَصَابَ كَنْزًا

وَمَنْ اسْتَكْتَرَ مِنْ جَمْعِ الْكُتُبِ وَقِرَاءَةِ الْعُلُومِ مِنْ غَيْرِ أَعْمَالِ الرِّوَاةِ فِيهَا يَقْوَاهُ، كَانَ خَلِيفًا أَنْ لَا يُصِيبُهُ إِلَّا مَا أَصَابَ الرَّجُلَ الَّذِي زَعَمَتِ الْعُلَمَاءُ أَنَّهُ اخْتَارَ الْمِقَازَ؛ فَظَهَرَ لَهُ مَوْضِعُ آثارِ كَنْزِهِ فَجَعَلَ يَحْفَرُ وَيَطْلُبُ فَوَقَعَ عَلَى شَيْءٍ مِنْ عَيْنِ وَوَرَقٍ، فَقَالَ هِيَ نَفْسِي؛ إِنَّهَا أَخَذْتُ فِي نَقْلِ هَذَا الْإِمَالِ قَلِيلًا قَلِيلًا، طَالَ عَلَيَّ، وَقَطَعَنِي الْاِسْتِغَالُ بِنَقْلِهِ وَأَحْرَانِي عَنْ التَّذَنُّعِ بِمَا أَصِيبُ مِنْهُ. وَلَكِنْ سَأَسْتَأْجِرُ أَقْوَامًا، يَحْمِلُونَهُ إِلَى مَنزِلِي، وَأَكُونُ أَنَا آخِرُهُمْ، وَلَا يَكُونُ بَقِيَّةٌ مِثْلِي شَيْءٌ، يَشْغَلُ فِكْرِي بِنَقْلِهِ. وَأَكُونُ قَدْ اِسْتَعْنَيْتُ لِنَفْسِي فِي إِرَاجَةِ بَدَنِي عَنِ الْكَدِّ بِبَسِيرِ أُجْرَةٍ، أُعْطِيهَا لَهُمْ. ثُمَّ جَاءَ كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ؛ فَجَعَلَ يَحْمِلُهُمْ مَا يُطِيقُ؛ فَيَنْطَلِقُ بِهِ إِلَى مَنزِلِهِ هُوَ، فَيَقْوُزُ بِهِ، حَتَّى إِذَا لَمْ يَبْقَى مِنَ الْكَنْزِ شَيْءٌ؛ انْطَلَقَ خَلْفَهُمْ إِلَى الْمَنزِلِ، فَلَمْ يَجِدْ فِيهِ مِنْ الْإِمَالِ شَيْئًا. وَإِذَا كُلُّ وَاحِدٍ مِنَ الْحَمَّالِينَ قَدْ فَارَ بِهَا حَمْلَهُ لِنَفْسِهِ. وَلَمْ يَكُنْ لِلرَّجُلِ مِنْ ذَلِكَ إِلَّا الْعَنَاءُ وَالتَّعَبُ، لِأَنَّهُ لَمْ يَفَكَّرْ فِي آخِرِ أَمْرِهِ.

- . Reading measures finished with the last fixation in the target area on the first pass of reading the text.
- . Regressions into the target area after the eye had left it were not counted.

Pilot testing

The experiment was run on a small number of volunteers to ensure that the setup was working, and to serve as a training period for this researcher to get familiar with the tools. This was done in Germany and Beirut prior to the start of the experiment. The pilot testing included running the experiment fully, but did not include any data analysis due to time restrictions.

Results

Tobii Studio provided raw data that could be cross-checked with the visualization tools, and this made up the basis for the extensive data analysis needed to analyze the reading measures of 72 subjects reading 6 paragraphs each. Each subject's data was manually checked, and the data points calculated and compiled and then fed into a global sheet that included all the reading measures that could then be analyzed for trends and effects. The analysis of the data collected is reported via descriptive¹⁵ and inferential¹⁶ statistics. To best illustrate what these statistics are about, below is a reference as to what the terms are relating to.

The 3 independent variables:

- . **Style (S, T, D):** This is the rotation of the 3 versions of Afandem and an effect here is related to the complexity of the typeface style. This is at the heart of the research and the most design-related. The hypothesis posits that the Simplified style will be faster to read so this is the place to look at in the speed-related dependent variables.
- . **Vocalization (on or off):** With this variable, one checks to see if the presence (v) or absence (uv) of vocalization marks had an effect on the independent variables. An effect here is most likely linguistic in nature due to the role of vowels in reading Arabic. In the following texts and table, the presence/absence of vowels is marked by adding the letter(s) v/uv after the style being tested. The hypothesis is expecting that the vowels will reduce the number of regressions.
- . **Age (13, 16):** The test was run with two age groups. An effect here would most likely relate to the difference in comprehension levels across the two ages. Note that both groups read the same texts, and that the texts were originally

¹⁵ Descriptive statistics (def.): "A general label for statistical procedures used to describe, organize and summarize samples of data. Basically, a descriptive statistic is a number that represents some aspect of a sample of data" (Reber et al., 2009).

¹⁶ Inferential statistics (def.): "Statistical procedures used to make inferences. Basically, they utilize the mathematics of probability theory to infer or induce generalizations about populations from sample data" (Reber et al., 2009).

meant for the 16 year old age group. One would expect the younger group to be slower in reading.

Inferential statistics provide the possibility to show interactions across the independent variables. An interaction¹⁷, by definition, shows an effect that is different from the expected combined effect of two independent variables. Possible areas of interaction are:

- . Vocalization x Style: An interaction here would mean that the addition of short vowels shows different results across the various styles. For example, Suv could be faster than Tuv, but Sv is much faster than Tv.
- . Vocalization x Age: An interaction here would mean that vowels help/hinder the two age groups differently. For example, the short vowels could help the younger group more than they help the older one.
- . Style x Age: A statistically reliable result here would mean that the two groups are reacting to the three styles differently. For example, the younger group could benefit from the simplified style more than the older group. The hypothesis expects that the younger group will benefit more from the Simplified style so this is where that effect would show.
- . Vocalization x Style x Age: This would be an interaction that is affected by all the variables.

As mentioned earlier, the manipulation of the variables of Style, Vocalization, and Age could be expected to have an effect of the various reading measures that eye tracking could provide, and these would make up the dependent variables in this experiment. These dependent variables are:

- . Total reading time
- . Number of fixations
- . Average fixation duration
- . Number of regressions
- . Average distance of forward saccades
- . Average distance of regressive saccades

A 3-way analysis of variance (ANOVA)¹⁸ was calculated for each of the dependent variables. Each ANOVA is a 2x3x3 with 2 levels of Age (13 and 16) x 3 levels of Style (S, T, & D) x 2 levels of Vocalization (un-vocalized and vocalized: uv and v

¹⁷ Interaction (def.): Reciprocal effect or influence... In statistical interaction the effects of two (or more) variables are interdependent; e.g. task difficulty and arousal often interact so that increased arousal increases performance on easy tasks but decreases it on difficult tasks" (Reber et al., 2009).

¹⁸ Analysis of variance (def.): "A statistical method for making simultaneous comparisons between two or more means. An ANOVA yields a series of values (F values) which can be statistically tested to determine whether a significant relation exists between the experimental values" (Reber et al., 2009).

respectively). The Style and Vocalization independent variables are within-subject and the Age variable is between-subject. In the results below, the descriptive statistics (average, standard deviation, standard error) are first reported and followed by the inferential statistics showing the *df* (degree of freedom), *F*, *p*¹⁹ (probability), and η_p^2 (partial eta squared).

A note here needs to be added regarding the directionality of the testing. When setting up an experiment, the hypothesis has the possibility to be directional i.e. to posit that the effect of a variable is specifically smaller or larger than the null hypothesis²⁰. The difference between a directional and non-directional hypothesis is one of logic and the same reported numbers could have more statistical reliability in a directional hypothesis (Lowry, 1998). In this experiment, the hypothesis is directional: It expects faster reading in the Simplified version, and less regressions in the vocalized conditions. However, the nature of the experiment with its 3 axes of variability requires the use of the ANOVA, which is by definition directionless.

Reading Time

The total reading time was calculated from the instance the subject's eyes moved into the target region in the paragraph until it left that area for the first time. Any time that was spent reading from the text outside of the target region was discounted, and returns into the target area after the first pass had moved outside were also not counted. In other words, reading time calculation stopped with the last fixation in the target area on the first pass of reading the text.

The averages, standard deviation²¹, and standard error²² are shown in Table 6.2. If one were to look at the averages for the group aged 16 when reading un-vocalized text, one could see a trend whereby the Simplified takes less time than Traditional which also in turn takes less time than the Dynamic. However, for this to be validated, and to make sure that these numbers are not occurring by chance, one needs to look at inferential statistics to see how much confidence one can put into these results, and what probability there is that these results occurred by chance.

A 3-way ANOVA was calculated (Table 6.3) and the values show a main effect²³ for the dependent variable Age with $p=0.007$ which translates to a probability of 0.7% that this result has occurred by chance²⁴. The group aged 13 years all spent more

¹⁹ P-value: "It is equal to the significance level of the test for which we would only just reject the null hypothesis. The p-value is compared with the actual significance level of our test and, if it is smaller, the result is significant. That is, if the null hypothesis were to be rejected at the 5% significance level, this would be reported as $p < 0.05$ " (Easton & McColl, 2012).

²⁰ Null hypothesis (def.): "A hypothesis of no difference, no relationship. In the standard hypothesis-testing approach to science one attempts to demonstrate the falsity of the null hypothesis, leaving one with the implication that the alternative, mutually exclusive, hypothesis is the acceptable one" (Reber et al., 2009).

²¹ Standard deviation (def.): "A measure of the variability of a sample from the mean of the sample." (Reber et al., 2009)

²² Standard error of the mean (def.): "The standard deviation of the theoretical sampling distribution of the mean. In practice it is used as an estimate of the degree to which the obtained mean of a sample may be expected to deviate from the true population mean." (Reber et al., 2009)

²³ Main effect (def.): "In statistical analysis of data, the basic relationship between a single independent variable and a single dependent variable." (Reber et al., 2009)

²⁴ By definition, for an effect to be statistically reliable it needs to have a degree of confidence that is higher than 95% i.e. for the calculated value *p* to be less than 5% or 0.05.

Reading Times: Descriptive Statistics							
Condition		Suv	Tuv	Duv	Sv	Tv	Dv
Average	13	56995	55633	56345	57546	56999	59312
Standard Deviation	13	16381	16358	12200	16480	13117	17400
Standard Error	13	2730	2726	2033	2747	2186	2900
Average	16	45832	48520	50804	47297	48657	48659
Standard Deviation	16	13137	13476	14732	14617	16146	15117
Standard Error	16	2190	2246	2455	2436	2691	2519

Table 6.2 The averages of the reading times of the two groups showing longer times for the younger group.

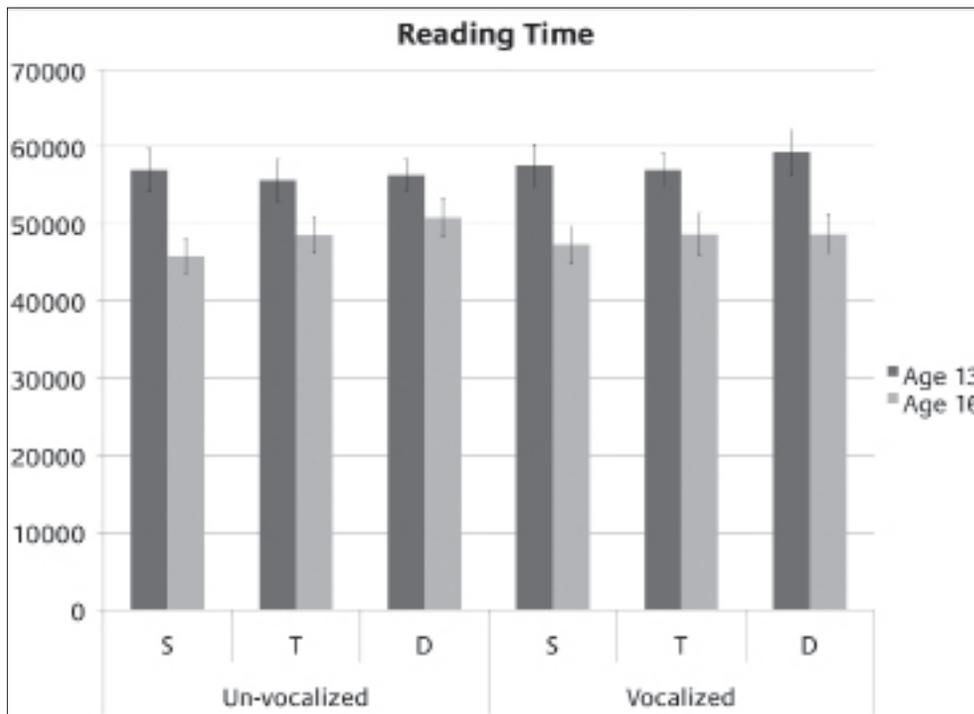


Chart 6.1 The averages of the two groups and the standard error.

Reading Times: Inferential Statistics				
Source	df	F	p	η^2
Vocalization	1	1.23	0.27	0.02
Style	2	1.66	0.2	0.05
Age	1	7.86	0.007	0.1
Vocalization x Style	2	0.05	0.95	0.001
Vocalization x Age	1	1.92	0.17	0.03
Style x Age	2	1.56	0.22	0.04
Vocalization x Style x Age	2	1.25	0.29	0.04

Table 6.3 Inferential statistics show a main effect for Age.

time reading than the 16 year olds. This is to be expected since they read material intended for an older age and as such, the result is not surprising.

The data did not show a statistically reliable effect for either the Vocalization or Style variables, nor did it show any reliable interaction. On the other hand, the data does not say that there is NO effect for these variables. With numbers like these, one can only say that the experiment did not find results that one can be confident about. It is always possible that testing with a larger group or with different methods would succeed where this experiment failed. If this were a test that was run with a stopwatch, it would have indeed been a disappointment to get results as these. However, eye tracking gives a host of other measures of reading and legibility, and these are just as interesting and as valid as variables to be studied.

Fixation Duration

The average fixation duration was calculated for every subject across the 6 different paragraphs (Table 6.4). This was the average of all fixations that occurred within the target area. Excluded are as mentioned earlier: fixations above 1 second, and fixations on white space that are less than 75 milliseconds.

The averages across the three different styles showed a rising trend that went hand in hand with complexity. The Simplified style had shorter fixation durations than the Traditional one, which also in turn had shorter fixation durations than the Dynamic style. The Vocalization variable also seemed to increase the average fixation duration for both groups. In line with the findings from the total reading time, fixation durations were longer in the younger group.

The 3-way ANOVA found several statistically reliable results for fixation durations (Table 6.5). The calculations showed a main effect for the Vocalization variable i.e. fixation duration was shorter in the un-vocalized texts with $p < 0.001$ which is an extremely reliable result as it gives more than 99.99% confidence that these results did not occur by chance.

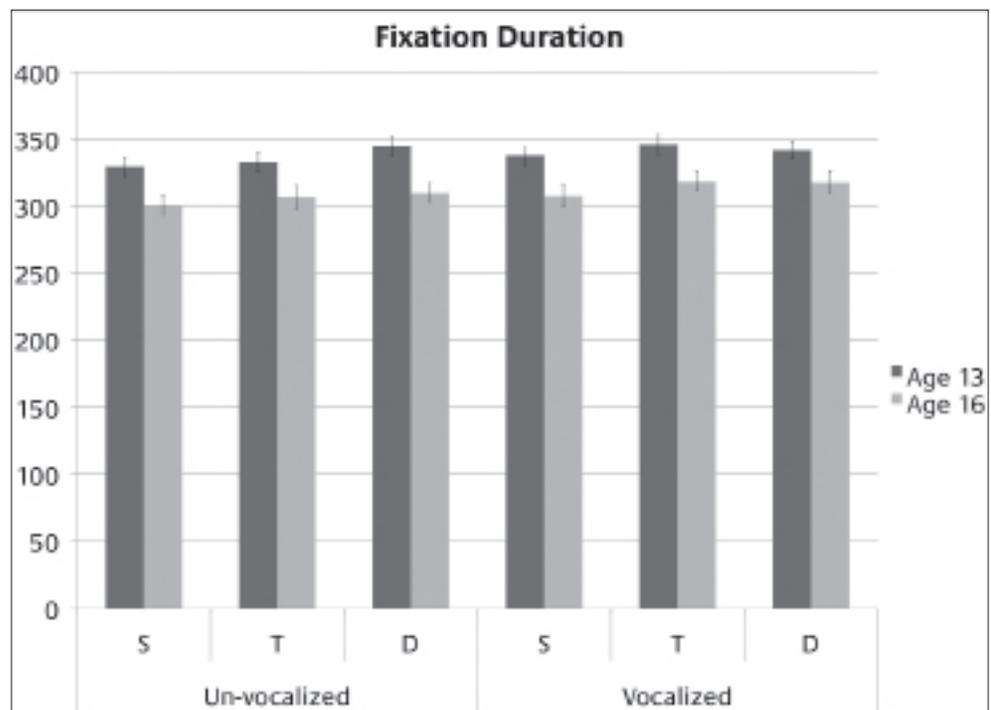


Chart 6.2 The averages of the two groups and the standard error.

Fixation Duration: Descriptive Statistics							
Condition		Suv	Tuv	Duv	Sv	Tv	Dv
Average	13	329.246	332.648	345.006	337.670	345.852	341.888
Standard Deviation	13	42.843	44.961	43.625	42.900	46.345	38.087
Standard Error	13	7.140	7.494	7.271	7.150	7.724	6.348
Average	16	300.476	306.627	309.661	307.597	318.243	317.538
Standard Deviation	16	44.993	49.802	43.032	45.834	43.989	51.201
Standard Error	16	7.499	8.300	7.172	7.639	7.331	8.533

Table 6.4 The averages of each group showing shorter fixation durations for the Simplified style and for the younger group in all conditions, and longer fixations for the vocalized conditions.

Fixation Duration: Inferential Statistics				
Source	df	F	p	ηp^2
Vocalization	1	25.34	<.001	0.27
Style	2	8.22	0.001	0.19
Age	1	8.7	0.004	0.11
Vocalization x Style	2	1.96	0.15	0.05
Vocalization x Age	1	0.82	0.37	0.01
Style x Age	2	0.26	0.77	0.01
Vocalization x Style x Age	2	1.01	0.37	0.03

Table 6.5 Inferential statistics show a main effect for Vocalization, Style, and Age.

A similarly reliable main effect could be seen for Age with $p=0.004$. Fixation durations were shorter for the 16 year olds as can be expected. The younger group fixated longer and this is most likely due to the difficulty of the test material and the less developed reading skills.

Finally, the results also show a strongly reliable main effect for Style with $p=0.001$. Readers in both groups had shorter fixation durations when reading in Simplified than in Traditional, and again fixation times during the reading in Simplified were shorter than in Dynamic. The results were not statistically reliable when comparing Traditional to Dynamic. No interactions were found.

Number of Fixations

The number of fixations was calculated for every subject across the 6 different paragraphs (Table 6.6). This was the sum of all fixations that occurred within the target area. Fixations on white space that are less than 75 milliseconds were excluded. Fixations that were above 1 sec were not excluded. The doubts around these extra long fixations stem from the possibility that these are in reality 2 fixations. They are so long, that they greatly skew the averages which is why they were not included in the Average Fixation Duration. However, whether they are 1 or 2 or even 3 fixations, they are still at least 1 fixation. Therefore, they are still included in this count.

No trend can be seen in the number of fixations across the different conditions. The numbers are close but without a clear direction.

The 3-way ANOVA did not return any main effect, even for Age (Table 6.7). There was a Vocalization x Age interaction with $p=0.04$. The 13-year-olds averaged 151 fixations over the un-vocalized conditions and 152 fixations over the vocalized ones, while the 16-year-olds averaged 146 fixations over the un-vocalized conditions and 140 fixations over the vocalized conditions. In other words, the 13-year-olds had similar fixations while reading fully vocalized text while the 16-year-olds had fewer.

Number of Fixations: Descriptive Statistics							
Condition		Suv	Tuv	Duv	Sv	Tv	Dv
Average	13	155	150	148	152	148	156
Standard Deviation	13	33	31	29	30	28	33
Standard Error	13	5	5	5	5	5	6
Average	16	142	146	150	140	140	139
Standard Deviation	16	30	30	31	26	36	30
Standard Error	16	5	5	5	4	6	5

Table 6.6 Averages for the number of fixations that each group had.

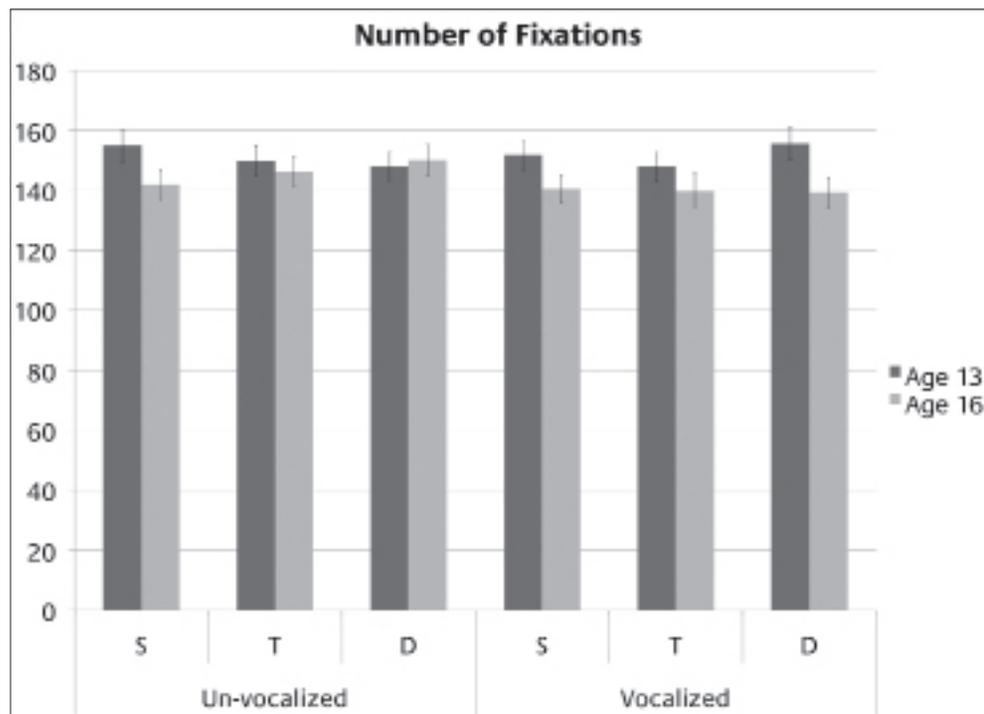


Chart 6.3 The averages of the two groups and the standard error.

Number of Fixations: Inferential Statistics				
Source	df	F	p	ηp^2
Vocalization	1	2.44	0.12	0.03
Style	2	0.56	0.57	0.02
Age	1	2.05	0.16	0.03
Vocalization x Style	2	0.22	0.8	0
Vocalization x Age	1	4.41	0.04	0.06
Style x Age	2	0.77	0.46	0.02
Vocalization x Style x Age	2	2.63	0.08	0.07

Table 6.7 Inferential statistics showing an interaction between Vocalization and Age.

Number of Regressions

The number of regressions was calculated for every subject across the 6 different conditions (Table 6.8). This was the sum of all backward saccades that occurred within the target area. Regressions on white space that are less than 75 milliseconds were excluded. Regressions that were above 1 sec were included for the same reasons as mentioned in the Number of Fixations earlier. These fixations were 1 or 2 or even 3 fixations, but at least one of them was a regression. Therefore, they are still included .

The 3-way ANOVA showed two interesting phenomena (Table 6.9). The first is as one can expect given the role that the vowels play in disambiguating texts: the numbers show a main effect for the Vocalization variable with $p=0.001$. Fewer regressions took place in the vocalized conditions. The second interesting phenomenon is the lack of a main effect for Age. The 13-year-old group was reading the same material as the older group, so the difficulty level is higher and the reading skills are lower. With such conditions, one usually expects a higher number of regressions, and yet the numbers did not show a reliable effect there. Of course, one cannot deduce that Age had no effect on the number of regressions, merely that the effect was not found.

Number of Regressions: Descriptive Statistics							
Condition	Suv	Tuv	Duv	Sv	Tv	Dv	
Average	13	25	23	23	23	21	24
Standard Deviation	13	13	13	13	13	13	13
Standard Error	13	2	2	2	2	2	2
Average	16	21	23	23	20	18	20
Standard Deviation	16	13	13	12	10	12	12
Standard Error	16	2	2	2	2	2	2

Table 6.8 Averages for the number of regressions showing fewer regressions in the vocalized conditions.

Chart 6.4 The averages of the two groups and the standard error.

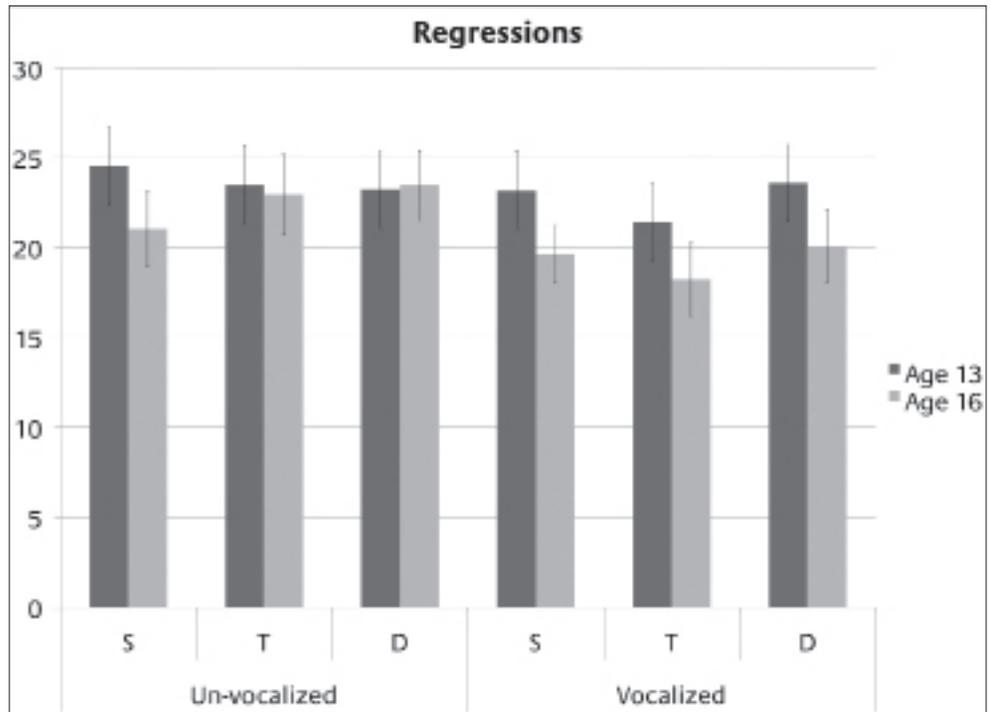


Table 6.9 Inferential statistics showing a main effect for Vocalization.

Number of Regressions: Inferential Statistics				
Source	df	F	p	η^2
Vocalization	1	11.47	0.001	0.14
Style	2	0.78	0.46	0.02
Age	1	0.88	0.35	0.01
Vocalization x Style	2	1.09	0.34	0.03
Vocalization x Age	1	2.98	0.09	0.04
Style x Age	2	0.48	0.62	0.01
Vocalization x Style x Age	2	0.66	0.52	0.02

Average Distance of Forward Saccades

The average distance of forward saccades was calculated for every subject across the 6 different paragraphs (Table 6.10). This was the average distance, in pixel measure, of all forward saccades that occurred within the target area in the first pass. Return sweeps, movements bring the eye to the beginning of the next line, were not included. Saccades that related to fixations that occurred after a regression were discounted up to the point where new text was being read. In effect, this average is measuring the subjects' forward eye movements while reading text for the first time.

In eye movement studies related to the Latin script, the unit of measurement of saccade length is in the number of characters covered where the text is set in a monospaced typeface with all characters being the same width. The typefaces used in this experiment are proportionately spaced, as are all Arabic fonts in use today. To try to make them into a monospaced version would have seriously affected the

Average Distance of Forward Saccades: Descriptive Statistics							
Condition	Suv	Tuv	Duv	Sv	Tv	Dv	
Average	13	77	73	73	75	73	72
Standard Deviation	13	10	9	9	11	8	8
Standard Error	13	2	2	1	2	1	1
Average	16	80	75	75	76	74	74
Standard Deviation	16	10	12	9	9	9	10
Standard Error	16	2	2	1	1	2	2

Table 6.10 Average distances for the forward saccades.

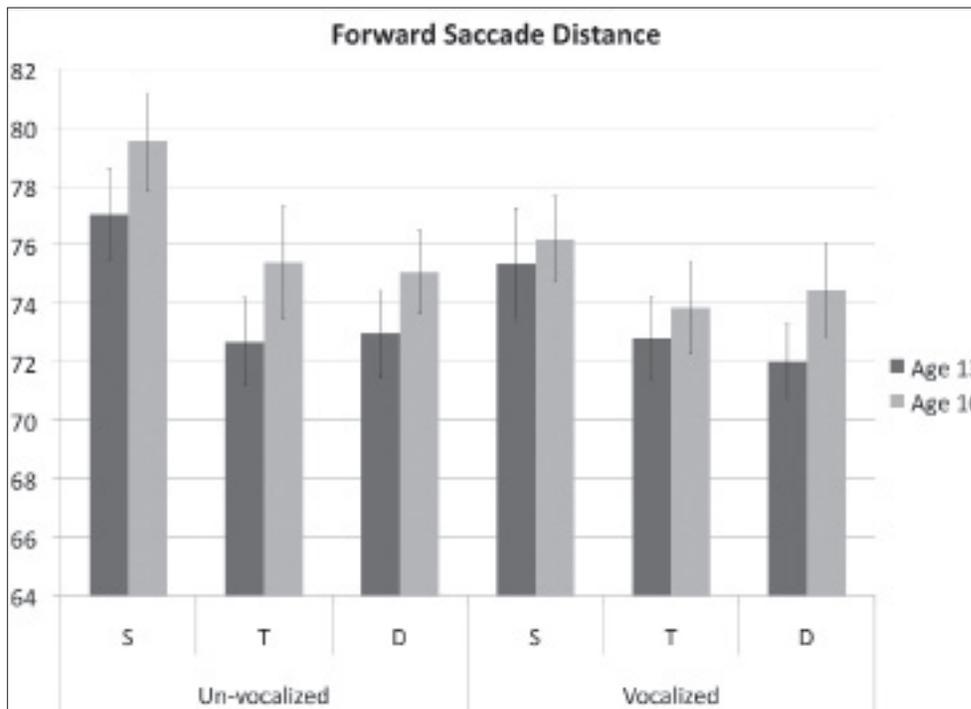


Chart 6.5 The averages of the two groups and the standard error.

Average Distance of Forward Saccades: Inferential Statistics				
Source	df	F	p	ηp^2
Vocalization	1	6.87	0.01	0.09
Style	2	21.84	<.001	0.39
Age	1	1.72	0.20	0.02
Vocalization x Style	2	1.44	0.24	0.04
Vocalization x Age	1	.07	0.80	0.00
Style x Age	2	.67	0.52	0.02
Vocalization x Style x Age	2	.98	0.38	0.03

Table 6.11 Inferential statistics showing a main effect for Vocalization and Style.

Table 6.12 Adjusted averages of forward saccades.

Average Distance of Forward Saccades: Descriptive Statistics - New Values							
Condition	Suv	Tuv	Duv	Sv	Tv	Dv	
Average	13	77	76	77	75	77	76
Standard Deviation	13	10	9	9	11	9	8
Standard Error	13	2	2	2	2	1	1
Average	16	80	79	80	76	78	79
Standard Deviation	16	10	12	9	9	10	10
Standard Error	16	2	2	2	1	2	2

Chart 6.6 The adjusted averages of the two groups and the standard error.

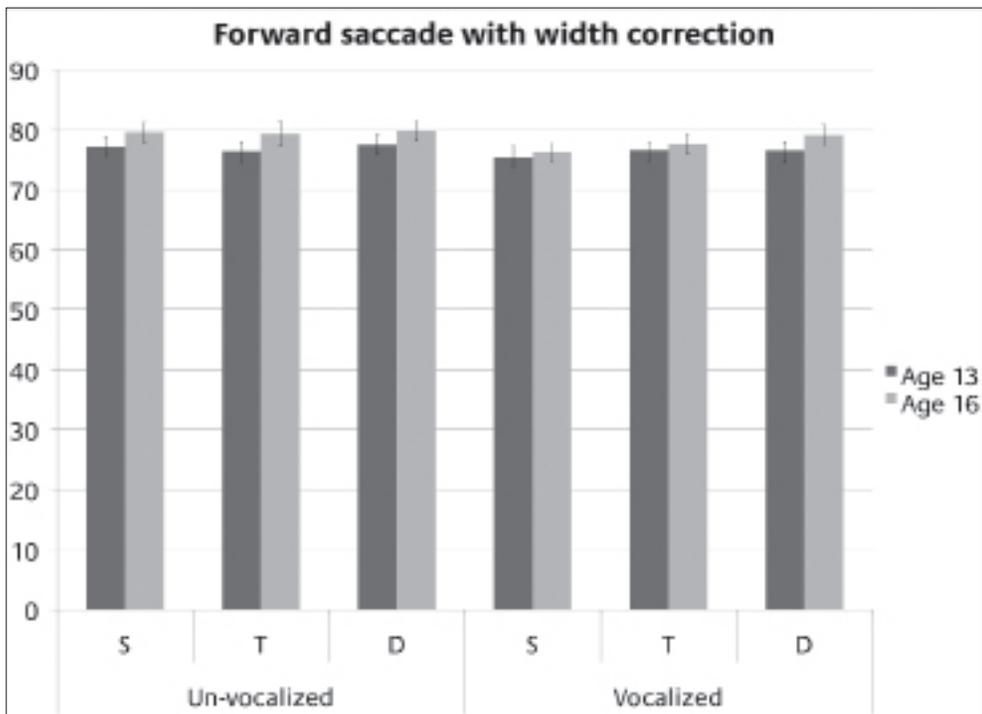


Fig. 6.13 The adjusted statistics show a main effect for Vocalization.

Average Distance of Forward Saccades: Inferential Statistics - New Values				
Source	df	F	p	η^2
Vocalization	1	11.768	0.001	0.144
Style	2	1.909	0.156	0.052
Age	1	0.928	0.339	0.013
Vocalization x Style	2	1.806	0.172	0.050
Vocalization x Age	1	1.646	0.204	0.023
Style x Age	2	0.184	0.832	0.005
Vocalization x Style x Age	2	0.623	0.534	0.018

integrity of the design and the ability of these typefaces to represent the genre of typefaces that they belong to. As such, a character count in Arabic would not be useful as that measure is in Latin. Still, a saccade length count in character would be interesting to do. The laborious nature of data analysis needed for that, manual fixation-by-fixation analysis and assignment, made it unfeasible within the time scale of this research. This is a recognized endeavor that is planned for further research.

The 3-way ANOVA shows a main effect for the Vocalization variable with $p=0.001$ where the presence of vowels reduced the length of the forward saccade (Table 6.11). The ANOVA also showed a main effect for Style with $p<0.001$ where the Simplified style had the longest saccade length.

Because the three styles differ in how letters are stacked, the word lengths end up being different across these styles. To compensate, the line lengths were measured and compared across the 3 typeface styles. Traditional took up 5.1% less space than Simplified, and Dynamic taking up 6.2% less than Simplified. So, the saccade lengths for Traditional and Dynamic were multiplied by 1.051 and 1.062 respectively (Table 6.12), and the ANOVA was re-calculated (Table 6.13). The new numbers still show a main effect for the Vocalization variable with $p=0.001$ since the correction has no impact on vocalization but only on the style. The re-calculated numbers are based on multiplying all T's and D's by their respective values and so do not affect the internal proportions of Tuv to Tv and Duv to Dv. The new calculations no longer show a main effect for Style and no interactions were found.

Average Distance of Regressive Saccades

The average distance of regressive saccades was calculated for every subject across the 6 different paragraphs (Table 6.15). This was the average of the length, in pixel measure, of all regressive saccades that occurred within the target area. Regressive movements that bring the eye from the beginning of a line to the end of the previous line, were not included. Saccades that were larger than 200 pixels, corresponding to 4 or 5 words, were discounted. The reason for this was that these were more likely a conscious decision to reread a large part of the text most likely due to comprehension reasons. These would have offset the averages and increased the standard deviation and errors in ways that make drawing conclusions that much harder. Another type of regressive saccades that were discounted was one where saccades went back across one or more lines. The pixel measure of these regressions (measured via the respective positions on the x-axis) is not indicative of the actual length of the saccade since it went over many lines. Therefore, this type was also not included in the regressive saccade calculation. Note that these regressions were still counted in the number of regressions data. As with the forward saccade calculation, a saccade length by character count is planned for future research.

The 3-way ANOVA showed a main effect for the Vocalization variable with $p=0.03$ where the presence of vocalization marks reduced the lengths of regressions (Table 6.16). As with the forward saccades, the statistics were recalculated to account for the difference in proportion (Table 6.17). The new results show a main effect for the Vocalization variable with $p=0.035$ (Table 6.18). Vocalization reduced the regressive saccade length for both age groups. The data also showed a main effect for Style with $p=0.014$. Post-hoc analysis reveals that Simplified had reliably shorter regressive saccades than both Traditional and Dynamic, but Traditional and Dynamic were not reliably different from one another. This effect for Style had not been present in the original analysis. No interactions were found.

(حسنة التائي)، وأكيس منها، وعاجزة. وكان
 . وبقره نهر جار. فاتفق أنه اجتاز بذلك
 . عا إليه بشباكهما، فيصيدا ما فيه من السمك.
 بت بقولهما، فخرجت من المكان الذي يدخل

(حسنة التائي)، وأكيس منها، وعاجزة. وكان
 . وبقره نهر جار. فاتفق أنه اجتاز بذلك
 إليه بشباكهما، فيصيدا ما فيه من السمك.
 بقولهما، فخرجت من المكان الذي يدخل

(حسنة التائي)، وأكيس منها، وعاجزة. وكان
 . وبقره نهر جار. فاتفق أنه اجتاز بذلك
 عا إليه بشباكهما، فيصيدا ما فيه من السمك.
 بقولهما، فخرجت من المكان الذي يدخل

زعموا أنّ غديراً كان فيه ثلاث سمكات: كيّسة (ح)
 ذلك الغدير بنجوة من الأرض، لا يكاد يقربه أحد
 النهر صيادان، فأبصرا الغدير، فتواعدا أن يرجعا
 فسمعت السمكات قولهما. فأما أكيسهنّ، فارتابت

زعموا أنّ غديراً كان فيه ثلاث سمكات: كيّسة (ح)
 ذلك الغدير بنجوة من الأرض، لا يكاد يقربه أحد.
 النهر صيادان، فأبصرا الغدير، فتواعدا أن يرجعا إل
 فسمعت السمكات قولهما. فأما أكيسهنّ، فارتابت

زعموا أنّ غديراً كان فيه ثلاث سمكات: كيّسة (ح)
 ذلك الغدير بنجوة من الأرض، لا يكاد يقربه أحد.
 النهر صيادان، فأبصرا الغدير، فتواعدا أن يرجع
 فسمعت السمكات قولهما. فأما أكيسهنّ، فارتابت بق

Fig. 6.14 A comparison of line width of the 3 styles.

Fig. 6.15 The average distance of regressions for both age groups.

Average Distance of Regressive Saccades: Descriptive Statistics							
Condition		Suv	Tuv	Duv	Sv	Tv	Dv
Average	13	-67	-66	-64	-61	-64	-63
Standard Deviation	13	10	14	12	10	11	12
Standard Error	13	2	2	2	2	2	2
Average	16	-68	-69	-68	-68	-67	-67
Standard Deviation	16	13	17	14	12	23	16
Standard Error	16	2	3	2	2	4	3

Chart 6.7 The averages of the two groups and the standard error.

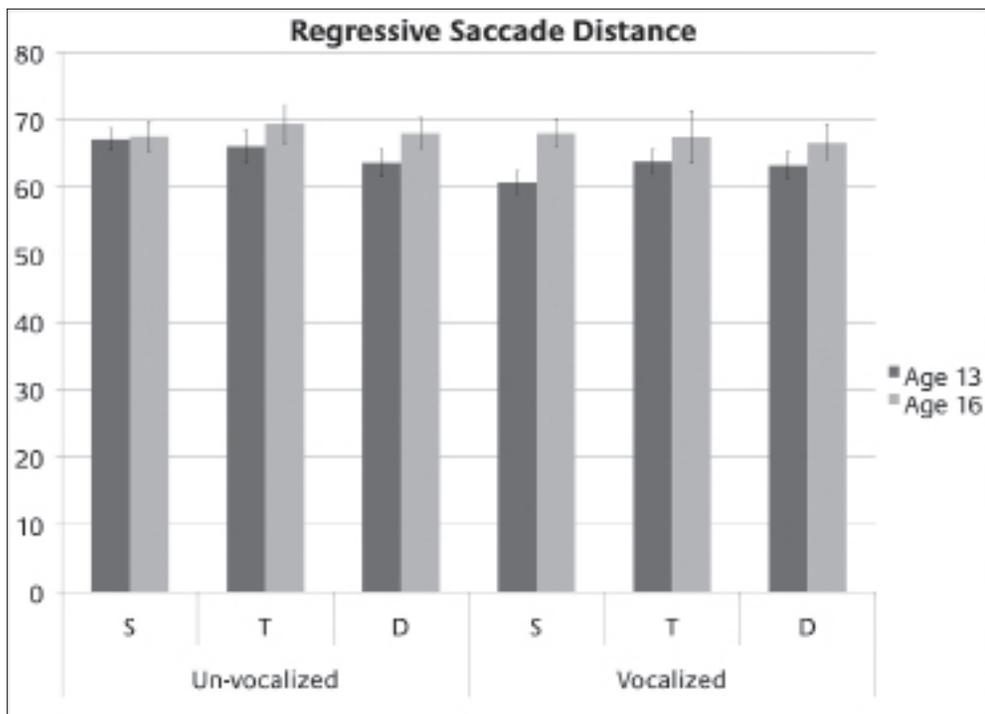


Fig. 6.16 Inferential statistics showing a main effect for Vocalization.

Average Distance of Regressive Saccades: Inferential Statistics				
Source	df	F	p	η^2
Vocalization	1	5.30	0.02	0.07
Style	2	1.28	0.28	0.04
Age	1	2.79	0.10	0.04
Vocalization x Style	2	0.08	0.92	0.00
Vocalization x Age	1	2.27	0.14	0.03
Style x Age	2	0.36	0.70	0.01
Vocalization x Style x Age	2	1.09	0.34	0.03

Average Distance of Regressive Saccades: Descriptive Statistics - New Values							
Condition	Suv	Tuv	Duv	Sv	Tv	Dv	
Average	13	-67	-69	-68	-61	-67	-67
Standard Deviation	13	10	15	13	10	12	13
Standard Error	13	2	2	2	2	2	2
Average	16	-68	-73	-72	-68	-71	-71
Standard Deviation	16	13	18	15	12	24	17
Standard Error	16	2	3	3	2	4	3

Fig. 6.17 The adjusted average distance of regressions for the two age groups.

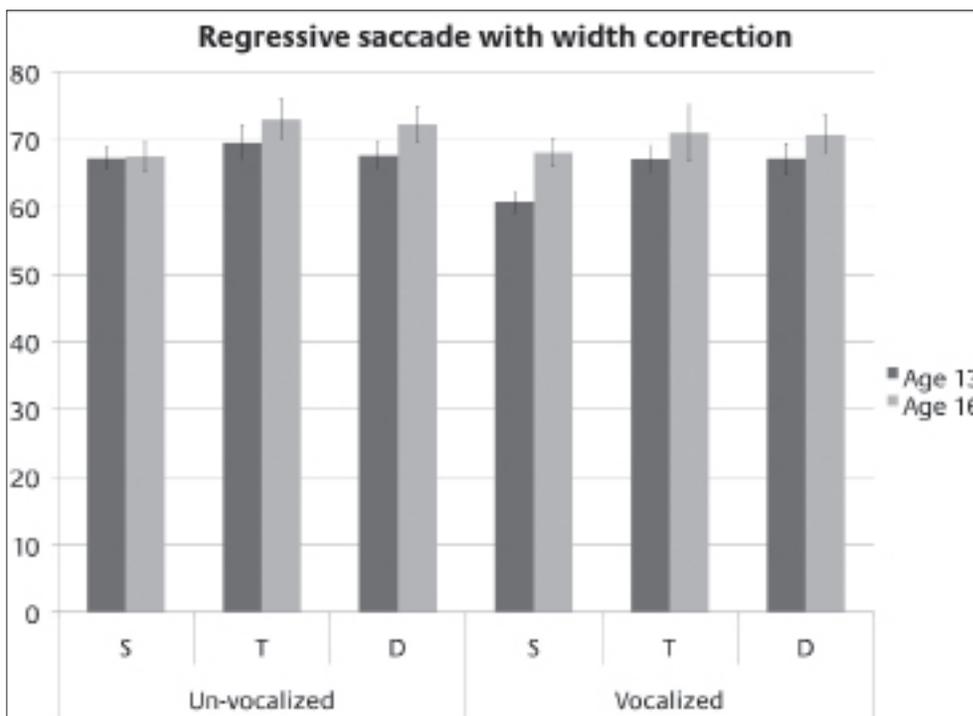


Chart 6.8 The averages of the two groups and the standard error.

Average Distance of Regressive Saccades: Inferential Statistics - New Values				
Source	df	F	p	ηp^2
Vocalization	1	4.624	0.035	0.062
Style	2	4.511	0.014	0.116
Age	1	2.089	0.153	0.029
Vocalization x Style	2	0.409	0.666	0.012
Vocalization x Age	1	1.228	0.272	0.017
Style x Age	2	0.015	0.985	0.000
Vocalization x Style x Age	2	1.838	0.167	0.051

Fig. 6.18 The adjusted inferential statistics showing a main effect for Vocalization and Style.

Discussion

Does the complexity of word formation affect legibility? Or, more accurately, how do the varying degrees of complexity of the Naskh typographic styles affect reading speed? Does the addition of short vowels aid or hinder reading? These questions are at the heart of this research.

The hypothesis posits that the complexity of the word form (as can be seen in the different interpretations of Naskh) decreases reading speed and that this effect is more evident in younger readers. The presence of vocalization marks removes ambiguity, and therefore the hypothesis expects a lower number of regressions in vocalized texts. So what do the results show?

The Effect of Style on Reading Measures

The results show that readers have shorter fixation durations while reading in the Simplified Style condition than in either the Traditional or Dynamic one. The data does not show an effect on reading time, number of fixations, or number of regressions. Do shorter fixation durations imply a more legible design? As proposed in chapter 5, legibility is the ease with which words are encoded.

Shorter fixation durations equate with faster encoding. Shorter fixation durations are then the result of a more legible typeface design, as has already been shown in chapter 5. The design of the three different versions of the Afandem typeface is made in a way that each typeface is representative of its genre. Simplified Afandem stands for Simplified Naskh, and so forth. It is then the case, that when all other design variables are held constant Simplified Naskh is more legible than both the Traditional and Dynamic versions of Naskh. This result is in line with the hypothesis, where one can clearly see longer fixation durations with the more complex styles. This result stands to reason: the more complex the visual stimulus, the longer one needs to encode it. This is also in keeping with the body of evidence discussed in chapter 5. Style also has an effect on the length of regressive saccades where the Simplified condition has shorter regressive saccade lengths, and the more complex versions have longer regressions. The implications of this finding are not really clear, but it is possible that the complexity of the word forms results in the need to regress to earlier parts of the text. As discussed in chapter 5, regressions often follow skipped words or mis-located fixations. It is possible, then, that the pattern of mis-located fixations and the complexity of the visual in the Traditional and Dynamic styles necessitate a longer step backwards.

The Effect of Vocalization on Reading Measures

The presence of vocalization marks is a linguistic research element rather than a design one, as the majority of this research is. Previous chapters have shown the intricate relationship between the Arabic script and the Arabic language. Consequently, the study of the Arabic script cannot be taken in isolation of the language it holds within. The role of the vowels in reading Arabic has been a subject of several research studies that have shown opposing results. The chance to investigate, via the intimate view offered by eye movement research, the effectiveness of their inclusion in long reading could not missed be out on and hence their inclusion as an independent variable.

This inclusion was done for several other reasons as well. One, the presence of the vowels is an added visual complexity that might react differently with the styles. It is possible that the vocalized Dynamic might get to be too crowded, and we would then see an interaction between Style and Vocalization. The vowels help to disambiguate meaning, and this is why the hypothesis expects fewer regressions in vocalized texts. There is also the cost of added visual complexity versus the benefit of clarity of meaning. Which one has a higher cost in terms of mental processing: Linguistic ambiguity or visual complexity? These are tantalizing questions that, though tangential to the role of word shape complexity, are worthwhile endeavors to be sought.

Vocalization was also shown to reduce the number of regressions, also with a strong reliability of $p=0.001$. This result is in line with the hypothesis and is very clear in logic: Vocalization adds a linguistic layer that clarifies the meaning by disambiguating homographs, and as such eliminates the role of guessing and context referencing that is involved in the reading of Arabic. As discussed in previous chapters, the context of a word plays a large role in its processing given the different array of meanings that a lot of words can have. A study had shown that reading Arabic involves more regressions than reading other languages like English (Gray, 1956). Once this ambiguity is removed, the need for so many regressions would no longer exist. In the reading of English, regressions amount to a total of 10–15% of total saccades. In the results here, the regressive fixations amounted to 15% of the total number of fixations, so it is similar to that in English but on the higher side.

As it turns out, Vocalization shows the largest number of effects. It increases fixation duration for all three styles, and the effect is very reliable with $p<0.001$ which translates to the result that there is only a 0.1% likelihood that this effect is due to chance. As far as statistics go, this is a very strong number. What does it mean? It basically says that the cost of more visual noise and complexity in the visual stimulus outweighs the benefits of the increased clarity of meaning that a fully vocalized text brings. Or, the addition of vowels made for an unfamiliar setting that tipped the scales. In either case, these results are in line with the one eye movement study that has dealt with the role of vowels in reading Arabic (Roman & Pavard, 1987), and also in line with the majority of studies mentioned in the previous chapter related to the role of vowels in the reading of Arabic text.

It is important to note the role of familiarity here. Though the students were selected from a network of schools where vocalized text was much more common than in other schools, it is still possible that the students were still somehow more familiar with reading un-vocalized texts. The experiment has tried to reduce the role of familiarity as much as can be reasonably expected within the typographic norms of today. For this factor to be neutralized completely, one would need to test with readers who are equally familiar with both conditions and within the current norms of today, the only type of fully vocalized text that one comes across is in poetry and the Quran. The selected school network offers religious teachings as part of its curriculum (Makessed, 2003) and therefore was as close to the ideal as possible.

Even if one were to say that students were more familiar with un-vocalized text, the result still goes to say that if one were to introduce full vocalization to readers today, the vowels will hinder reading rather than aid it. It is the norm in teaching Arabic that readers are first introduced to vocalized text and are then gradually weaned off it so as to arrive to a point where text is un-vocalized. Confusing as this might be, this is common practice across the region. Historically speaking, Arabic texts were originally un-vocalized. This is seen in early Quranic manuscripts in the 8th century. Vocalization was introduced when the Arabic language spread out of the peninsula and non-natives started to speak it, and one can see its early forms, as

dots, in the manuscripts from the late 8th or early 9th century (Déroche, 1992, p. 32). Is it possible then, that the early Arabic writers were on to something?

It is interesting here to look at the overall proportion of homographs in Arabic texts, though that number is not readily available. It is possible that their number is not so high, and this is why the benefit of clarity is not making up for the overall cost of more visual complexity.

Vocalized conditions also show shorter forward and regressive saccades. This again goes back to the complexity of the visual: The more complex the visual, the smaller jumps that the eye can make. This is in keeping with the body of evidence shown and discussed in chapter 5.

The Effect of Age on Reading Measures

The experiment shows strong results for the effect of age. The younger group has longer reading times and longer fixation durations. This is to be expected: the younger group was reading the same material as the older one and so the difficulty level is greater for them. Their reading skills are naturally less advanced and that also contributes to slower reading times. The lack of an effect of Age on the number of regressions was surprising, as one would expect that due to the higher difficulty level.

The inclusion of a younger group was meant for one specific purpose: to check if the changing of the Style or Vocalization has different effects across different age groups. The results showed limited interaction there and could only be seen in the number of saccades. If this were a study that primarily dealt with investigating the effect of Age on reading measures then the difficulty level would have been a possible confounding factor, as it was not maintained constant. However, the purpose here is to look for interactions with the other independent variables. It does not matter that the younger readers are slower, what matters is if the Style and Vocalization are interacting to give different results for the two groups, i.e. if the effects of these two variables are not simply additive. As such, only one interaction is statistically reliable, and that is the interaction of Age and Vocalization in the results of the number of fixations.

Interactions

The interesting aspect for testing with several independent variables is the possibility of observing how these interact. Given the experiment set-up, there are four possible interactions: Style and Vocalization, Style and Age, Vocalization and age, and Style and Vocalization and Age. The hypothesis expects an interaction between Style and Age, but that is not seen in any of the reading measures. This could imply that both age groups react in similar fashion to the change in styles i.e. Simplified is aiding legibility for both groups without a bigger impact on one or the other. However, the safe conclusion here is to say that no interaction is found for Style and Age, rather than there is no interaction at all. This is also the case for an interaction between Style and Vocalization, and between Style and Vocalization and Age.

Though not specifically outlined in the hypothesis, one could also expect that perhaps the addition of vocalization marks would play different roles across the two age groups. The younger group has less developed reading skills and is reading material that is of a higher difficulty level. The vocalization could possibly help clarify the text. This interaction is observed between Vocalization and Age in the number of fixations. The experiment shows similar fixation numbers for the 13-year old group in

both vocalized and un-vocalized texts, but the 16-year-old group has fewer fixations in un-vocalized conditions. In other words, the effect of Vocalization changes as the subjects get older and this is reflected in the number of fixations they need to read vocalized and un-vocalized texts.

The Hypothesis Revisited: The Findings

Going back to the original hypothesis, it comprises three proposals:

- . The complexity of the word form (as can be seen in the different interpretations of Naskh) decreases reading speed.
- . This effect is expected to be more profound in younger readers.
- . The presence of vocalization marks removes ambiguity and therefore decreases the number of regressions.

In terms of the first proposal, the shorter average fixation duration for Simplified goes to prove this point. The more complex styles (Traditional and Dynamic) require longer fixations. This is true for both age groups. The results are not statistically reliable when comparing the averages for Traditional and Dynamic, though when one looks at the results, one can clearly see a trend that the averages for Traditional are smaller than those for Dynamic.

As for the second proposal, that statement can not be proven with the current data. The inferential statistics show no statistically reliable interaction between Style and Age. As mentioned earlier, this does not mean that there is no interaction, but merely that it could not be found. The reason that this distinction is relevant is that perhaps this experiment set-up was not the perfect way to test for this effect. It is possible that if one were to test with a much younger age group who are just starting out to read, then the effect would be more profound. Or simply test with a larger group. As it stands, the fact that all the subjects aged between 13 and 17 benefitted from a less complex style would lead one to expect that to be the case for the beginner readers as well. The point of the second proposal is that the benefit would be more marked. Still, the fact that that benefit exists, i.e. the fact that added complexity has a cost in terms of mental processing, is already enough for one to draw conclusions and make recommendations for typographic practices.

The third proposal is strongly confirmed with a marked decrease in the number of regressions in vocalized texts, for both age groups. The hypothesis itself does not specify what expectations to have regarding the role of vowels in reading, other than the number of regressions, though the literature review does clearly point to the same conclusions as found in this study. Vocalization has the largest number of effects, from fixation duration to the number of regressions to both forward and regressive saccade lengths. In effect, Vocalization turns out to be intimately involved with the mechanics of eye movement while reading Arabic. This goes to prove how intricate the relationship between the Arabic language and its vessel is.

The Question of Authenticity

Results that show that a simpler design is easier to read are, in a way, not so surprising. So why are these findings so important in the context of Arabic type design?

There are two reasons for that. The first is related to the effect of technology on the development of typographic forms. The trend, as discussed in previous chapters, has been to simplify the complexity of the manuscript forms in order for them to be represented in print. The varying levels of vertical alignments and the large number of forms that each character can take presented technical challenges that were often hard to resolve. This is how the shaping of Arabic Naskh typefaces has evolved into its most common form: four basic shapes per letter that always connect to each other at the same vertical level, plus a handful of extra ligatures. These technological constraints are no longer there and current technology allows us far more sophistication in how words can be shaped. So we are then faced with the question: do we go back to the manuscript models of shaping or not?

This brings us to the second point. This is the question of what is authentic Arabic? It is a question being raised by prominent type designers today and it boils down to: Are typefaces that are based on manuscript models more authentic, more Arabic, than ones that are simplified? There is a controversy in the world of Arabic type design today and that is pertaining to the question of authenticity of reference.²⁵ The modern designs are generally simplified but there are many who believe that we need to return to classical manuscript models. The findings of this legibility research can guide us in answering that question.

Implications for Design and Reading

Looking back at all the results, one can find common themes running across and those are costs and opportunities. Reading is a process that is facilitated by a host of factors such as language skills, intellect, eyesight, and reading conditions such as the clarity of the stimulus, the lighting conditions, etc. The aim of reading is to assimilate written information, whether that is for entertainment, practical purposes, education, or any of many other reasons. The point is, reading is an act that requires the mental processing of a visual stimulus. In the case of reading Arabic, the complexity of the word shaping, as well as the presence of full vocalization, bring an added cost to that mental process. They add an extra cost to word encoding. As such, one can infer that Simplified Naskh is more legible than the traditional versions of Naskh, and that non-vocalized texts, even with the occasional ambiguity with homographs, are still more legible than fully vocalized ones.

The results also show the wealth of opportunities available for the design and setting of Arabic texts. The speed benefits of Simplified Naskh can be very advantageous in situations where speed of word reading is essential: on a highway, on a sign, and even in newspapers. The simplification of Naskh was driven by technology, but with this simplification came enhanced speeds of word encoding. The cost of complexity is then offset by the opportunities made possible by Simplified Naskh.

The implications of such findings are important on three levels: aesthetics, linguistics, and information design. The question of aesthetics comes down to that of typographic practices and preferences in design. Looking at the three versions of Afandem, it is not hard to say that the Dynamic version brings with it a level of elegance and beauty of construction that is missing from the Simplified version. This is not to say, that a simplified Naskh cannot look good. That is untrue. However, the poetic and fluid movement of the Dynamic forms cannot be reproduced in the simplified version. That is by nature of its construction. That beauty though comes at a

25 This is reported in various contemporary sources such as Mulder (2007)

cost of up to 4.7% increase in fixation times. The question then is, when would that form of beauty justify such a loss in word reading speed?

Such a question goes back to the function of the written text. If that was the kind of text that one can take time to read, say a novel or a literary book, where the nature of the text calls for that mood of fluid elegance, then yes, that cost is justified. When the nature of the text is less about leisurely enjoyment and more about the acquisition of information, then the simplified forms are more appropriate. When it comes to education, there is a clear benefit in presenting information that students can easily process and again the simplified versions would work best. This is not to say that the different styles cannot be mixed to present different kinds of typographic textures related to how the reader needs to read the text. The use of more complex forms could signify that that specific text is one that the reader needs to stop by and mull over. It could take on the kind of usage that italics have in Latin texts. Slope aside, true italics in a serif typeface bring with them an added complexity of shape though not of word construction. It is perhaps not a coincidence that Latin typographic norms assigned to the italic a secondary role to that of the upright.

Another situation to consider would be that involving digital displays such as a phone or an e-reader. Screen resolutions are increasing dramatically, but if one were to look at user interfaces, one is still constrained by a limited number of pixels both vertically and horizontally. A simpler design would be more applicable in such a case as the speed of interaction is quite a decisive factor, and the constraints of pixel rendering work in favor of the simpler styles. This is again an opportunity for the design of typefaces that are tailored to improve the reading experience.

In any case, the different versions of Naskh offer a wealth of variation in typographic texture that would be very interesting to investigate. The question asked at the beginning of this chapter: What form of Naskh would one recommend for a book can now be answered not purely on the grounds of beauty and aesthetics and personal preference but also in terms of functional properties that have been scientifically proven.

With regards to linguistics, and there is an educational element to that, the presence of full vocalization brings a cost that outweighs the clarity of meaning. The results showed fewer regressions but longer fixation durations for vocalized texts. This is in itself puzzling, and it is hard to ascertain whether this is due to reading habits or to the complexity of including an additional visual layer to be processed in parallel to that of the running text. The added vowels seem to be more of a distraction than an aid. So what can one deduce from that? Typographic norms are as they are. It is an interesting hypothetical exercise, say if one were to educate a group of people to read with fully vocalized text at all times, would that make a big difference in language acquisition and proficiency? Would it make learning Arabic easier? Probably yes.

As mentioned in previous chapters, children go into schools knowing only spoken Arabic, and there they are confronted with Modern Standard Arabic that is practically a second language to them. It is possible that they had heard it in on TV and maybe been read to in it, but it is still new nevertheless. Any extra help to bridge that gap is helpful, and texts for young children are fully vocalized in any case. It would be interesting to see what would happen if children never needed to be weaned off the vowels. Would that make reading Arabic less of a chore?

These are questions for educators and policy makers. Such issues need more research and more participating institutions that support such investigations. One research paper is not enough to bring about serious improvements in the teaching of Arabic. One needs a culture of research and scientific investigation, rather than a rehashing of old arguments. Language is as valid as an experiment topic as any other entity. Given

the volatility of politics in the Middle East, and a lack of reading culture, the possible benefits of further research in this domain cannot be clearer or more pressing.

Lastly, one gets to information design. This issue will be discussed separately from aesthetics for the very simple reason: Information design is about the speed of communication through text and graphics. Such areas of visual communication are of course governed by the same aesthetic considerations as any other form of design such as typographic treatment, color, proportion, layout, etc... However, such areas of design are not only governed by how well they look, but also by how well they communicate information. If one were in an airport terminal and rushing to catch a plane, one's main concern is how to get there as soon as possible. A directional sign that does not serve that purpose fails as a design, no matter how great it looks.

In situations as these, time is a deciding factor. This is no longer about the fluidity of form or authenticity to manuscript forms. This is about the fast communication of relevant information. In such cases, simplified Naskh is the typographic choice to go for. A 3% or 4% increase in the time needed to read a word can be deadly on a highway. Typography, like language, is there to serve a purpose, and in the cases like these, using simplified forms is more a duty rather than an aesthetic choice. It is again an issue of cost: that of losing a flight, or missing an exit. It is also the opportunity of saving lives.