The influence of the first script on spelling in a second script

A study of Eritreans and Syrians writing Dutch

Laura Griffioen

Faculty of Arts
Master General Linguistics
Radboud University Nijmegen

Master thesis

Supervisor: Dr. H.A.K. Klatter-Folmer
Second supervisor: Drs. E. Krikhaar

August 2018

መገዲ ትዕና ዓመት ኪዳ

"Only with aim you reach your goal."

(Tigrinya saying)

السلامة التاني في و الندامة العجلة في

"In haste there is regret, but in patience and care there is peace and safety."

(Arabic saying)

Preface

Deze scriptie is het resultaat van een soms lastige, maar vooral bijzonder leerzame periode van mijn studententijd. Het heeft uiteindelijk wat langer geduurd dan mijn bedoeling was, maar ik ben blij dat ik de tijd kon nemen om er iets moois van te maken. Aan het begin van het jaar wist ik alleen dat ik graag onderzoek wilde doen naar schrijven en schriftsystemen. Ik heb lezen en schrijven altijd interessant gevonden, en toen we een keer college hadden over analfabetisme en laaggeletterdheid ben ik geschrokken van de grote hoeveelheid mensen die moeite heeft met lezen en schrijven. Hier wilde ik graag iets mee doen, want het is zo belangrijk in onze samenleving om deze vaardigheden te beheersen. Veel van de mensen in Nederland die hier moeite mee hebben zijn immigranten. Mijn interesse ging vooral uit naar de groep mensen die wel in een ander schrift zoals Arabisch kunnen schrijven, maar het Latijnse schrift niet goed beheersen. Mijn begeleider verwees me na ons eerste gesprek door naar Jeanne Kurvers, die veel onderzoek gedaan heeft naar deze zogeheten anders gealfabetiseerden. Zij heeft me vervolgens in de richting van het schriftsysteem van Eritreeërs gestuurd, omdat ze van ISK scholen te horen kreeg dat vooral deze leerlingen problemen hadden met schrijven. Ik wil haar hartelijk bedanken dat ze tijd wilde maken voor een gesprek en me vervolgens op weg geholpen heeft met verschillende artikelen.

Er zijn diverse mensen die ik dankbaar ben voor hun steun in het schrijfproces. Allereerst wil ik mijn eerste begeleider Jetske Klatter van harte bedanken voor haar ondersteuning en de aanwijzingen die ze me steeds gegeven heeft. Ik heb het enorm gewaardeerd dat je altijd tijd wilde maken voor een gesprek. Je enthousiasme heeft ervoor gezorgd dat ik zelf ook enthousiast bleef en vertrouwen kreeg in het eindproduct. Daarnaast wil ik Evelien Krikhaar bedanken voor haar bereidheid om de tweede lezer te zijn. Ook dank aan prof. dr. Rob Schoonen, die mij als een tweede begeleider vooral bij het maken van de toets erg geholpen heeft. Door zijn goede adviezen is de toets aanzienlijk verbeterd. Bedankt dat je me uitgedaagd hebt om er meer van te maken.

Dank ook aan het Pontem College voor het feit dat ik van de lestijd gebruik mocht maken om hun leerlingen te testen. In het bijzonder Raissa van Bergendael, de docent van de AA klas en de R1 klas, enorm bedankt voor het openstellen van je lessen. Ik heb ervan genoten om te zien hoe je met de leerlingen omgaat en hen het schrijven in het Nederlands bijbrengt. Ik hoop dat de resultaten die het onderzoek oplevert nog van nut kunnen zijn voor lessen in de toekomst.

Carlien Alferink van Radboud in'to Languages, dankjewel voor het inspreken van de toets. Ook de andere collega's van Radboud in'to Languages wil ik bedanken voor hun begrip en voor alle succeswensen gedurende het jaar. Dat heeft me echt goed gedaan.

Verder wil ik mijn vrienden en familie in mijn directe omgeving bedanken dat ze er steeds voor me waren. Aan het thuisfront in Woudenberg en aan de Stuporianen: dank voor de afleiding, de aanmoedigingen, voor een luisterend oor en vooral voor jullie vertrouwen in mij. Elise en Fieke in het bijzonder bedankt, gedeelde smart is ten slotte halve smart. En vooral God, dank dat U er altijd bent en mij alles gegeven hebt wat ik nodig had om deze scriptie te kunnen schrijven.

Abstract

Many people who immigrate to the Netherlands have problems with writing Dutch. Among the biggest groups coming to our country in the last years are people from Syria and Eritrea. Most of them are well able to write in their own Arabic (Syria) or Ge'ez script (Eritrea), but they are less familiar with the Latin script. In this thesis the effect of the first scripts of mother tongue speakers of Arabic and Tigrinya on their spelling abilities was investigated. The aim was to find out whether the mistakes that these people made could be explained from the characteristics of their first script. A dictation test was created for them that included Dutch words and non-words with elements that were expected to cause difficulty. The test revealed that Syrians more often omitted vowels than Eritreans, which can be explained from the practice of vowel omission in Arabic. Eritreans had more difficulty with consonant clusters, which do not exist in their syllable-based script with syllables that have a CV structure. They performed similarly on a difficult spelling element where no effect of their first script was expected. Eritreans generally seemed to have more difficulty with transposing the overheard words to letters, which is interpreted as a difference in phoneme awareness. It is assumed that Eritreans may be less aware of the phonemes, because their first script has a courser granularity than Dutch. Extra class instructions that focus on the differences between the first and the second script are suggested as a solution to improve the writing products of Syrians and Eritreans.

Contents

1. Introduction

Many people who immigrate to the Netherlands have problems with reading and writing in Dutch. For some of them this is because they received little to no education in their home country. These people need to learn a new language, and on top of that they must acquire the cognitive skills and motor skills that are essential for reading and writing. Many others did learn how to read and write but are less familiar with the Latin script. Coming from a country where for example Chinese or Arabic characters form the main script that is taught in school, our Dutch alphabet can be quite a challenge. Chinese characters are namely more based on the level of the word and in Arabic vowels are not always written out in the way that we are used to. The latter might form less of a problem, but there is little research on how exactly the knowledge of the Arabic system would influence learning to read and write in the Latin script.

Among the biggest groups of immigrants coming to our country in the last two years are people from Syria and Eritrea (Centraal Bureau voor de Statistiek, 2017). This is evidently the result of the troubled situation in the two countries. Many of these immigrants are used to respectively the Arabic script and the Ge'ez script. The Ge'ez script is used for two of the main languages of Eritrea: Tigrinya and Tigre. When these immigrants come to the Netherlands they first need to complete an alphabetization course in order to integrate. It turned out that many of them have problems with writing and especially spelling in the Latin script, because it is so different from the script they are used to in daily life. The research in this thesis will focus on these two groups of immigrants, because they make up for a large part of the alphabetization courses. Knowing more about the way their first script influences their writing skills in Dutch will provide important insights into the problems that these people run into, and this again might help to adapt class instructions to their situation.

For this thesis both people with a background in Arabic and people with a background in the Ge'ez script were tested. The main question of this thesis is the following: ***How do the first scripts of native speakers of Arabic and Tigrinya influence writing in Dutch as a Second Language with the Latin script?*** To answer this question, people with the Ge'ez script and Arabic script as their first systems of literacy have been tested. Two sub-questions had to be answered:

1. Is there a clear relationship between the first script and the type of mistakes that are made in spelling?
2. How does the influence of the first script on the new script change overtime?

A writing test was created to answer the first question, and the answers of participants were investigated on characteristics of the first script. To answer the second question, other Eritrean and Syrian participants were tested who also started in alphabetization class and had received education for a longer period of time.

In the second chapter of this thesis the case of learning a new script is discussed and substantiated in literature. Also, the two scripts that are subject of this study and the educational background situations of both Eritrea and Syria are explained. In chapter 3 'Methodology' the actual research is described. Chapter 4 covers the results of the test and statistical analyses that were performed. Subsequently, chapter 5 contains the discussion about these results. The trends in the data are examined and interpreted. The final conclusions that can be drawn from the results are stated in chapter 6. This last chapter also discusses the possibilities for future research that follow from this study.

2. Learning a script

2.1 Early literacy

To understand what happens when adults learn to read and write it is important to go back to the way children first acquire these skills. That way, we have a clear overview of the basic learning pattern of writing and the skills that become a part of us already at a young age. It also enables us to make relevant comparisons between children and adults. Learning to read and write is one of the most important skills children learn in school, as it will be of great influence for the ways in which they can communicate and for their position in society. Our society is built in such a way that it is difficult to function normally in it without being literate. There have been multiple studies that investigated the way this skill emerges in younger children. An important contribution to our understanding of the process of learning to read and write is the work of Gibson & Levin (1976), who selected and summarized fourteen years of research in this area and created an overview from the results. They show which skills are present when children start to learn, and which are still in development. They also show that there are several stages that a child goes through while learning how to write.

According to Aarnoutse (2004) the process of learning how to read and write can roughly be divided into three stages, namely evolving, beginning and advanced literacy. The first phase can be located before children enter primary school and it is the time in which children acquire the basic principles of their language, such as communicating vocally, speaking and listening. Then the second phase includes the first three years of education in which children are taught the basic principles of reading and writing. They learn how letters represent sounds (the so-called alphabetic principle) and that words are in fact different combinations of the same letters. In the third phase children learn to further develop these basic skills so they will be able to recognize words more quickly. They also learn to communicate via written messages and they learn to master more advanced techniques such as reading strategies and grammatical knowledge (Aarnoutse, 2004).

We will now look at the development during these stages in a more detailed way. In general, by the time children start to read and write they have received a lot of input in the language they will use: the average six-year-old has a vocabulary of about 2500 words (Gibson & Levin, 1976:129). Children are familiar with the sounds of their language and have intuitions about the sound combinations that are possible in that language. English-speaking children know for

instance that 'wug' could be a word in English, while 'wgu' could not. At about seven years of age they are also quite able to divide a sentence into words, but they still have some difficulty with function words (Gibson & Levin, 1976). Up to the age of seven, children seem to prefer to divide a sentence into a subject part and a predicate part, rather than using the smaller unit of the word (Karpova, 1966, in Kurvers, Van Hout & Vallen, 2007). However, they do know how to divide a sentence into words, as has been shown by studies that used storytelling and asked children to tell the story word-for-word, so the researcher could write it down. These children were very well able to separate words (Chaney, 1989). Considering morphology, the situation is a little different, since this is still in development when children enter primary school. There is a well-known study by Berko (1953, in Gibson & Levin, 1976) that clearly shows that English first-graders are quite well able to create plurals ending in [s] or [z] for nonsense words (90% correct), while preschoolers make many more mistakes (75% correct). Still, both groups made mistakes with words ending in [ɪz], as in the real word *glasses*, which shows that morphology is not yet fully developed when children go to school and start to learn reading and writing. Grammatical knowledge is another aspect that is still in development as children enter primary school. Although the basic grammatical structure of the language is fully developed, children do make mistakes in less frequent and more difficult structures (Gibson & Levin, 1976). Nevertheless, the basic structure is ample knowledge for starting to write sentences.

To be able to actually start writing the child first needs to understand that there are smaller units of sounds in a word. This starts with phonological awareness, which is the realization that words consist of smaller units such as syllables, and that they have a beginning and an ending (e.g. /str/, /eet/) (Braams & Bosman, 2000). A step further than this is phonemic awareness or awareness of the phonemic principle. This is awareness of the smallest units of sound in a language that can change the meaning of a word (e.g. /c/, /a/, /t/ are phonemes in the word /cat/, because you could change the /c/ to a /b/ and you will have a different word meaning). As soon as a child is aware of the different phonemes that a word is made of writing becomes a lot easier, because phonemes are often associated with letters. Phonemes should not be confused with letters however, because for example /ck/ in English is one phoneme. But phonemic awareness does aid acquisition of the aforementioned alphabetic principle (the fact that letters represent sound) (Gibson & Levin, 1976:174-182; Morais, 1991). Morais (1991) points out that it is very important to distinguish between phonemic awareness and the initial phonological awareness. Phonological awareness namely emerges spontaneously, while phonemic awareness needs more instruction. This can be seen from the fact that children do start rhyming on their own

initiative, whereas most children do not tend to make their own phonemic changes before being instructed in some way (Gibson & Levin, 1976:120; Morais, 1991; Bertelson & de Gelder, 1991). Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh and Shanahan (2001) conducted a meta-analysis on phoneme awareness instruction and they found that children largely benefited from this, both for reading success and spelling success. Especially when taught in combination with letters, phoneme awareness instruction was very effective (Ehri et al., 2001). The ability to distinguish between phonemes is also known to be a better predictor of children's future performance in reading and writing than phonological awareness (Braams & Bosman, 2000). These findings show that phoneme awareness is an important aspect of learning to read and write, and thus also an important competence to take into account when investigating the emergence of these skills.

As we have seen a lot of knowledge of the language and its structure is already present when a child starts to learn its written form and this knowledge facilitates the process. Now we will consider the actual process of learning to read and write. At the very first beginning children just scribble on paper. Even this is part of the process of learning though, because they get some feeling for the way to hold a pen and for the different shapes that can be created. Of course, education of writing differs across languages, because the sounds and scripts are different from each other. Ziegler and Goswami (2005) observe that children with different mother tongues often already have a different starting point. For example, Turkish children have above average phoneme awareness prior to literacy, simply because Turkish has vowel harmony and they need this knowledge to create plurals. Likewise, children who speak a language that has a simple syllable structure are more aware of syllables. Ziegler and Goswami (2005) argue that three factors play a role in the way children acquire writing skills, namely availability, consistency and granularity. Availability is defined as the ability to consciously access phonological units. Not all phonological units are accessible when children start literacy acquisition. Consistency is about the directness of the script, such as whether it has multiple pronunciations for the same written unit. Granularity involves the size of the linguistic units that the script works with, so whether the word, the syllable or the phoneme forms the smallest written unit (Ziegler & Goswami, 2005). Large units are considered the most accessible phonological units for beginning readers (Ziegler & Goswami, 2005). This means that children who acquire a script with larger units initially have an advantage. We will come back to this phenomenon when discussing writing in second language acquisition. By all means, the discussed findings reveal the extent to which the mother tongue triggers certain awareness prior to reading.

Although this thesis is dealing with writing, we will also need to talk about reading shortly. When learning to write one inevitably also learns to read, so these two processes cannot be separated. Reading and spelling seem to depend on the same phonetic and orthographic awareness, as a study by Gill (1989) also indicates. Gill presented children with words with correct spelling and the same words with the spelling of the children's own writing products and found that they were able to identify both with equal performance. From this he concluded that apparently the same orthographic and phonological knowledge is at work in reading and writing (Gill, 1989). The errors that children make in the process of learning to read also give us a better understanding of the strategies that they use for the way they view words in general. When children first begin to read the most frequent error is substitution, which is that they come up with their own word instead of reading the one that is printed. They base their assumed word on context and on the graphic features of the printed word (Gibson & Levin, 1976). This type of error shows that children really make use of the knowledge that they have of their language, which again reveals the importance of this knowledge when starting to read and write. In fact, there are three stages of learning to read:

1. Substituting any word that makes sense;
2. Inspecting the graphic display and saying nothing if you cannot decode it;
3. Basing the words read on the graphic features and the logic in the context.

(Gibson & Levin, 1976)

After a while, children learn to recognize more and more letters and are able to decode words better. As they learn to recognize all the letters that words are made up of, this also improves their ability to write. A child namely learns to read by first dividing the word into letters, then converting the letters to sounds and connecting the combination of sounds to a spoken word with a certain meaning. For writing the order of the process is reversed: they have to start with a word that needs to be divided into separate sounds, then these sounds need to be converted to letters and in the end the letters form a written word (Verhoeven & Aarnoutse, 1999; Verhoeven, Aarnoutse & Van Leeuwe, 2000; Aarnoutse, 2004). The described processes of decoding and converting are really acquired at the same time, which shows how closely related the two processes are. This underlines the importance of research in reading for investigations of writing ability, and also shows that it is good to take reading ability into account when investigating writing ability.

Another important aspect of writing to discuss is the way it is learnt in school. The way children learn to spell really differs per language, per country, and even per teacher. Johnston (2000) investigated the different practices that occur when teaching spelling in English. For that she used the three perspectives on spelling as determined by Nelson (1989). These three perspectives are also relevant for discussing the differences that there are across countries. They are:

a) *Rote visual memory*

   This is the vision that is based on the idea that the orthography of a language is very irregular. It is assumed that each word stands on its own. Learning to spell from this perspective includes memorizing and practicing with word lists.

b) *Generalization*

   From the generalization perspective it is assumed that a language has an underlying structure or logic that can be learned. The way words are presented in the teaching practice that follows from it is more in categories of words that have similar orthography.

c) *Developmental*

   The developmental perspective follows from the generalization perspective. It adds the element that children can be placed on a scale of development in spelling, with certain orthographic features being acquired later than others.

(Nelson, 1989; Johnston, 2000)

In the chapters about language background in Eritrea and Syria the teaching situation in those specific countries will be discussed in more detail. What is important to remember is that not all perspectives on spelling instruction are equal. This may also affect the learner's perspective on learning to write and spell in a second language.

In summary, we have seen in this chapter that as a child we go through lots of processes and phases when learning to read and write. Before all else children acquire the basic principles of language. When they enter primary school basic knowledge about communication, possible sound combinations, and the vocabulary of their language are already present. Morphological and grammatical knowledge is also largely present, although children have some difficulty with less frequent structures as this knowledge is still in development at young age. The two important principles a child has to learn in order to start writing are the alphabetic principle (the fact that letters represent sounds) and the phonemic principle (words are made up of different phonemes). We have seen that these two principles need to be taught explicitly, while knowledge of the language and its structure has already been acquired earlier. It is very relevant

to recognize that instruction is required for this, because that makes it easier to understand the problems that adults experience when becoming literate in a second language. Another important matter is that the knowledge prior to becoming literate may differ across languages, because of their different structure. Therefore, children of the one language are more aware of phonological units or syllables than children with a different mother tongue. Considering actual writing development, we know that children start by creating random scribbles on paper and from there they go on to learning to write actual letters. Furthermore, we have seen that reading and spelling appear to depend on the same orthographical knowledge, which means that results from studies on reading can also provide insights into writing. Children's mistakes in reading for example show how they make use of the general knowledge that they have of their language. Lastly, the way children learn to write and spell in school can really differ per country, depending on the perspective on spelling that is taken. In conclusion we can say that development of literacy skills in children most of all shows the relevance of knowledge of the language and the importance of acquiring phonemic and alphabetic knowledge. These are the important qualities that we need to keep in mind when studying adult literacy acquisition.

2.2 Late literacy

In this chapter, another important group of people will be discussed: the illiterates. Research on the illiterate mind namely reveals how exactly literacy influences our way of functioning and thinking about language. Knowing this will help to see which advantages literate learners of a second language have, although they might have learned a different script. Sadly, approximately 13% of all people never acquires the very important skills of reading and writing before reaching adulthood (UNESCO, 2016) and thus remains illiterate. For these people, learning a script at an older age is incredibly difficult. This is partly so because the ideas that those illiterates have about written words and sentences are very different from the literate way of thinking and more like that of children. Just like children, they might for instance have the idea that the written word for tower is longer than that for house. In her book 'Met ongeletterde ogen' ('With illiterate eyes') Kurvers (2002) gives many examples of similar misconceptions. For example, when asked to indicate which word is bigger, 'television' or 'room', an illiterate person is likely to answer that a room is much bigger than a television. Their concept of words is simply different and therefore many illiterates find it hard to explain what exactly a word is and have difficulty providing their own examples of long or difficult words (Kurvers, 2002). Illiterates also have difficulty with determining what the units in a sentence are. Bossers, Kuiken and Vermeer (2015:252) give an example of a woman who was unable to divide the unit 'the old man' into three parts. Another striking example is that of a woman who asked someone to read a letter to her, but to close her ears so she would not hear the content, because that was secret (Bossers, Kuiken & Vermeer, 2015:252). This difference in conceptualization of the written word makes the process of learning to read and write more complex for adults than for children, because their divergent idea of the written word has been persistent for a longer period of time.

Several studies (e.g. Morais et al., 1979; Morais, 1993; Read et al., 1986; Reis & Castro-Caldas, 1997) have revealed that illiterate adults also find it difficult to divide spoken words into smaller units of sound. As mentioned before, the awareness that words are made up of such smaller units of sound is known as phoneme awareness. Just like children, most illiterate adults are unaware of the fact that each word consists of a number of different sounds. Consequently, illiterates have trouble with tasks in which pseudo words are used. The use of pseudo words namely stimulates different processing in the brain, because the word is not stored anywhere as a lexical unit. For good performance on a task like this phonological processing is required and this has been proved to be difficult for illiterates. Castro-Caldas, Petersson, Reis, Stone-Elander

and Ingy (1998) provide a clear example of that with their study on brain activation. They tested both literate and illiterate participants and asked them to repeat actual words and pseudo words. When real words were presented, literate and illiterate participants performed identically: both groups had good results and the same areas in the brain were activated. For the pseudo words however, the illiterate subjects made four times more errors in the repetition. Also, different neural structures in the brain were activated than for the literate participants (Castro-Caldas et al., 1998). Because the subjects that were used had comparable backgrounds, the big difference between the two groups could not be explained by factors such as lexical knowledge. From this the authors concluded that apparently the ability to read makes changes to the way the human brain is functionally organized (Castro-Caldas et al., 1998). Knowing that helps us understand the extra difficulty that adults have when becoming literate. New functional categories must be created in their brain and this takes time, especially at an older age.

We also know that literacy influences non-linguistic processing to some extent. Petersson, Reis and Ingvar (2001) found similar results as the aforementioned authors with their study on pseudo words. Besides reaffirming the difference in verbal working memory between literate and illiterate subjects they further examined the cognitive difference between the two groups. They gave their subjects two different picture naming tasks, one task with 2D pictures and the other with 3D pictures. The idea was to find out whether literate subjects would be better at naming the 2D pictures, as they are more familiar with figurative representation because of the education they have received (Petersson, Reis, and Ingvar, 2001). By analyzing interactions between areas in the brain the authors found that indeed literacy not only provides skills that are related to language, but it also affects other cognitive skills and with that other brain areas. The illiterate subjects namely had more difficulty with naming the 2D pictures, but not the 3D pictures. This study shows how literacy does not only affect language processing, but also other cognitive processing that has to do with 2D material. Getting familiar with 2D representation is yet another part of the skill set that still needs to be acquired by illiterates at an older age.

Another area of research within the subject of illiteracy is the awareness of word boundaries. People used to think that illiterates could not determine where word boundaries are, but several studies (e.g. Scribner & Cole, 1981; Hamilton & Barton, 1983) have shown that illiterates are quite well able to do this, for example when asked to dictate a story word-for-word to someone who needs to write it down. It still is a point of discussion among linguists whether literacy does have an influence on the way word boundaries are determined. Kurvers, Van Hout and Vallen (2007) discuss the results of a study that compares adult illiterates, pre-

school children, and adult readers on their ability to segment sentences. Their participants who could not read (both children and adults) hardly ever segmented a sentence into isolated words, while the literate adults were likely to do this. An interesting finding was that the children preferred to divide the sentence into syllables, whereas the illiterate adults talked about the content of the sentence or grouped nouns and function words (Kurvers et al., 2007). From these results it seems that illiterate adults do differ from children in the way they think about words. This might be because young children are very busy with the process of language learning, while illiterate adults have grown more accustomed to the language and do not actively think about its smaller units. Kurvers et al. (2007) conclude from their experiments that the competence to mark word boundaries strongly depends on knowledge of the written form of that language.

How big the difference between literates and illiterates in cognitive processing actually is can only be proved by testing subjects with online tasks. This means that participants are studied as they are processing unconsciously, without actually having to perform a task such as repeating words. That way, there is less interference of the task and use of other skills that might differ per participant. An online task would for example be to measure brain activity, while offline studies include word repetition tasks. Veldhuis and Kurvers (2012) conducted such a study to separate online and offline performance on word boundaries. They created several tasks that differed in the extent to which they were online or offline. What they found was that the influence of literacy on segmentation was strongly present for the more offline tasks, but for the more online tasks the difference between literate and illiterate participants was a lot less clear (Veldhuis & Kurvers, 2012). Apparently, the difference between literates and illiterates is not so big in the way they process existing words and determine word boundaries in the brain. The difference seems to lie more in performance, when metalinguistic knowledge is required.

Now that we have a basic idea of the way illiterate adults view language, it is important to discuss the way they learn to read and write in a new language. Sadly, there is little research on the way the emergence of this skill in adults differs from the process that children go through. Kurvers and Van der Zouw (1990) assume that at least the way adults learn to read is comparable with the way children learn to read. As they note, children start out as impressionistic readers, initially only pronouncing the words they know. After that they slowly become analytical readers who start to recognize the correspondence between the graphemes and their sounds, a process that gets more and more automatized as they get used to reading (Kurvers and Van der Zouw, 1990:260-261). A study that did compare the stages of learning that adults and children go through when becoming literate is the study by Viise (1996). Viise

compared adult literacy learners to children in the classroom who were at comparable levels of learning. Viise studied and compared their spelling errors by looking at 24 different word features such as phonetic patterns, visual patterns, and word relationships like beginning and ending consonants or vowels in stressed and unstressed syllables. Her most important finding was that adults and children appeared to have a similar developmental pattern in spelling. Children and adults were comparable in their mastery of recognition of initial consonants, end consonants, and short vowels. Also, both groups were evenly likely to omit vowels. These findings were persistent, despite the other differences there were between the groups such as length of education and differences in ethnicity. There were some differences in mastery of the different features that were tested. Adults for example had better understanding of the word concepts and were better at identifying visual patterns of the words that were used, whereas children were better at syllable representation of phonetic features (adults were especially more likely to omit or change word endings). However, the overall development seemed identical in adults and children, particularly in the early levels of spelling. Both had difficulty with vowels and consonant clusters. The difference between the two groups was more than 10% in only 7 of the 24 features that were tested, which indicates a strong similarity in development (Viise, 1996). Considering the order in which skills emerge, the study by Viise further shows that both adults and children acquire spelling skills in a specific order, starting with the simpler features and later acquiring the more difficult features such as double consonant clusters (Viise, 1996).

Bossers, Kuiken and Vermeer (2015) discuss three phases that adults go through when becoming literate. The first phase is the logographic phase. They learn to recognize written words from certain visual features, for example the dot above the 'j' helps them recognize the word 'jam'. In this phase they guess a lot and view the words more or less as pictures. The next phase is the phase of reading by spelling. There the learner is mastering how to decode a word by recognizing its separate units. This is the important alphabetic principle that we have discussed earlier. The third phase is the orthographic phase, in which one learns to automatize the alphabetic principle and recognize words in an instant. Bossers, Kuiken and Vermeer (2015) note that this phase might take much more time for second language learners, because their lexicon of Dutch is not big enough to know which of them are real words in Dutch.

Although the ways adults and children learn to read and to master the match between orthographic representation and sound may be comparable to that of children, the first stages of learning to write are considerably different and in fact do not apply to adults. In general adults do not start by randomly scribbling on paper, since they usually begin to learn in the classroom. Instead, they immediately learn to make shapes that will later help them to create

letters. Some schools even start with the letters without first practicing with shapes. There are few studies that have looked into the way the practice of writing emerges in adults. In her book on second language acquisition Saville-Troike (2012) devotes a chapter to learning to write in a second language. Some learners are said to start by copying or even tracing over example words and symbols or try to make their own depiction of spoken words that they have heard. Bossers, Kuiken and Vermeer (2015) also briefly discuss the way adults learn to create letters. Writing practice in the Netherlands is often built up from writing between contours to tracing over, partly tracing over, copying, and finally writing individually. There may be some difficulty with mastering the required hand-eye coordination and adults need to learn what the distinctive features of letters are. In writing courses most teachers also try expose their students to different appearances of written language, in order to get them used to the different fonts and ways of writing that exist beside the type they learn in class (Bossers et al., 2015:266).

In this chapter we have seen that literacy has a considerable influence on the way people think and the way they process information in certain regions in the brain. The existing research shows that becoming literate at an older age requires people to acquire a bigger set of skills that are all new to them, such as 2D representation and dividing words into smaller units at phoneme level. This demands a lot from the brain that is less adaptable at an older age. Considering word boundaries, illiterates seem quite well able to determine these when asked to dictate word for word. And although illiterate adults do perform a lot worse in offline tasks, their online performance does not seem to differ much from that of literates. The difference between literates and illiterates seems to lie more in performance, for which metalinguistic knowledge is required. Although research on the emergence of reading and writing skills in adults is limited, we can carefully conclude from what research there is that at least the process of learning to read is quite comparable to that of children. They seem to follow the same pattern in the order in which they acquire the relevant features of reading. There has been less research on writing development, but from classroom situations we can conclude that adults do not begin to write from the same starting point as children do, as they skip the childish state of drawing and scratching. Other than that, the process seems to be comparable for adults and children. Most importantly, all the aspects in which illiterate adults are at a disadvantage show how big the head start of people who already know a script is. This group of language learners will be discussed in the next chapter.

2.3 Learning a new script

The last group of learners that is considered here form the focus of this thesis. In this chapter we will look at those learners that have become literate in their own language but need to learn a new script as a second language learner. Learning a new language already is a complex task. A new language often has a different phonology, different grammar, and different word order and there are many other concepts and characteristics that the languages the learner is familiar with might not have. The task of learning a new language becomes even more complex when this language has a script that the learner is not familiar with. When a learner already has knowledge of some other writing system, this is of course an initial advantage: he or she has the same basic idea of what words and sentences are and knows that these can be written down. The basic concept of the way symbols transfer to language is applicable. That is a great benefit as opposed to illiterates who have to start from scratch (as we have seen in the previous chapter). But beside this important cognitive advantage, the first script (hereafter also indicated as S1) of the learner may either have a more positive effect or a less positive effect on the learning pattern. The concept that the learner has in mind of characters, letters, and syllables may be very different for a new script. For example, someone who knows Japanese characters is more familiar with characters on word level and has a less specified idea of individual sounds within these words. The assumptions that someone has about written language are likely to interfere with the actual characteristics of the new script. It has been a common practice in L2 education to view L1 literacy as a barrier rather than an advantage (Edelsky, 1982). However, the positive influence of L1 literacy has also been defended and nowadays most scholars recognize that functional and contextual experience with writing has its advantages for the L2.

Especially the influence of Asian scripts on learning a second script (S2) is a well-studied topic in this area. There are many examples of studies that show how adults who can only read the Chinese script are less able to recognize individual sounds in spoken words. De Gelder, Vroomen and Bertolson (1993) for example, tested bilingual Chinese adults who had lived in the Netherlands for several years. The hypothesis from which they started is that the way phonological information is stored in the brain depends on the phonological level at which a language's orthography is functioning (De Gelder, Vroomen & Bertelson, 1993). If this were indeed the case, their Chinese subjects would have to perform worse on segmentation of words into smaller units, since Chinese characters function on word level. De Gelder and colleagues compared two groups of individuals: the first group only had knowledge of the Chinese script and the second group was able to read in Dutch as well. The participants had to complete three

different tasks. In the first task they had to delete the first consonant of a Dutch pseudo-word. The second task was a progressive fragmentation task in which they had to divide sentences into ever smaller units until they could no longer make them smaller. The third was a rhyme judgement task. The alphabetically trained Chinese subjects performed significantly better on the consonant deletion task and on segmentation. However, all participants performed equally well on the rhyme judgement task. This shows that literacy in the second language mainly influenced the ability to manipulate word segments and to distinguish smaller units, which is not a property of the Chinese script (De Gelder, Vroomen & Bertelson, 1993).

Because there are such big differences in the way scripts are structured, one can imagine that this also has an influence on the way people learn a new script. A differently structured S1 may result in a different learning process of learning an S2. Where an S1 like Chinese might slow down the process of learning a script that is more alphabetic, a different phoneme-based script may very well facilitate the process. A study by Mishra and Stainthorp (2007) provides more clarity about the differences between first scripts, and how one script can facilitate reading and writing while another may slow down the process. They tested phonological awareness and word reading in the first and second language of fifth grade children. Recall that awareness of the phonology of a language facilitates reading in that language. The languages that were tested were English and the syllable-based script of Oriya. The fact that Oriya has a syllable-based script means that the way words are written down goes per syllable rather than per phoneme. Thus, the script of Oriya is made up of larger units than the script of English. For some of the children the script they learned in school was that of their first language, for others it was not. What Mishra and Stainthorp (2007) found was that children who learned to read Oriya as their first literacy language had advantage of their awareness of large phonological units, but children who learned it as their second literacy language did not benefit from this knowledge. For English on the other hand, awareness of the phonemes in English contributed to English word reading, also when it was the second literacy language of the child. The authors connected this information to the notion of grain size, which means the size of the units that are used in a script. Thus, Oriya is a large grain-size script, where English is a small grain-size script. They concluded that awareness of smaller units (like phonemes) can facilitate reading in a small grain-size script and it might also facilitate reading in a larger grain-size script. However, this does not apply the other way around: awareness of bigger units seems insufficient to facilitate reading in smaller grain-size scripts (Mishra & Stainthorp (2007).

Other terms that are important when considering the characteristics of scripts are granularity and transparency. Wydell and Butterworth (1999) provide a hypothesis on this

15

account, called the hypothesis of granularity and transparency. Granularity is a different term for the grain size of a script, and thus means the size of the linguistic units that a script works with. Transparency is the degree to which a script is transparent in displaying sound. According to Wydell and Butterworth (1999), orthographies can be described by using two orthogonal dimensions. The transparency dimension displays the degree to which the orthography maps to the sound directly. For example, in a language like Turkish each orthographic unit can directly be mapped onto a certain sound, whereas the orthography of English is less transparent with words like 'thorough'. The granularity dimension is connected to the size of the linguistic units of a language (e.g. word, syllable, letter), finer granularity being smaller units of sound, see Figure 1. Wydell and Butterworth (1999) used their hypothesis to indicate that people who learn languages with a low transparency and a fine granularity are more likely to develop forms of dyslexia, because an opaque orthography and smaller units of sound would be more difficult to learn. The same effects of transparency and granularity may be found in second language learning. Asfaha, Kurvers and Kroon (2009) use the same granularity and transparency hypothesis to explain the difference in difficulty between learning one second language or another.
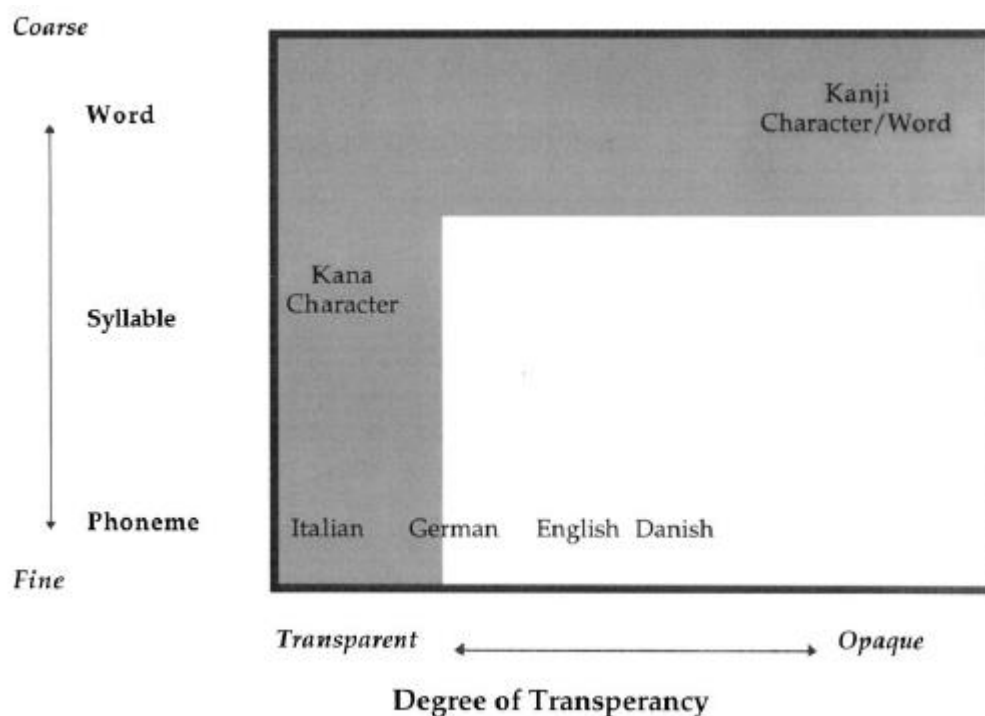
Figure 1: The hypothesis of granularity and transparency. Languages in the grey area are almost 100% transparent. (Wydell & Butterworth, 1999:280)

16

What is also important is to know which factors play a role in learning to read and write in a new language. In the previous chapters about children and adults we have seen the importance of phoneme awareness and the alphabetic principle. Depending on the granularity and transparency of the S1 of the learner he or she is either more or less familiar with these principles. A language with a coarser granularity namely makes less use of smaller units such as the phoneme, and thus phoneme awareness may be less developed. But for people who learn an S2, there are more competences that may affect their learning ability. Asfaha, Beckman, Kurvers and Kroon (2009) investigated this, and they show how reading in a second language is for example strongly influenced by both reading comprehension in the first language (L1) and proficiency in the second language (L2). Furthermore, they did not find a significant effect of L1 script on L2 reading comprehension. The authors could not explain why this would be the case and concluded that more research is needed to investigate the relationship between L1 script and L2 reading performance. That is exactly what this thesis study aims to do, although the focus will only be on writing in the second language.

There are two important hypotheses about the way L1 competences can influence L2 competences. According to the Linguistic Threshold Hypothesis (LTH), a certain level of competence in the second language is necessary to be able to read in that language. The first language reading knowledge that a person has, cannot help them with reading in their second language before some grammatical and linguistic knowledge is achieved (Bernhardt & Kamil, 1995). The same will hold for writing in that second language: this cannot be done before a certain amount of knowledge in the language is acquired. Another hypothesis about this is the Common Underlying Proficiency hypothesis (CUP) or Linguistic Interdependence Hypothesis (LIH), which states that 'reading performance in a second language is largely shared with reading ability in a first language' (Bernhardt & Kamil, 1995:17). What this hypothesis expresses, is that knowledge of reading or writing would not have to be acquired all over again, but the knowledge is instead present and available for any different language. It simply has to be used differently for the L2.

There are few studies that have investigated the process of learning to write in an L2. One example is a study by Edelsky (1982) who studied text writing by first, second, and third graders in a unique classroom situation where the children got all freedom to write in whatever language they preferred. She compared their L1 Spanish texts to their L2 English texts and found that the children generally made much use of the Spanish orthography when spelling in English. Children for example wrote 'chiquen' (for 'chicken') and 'chi lismi siet' (for 'she lets me see it'). She comes to the conclusion that, to write in their L2, children make use of "potentially

17

anything [from their L1] - from directionality of print to spelling hypotheses to general principles and processes." (Edelsky: 1982:225). According to Edelsky (1982), the following factors might influence writing performance in the L2:

1. The nature of the written system of the two languages;
2. The writer's proficiency in the L2;
3. The nature of the literacy experience;
4. The nature of the writing process itself.   Edelsky, 1982:211)

The factor of L2 language proficiency had the most direct influence on complexity of the English texts that the children wrote (Edelsky, 1982). This is something that we will have to be aware of when studying writing performance. Some people are faster learners and will perform better, even though their S1 may be more distinct from their S2. This could simply be because they have a higher level of proficiency in the L2.

There are even fewer studies that have investigated the actual process of learning to write in an S2. Saville-Troike (2012) discusses this shortly in her book on second language acquisition. She writes that "transfer of effective language-specific writing processes that have been acquired in L1 to L2 is not possible until a threshold level of L2 structural knowledge has been reached" (Saville-Troike, 2012:174). This again stresses how important knowledge of the L2 is when starting to write in the language. However, Saville-Troike (2012) further argues that this is not all that is needed; learners benefit also from their general knowledge of content and context that can be transferred from L1 writing skills. She suggests a way of learning that focusses on the content and context, because the linguistic forms of the L1 are already present. From there learners could shift more easily to the L2 graphic forms of expressing these concepts.

In conclusion we can say that there are two sides to being literate in a different language. There is the positive influence of already being literate and having acquired the basic principles of reading and writing and the way a script works, which is a substantial advantage. On the other hand, the first script can have an inhibitory influence, because the literate learner might have different assumptions about scripts and their orthography. It might actually be an advantage for illiterates that they first develop phoneme awareness with the new language, as for them there is no interference of a first script with a different view on phonemes. How exactly this positive or negative influence would function is not clear. Following the hypothesis of granularity and transparency by Wydell and Butterworth (1999) we would expect languages

with a transparent orthography and a coarse granularity to be easy to learn. However, they might complicate later learning of a script with a finer granularity and lower transparency. The granularity and transparency of a script interact with phoneme awareness and the alphabetic principle that we have seen before. It is thus important to discuss the granularity and transparency of Arabic and the Ge'ez languages when investigating these language backgrounds. This subject will be treated in the following chapters. Additionally, Asfaha et al. (2009) already found that the level of reading comprehension and proficiency in the second language has a positive influence on reading abilities in an L2. One would expect the same relation to apply for writing in a second language: L1 writing proficiency and proficiency in the L2 might predict L2 writing proficiency. As Asfaha et al. (2009) already pointed out more research is needed to investigate the relationship between L1 script and L2 performance in reading and writing, because they surprisingly did not find an effect of L1 script. Other factors that may influence writing proficiency in an L2 are the nature of the literacy experience that a person has had and the nature of the writing process itself. This is something that we should also be aware of when comparing the results of different subjects.

2.4 Learning a script in Eritrea

In this chapter the language background of people from Eritrea will be discussed, as they are the first group of people that are the subject of this study. Eritrea is a small country near the coast in the north east of Africa. It has been colonized by several countries, successively Italy, Great Britain, and Ethiopia. Of course, these different reigns have each had their influence on the literacy and the languages that are spoken in Eritrea. Additionally, missionaries have had some influence on language use and literacy. During the Ethiopian colonial rule all Eritrean languages were prohibited in public domains. This resulted in a very open language policy after the Ethiopian colonization (Bereketeab, 2010). Nowadays, each of the country's nine different languages and their three different scripts are accepted for primary education and are regarded of equal status (Asfaha, Kurvers & Kroon, 2008). In Table 1 an overview can be seen of the languages and scripts that are used in Eritrea and the percentages of the population that speak those languages.

Table 1: The languages that are spoken in Eritrea with their share of the population, language family, and script, as published by Asfaha, Kurvers and Kroon (2008:225), based on estimates by Alders and Abbink (2005).

Language	Share	Language family	Script
Afar	6%	Cushitic	Latin
Bidhaawyeet	3%	Cushitic	Latin
Bilen	2%	Cushitic	Latin
Kunama	3%	Nilo-Saharan	Latin
Nara	2%	Nilo-Saharan	Latin
Arabic	1%	Semitic	Arabic
Saho	5%	Cushitic	Latin
Tigre	28%	Semitic	Ge'ez
Tigrinya	50%	Semitic	Ge'ez

As Table 1 shows, the two biggest official languages of the country are Tigrinya and Tigre. Both are written in the Ge'ez script. It is an old script that was originally used to write the Semitic Ge'ez language, an old language that is still used in the liturgy of the Orthodox Church of Eritrea (Asfaha, Kurvers & Kroon, 2009). The Ge'ez script is an alphasyllabary, which means that it makes use of the syllable as the basis for orthographic units but has a way of

writing symbols for independent phonological parts which is alphabetic. It is a therefore a system that shares features of both alphabets and syllabarys (Bright, 2000). An alphasyllabary assigns a different status to consonants than to vowels, as the consonants are the basis on which the syllable is built. The Ge'ez script has a specific system of symbols for combining consonants and vowels. A Ge'ez letter, or *fidel*, is made up of a basic consonant and a diacritic vowel. This means that a small change is applied to the consonant base of the fidel to indicate the vowel. An example of this can be seen in Figure 2.

Figure 2: Ge'ez fidel symbols of consonants and their vowel changes.

The system is very stable: the same fidel always represents the same sound, although there are some graphic variations in the diacritics. The three possible syllables in Tigrinya and Tigre are vowel (V), consonant-vowel (CV) and consonant-vowel-consonant (CVC) combinations. Consonant doubling is not possible, but sometimes a CCV syllable can be found in the Tigre language. This practically only happens with the [t] sound, e.g. in a word like *tgasa*, which means 'he sat down' (Asfaha, Kurvers & Kroon, 2009). Tigrinya has 35 consonants and seven vowels in total and Tigre has 25 consonants and also has seven vowels. 35 consonants that can be combined with seven different vowel results in a total amount of 245 fidels. That means Eritrean children have many more letters to learn than children who learn the Latin alphabet with about 26 letters.

Learning to write in a syllabic language is quite different from learning an alphabetic language. For the Ge'ez script there is a traditional ordering of the consonants in rows, so all children learn them in the same order (Asfaha et al., 2008). The order is established as can be seen in Figure 2, so in school children learn to write down such rows of fidels from the fidel table. Just like in the Netherlands schoolbooks start by showing a picture of a concept combined with the fidel that the concept starts with. Then their knowledge of fidels is slowly expanded by adding more and more fidels and consequently practicing with the fidel table (Asfaha, Kurvers & Kroon, 2009). In Eritrea, chanting after the teacher is still a very common practice

of teaching. After the introduction of fidels children are taught how to blend the syllables into a word. It is important to note that a child does not learn for example the [k] sound as a separate letter, but rather it is introduced with the syllable *ka*. And with the introduction of *ka* the row of syllables that can be made with that consonant is immediately introduced (e.g. *ka, ku, ki, ke, ko*). The practice of syllable blending is again followed by learning short sentences (Asfaha et al., 2008). The perspective on the teaching of spelling in Eritrea could be determined as the generalization perspective (see chapter 2.1: Johnston, 2000), because it is assumed that the language has a clear structure that can be learned. Still, the Eritrean way of learning to spell is slightly different from the way we acquire spelling in Western countries, as children really learn to study rows of syllables by heart. However, this is exactly the learning practice that the big amount of fidels and the regularity within the fidels demand and it seems to be effective.

When becoming literate, all learners across the world acquire some system for the correspondence of sound and orthographic representation. The Ge'ez script is a larger grain size script, because it uses the syllable as its basis. As we have seen in the previous chapter the grain size of a script really matters for someone's view on written language. The question is how practical and effective a system with syllables is. There are studies that suggest that syllables are more accessible than phonemes. This was for example investigated by Asfaha, Kurvers and Kroon (2009). They conducted a study in which they investigated the effect of grain size on the degree of difficulty to learn a certain script. They tested 385 grade 1 children in Eritrea, using self-developed tests on letter knowledge, word reading, and spelling. What they found was that the larger grain size scripts were easier to learn indeed, even though children need to learn a larger number of basic units. Although children who learned a Latin script had better knowledge of letters, this did not have a positive influence on their results in spelling and reading. Children who used the Ge'ez script were better at the spelling and word reading task, and they also appeared faster at learning to read and spell (Asfaha, Kurvers and Kroon, 2009). Apparently, the larger grain size script is easier to acquire than the Latin one. This is in accordance with the observations of Ziegler and Goswami (2005) that large units are the most accessible units. Ziegler and Goswami (2005) argue that three factors play a role in the way children acquire writing skills, namely availability, consistency and granularity. Considering consistency, the Ge'ez script languages also score high. Even the pattern of the vowel parts of the fidels is very consistent. Availability is more difficult to determine, as it involves the conscience and the accessibility of phonological units. It seems that there is less need for these to be accessible in order to start writing in the Ge'ez script.

Because the Ge'ez script is alphasyllabic one might think that learners do not become familiar with the level of the phoneme at all. This is not true however, because as can be seen in Figure 2 each syllable unit consists of a consonant and a vowel part. Therefore, to decode these, children who learn an alphasyllabary need to have phonemic awareness (Reddy & Koda, 2013). Still, the awareness does take longer to emerge, because the unit of the phoneme is less directly present than in an alphabetic script (Nag, 2007). Since both consonants and vowels are part of a single unit they are less perceived as separate units. Because Eritrean students acquire the fidels by learning them by heart, it may be so that they do not really think about their smaller parts. This is logical, because once you know all fidels by heart there is no need to think about the separate signs that they are made of. This may be an inhibitory factor in acquisition of a second language like Dutch, because phoneme awareness is very important for writing in a language that makes use of smaller linguistic units.

The larger grain size of a script such as Ge'ez indeed appears to have its negative side in second language acquisition. Asfaha, Beckman, Kurvers and Kroon (2009) investigated the way Eritrean L1 reading ability affects proficiency in English. English is important in Eritrea, because from the sixth year of school all education is in English. Missionaries first brought English to Eritrea, and later it gained more influence during the British colonial reign (Asfaha et al., 2009). Asfaha et al. (2009) discuss a national survey held in 2002 that is examined by Walter and Davis (2005). The survey revealed that Eritrean children who had acquired the Ge'ez script as their S1 had considerable difficulty with writing and reading in the English language. Walter and Davis mention several explanations for this. First, little time is set aside for teaching English. Secondly, teachers have low levels of proficiency in English. Thirdly, other than in education English is hardly used, except for international communication (Walter & Davis, 2005). Furthermore, Asfaha et al. (2009) also mention a study by Wright (2002) that explains how English is taught using a teaching method that is different from the traditional way of teaching that both teachers and students are used to. This complicates both teaching and learning. Anyway, Eritrean students appear to have difficulty in learning English as a second language and with a second script.

In conclusion, the Ethiopian Ge'ez script that is used in Eritrea is considerably different from the Latin script. The basis is the bigger unit of the syllable. This has its advantages, because children appear to learn the syllable system quicker. However, as the study by Asfaha et al. (2009) shows, Eritrean children later have trouble learning a system that is different from their own and has a smaller grain size. This may have something to do with differently developed phonological awareness. It is not the case that Eritrean students are not at all aware

of phonology smaller than the syllable, because each of their syllable fidels exists of consonant and vowel combination signs. Still, it seems that they are less aware of the smaller phonological units, because in fact there is no real need for that.

2.5 Learning a script in Syria

The second group of people in this study are those from Syria. Syria is located on the map in western Asia, just below Turkey. The official language of Syria is Arabic. In the 19th century the country gained independence after the Ottoman and French reigns. After this independence all Arab states endorsed an Arabization policy, which meant implementation of the Arabic language, followed by more Arabic culture and Arab and Islamic identity. To be Arab is to speak Arab was the idea, and many influential non-Muslim intellectuals supported the implementation of the new policy (Miller, 2003). They thought it might help to overcome religious differences. A common language would create more unity and a common identity. Fortunately, in Syria implementation of the new policy was not as problematic as in some other countries, because many already spoke Arabic (Miller, 2003).

Arabic, like Tigrinya and Tigre, is a Semitic language (see also: Versteegh, 2001). Most Semitic languages are so-called *ajabads*, which means that the consonants are the main carriers of meaning. Therefore, often some or all the vowels are omitted in the written form. Instead, in Arabic small scrawls on top or below the consonants are used to indicate the short vowels. This is different from the Ge'ez languages, which are in fact *abugidas* (see Daniels, 1996; Bright, 2000). That means they have obligatory diacritic vowel marks on the consonants, while in an *ajabad* use of vowel marks is optional and depending on the context (Bright, 2000). Arabic words are built from roots of consonants, and many of these are very old and can also be found in Hebrew. Most are trilateral verbal roots with three consonants (Bachra, 2001; Versteegh, 2001). So, the way this works is as follows: For example, k-t-b is the root of the verb 'write'. McCarthy (2004:427) gives the following examples (1-7) of the morphologically related forms that share this root:

1. kataba 'he wrote'
2. kattaba 'het caused to write'
3. aataba 'he corresponded'
4. kitaabun 'book'
5. kuttaabun 'Koran school'
6. kitaabatun 'act of writing'
7. maktabun 'office'

These relations can be very effective when trying to understand a text, but they can also be confusing when vowels are omitted, because it often results in ambiguity.

Arabic has 28 consonant phonemes (see e.g. Bachra, 2001, Abu-Rabia, 1998). Those are the basis of the script. They can be seen in Figure 3.

ي	و	هـ	ن	م	ل	ك	ق	ف	غ	ع	ظ	ط	ض	ص	ش	س	ز	ر	ذ	د	خ	ح	ج	ث	ت	ب	ا
y	w	h	n	m	l	k	q	f	gh	'	ẓ	ṭ	ḍ	ṣ	sh	s	z	r	dh	d	kh	ḥ	j	th	t	b	ā

Figure 3: The Arabic alphabet

Then there are the long vowels that also have their own symbol. There are five long vowels with three corresponding symbols: [a:] (ا), [o:] and [y:] (و), and [e:] and [i:] (ى) (Abu-Rabia, 1998). Short vowels are not part of the alphabet. For writing the shorter vowels, the system of adding small scribbles on top and below the consonants is used. The four most frequent symbols are the *fatha* (َ)[1] for [ɑ], *kasra* (ِ)for [ɪ], *damma* (ُ) for [ɔ] or [u], and *sukon* (ْ) to indicate that a certain consonant is not followed by a vowel. *Sukon* may also be used where we would pronounce a schwa sound. The complex thing about Arabic is that words are pronounced differently based on their position in a sentence, because their function changes. The intended pronunciation can be derived from the position of a word in the sentence, and the diacritic vowel marks help to determine the function (Abu-Rabia, 1998). However, in most modern texts the vowel signs are not displayed, and the correct reading must be inferred from context and prior knowledge (Abu-Rabia, 1998).

Arabic has a very specific syllabic structure, that can be summarized in this way: CVCC. The onset in principle consists of one consonant (C), which can be any consonant of the alphabet. Then the rime is made up of a maximum of three segments. The consonant is followed by a vowel. This vowel could optionally be preceded by a so-called semivowel, which is like a vowel except it functions as the boundary of a syllable. Examples of Arabic semivowels are [j] and [ʊ]. The vowel can be followed by one or two consonants, that can both also be any consonant of the alphabet. This leaves the official syllable inventory of Arabic to be CVCC, CVC and CV (Kenstowicz, 1986).

Some scholars consider Arabic a case of *diglossia*, which is defined as "a stable linguistic state which includes different spoken dialects and a totally different literary language version, which is usually grammatically complicated, as distinct from the different spoken

---

[1] For all examples of symbols: the dotted circle that is put with the small symbol indicates the consonant that would be below or above it.

dialects, and includes a respectable written literature" (Abu-Rabia, 2000:148). Although most people in Syria grow up with Arabic as their first language, there is a linguistic gap between the mother tongue of many Syrians and the Arabic that is used officially. The literary version of Arabic is in fact only learnt in school and used formally. It is not acquired as the mother tongue in a natural situation. Also, pre-school children hardly get exposed to literary Arabic, because it is generally believed to be too difficult for them (Abu-Rabia, 2000). It could therefore be said that literary acquisition of Arabic happens in a second language situation; the Modern Standard Arabic (MSA) that children learn in school is so different from their own dialect (known as colloquial Arabic) that it can be compared to learning a second language. Because this is the case, education of Arabic has very much overlap with teaching syntax and grammar. About the learning situation in school, several other differences can be noted. In Arabic schools, oral reading is a very important part of the teaching process. Reading is viewed as a holistic process, because the reader needs to guess vowels, disambiguate words that have an identical form and understand the inflections at the end of a word, all to determine its meaning and function in a sentence (Hussien, 2014:89; Abu Rabia, 1998:107). It is extra demanding to read in Arabic when the diacritic symbols are not shown, and reading becomes an even more holistic process. Considering second language acquisition in Syrian schools, nowadays children are taught Modern Standard Arabic in primary school and get introduced with foreign languages late in primary school or else in secondary school. Some of the university studies are taught in English or French, but for most Modern Standard Arabic is the main language of education. Therefore, it is possible to grow up in Syria and not have the need to write in the Latin script in daily life because Arabic is used everywhere.

There are quite a few other characteristics that distinguish writing Arabic from writing Dutch. For example, Arabic letters can take different shapes according to their position in a word. There can be different forms for the beginning of the word, the middle position of the word and the end of the word (Abu-Rabia, 1998). Also, Arabic has a lot of homographs, especially when the diacritic vowel marks are not written out. Several words might have the same orthographic representation but need to be pronounced differently and carry distant meanings (Abu-Rabia, 1998). Furthermore, in principle Arabic has a one-to-one correspondence between the used symbols and their sounds (Hussien, 2014). This is not the case for Dutch, where for example the letter 'e' can be represent [e], [e:], [ə] and [ɛ]. So, at first sight, Arabic could be identified as a small grain size script with a transparent orthography. However, the optional display of diacritic symbols in Arabic makes the difference whether the orthography of the language is transparent. Arabic is quite transparent when the diacritics are

27

shown, but rather opaque if these small symbols are not added to the letters. In daily life and writing in dialects, these extra symbols are usually omitted. Only in official Modern Standard Arabic the diacritic symbols are always written out.

Another important characteristic to note is that most languages are written from left to right, but Arabic is written from right to left. This is just a matter of which agreement was made in the past: one of the systems cannot be declared more 'natural' than the other (Gibson & Levin, 1976:170). However, studies do report effects of directional reading habits. Participants are better at recognizing words that are written on the side that the script starts from, which can be explained by the focus they have habitually (Gibson & Levin, 1976). Therefore, this preference for the right side of people with Arabic as their S1 could impede the learning process of a left to right script to some extent.

In conclusion we can say that the Arabic system is very different from the Dutch system, and also very different from the Eritrean system that has been discussed in the previous chapter. It can be identified as a small grain size script with a rather transparent orthography. At first sight this is an advantage for language learning. However, depending on the type of Arabic that is written, the short vowels are either displayed or omitted. With omission of vowels, the orthography of Arabic becomes opaquer. Because the display of vowels is only mandatory in Modern Standard Arabic and Modern Standard Arabic is only used in formal situations, most mother tongue speakers of Arabic are very much used to omitting the short vowels and thus have a written language with an opaque orthography and few units of sound. Homographs make understanding Arabic extra complicated. Context is therefore very important and acquiring Arabic is seen as a holistic process, which is very different from acquiring Dutch.

2.6 Conclusion

We have seen in this second chapter that writing is a relatively little studied topic, since most studies focus on reading. Writing is very important however, because it is an important requirement to be able to function well in our society. The problem with acquisition of Dutch that was signaled by ISK schools arose more in writing than in reading, and therefore this study focusses on writing development. The question in the field of writing that this thesis aims to answer is: How do the first scripts of Arabic and Tigrinya influence writing in Dutch as a second language with the Latin script? These two scripts are in focus, as they are the first scripts of many newcomers in the Netherlands who are learning Dutch as a second language. The aim is to find out how the knowledge of a first script that is an *alphasyllabary* or *ajabad* can facilitate or inhibit writing acquisition in Dutch.

This literature section was started from the way children first acquire the skills of learning to read and write. Phonological and phonemic awareness appeared to be important factors in acquisition, phonemic awareness being something that must be taught explicitly. That means that when one's first language does not use separate symbols for small phonemic units, phonemic awareness is not developed to the same level. We have discussed studies that show how children who learn scripts that work with larger units do develop phonemic awareness, but this acquisition happens in a later stadium and their awareness of phonemic units is less explicit. Therefore, learners of a script with bigger units, like Ge'ez, may experience difficulty when learning a second language that makes use of smaller phonemic units in writing. Their performance in writing will be very different from that of illiterates. Without previous experience in writing, there is nothing that could influence their writing products but the language that they know and its sounds. As we have seen in chapter 2.2, many new adaptations need to be made in their brains, because literacy entails a whole way of thinking about the world that has never been assessed. Second language learners that are literate in a different script therefore have a great advantage. However, their phoneme awareness is already developed in a different way, where illiterates can develop it with the new script. The first language competences could also inhibit learning to spell in a different script.

It is essential to determine how we think this influence of first language competence on second language competence works. According to the Linguistic Threshold Hypothesis that has been discussed, a certain amount of grammatical and linguistic competence in the second language is requisite to be able to read and write in that language. Before that, their first language reading knowledge cannot help them. When this requirement is met, the literacy

29

knowledge from the first language can become effective. Following the Linguistic Interdependence Hypothesis, we believe that Eritrean and Syrian learners do not need to acquire an entirely new competence when learning to write Dutch with the Latin alphabet. Rather, it is assumed that they can use the knowledge that they already have from their first scripts. However, there are some differences between their S1 and S2 that may cause difficulty.

We have furthermore seen that the Ge'ez script and the Arabic script both have a different granularity than Dutch, with overall bigger (Ge'ez) or fewer (Arabic) units of sound that are expressed in their written form. From the section about illiterates it became clear that in case this knowledge of smaller units is not present, new functions need to be created in the brain to process such smaller units of sound. That is something both learners with an Eritrean background and a Syrian background must go through when learning a second language such as Dutch. Based on their background, we could expect the Eritrean learners to need more time for learning, because their script has a coarse granularity and higher transparency. As research has shown, a bigger unit script does not facilitate reading in a script with smaller units. The Syrian learners have the advantage that their script has a slightly finer granularity, with separate characters for some of the vowels, and separate signs for the short vowels that are not part of the bigger unit of the syllable. Thus, the expectation is that Eritrean students who have been learning the Latin script of Dutch for a longer period of time will generally make more mistakes than the Syrian students who have been learning the Latin script for a longer period of time.

Considering the type of mistakes that Syrians and Eritreans make in learning Dutch, we may expect difficulty with vowels for both groups. The Eritreans are not familiar with using an independent character for a vowel. The Syrians are more familiar with characters for vowels, but still their system does not use separate characters for all the phonemes that we distinguish. The alphabetic principle of Dutch thus forms a challenge for both groups. Because Arabic does have separate characters for vowels, we may in general expect more difficulty with the vowels for the Eritrean students. However, more omission of short vowels is expected in the performance of Syrian students, as they are used to omitting these in daily life.

Because of their different language backgrounds, there may be a difference in phoneme awareness between Eritreans and Syrians. We have seen that Eritreans have the fidels as the smallest unit. Never do they explicitly learn that fidels have a consonant part and a vowel part. In education of Arabic however, all letters and diacritic symbols are separately acquired. This results in explicit phoneme awareness. It might be so that people with an Eritrean background have less explicit awareness of phonemes, because they were only explicitly taught bigger units

of sound. As explained before, this awareness always emerges to some extent, and is very important because it is closely related to writing performance.

Now, if we wish to investigate the influence of the first scripts of Ge'ez and Arabic we will need a test that discriminates between the type of mistakes that people with different language backgrounds make. To find out whether their mistakes can really be explained by their S1, performance of Eritrean and Syrian learners must be compared. The influence of characteristics of the L1 should also be considered in the investigation. If we can find out how strong the influence of characteristics of the S1 on assumptions about spelling in an S2 is, this will help us to understand the difficulty that second language learners of Dutch experience. Finding out more about the relationship between S1 and S2 may give insights in the aspects of the structure of the Dutch language that should get focus in second language education.

3. Methods

3.1 Start-up

Before creating a test, it was necessary to become familiar with the type of education that participants were receiving and with their level of Dutch. Therefore, I visited the Pontem College, one of the ISK schools in Nijmegen. ISK stands for 'Internationale SchakelKlas' which translates to 'international linking class'. In ISK schools, newcomers receive education in Dutch. The purpose of these schools is to form the bridge between the latest education that pupils have received and their future education in the Netherlands. At ISK schools they are prepared for functioning in regular education or work. I observed at two hours of class of two different groups of pupils. The first group was in the alphabetization class (AA) that had started after summer holidays. The second group had started alphabetization in the previous year (R1). During the first lesson that I observed the students had their very first dictation test. With some help of the teacher they were to write down words that she read out. All the words that they heard were words they had seen and practiced before. In the second hour of class I observed how the students received a test where they had to translate sentences to English by putting words in the correct order. Since this class was in English it was slightly less informative for me, but it still gave some idea of the level at which the students were functioning and the type of education that they are used to.

3.2 Participants

The participants in this study were 16 students from the alphabetization class of the Pontem College. 24 students were tested in total, but only the test results from people with Eritrean or Syrian background were selected for analysis. The average age of the participants was 16,5 years old; the oldest participant was 19 years old and the youngest participant was 13 years old. There were 9 male and 7 female participants. Two different groups of students were tested. The first group had started education in writing Dutch approximately half a year ago. Some of them had started the lessons only three months ago, others had attended classes for almost a year at the moment of testing. The second group had already been learning to write Dutch in the previous school year, so many of them had had almost 2 years of education in writing Dutch. 5 students from the first group had Eritrean backgrounds, 5 students had Syrian backgrounds.

From the second group, 4 students had Eritrean backgrounds, 2 students had Syrian backgrounds.

The participants that had received the least education in writing Dutch were an important point of focus for this study. At the moment of testing, these participants had already acquired the alphabetic principle, so they had learned that letters represent certain sounds. Their Dutch vocabulary however was rather small, as Dutch was their second language and they were still learning it. This is important to acknowledge, because as we have seen in the literature section children who learn how to write really make use of their word knowledge. Second language learners do the same, but they must rely on a smaller vocabulary. Therefore, knowledge of the alphabetic principle and phonemic principle is extra important for them. The participants were still in the process of acquiring phonemic awareness in Dutch, the awareness of the smallest units of sound in a language that can change the meaning. Recall that phonemic awareness is very important, because often a letter represents a phoneme and thus this awareness really facilitates the process of learning to write. As second language learners they were not yet used to some of the sounds of the Dutch language. Therefore, their ideas about which letters represent certain phonemes could be wrong. As discussed in the previous chapter, there might be a difference in the level of phonemic awareness in Dutch based on their language backgrounds. This difference is expected to emerge from their writing products. The participants that had received education for a longer period were tested to see how the learning process develops over time.

3.3 Test design

In this study the aim was to distinguish between specific characteristics of the scripts that we are interested in. Previous research shows us how errors in spelling can be informative about those aspects of a new language that are most difficult for the learners. Protopapas, Fakou, Drakopoulou, Skaloumbakes and Mouzaki (2013) for example, state this clearly when they say that "spelling errors can be informative about the aspects of the orthographic system that pose the greatest difficulties for developing spellers" (Protopapas et al., 2013:640). The choice was made to create a specified test that could elicit word spelling of those characteristics that were relevant for the target group of this study. Naturally, there were completed writing tests of the students available at school. However, these often included words that they had been practicing during the lessons. By creating a specified test, we could examine exactly those characteristics of the Latin script that were expected to cause difficulty for the Eritrean and Syrian subjects.

Studying the difficult aspects would enable us to create a clearer overview of the problems that occur when acquiring Dutch, and the possible relationship between those problems and the S1 of the participants.

3.3.1 S1 characteristics

In this part the characteristics of the scripts of Ge'ez and Arabic that were considered relevant for setting up the test will be discussed. This will help to understand the basis of the test. The first point of focus of the test had to be the phonemes. Because although the Arabic language and Tigrinya and Tigre do have spoken phonemes, the biggest difference between the S1 and S2 lies in the way these are represented in the writing system. Thus, to answer our first sub-question of what exactly the effect of the first scripts is, phonemes and more specifically vowels were important characteristics to be tested. As mentioned earlier, the two Eritrean scripts that were studied only make use of fidels that are combinations of vowels and consonants. Therefore, it could be expected that Eritreans would have difficulty with writing vowels with a separate symbol. Considering vowels in Tigrinya and Tigre, these Eritrean scripts have the following sounds that can be part of a fidel: [a], [u:], [i:], [a:], [e:], [ə], [o:]. Although Eritrean learners are familiar with these sounds, they might still form a problem when written with a separate vowel symbol, and thus had to be included in the test. The Syrian learners have an advantage with vowels, because their S1 does have separate symbols for some vowels. However, in Arabic only the long vowels ([a:], [e:], [i:], [o:], [u:]) are written with a separate symbol, so it might be the case that they have more difficulty with separate symbols for short vowels. For the Syrian participants, difficulty was mainly expected with words that end with an unstressed syllable that contains a schwa ([ə]), such as 'tafel' (table). This is because in their own language there would not be a symbol for the 'e' that we hear in such contexts. Instead, the *sukon* diacritic would be used, that indicates the absence of a vowel.

Another important aspect of the two scripts are the orders that are possible in phonetics. As mentioned before, the Eritrean languages mostly have CV patterns and only some CVC patterns, because fidels are made up of a C and V combination. Possibly, Eritrean learners perform better with words that conform to that pattern. Therefore, in the test a difference had to be made between words with CV syllables (open syllables) and words with CVC syllables (closed syllables). This distinction between open and closed syllables is also relevant in a different way for those with a background in Arabic. Their first language has more frequent occurrences of CVC and even has syllables that end with double consonants. Only, the vowel is often omitted in the written version of the language. Furthermore, Arabic has many

possibilities in combining consonants that cannot be put together within a syllable in Dutch. With the same combination of sounds Dutch has a schwa ([ə]) sound that we write down as e, where there would not be a character in Arabic. It may be possible that the omission of schwa sounds is influenced by stress. It can be imagined that the schwa sound would be written down in contexts where it is part of a stressed syllable, because in that case the sound gets more focus. In unstressed syllables the schwa sound is less present and might be more likely to get omitted, where people with the Ge'ez script as their background would write it down because they are more likely to expect a vowel in that position. Therefore, the aspect of stress also had to be included in the test set-up.

Additionally, it is known that diphthongs form a problem for all learners of Dutch. For this research it was assumed common knowledge that all people with a first language that does not distinguish the sounds of [ɛi], [œy], [ɑu] (respectively 'ei', 'ui', 'ou') et cetera have trouble acquiring their written form. Because this would not be a problem that is specifically related to Syrian or Eritrean backgrounds, the choice was made to leave those sounds out of the test. Instead, it was determined that the test should focus on double vowels. The sounds of double vowels such as 'aa' can be found in both the Eritrean languages and in Arabic, but they do not have the same notation as we have in Dutch. In Dutch namely, the sound of the long a ([a:]) can either be written as a single letter 'a' or a double letter 'a', depending on context. The question is if this causes difficulty that can be explained by the S1. Because Syrian learners are used to characters for the long sound vowels, long sound vowels are less likely to be omitted by them. Eritreans are also not likely to omit these vowels, because the long sound vowels can be part of a Ge'ez fidel. Therefore, in writing long vowels the Eritrean and Syrian participants were expected to have the same advantage and disadvantage. Both Eritrean and Syrian participants are namely not used to having a notation with two letters in such contexts. By including the difference between double and single vowels with long sounds, the general difference in performance of the Eritrean and Syrian participants could be investigated. Also, the long sound single vowels could be compared with the short sound single vowels, to see whether these are indeed less likely to be omitted.

Another important aspect of spelling in Dutch are the consonant clusters. Consonant clusters may form more of a problem for the Eritrean participants than for the Syrian participants. This is so because the Ge'ez script in principle does not have consonant clusters, while Arabic does have consonant clusters, be it that a separate scribble is often used to indicate that this is the case. If the S1 indeed has a direct influence, there would have to be a clear difference in performance on consonant clusters between Syrian and Eritrean participants. The

expectation was that Eritreans would have considerably more difficulty with consonant clusters. Thus, in the test, the number of consonants at syllable endings had to play a role, and also the number of consonants in the first following syllable. The latter choice was made, because there is a more complex cluster of consonants when the next syllable also starts with a consonant. A schematic overview of the expectations of difficulty for each aspect per background can be seen in Table 2 below.

Table 2: Characteristics and estimation of expected difficulty per language background.

Characteristic	Syrian	Eritrean
Single vowel (V) short sound	-	-
Single vowel (V) long sound	+	-
Double vowel (VV) long sound	-	-
Vowel in open syllable (CV)	-	-
Vowel in closed syllable (CVC)	-	- -
Stressed vowel	-	-
Unstressed vowel	- -	-
Syllable ending in C	+	-
Syllable ending in CC	+	- -

3.3.2 Test structure and set-up

The choice was made to structure the test as a dictation. By dictating words for participants to write down, it was possible to test their ability to directly transfer the sounds that they heard to written letters. This way the functioning of the phonemic principle was tested. The fact that someone can write down an overheard word namely shows that this person is able to process a sound signal and transfer it into the separate phonemes and again to the spelling of the language.

The test was made by creating three main categories, that tested three important characteristics of spelling in Dutch. These categories were as follows:

1. vowels
2. long sound vowels (single and double)
3. consonant endings

Within these categories, sub categories were needed. For testing the vowels, we had to distinguish between three aspects: 1) the syllables (monosyllabic and bisyllabic words, first or second syllable), 2) stress and 3) open or closed syllable. For the double vowels we had to discriminate between 1) single or double vowels, 2) syllables, 3) stress and 4) open or closed syllable. For the consonant endings 1) the number of consonants and 2) the number of consonants in the first following syllable were considered relevant. Like this, several subcategories emerge, as can be seen in the matrices that are included in Figures 4 to 6. For

complete matrices with all words per category that were included in the test, see Appendix I. The matrices include a schematic overview of the subcategories that were created per category and give an idea of the way each aspect was graded. This will also be explained in more detail after the matrices.

Category 1: Vowels				Item Count
Rating on two aspects	1. vowel written (1) / not written (0) 2. correct vowel (1) / incorrect vowel (0) a double vowel is graded as an incorrect vowel ➔ for each category, only the mentioned syllable gets graded. e.g. category 'first syllable stressed open', only the first syllable is graded on these two aspects			Total number of items in this category: 63 (126 points possible) Per subcategory min. 6 high frequency words. If no 6 high freq available, add up with low freq. Strive for 10 words, max. 10 per category. (high freq. = frequency in CELEX>100 per 1,000,000 words). (less freq. = frequency in CELEX>30 and <60 per 1,000,000).
Item restrictions	- no double vowels, only a, e, i, o, u - the graded syllables are only syllables that have a CV (open) or CVC (closed) pattern - try to keep ungraded syllables simple (try to avoid diphthongs) - no verbs - no proper names, place names, etc.			
Subcategories	**Syllables**	**Stress**	**Syllable type**	**Count**
	Monosyllabic words	Stressed	Open	- (removed for lack of relevance and to shorten the test)
			Closed	10 items
	Bisyllabic words	1st syllable stressed	Open	10 items
			Closed	10 items
		1st syllable unstressed	Open	9 items
			Closed	6 items
		2nd syllable stressed	Open	- (removed for lack of data)
			Closed	6 items
		2nd syllable unstressed	Open	6 items
			Closed	6 items

Figure 4: Test matrix of the items that test category 1: vowels

Category 2: Long sound vowels					Item Count
Rating on two aspects		1. correct number of vowels (1) / incorrect number of vowels (0) 2. correct vowel(s) (1) / incorrect vowel(s) (0) a double vowel where a single vowel is expected gets graded as an incorrect vowel and vice-versa. For each subcategory, only the mentioned syllable gets graded. (e.g. for the category 'single vowel 1st syllable stressed open', only the vowel in the first syllable of the word is graded on the two aspects.)			Total number of items in this category: 54 (108 points possible) Per subcategory min. 6 high frequency words. If no 6 high freq available, add up with low freq. Strive for 10 words, max. 10 per category (high freq. = frequency in CELEX>100 per 1,000,000 words). (less freq. = frequency in CELEX>30 and <60 per 1,000,000).
Item restrictions		- the graded syllables are only syllables that have a CV or CVV (open) or CVC or CVVC (closed) pattern. Try to avoid double consonants. - try to vary vowels as much as possible - try to keep ungraded syllables simple - no verbs - no proper names, place names, etc.			
Subcategories	**Number of vowels**	**Syllables**	**Stress**	**Syllable type**	**Count**
	Single vowel (e.g. a)	Monosyllabic	stressed	Open	- (removed for lack of relevance and to shorten the test)
				Closed	- (always double vowel)
		Bisyllabic	1st syllable stressed	Open	10
				Closed	- (always double vowel)
			1st syllable unstressed	Open	6
				Closed	- (always double vowel)
			2nd syllable stressed	Open	- (removed for lack of data)
				Closed	- (always double vowel)
			2nd syllable unstressed	Open	- (removed for lack of data)
				Closed	(always double vowel)
	Double vowel (e.g. aa)	Monosyllabic	stressed	Open	3
				Closed	9
		Bisyllabic	1st syllable stressed	Open	- (only in compositions)
				Closed	7

			1st syllable unstressed	Open	-
				Closed	3
			2nd syllable stressed	Open	- (removed for lack of data)
				Closed	10
			2nd syllable unstressed	Open	-
				Closed	6

Figure 5: Test matrix of the items that test category 2: double vowels

Category 3: End consonants				Item Count
Rating on two aspects	1. correct number of end-consonants (1) / incorrect number of end-consonants (1) 2. correct end-consonants (1) / incorrect end-consonants (0)			Total number of items in this category: 60 (120 points possible) Per subcategory min. 6 high frequency words. If no 6 high freq available, add up with low freq. Strive for 10 words, max. 15 per category (high freq. = frequency in CELEX>100 per 1,000,000 words). Words in blue are less frequent words (less freq. = frequency in CELEX>30 and <60 per 1,000,000).
Item restrictions	- try to vary the consonants as much as possible - no diphthongs as graded consonants - try to keep ungraded syllables simple, try to avoid diphthongs - no verbs - no proper names, place names, etc.			
Subcategories	**Number of consonants**	**Syllables**	**Next syllable starts with consonant?**	**Count**
	Single end consonant	Monosyllabic	-	10 items
		Bisyllabic (first syllable)	No	4 items
			One C	11 items
			Two C	10 items
	Double end consonant	Monosyllabic	-	10 items
		Bisyllabic (first syllable)	No	- (removed for lack of data)
			One C	3 items
			Two C	3 items

Figure 6: Test matrix of the items that test category 3: consonant endings

As can be seen in the matrices, a few restrictions had to be met for each category. For category 1 'vowels' no use of double vowels was allowed, as these were part of category 2. Furthermore, only syllables with CV or CVC pattern were tested. This choice was made, because the test need not be unnecessarily complex. Keeping just these two options enables us to distinguish the relevant difference between CV and the CVC syllable that is less frequent in the Eritrean script and more frequent in the Syrian script. It was attempted to keep ungraded syllables as simple as could be and to avoid diphthongs. That way, the graded syllables could be kept as constant as possible, and potential interference of other aspects of the words could be reduced. Besides, no proper names and place names were included, because these do not always follow the same pronunciation rules and there might be interference of specific background knowledge. Also, no verbs were included, because this might activate grammatical awareness, which is not what we intended to test here. The words of category 1 were rated on two aspects: the fact whether a vowel was present or not and the correctness of the written vowel. For each aspect 1 point could be achieved. Category 2 'long sound vowels' included only syllables with CV and CVV (open) or CVC and CVVC (closed) patterns for grading. The words had the same restriction regarding avoidance of verbs and place names as category 1. Here also it was attempted to keep the ungraded syllables uncomplicated. Furthermore, use of double consonants was reduced as much as possible, because this might further complicate the words. This was not always possible however, because we were working with a limited amount of words from a database, as will be explained later. For category 3 variation of consonants was one of the restrictions. There might namely be a difference between consonants that are more familiar to the participants and consonants that are less familiar to them. This is true for both familiarity in the S1 and in the S2 experience. Here also, ungraded syllables had to be kept uncomplicated, and verbs and proper names must be avoided.

All words that were included in the test were extracted from the WebCelex database (WebCelex, 1995). This is a database that is part of the Max Planck Institute of Psycholinguistics. It has a rich variety of Dutch wordforms that are stored online, with specific information per word. Stored in the database are words from the Dutch Van Dale Dictionary (Geerts & Heestermans, 1984), words from the list of the Dutch language 'Het Groene Boekje' (English: 'Word list of the Dutch language') (De Vries & Te Winkel, 1954) and the most frequent words from the text corpus of the Institute for Dutch Lexicology (INL) (SoNaR, 2013). The INL corpus includes 930 entire fiction and non-fiction books and based on the contents of these books frequency estimations were created. It is therefore possible to search the CELEX database for words with a certain frequency. The words in the test had to be quite frequent

words, because for learners of the Dutch language it is relevant to be able to write down those words that occur in daily conversations. If learners would have problems with writing the more complex less frequent words, this would be no more than logical. It would not tell us anything important for their immediate benefit as language learners. The frequency of the words that were used in the test was determined at an INL frequency of a minimum of 100 per 1,000,000 words. Furthermore, CELEX enables its users to determine the number of syllables of the words they look for. This was set to a maximum of three syllables, because originally the idea was to use words of three syllables as well. The lexicon that was created was set to contain a maximum of a thousand words. Because we needed to distinguish between syllables with stress and syllables without stress, this function was activated to be included in the outcome table as well. The search restrictions that were used can be seen in Table 3.

Table 3: Search restriction in CELEX for frequent and less frequent words.

	Frequent words	**Less frequent words**
Selected columns	- Dutch wordforms - INL frequency (1,000,000) - PhonStrsCLX - WordSylCnt	- Dutch wordforms - INL frequency (1,000,000) - PhonStrsCLX - WordSylCnt
Used query	`(WordSylCnt < 3) && (INLmln > 100)`	`(WordSylCnt < 3) && (InlMln > 30)`

The first query of frequent words resulted in a basic list of 852 words. From this list words were selected that fit well into the three main categories. There were also words that were placed in more than one category. Words were never in more than one sub-category of the same category. The double classified words thus got graded for two aspects. This was done to reduce the total number of items in the test. After this, words that were highly frequent were removed from categories with many words, so they had no more than 10 items. As can be seen in the matrices, for all categories it was strived for having a minimum of 6 words and a maximum of 10 words in each subcategory. Sometimes, the list of frequent words did not yield enough words for a certain sub category. If no 10 words could be included from the list of frequent words, the category was supplemented with words from the list with less frequent words. Which words these were can be seen in the overview of test words in Appendix I.

After the first selection it became clear that there were some categories that still contained very few words. Little data for a certain aspect means little material to compare other results with. Assumingly, constructions that are very infrequent are also less relevant for an

indication of general performance. Therefore, these subcategories were emptied as can been seen in the matrices. However, some categories with few words were kept. In category 2 the open, stressed, monosyllabic double vowels were left in the test even though there were only three such words in the corpus. This decision was made because there were already few test items for category 2, and even though there were few such words in CELEX, in daily life these words are not very rare. They are words like 'twee' and 'zee'. The three words from the subcategory with closed, unstressed, first syllable bisyllabic words with a double vowel were decided to be removed from the test, but accidentally they were not. In category 3, the few bisyllabic words with a single end consonant that were followed by a syllable that started with a vowel were also kept in the test. These are words like 'daarom' and 'vanaf'. This subcategory was considered too interesting to leave out. If the next syllable namely does not start with a consonant, the end consonant is the only consonant that separates the vowels. However, the consonant is still located at the end of the first syllable. It would be interesting to see if subjects that have difficulty with end consonants performed better on these items. The double end consonants that were followed by a syllable that started with one or two consonants were also left in the test, but for qualitative study. These are namely the more complex words and there were too few to rate them quantitatively, but the data could provide an interesting insight in the performance of individuals.

After this first clearing, still quite many words were left. The test should not be too long, to prevent effects of boredom. Therefore, in the end it was decided that monosyllabic open syllable words (such as 'ga', 'nu', 'zo') would be left out of the test. These would not give much extra information that the bisyllabic words could not provide and would thus only take up extra time in the test. This resulted in a list of 137 words in total, so the full test contained 137 words. The final list of words with their CELEX frequencies is included in Appendix II.

Another important decision was that the test also had to contain pseudo words. In the lessons the students namely learn to write down words starting from concepts and pictures. Many of the frequent words that we tested would probably be already familiar to them. At the moment of testing many of the participants had received education in Dutch for several months. Therefore, their background knowledge of the way these familiar words are written could interfere. Participants might not make the same mistakes they would have made when they just started to acquire Dutch spelling. Creating non-existent words with the same characteristic could provide us with the information we needed. It was the only way to test if the match between a heard sound and a written word had actually been made in their brain (the functioning of the phonemic principle), without interference from background knowledge. The pseudo

words were created by making small changes to the original words in the test. The vowels and consonants that got graded had to remain the same; only those consonants that did not get graded were altered. A few examples are given below in Table 4.

Table 4: Examples of words in the test and the pseudo words that were created from them. (The bold part of the word is the part that got graded.)

Original word	Pseudo word
*elk**aa**r*	ark**aa**r
*an**t**woord*	an**t**wolg
*b**a**sis*	g**a**sis

The full list of pseudo words can be found in Appendix III, in which the full list of test words for both versions is included.

3.3.3 Recording and composition

The final step was the recording of the words and pseudo words. This was done by putting all the words in a list in alphabetical order. The pseudo words were always placed right after the original word that they were based on. That way, the person who read the words would know the intended pronunciation and stress pattern. The words were voiced by an employee of Radboud in'to Languages, the Center of Expertise for Language and Communication in Nijmegen. She was working at the Department of Testing and Assessment and had experience in recording Dutch language tests. The recordings were made in a closed recording booth, so background noises would be filtered out. The program that was used for this was Audacity (version 2.2.2). Audacity is a multi-track audio editor and recorder that is suitable for use with several operating systems. Words for which the pronunciation was wrong or unclear were recorded again, until all words and pseudo words were recorded. After this, two versions of the test were created. The alphabetically ordered list of normal words was used a basis for this. First, the order of this list was randomized. Then, the first half of that randomized list (69 words in total) were the normal words for version 1. The second half of the words on the list were replaced with their connected pseudo word. Then the order was randomized again, and this was the order of words for version 1 of the test. Likewise, the second half of that randomized list (68 words) were the normal words for version 2. For version 2, the first half of the words was replaced with their connected pseudo word and then the order was randomized. Now there were

two lists of words, that consisted for 50% of real words and for 50% of pseudo words. For each version, an audio file was created in which the recorded words were put in the right order. Each word could be heard twice and was preceded by a beep sound. After the second time, each word was followed by a period of silence to write it down. The length of this silence varied: words with one syllable were followed by a 10 seconds silence, words with two syllables had a 12 seconds silence. Very complex words had a silence of 13 seconds. The total length of the audio files was approximately 38 minutes.

3.4 Test taking

The test was given during a normal school day of the AA class and R1 class. First the R1 class was given the test, in the next hour the AA class was given the test. The group was divided in half: always one half of the group stayed in the classroom and made the first version of the test; the other half went to another classroom to make the second version of the test. The teacher made sure that the Eritrean and Syrian participants were divided over the two versions. The audio was played from an MP3 file using the loudspeakers of the classroom. Participants were informed about the length and set-up of the test. They were instructed to listen carefully and to try and write down the sounds that they heard as precisely as they could and that it was better to just try and write something than to leave a question open. Also, when they were not finished, they were instructed to move on to the next word. They were not informed that there were words in the test that were not real. Instead, they were informed that there were strange words in the test that they might not have heard before. The test taker was informed that the audio could be paused a little longer, should the participants have too little time. After the test, participants were awarded with sweets.

4. Results

4.1 Analysis and coding

When the tests were completed by the participants, their answers were analyzed. Unfortunately, not all participants were able to finish the entire test. The R1 group that made version 2 of the test started somewhat later because the reserved room could not be used. For that reason, they could not complete the full test within the hour of the lesson. They reached question 131 of 137. Furthermore, the test turned out to be too long for the AA group of which some had started writing lessons only 3 months ago. Therefore, the AA group that made version 1 of the test only reached question 48. The group that made version 2 reached question 88, but many quit writing earlier. Also, with version 2 of the AA group the teacher accidentally started the wrong version and switched to the correct version after 27 words. Therefore, there were more results for version 1 of the AA group than for version 2.

The analyzation process was as follows. First, each participant was anonymized by naming them AA.1, AA.2, R1.1, R1.2 et cetera. That way their names would afterwards be left out of the data. Then the tests were graded on the relevant aspects that have been discussed in the previous chapter. This was done by creating an Excel sheet with all the words per version listed. Then, subsequently each word was graded on the categories that it was included in. Only that part of the word was graded that was mentioned in the test matrix. Answers were coded with 1 or 0 points or an X for missing information when a word had not been filled in. After all the results had been entered in the Excel sheet, the information about the country of origin of each participant was connected to their answers. By doing this after the grading, bias in the grading process could be prevented as much as possible. Also, for each category the relevant aspects of the word were added in separate columns. For category 1 words were coded on whether the graded syllable had stress and whether it was a closed or open syllable. Likewise, words in category 2 were coded for stress, closed or open syllable and whether they had to be written with a single or a double vowel and words in category 3 were coded for the number of consonants that they contained and the number of consonants of the next following syllable.

4.2 General test results

The results were analyzed by conducting several statistical tests. The dependent variable that was studied was always the amount of points that participants scored for a certain category. The

45

independent variables were class (with two levels: AA and R1), origin (with the levels Eritrean and Syrian), education (the number of months that they received education), version (1 or 2) and word type (with two levels: word and non-word).

First, a one-way MANOVA was conducted over all the participants test results. The MANOVA was used to see if there were general differences between groups for the three categories that were tested. Should this be the case, then further smaller statistic tests such as ANOVA's could be conducted for the data of each category. In this MANOVA, the points that participants scored per category were the dependent variables. It would have been impossible to simply use the scores of participants as the dependent variable, because not all participants had completed the full test. Therefore, the percentage of correct answers was used as the value of measurement instead. We do have to realize that percentages are in themselves a measurement that is also very sensitive to size. The group that participants belonged to was the independent variable that was measured. The variable group in this analysis had 4 levels: 1) AA Eritrean, 2) AA Syrian, 3) R1 Eritrean and 4) R1 Syrian. The percentages per participant per group that were used for this analysis can be seen in Table 5.

Table 5: The scores of each participant per group (1: AA Eritrean, 2: AA Syrian, 3: R1 Eritrean, 4: R1 Syrian) with their writing education displayed in months and scores displayed with percentages per category.

Group	Participant	Education	C1_written	C1_correct	C2_number	C2_correct	C3_number	C3_correct
1	AA.1	3	78,95%	42,00%	43,48%	21,74%	55,00%	45,00%
1	AA.2	3	73,68%	36,84%	13,04%	56,52%	55,00%	45,00%
1	AA.3	3	94,74%	42,11%	34,78%	39,13%	55,00%	65,00%
1	AA.4	3	94,74%	36,84%	21,74%	39,13%	50,00%	60,00%
1	AA.7	9	56,67%	26,67%	23,33%	33,33%	57,69%	24,04%
2	AA.5	12	73,68%	43,37%	47,83%	82,61%	85,00%	85,00%
2	AA.6	4	89,47%	52,63%	56,52%	95,65%	90,00%	85,00%
2	AA.8	9	46,67%	30,00%	33,33%	53,33%	53,85%	50,00%
2	AA.9	7	80,00%	50,00%	40,00%	86,67%	88,46%	84,62%
2	AA.10	3	40,00%	16,67%	10,00%	26,67%	46,15%	38,46%
3	R1.1	24	80,65%	38,71%	33,93%	85,71%	76,31%	71,15%
3	R1.3	30	82,26%	69,35%	53,57%	85,71%	71,43%	71,15%
3	R1.5	24	98,39%	64,52%	31,37%	64,71%	76,47%	78,43%
3	R1.6	24	93,55%	53,23%	45,10%	78,43%	76,47%	66,67%
4	R1.2	24	86,84%	60,53%	41,67%	72,22%	77,78%	74,07%
4	R1.4	24	93,55%	66,13%	56,86%	78,43%	92,16%	86,27%

There was a statistically significant difference in performance on the three categories, based on the group that participants belonged to, $F_{(18, 20)} = 2.28$, $p < .05$; Wilk's $\Lambda = 0.044$, partial $\eta^2 = .65$). Two other multivariate tests that were performed (Hotelling's trace and Roy's largest root) were also significant, but Pillai's trace was not significant, $F_{(18, 27)} = 1.67$, $p = .11$; Pillai's trace $= 1.58$, partial $\eta^2 = .53$. However, Pillai's trace is a test that is very sensitive to sample size, and since there were only 16 participants for this study it is unsurprising that this multivariate test was not significant. Because the sample size was small and the other tests were still significant, it was considered possible to proceed to conducting further small tests such as ANOVA's with the data, keeping in mind the fact that we had a small sample size. Because so many comparisons were made, a Bonferroni correction had to be applied. The expectation was to find further significant differences between groups within the categories, but unfortunately this did not follow from the MANOVA after a Bonferroni correction was applied. Even without the Bonferroni correction, only use of the correct vowel in category 1 words revealed a significant difference ($p < .05$) between AA Eritreans (M= 36.89, SD=15.96) and R1 Eritreans (M=56.45, SD=13.62) and a significant difference in category 2 for using the correct vowel between AA Eritreans (M=37.97, SD=12.57) and AA Syrians (M= 68.99, SD=28.48), and AA Eritreans and R1 Eritreans (M=78.64, SD=9.90), $p < .05$.

These were not exactly the main differences that were expected, so a closer look at the data was required. Still, some general observations can already be made from Table 5. Category 1_written in this table shows how often omission of vowels occurred. The percentage is the percentage of words for which a vowel was written. We can see that most Eritreans (groups 1 and 3) have high scores for this category, most of them scoring higher than 90%, while Syrians (groups 2 and 4) more often have low scores here. For category 2 the scores differ very much, but an interesting finding is that all participants (but one) have higher scores for C2_correct than for C2_number, which means they were better at using the correct vowel (e.g. 'a') than at using the correct number of vowels ('a' or 'aa'). Category 3, which measured consonants, shows that in general scores of Eritrean beginner participants (group 1) lie at an average around 50 to 60%, while percentages of Syrian beginners are more at a level around 80 to 90% correct consonants. The participants that had received more education clearly score higher here, except for one Syrian participant.

Before proceeding with smaller statistical tests, another MANOVA was conducted in which the variable of group was split into 2 independent variables: class (with levels AA and R1) and origin (with levels Eritrean and Syrian). Only for category 3 (number of consonants written) the output showed a significant effect of origin, that is a difference between Eritreans

and Syrians, $F(1, 19) = 33.19$; $p < .005$; partial $\eta^2 = .89$). None of the other categories had any general significant differences that could be explained by class or origin. But because we were dealing with small sample size and looking at percentages of groups that had made either a bigger or smaller part of the test, this was not surprising. Instead, it was decided to examine the data more closely, to see which tendencies and effects could be found then. Although the tests of between subject effects of the performed MANOVA's did not reveal any further significant differences, several important tendencies could be observed in the data. These tendencies are shown in Figures 7 to 12.
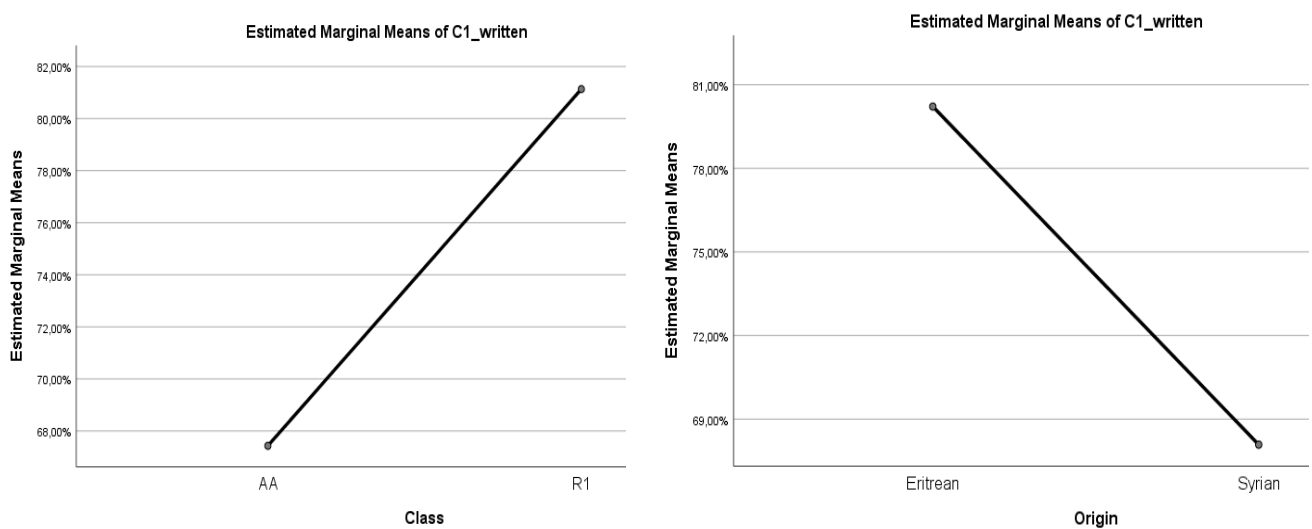
Figure 7: The trend of the data for 'category 1 written', which shows whether a vowel was written down or not, by participants from classes AA and R1 and participants with respectively Eritrean and Syrian backgrounds.

As can be seen in Figure 7, the trend is that participants from the higher class R1 that have received education for a longer period of time performed better (averagely 81% correct) than participants from the beginners class AA (averagely 67% correct). Also, in general Eritrean participants wrote down a vowel more often (on average in 80% of the cases) than Syrian participants (68%).
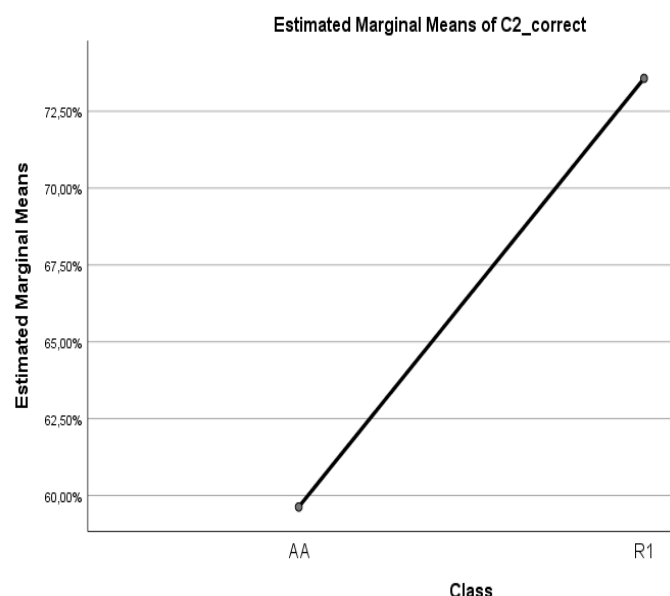
.
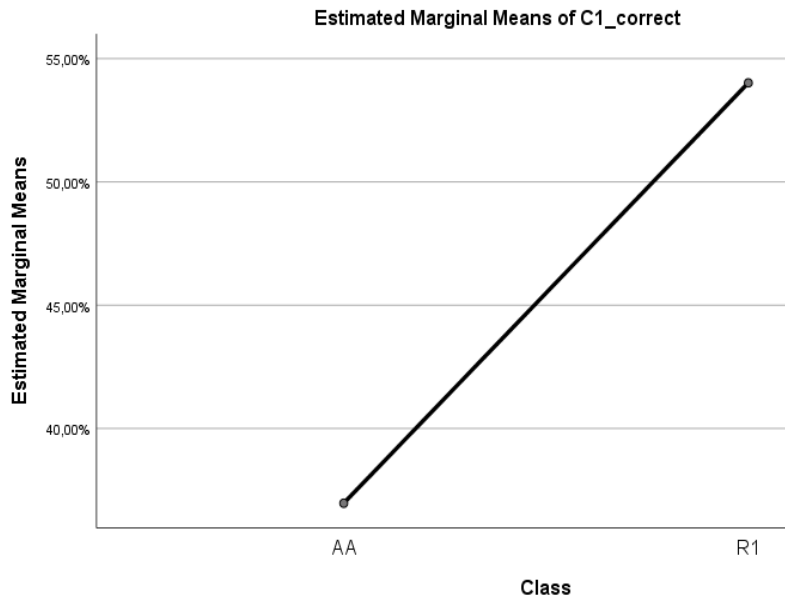
Figure 8: The trend of the data for 'category 1 correct', which shows whether the correct vowel was written or not by participants from classes AA and R1.

Figure 8 shows that here too, participants from class R1 had more correct vowels (54%) than the participants from class AA (36% correct). On average there was not a big difference in use of the correct vowel between Eritrean participants (approximately 47%) and Syrian participants (approximately 42%).
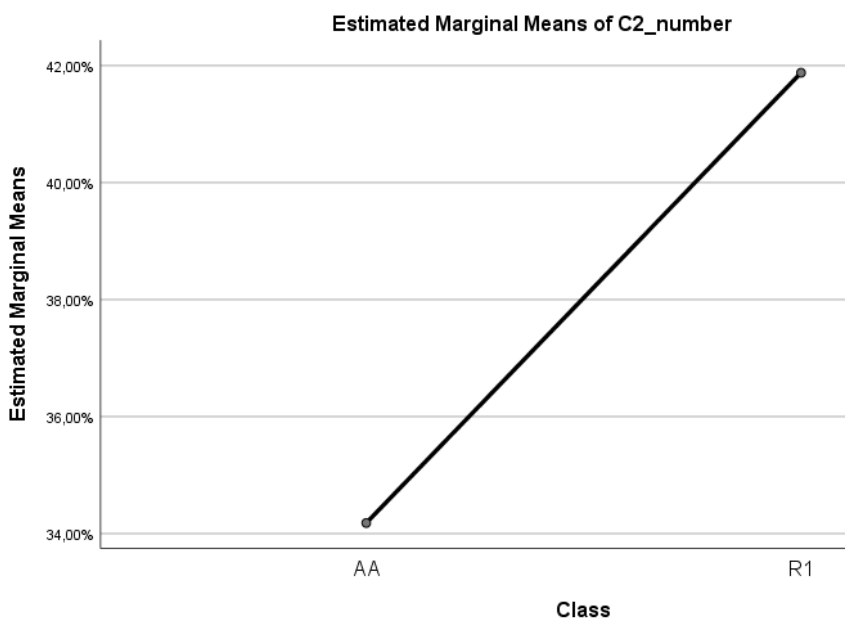
Figure 9: The trend of the data for 'category 2 number', which shows whether the number of vowels that were written by participants was correct, for participants from classes AA and R1.

In Figure 9 it can be seen that the R1 group overall performed slightly better (42% correct) with long sound vowels, using a correct number of vowels more often than the AA group (34% correct). Also, on average there was not a big difference in use of the correct number of vowels between Eritrean (35%) and Syrian participants (39%).
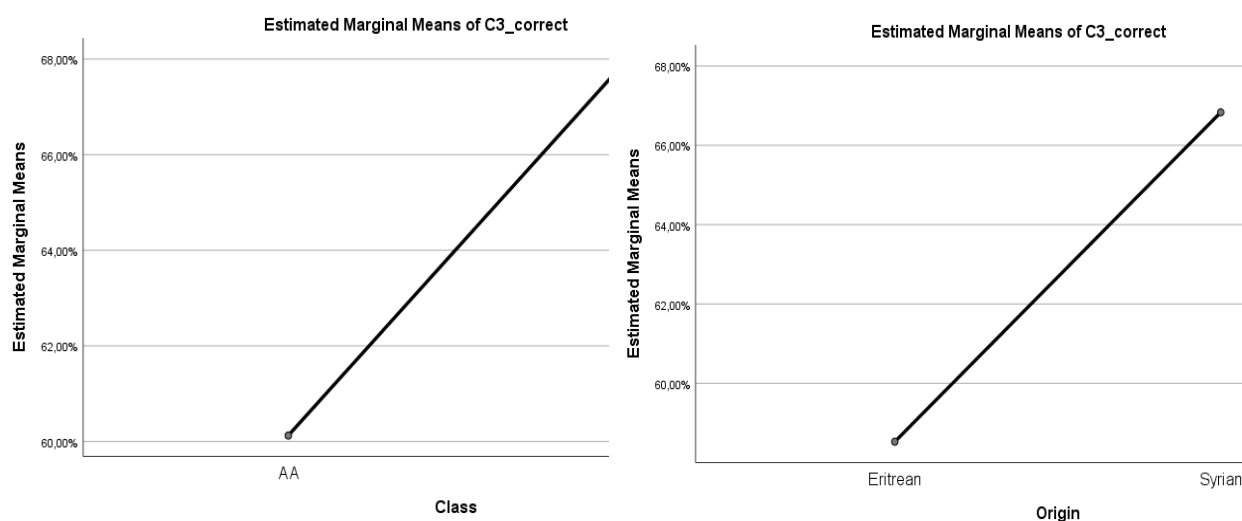

Figure 10: The trend of the data for 'category 2 correct', which shows whether the vowels that were written by participants were correct, for participants from classes AA and R1 and participants with respectively Eritrean and Syrian backgrounds.

Figure 10 displays that here as well the participants from the R1 class performed better (73%) than the AA class participants (60%), using the correct vowel more often. On average, there was not a big difference between scores of Syrian participants (67%) and Eritrean participants (63%).

Figure 11 shows how participants from the R1 class again performed better with the correct number of consonants at syllable endings (approximately 71%) than participants from the AA class (68%). Syrian participants had slightly higher scores (71%) than Eritrean participants (67%).
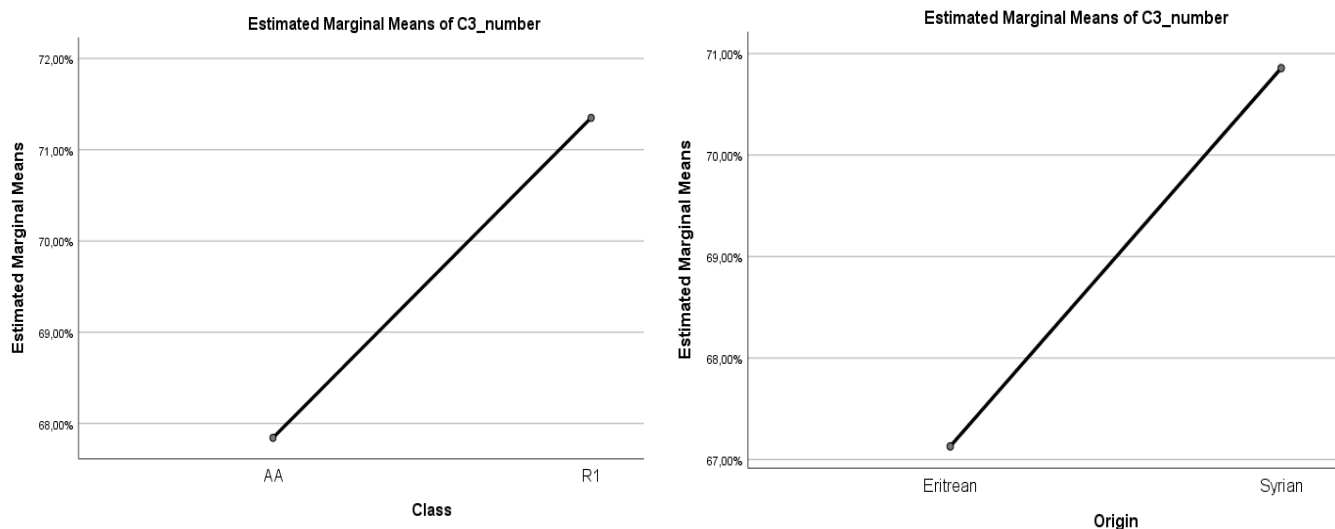
Figure 11: The trend of the data for 'category 3 number', which shows whether the number of consonants that were written by participants was correct, for participants from classes AA and R1 and participants with respectively Eritrean and Syrian backgrounds.

Figure 12: The trend of the data for 'category 3 correct', which shows whether the consonants that were written by participants were correct, for participants from classes AA and R1 and participants with respectively Eritrean and Syrian backgrounds.

Finally, Figure 12 shows that R1 participants had higher scores (68%) than AA participants (60%), using the correct consonants more often, and Syrian participants had more correct items on average (67%) than Eritrean participants (59%). The interpretation of these differences will be described in the Discussion section in the next chapter.

After this general analysis of effects and differences, Factorial ANOVA's were conducted for each category, to see whether the sub-levels of the categories would reveal important differences. These sublevels were the previously discussed factors such as stress, open or closed syllables, et cetera.

4.3 Results Category 1: vowels

For category 1, the effects of the independent variables Stress and Syllable Type were estimated. First, a factorial ANOVA with the independent variables Origin and Stress was performed, with the score for C1_written (i.e. whether a vowel was written down) in percentages as the dependent variable. The results are displayed in Figure 13. There was a significant main effect of Origin on C1_written, $F_{(1,28)} = 7.35$, $p < .05$, partial $\eta^2 = .208$. Eritreans wrote down the vowel significantly more often (M = 89,8, SD = 11,57) than Syrians did (M= 73,1, SD = 22,39).[2] There was no significant effect of Stress, which indicates that participants did not perform significantly worse or better with syllables that had stress. There was also no significant interaction effect of the two factors on the score for use of written vowels, which indicates that Eritreans and Syrians were not affected significantly differently by stress.
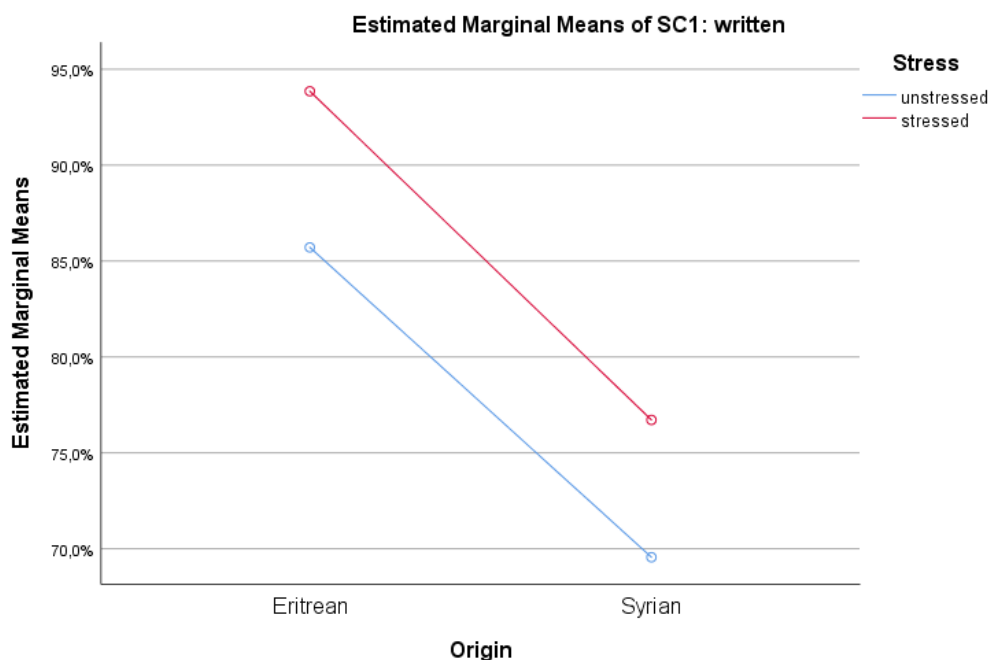

Figure 13: The scores for C1_written of Eritrean and Syrian participants for both stressed and unstressed variables.

---

[2] All numbers from the statistic analyses are percentages, as also mentioned earlier. For reasons of clearness of the text the choice was made not to use the % sign with all numbers.

A second factorial ANOVA was performed to find out the effect of stress, this time for the dependent variable C1_correct, which represent the score for correctness of the vowels instead of just use of any vowel. This test revealed that there was no significant main effect of the variables Origin and Stress on use of the correct vowel, which shows that neither group was much better than the other at using the correct vowel. Instead, both Eritreans (48.0) and Syrians (45.4) appeared to have difficulty with using the correct vowels, as both groups used the correct vowel in on average less than half of the cases. Although the effect of Stress was not significant, the data seems to show a tendency of a difference in performance on syllables with stress (Eritreans M = 51.9, Syrians M = 50.9) and syllables without stress (Eritreans M = 44.1, Syrians M = 40.0), as both groups clearly had higher scores for the stressed syllables. The results are visualized in Figure 14.
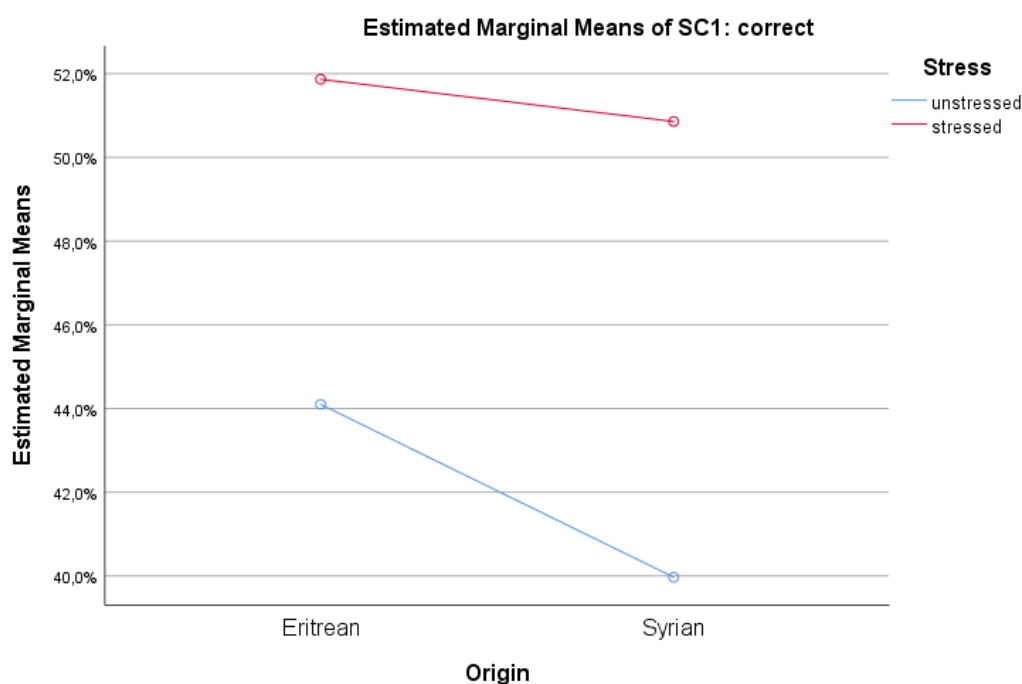

Figure 14: The scores for C1_correct of Eritrean and Syrian participants for both stressed and unstressed variables.

Furthermore, the effect of Syllable Type was tested with Factorial ANOVA's. Here too, there was a significant main effect of Origin on the dependent variable score for C1_written in open and closed syllables, $F_{(1,28)} = 4.88$, $p < .05$, partial $\eta^2 = .149$. Specifically, Eritreans wrote down the vowels more often in open and closed syllables (M = 87.4, SD = 13.1) than Syrians did (M = 71.3, SD = 21.3). No significant main effect of Syllable Type was found, which means that writing vowels in open syllables was not significantly easier for the participants than writing them in closed syllables and vice versa. Also, no significant interaction effects were

found, indicating that Eritreans and Syrians were not affected significantly differently by open and closed syllables in their use of vowels. Another Factorial ANOVA that tested the effect of the variables Origin and Syllable Type on the score for C1_correct revealed no significant effects of Origin and Syllable Type. Thus, although Eritreans wrote down the vowels more often in both open and closed syllables, they did not use the correct vowel significantly more often than the Syrian participants. The fact whether a syllable was open or closed did not influence the correctness of the vowel.

4.4 Results Category 2: long sound vowels vowels

For category 2 the effects of the independent variables Stress, Syllable Type and Number of Vowels (i.e. whether they had to write a single or a double vowel) were estimated. First, factorial ANOVA's with the scores for C2_number in percentages (i.e. the score for whether the number of vowels was correct) as the dependent variable were performed. Whether participants used the correct number of vowels did not appear to be influenced by their origin or by stress, as the effects of Origin and Stress were insignificant. A significant effect of Syllable Type on C2_number was found, $F(1,28) = 58.25$, $p <.001$, partial $\eta^2 = .675$. Participants scored significantly higher with the correct number of vowels in open syllables (M = 65.7, SD = 18.2) than in closed syllables (M = 23.8, SD = 12.8). They also scored significantly higher with the correct number of vowels for words in which a single vowel had to be used (M = 69.0, SD = 20.3) than for words that required double vowels (M = 26.5, SD = 13.0), $F(1,28) = 48.75$, $p <.001$, partial $\eta^2 = .635$. These results indicate that both groups had difficulty with writing double vowels for vowels that had a long sound.

The next factorial ANOVA's were performed to investigate the effects of Stress, Syllable Type, and Number of Vowels on the score for C2_correct (i.e. the score that represents whether the correct vowel was used). None of the factors Stress, Syllable Type, and Number of Vowels had a significant effect on C2_correct, which means that use of the correct vowel was not significantly influenced by stress, open or closed syllables or the number of vowels that had to be used. Initially the idea was also that the results for long sound single vowels should be compared with those of the short sound single vowels from category 1, to find out whether Syrians performed better on the long sound vowels than on the short sound vowels. However, the words that were used in the test were not coded for these aspects and it was too time-consuming to recode all words and run new analyses. Therefore, unfortunately this comparison could not be made.

4.5 Results Category 3: end consonants

For category 3, end consonants, the effects of the independent variables Number of Consonants (single or double) and Number of Consonants in the Next Syllable on C3_number (i.e. the score for use of the correct number of consonants) were estimated. A factorial ANOVA with the C3_number as the dependent variable revealed a significant effect of the number of consonants on participants' score for the number of consonants that they used, $F (2,28) = 7.47$, $p <.001$, partial $\eta^2 = .348$. Participants used the correct number of consonants significantly more often in words that ended in one consonant ($M = 81.64$, $SD = 18.10$) than in words that ended in two consonants ($M = 68.51$, $SD = 24.73$). There was also a significant main effect of Origin on C3_number, $F (1,56) = 4.98$, $p <.05$, partial $\eta^2 = .082$. Syrian participants used the correct number of consonants significantly more often ($M = 80.51$, $SD = 24.91$) than Eritrean participants ($M = 67.92$, $SD = 22.00$). Besides, there was a significant main effect of Number of Consonants in the Next Syllable on C3_number, $F (3,56) = 3.542$, $p <.05$, partial $\eta^2 = .160$. Pairwise comparisons show that participants used the correct number of consonants significantly more often ($p < .05$) in syllables that were followed by a syllable that did not start with a consonant than with syllables followed by two consonants ($MD = 20.61$, $SE = 7.98$). The number of consonants was also correct more often with syllables that preceded a syllable with 1 consonant than with a syllable with 2 consonants ($MD = 18.65$, $SE = 7.89$).

The factorial ANOVA with C3_correct (i.e. the score for correctness of the used consonants) as the dependent variable also revealed a significant effect of Number of Consonants on the correctness of the consonants, $F (2,28) = 12.216$, $p <.001$, partial $\eta^2 = .466$. There was no significant effect of Origin here, nor a significant interaction effect, which means that Eritreans and Syrians were not affected significantly differently by the number of consonants in their correct use of consonants. A significant main effect of Number of Consonants in the Next Syllable was found, $F (3,56) = 6.430$, $p <.001$, partial $\eta^2 = .256$. The pairwise comparisons revealed that participants scored significantly higher ($p < .001$) with consonants in syllables that were not followed by any consonants, than with words that were followed by a syllable that started with two consonants ($MD = 27.19$, $SE = 6.96$). They also scored significantly higher ($p < .05$) with consonants in syllable endings that were followed by a syllable that started with one consonant than with those followed by a syllable that started with two consonants ($MD = 14.40$, $SE = 6.96$).
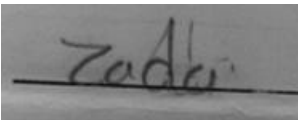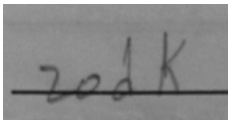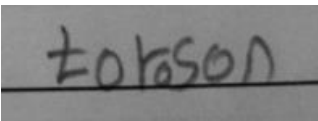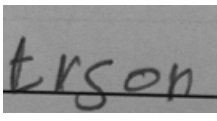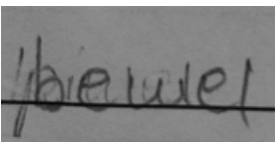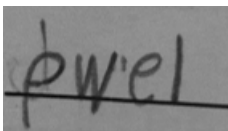
5. Discussion

The main question that we intend to answer is: 'How do the first scripts of native speakers of Arabic and Tigrinya influence writing in Dutch as a second language with the Latin script?' The test was structured in such a way that those characteristics of the first script that might cause difficulty in written Dutch were included in the grading per category. The statistical analyses revealed several trends and effects that give an insight in how this influence might work, but of course there are other factors that may have influenced the performance of the participants. In this chapter I will try to interpret those differences that were found, coming back on the hypotheses that were formulated earlier in this thesis in chapters 2.6 and 3.3. The structure of this chapter is based on the structure of the previous chapter, which follows the categories that were used for creating and grading the test. For each category the relevant findings are discussed and where possible placed in perspective of scientific knowledge.
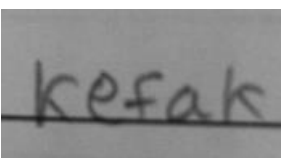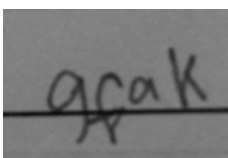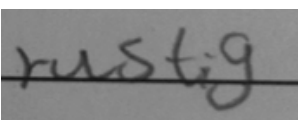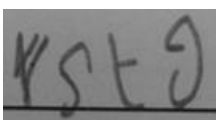
5.1 Discussion category 1: vowels

Difficulty with vowels was expected for both groups. This was indeed the case: many participants made mistakes with the vowels, as can also be seen in Table 5 in the previous chapter. The highest score was achieved by a participant from the R1 group who had used correct vowels in 69% of the cases. The expectation was that Syrian students would more often omit vowels, while Eritrean students were expected to make mistakes with them, but not to omit them. This expectation is supported by the data. The percentages are a less reliable measure here, because 0 points were also awarded for the aspect 'written' when half of the word was missing. Therefore, we need to examine the actual data in more detail. We find there that even Eritrean learners from the AA group were less likely to fully omit use of a vowel. In most cases they would just write some random vowel or skip half of the word when they did not hear it correctly. Writing products of Syrian learners on the other hand showed more instances of words that had no vowels at all. This omission is a good example of influence of the script rather than the language, because of course in Syrian Arabic one does pronounce the vowels; it is the script they are used to that makes them omit vowels. A few examples that show this difference are given in Table 6.

Table 6: Examples from writing products of Eritreans and Syrians that show omission of vowels by Syrians.

Word	Eritrean	Syrian
non-word 'zodak'	'zada' [AA.2]	'zodk' [R1.2]
non-word 'tersoon'	'toroson' [AA.4]	'trson' [AA.5]
word 'bevel'	'bewel' [AA.3]	'pwel' [AA.5]
non-word 'gevak'	'kefak' [AA.4]	'gfak' [R1.2]
word 'rustig'	'rustig' [R1.6]	'rstg' [AA.10]

Of course, there were also instances of omission of vowels by Eritreans. I would be inclined to interpret those instances as either a) words where the participant thought in the frame of the second language and perceived a consonant cluster, or b) words where the participant did not realize that the single character they used did not stand for both a consonant and a vowel, such as in their S1. The latter could only occur with participants of the AA group that had just started, because they would learn this soon enough in the writing lessons.

The expectation was also that Eritreans would make more mistakes with vowels than Syrians, because although the Syrians are likely to omit vowels, they do have separate characters for the sounds of vowels. In principle, their script has a fine granularity, which should make writing vowels easier for them if we follow the hypothesis of Wydell and Butterworth
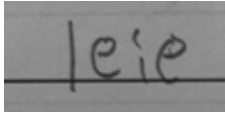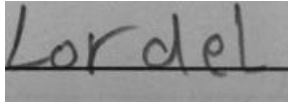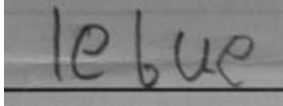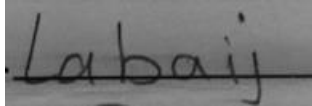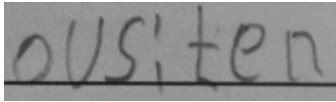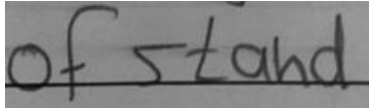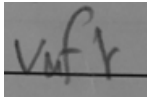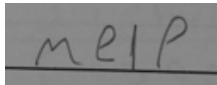
(1999). It was therefore expected that they would be better at writing the correct vowel that matched the sound that they heard. On average however, no clear difference was found. The results for correct or incorrect vowel use are represented by the score for C1_correct. The data show that both groups had difficulty with vowels, but scores were not explicitly lower for Eritreans than for Syrians, as can also be seen in Table 4 in the previous chapter. The results of the performed ANOVA indicate that both groups had similar difficulty with using the correct vowel, as participants from both groups had average scores below 50%. The reason for this could be that there are several factors that affect use of vowels in writing. The S1 plays a role, but so does the L1. Dutch has sounds that are familiar to the participants and sounds that are unfamiliar to them. When taking a dictation test, participants perceive the words with a mind that is trained to decode sound in their L1. It is understandable that they listen for sounds that are familiar to them and tend to write down those more often. The fact that the test contained non-words made listening even more complex, because participants could not rely on background knowledge of the L2 words.

Additionally, it was expected that Eritrean students would have less feeling for phonemes in general, because the way they developed phoneme awareness in their first language is a lot less specific and explicit. It is therefore remarkable to see that Eritreans are overall able to separate the sounds of consonants and vowels, and to understand that the Latin script uses two characters where their S1 has one character. The fact that they omit vowels less often shows that they have developed enough phoneme awareness to split a syllable into its smaller parts. This supports the view of Reddy and Koda (2013), who stated that children who learn an alphasyllabary need to have phonemic awareness to decode the consonant and vowel part of their syllables. Although the smaller phonemes are less directly present, it seems that the Eritreans that took the test had enough phoneme awareness to separate the characters.

If there would be some lack of feeling for phonemes with Eritrean learners, we would expect to find examples of this mostly in the results of Eritreans of the AA group. The data do provide some support for this hypothesis. Where it is mostly clear from the results of AA Syrians what the dictated word was, this is more often not the case for Eritrean writing products. Sometimes it seems that the Eritreans do not really have the right feeling for the order of the letters or write letters that are not in the word at all. A few examples of this are given in Table 7. The fact that the Syrian participants did not write down such distinct letters is an indication that they have a better feeling for phonemes. This is support for the way Asfaha, Kurvers and Kroon (2009) use the hypothesis of granularity and transparency: the language that has a finer

granularity does aid writing in another language with a fine granularity, whereas the language with a coarser granularity is less supportive in that sense.

Table 7: Examples from writing products of Eritreans and Syrians that show lack of feeling for phonemes with the Eritrean participants.

Word	Eritrean	Syrian
non-word 'loordeel'	'leie' [AA.1]	'lordel' [AA.5]
word 'lawaai'	'lebue' [AA.1]	'labaij' [AA.6]
word 'afstand'	'ovsiten' [AA.1]	'ofstand' [AA.6]
non-word 'mulp'	'vufr'? [AA.2]	'melp' [AA.9]

An effect of stress on correct use of the vowel was expected mainly for the Syrian participants. Because in their own language vowels without stress are often not present, it was considered likely that they would omit these more often. However, the statistical analysis revealed no significant effect of stress on the variable C1_written and no significant interaction effect. Apparently, stress did not determine omission of vowels in a strong way, and Eritreans and Syrians were not influenced in a different way by stress. Still, the average scores of both Eritreans and Syrians seemed visibly higher for syllables with stress, although this trend was not significant. The effect of stress is a logical phenomenon that can be explained by the fact that the stressed part of a word is often pronounced more clearly. To draw vast conclusions about the effect of stress, a bigger analysis would have to be conducted. Here namely only a maximum of 63 syllables was tested on stress, given that participants had filled in all the words. Perhaps the trend would be significant if more participants and more words were tested.

Furthermore, we had predicted an effect of open and closed syllables. The expectation was that Eritreans would be better with open syllables, since they conform more to the pattern of their S1. For Syrians, no specific difference between open and closed syllables was expected because they are used to writing consonants at the end of syllables. Contrary to expectations, no difference was found between Eritreans and Syrians in their scores for open and closed syllables. The question is why no effect was found here. A plausible explanation for this might be that the end consonants of closed syllables do not necessarily confuse participants in their use of the vowel that precedes it. So, when an Eritrean has to write a CVC syllable, the first consonant and vowel may be more connected to each other and he will write them down correctly. The actual confusion would then lie only in the end consonant itself rather than the vowel that precedes it. The end consonant is supposedly the more difficult part, especially when there are multiple end consonants. For more details on the difficulty of end consonants we should look at the scores for correctness of end consonants, which are discussed in the later paragraph about category 3 effects.

5.2 Discussion category 2: long sound vowels

Additionally, long sound vowels were tested. There were vowels with a long sound that had to be written with a single letter, and there were vowels that had to be written with a double letter. The aim of testing this was to find out whether there would be any difference between the Eritreans and Syrians on the acquisition of a rule in Dutch words. If there were, then they might have a different learning pattern that could not be explained by their language background or first scripts. Recall that both Syrian and Eritrean scripts are familiar with the sounds of these vowels, but for neither there is need to write them with a double vowel when the syllable is closed, as is the case in Dutch. The performed ANOVA's confirmed the expectation that there was no effect of origin or stress. However, there was a significant effect of syllable type on the score for the correct number of vowels: participants scored much higher for the open syllables. This is unsurprising, because the open syllables were cases of single vowels, while the closed syllables could also contain double vowels. Apparently, most participants had not quite understood the rule for vowel doubling in Dutch yet. It happened much more often that a participant wrote a single vowel where a double vowel was expected than the other way around. The results from this category show that there are no significant differences between the two groups when an aspect of writing in Dutch is tested that has nothing to do with their background. This is exactly what we hoped to find.
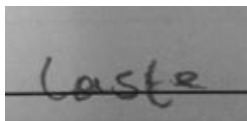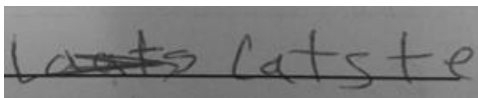
In addition, the idea was to use the results of the long sound vowels and compare them to the short sound vowels from category 1, to see whether participants scored better for sounds that their own language had too. The expectation was that Syrians would perform better with single long sound vowels than with the short sound vowels, and that Syrian would perform better with long sound vowels than the Eritreans, because in Arabic they do have separate characters for long sound vowels. However, the words from category 1 had not been coded for long or short vowels. Unfortunately, there was not enough time to code these items and re-analyze the results to make new comparisons. For this reason, there is no statistical comparison that can tell us whether there was an effect of sounds from the first language, and no evidence that Syrians were better with vowels that Arabic has separate non-diacritic characters for. This might be an interesting aspect for further study.
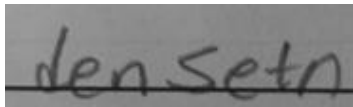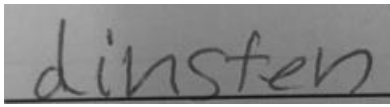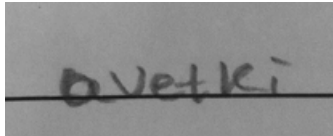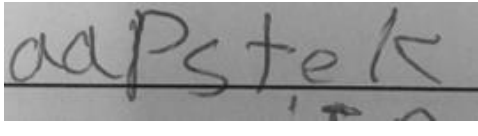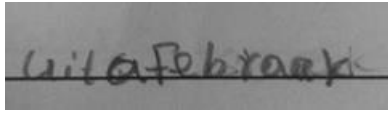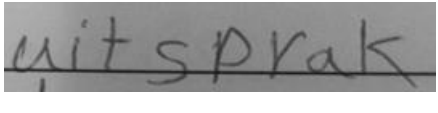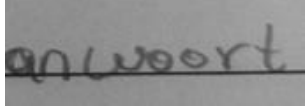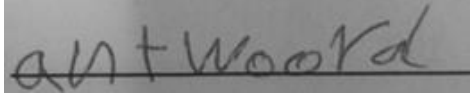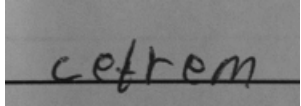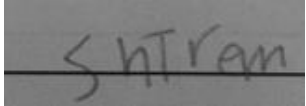
5.3 Discussion Category 3: end consonants

Finally, end consonants were tested. In general, there was an effect of the number of consonants. Naturally, participants found it easier to write single end consonants than double or triple ones. What is more interesting is that a significant difference was found between Eritreans and Syrians in their use of consonants. Syrians were clearly better at writing down the correct amount of end consonants, which means they used two end consonants when two were expected and one where one end consonant was expected. When we look at the data, it becomes clear that Eritreans more often leave the end consonants out or add vowels in between, whereas Syrians are overall better at writing down all the consonants in a cluster. Their mistake is sometimes even that they write down too many consonants. A few examples are given in Table 8.

Besides that, the number of consonants that the next syllable started with influenced the number of end consonants in the preceding syllable. When there were no consonants in the next

Table 8: Examples from writing products of Eritreans and Syrians that show the difference in performance with consonants.

Word	Eritrean	Syrian
word 'laatste'	'laste' [R1.6]	'latste' [R1.4]

61

word 'diensten'	[R1.6]	[R1.4]
	densetn	dinsten
non-word 'afstik'	'avetki' [R1.6]	'aapstek' [R1.4]
	avetki	aapstek
word 'uitspraak'	'uitafebraak'[R1.6]	'uitsprak' [R1.4]
	uilafebraak	uitsprak
word 'antwoord'	'anwoort' [R1.6]	'antwoord' [R1.4]
	anwoort	antwoord
word 'centrum'	'cetrem' [R1.3]	'shtrem'[R1.2]
	cetrem	shtrem
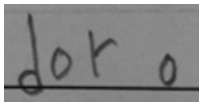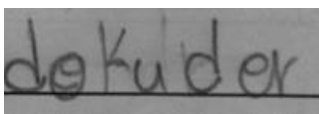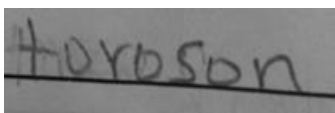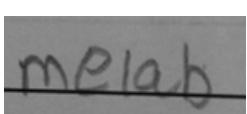
syllable, in general participants found it easier to use the correct number of end consonants. This can be explained by the simple principle that writing more letters is more difficult then writing fewer letters. Every separate sound must be heard, and more choices must be made in the time that a participant must write down the word.
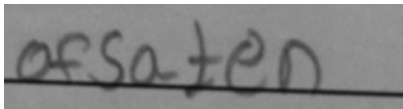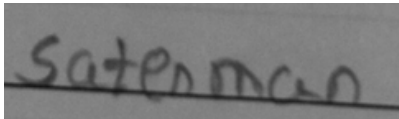
Yet, no significant difference was found between Syrians and Eritreans in their correct use of end consonants. This is surprising, because Eritreans are not used to consonant clusters at all, while for Syrians these are very frequent. This is an effect that would not only be explained by the first script, but also by the first language. Still, Syrians did not seem to perform much better than Eritreans. This may have something to do with the small sample size: There were only two R1 Syrian participants. Outliers in the Eritrean group may also have had an influence. The Eritrean participant R1.3 for example was very good with end consonants. For a deeper investigation it would be good to know which other experience with language may have influenced this performance. Other than that, it is hard to explain why the difference between Syrians and Eritreans was not bigger on this characteristic. Elements like this point out how important it is to be aware of a participant's other capacities, as has been discussed in

chapter 2.3 from the findings of Edelsky (1982). If this topic would be investigated further, this knowledge should definitely be included as a factor.

Anyway, we do not know how much of the difference in performance with end consonants can be explained by the first language and how much by the first script. It is known that the first script and the first language strongly interact with each other. It is therefore difficult to make a clear separation between the two, because the script of a language is strongly influenced by its form. The Ge'ez script for example is mostly structured in CV and CVC syllables, as its spoken form also consists such syllables. When we look at the Eritrean participants, they are used to writing in syllables in their first script. For that reason, they are also more likely to perceive an overheard word in the syllables as they are used to (CV or CVC) and then write it down from that point of view. In chapter 2 it has been explained how the mother tongue determines which characteristics a writer is aware of. Words like 'toroson' for 'tersoon' (table 4) occurred quite frequently and represent an effect of the first language as much as an effect of the first script. They namely show that the word was perceived as a three-syllable word that confirms to the Eritrean pattern of consonants and vowels. There are many words in the data that hint to this. Some other examples are displayed in Table 9.

Table 9: Examples from writing products of Eritreans that show their tendency for CV and CVC, instead of using double consonants.

Word	Eritrean writing product
word 'daaruit'	'doro'[AA.1]
word 'dokter'	'dokuder'[AA.2]
non-word 'tersoon'	'toroson'[AA.3]
non-word 'mulp'	'melab' [AA.4]

word 'afstand'	'ofsaten' [AA.4]
	ofsaten
non-word 'standmuns'	'satenman' [AA.4]
	satenman

Whether it is an effect of the script that makes Eritreans perceive words in syllables, or an effect of the L1 that is mostly structured this way cannot be concluded only from writing products. For that, we would need to know which considerations participants make when writing down a word. It would be interesting to know whether they think of the letters in their first script and convert these to the new script, or rather perceive a word in the way they write it down. This may be a good topic for future research.

6. Conclusion

In this final paragraph the conclusions that can be drawn from the results will be explained. First the answers to the two sub-questions will be explained and then the answer to the main question is discussed. Finally, the restrictions of the current study will be addressed and suggestions for future research shall be made.

6.1 Answering the sub-questions

The first sub-question was: 'Is there a clear relationship between the first script and the type of mistakes that are made in spelling?'. We can answer this question affirmatively for the cases of the two scripts that were studied. The test results reveal clear effects of the first scripts of the participants. Considering vowels, both Syrian and Eritrean participants had difficulty with using the correct vowels. The interesting finding is that Syrian participants often omitted vowels, just like they are used to in their first script. Eritrean participants also had difficulty with the vowels but were less likely to omit them. Even beginning Eritrean learners were less likely than Syrians to fully omit use of a vowel. They would rather write some random vowel or skip half of the word when they did not hear it correctly. Writing products of Syrian learners on the other hand showed more instances of words that had no vowels at all. This omission is regarded as a clear example of influence of the script rather than the language. On the other hand, Eritrean participants generally had more difficulty with consonant clusters. They tended to leave some end consonants out or added vowels in between, whereas Syrians were overall better at writing down all the consonants in a cluster. However, no significant difference was found in the correctness of the used consonants. When Eritreans had written down the correct number of end consonants, they did not make more mistakes in them then the Syrian participants. Still, the fact that Syrians had less difficulty with writing down the consonants results from the fact that they are used to writing them down, whereas Eritreans are not used to this. It must be noted that the first language may also play a role here. Furthermore, whether participants used the correct number of letters for long sound vowels did not appear to be influenced by their origin or by stress. The fact that there were no significant differences between the two groups when an aspect of writing in Dutch was tested that had nothing to do with their background, extra underlines the influence of the first script in the other cases. In general, there is enough reason to conclude from the answers of the participants that their first script had influenced them in their mistakes in spelling.

How then do we interpret these S1 effects that were found? On the one hand, the mistakes that are made could be viewed as a barrier for learning, because clearly a certain type of mistakes results from specific scripts. Yet, it is undeniable that even learners that had been learning to write Dutch for only 3 months were quite good at spelling in their new language. This is surely a positive effect of earlier writing experience. The fact that the first script can explain many mistakes that are made can be viewed as a positive finding, because it also makes it easier to adapt class instructions to these specific problems. When we know what causes the mistakes that learners make, it will be easier to explain to them what is correct in Dutch by pointing out the difference with their S1.

The second sub-question of this thesis was formulated as follows: 'How does the influence of the first script on the new script change overtime?' The expectation was that the Eritrean learners would need more time for learning, since their script has a coarse granularity and higher transparency. Because their phoneme awareness is developed in a different way, this would supposedly make writing in a language with a finer granularity more difficult for them. It appeared from several earlier studies that awareness of bigger units seemed insufficient to facilitate reading in smaller grain-size scripts. Therefore, Eritrean participants from the R1 group were expected to be slower learners and have lower scores than the Syrian participants from the R1 group. However, no significant difference between these two groups of learners resulted from the MANOVA that was performed to investigate this. This finding can partly be explained by the fact that there were only two Syrian participants in the R1 group, which does not make for a strong statistical difference. When there would have been more participants there might have been a significant difference, but from this data we cannot conclude whether there was any difference in performance between Eritreans and Syrians who had been learning to write in Dutch for a longer period. Also, the fact that we had to work with percentages must have troubled the outcomes to some extent for all statistical tests that were performed. When the participants from the AA group had also completed the entire test, the comparison between AA learners and R1 learners would have been more reliable. From the general analysis, no further conclusion can be drawn to answer the second sub-question.

6.2 Answering the main question

So, how do the first scripts of native speakers of Arabic and Tigrinya influence writing in Dutch as a second language with the Latin script? It can be concluded from the test that some of the most basic mistakes that Syrians and Eritreans make in spelling in Dutch can be explained from

their S1. Clear differences were found such as more instances of omission of vowels by Syrians and trouble with writing down end consonants for Eritreans. The effects from the S1 that were found in the data do also provide some evidence of a difference in phoneme awareness between Syrians and Eritreans. These findings further highlight the relevance of explicit phoneme instruction, that Ehri et al. (2001) and Braams and Bosman (2000) already pointed out. More awareness of phonemes has been proved to increase both reading and spelling success and would thus have a positive influence on all language learners. Of course, phoneme instruction is always part of writing lessons, but the results from the dictation test show that it can be very relevant to make these instructions more specific to a language learner's background. Overall namely, the results of the test revealed a clear relationship between the first script (and with that of the first language) and the mistakes that were made in writing Dutch. Therefore, we have reason to believe that instructions that focus on the differences between the S1 and the S2 could improve writing products of both Syrians and Eritreans. Perhaps, specified extra instructions could reduce the gap in explicit phoneme awareness that Eritreans may have because of their script with a coarser granularity. The focus would have to be on the way each letter stands for one sound, where Ge'ez characters stand for a syllable. If learners are aware of this difference it could really improve their writing. The further focus would have to be on aspects that their language does not have, such as consonant clusters and the sounds that their L1 has no characters for. Syrians could also benefit from instructions that highlight the differences with Arabic, as this may help them to become more aware of the vowels that need to be written between the consonants. If there is more focus on the way the Dutch representation of sound in letters is different from their S1 (each vowel must be written), they will also try to listen for such vowels.

6.3 Restrictions and options for future research

Although interesting results were found, this study also has its weak points. There were few participants, so for a further investigation it would be good to carry out a larger scale study with a bigger sample. Also, many participants did not complete the entire test, which made the analysis more difficult and less reliable as we had to work with percentages. Another problem was that not all elements of the words were coded for, so not every initially intended comparison could be made.

This research leaves enough interesting aspects for further investigation. Now only vowels and end consonants were investigated, but there might for example also be effects in consonants in the onset. There are also characteristics of the scripts left that have not been

investigated in this study, such as the shape of the letters themselves. There might be an effect of the directionality of Arabic, or perhaps of the shape of the Ge'ez fidels. The difference in writing familiar and unfamiliar sounds could also be examined in more detail. And in this study participants' ability in L1 skills was not included in the analysis. It is possible that some of the outliers are also more skilled writers in their L1. As mentioned before, it would be interesting to get more insight in what people think when they write in their S2. That way, we could further discriminate between the L1 and the S1.

As a final remark I would like to point out that the previous experience of the participants with languages other than Tigrinya and Syrian Arabic, and scripts other than Ge'ez and Arabic, were not included in this study. I expect that there would be a clear effect of such experience as well. Perhaps, the effect of previous experience with for example English on writing in Dutch would be a good topic for further investigation of the effects of scripts. Maybe language learners could also benefit from instruction from that point of view.

References

Aarnoutse, C.A.J. (2004). *Ontwikkeling van beginnende geletterdheid* [Lecture]. Nijmegen: Katholieke Universiteit Nijmegen. Retrieved from: http://hdl.handle.net/2066/19451

Abu-Rabia, S. (1998). Reading Arabic texts: Effects of text type, reader type and vowelization. *Reading and Writing: An Interdisciplinary Journal,* 10*(2),* pp. 105-119.

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing: An Interdisciplinary Journal,* 13, pp. 147-157.

Asfaha, Y.M., Kurvers, J. & Kroon, S. (2008). Literacy and script attitudes in multilingual Eritrea. *Journal of sociolinguistics, 12(2),* pp. 223-240.

Asfaha, Y.M., Beckman, D., Kurvers, J. & Kroon, S. (2009). L2 reading in multilingual Eritrea: The influences of L1 reading and English proficiency. *Journal of research in reading, 32(4),* pp. 351-365.

Asfaha, Y.M., Kurvers, J. & Kroon, S. (2009). Grain size in script and teaching: Literacy acquisition in Ge'ez and Latin. *Applied Psycholinguistics,* 30, pp. 709-724.

Bachra, B.N. (2001). *The Phonological Structure of the Verbal Roots in Arabic and Hebrew.* Leiden: Brill.

Bernhardt, E.B. & Kamil, M.L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics,* 16, pp. 15-34.

Bereketeab, R. (2010). The Complex Roots of the Second Eritrea-Ethiopia War: Re-examining the Causes. *African Journal of International Affairs*, 13, pp. 15-59.

Bertelson, P., & Gelder de, B. (1991). The emergence of phonological awareness: Comparative approaches. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception*, pp. 393-412. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Bossers, B. & Kuiken, F. & Vermeer, A. (2015). *Handboek Nederlands als tweede taal in het volwassenonderwijs.* In S.J. Andringa, R.T. Bakker & C.I. Hosea (Eds.). Bussum: Coutinho.

Braams, T., & Bosman, A.M.T. (2000). Fonologische vaardigheden, geletterdheid en lees- en spellingsinstructie. *Tijdschrift voor Orthopedagogiek*, 39, pp. 199-211.

Bright, W. (2000). A matter of typology: alphasyllabaries and abugidas. *Studies in the Linguistic Sciences*, 30*(1)*, pp. 63-72.

Castro-Caldas, A., Petersson, K.M., Reis, A., Stone-Elander, S. & Ingy, M. (1998). The illiterate brain. Learning to read and write during childhood influences the functional organization of the adult brain. *Brain,* 121, pp. 1053–1063.

Centraal Bureau voor de Statistiek. (2017, July 31). *Migratie naar achtergrond,1e helft 2016 en 2017* [Dataset]. Retrieved from:

https://www.cbs.nl/maatwerk/2017/31/migratie-naar-achtergrond-1e-helft-2016-en-2017

Chaney, C. (1989). I pledge a legiance to the flag: Three studies in word segmentation. *Applied Linguistics,* 10, pp. 261-281.

Daniels, P. T. (1996). The study of writing systems. In P.T. Daniels & W. Bright (Eds.), *The World's Writing Systems*, pp. 579-586. New York: Oxford University Press.

Edelsky, C. (1982). Writing in A Bilingual Program: The Relation of L1 and L2 Texts. *TESOL Quarterly*, 16*(2)*, pp. 211-228.

Ehri, L.C., Nunes, S.R., Willows, D.M., Schuster, B.V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 36, pp. 250-287.

Geerts, G. & Heestermans, H. (1984). Van Dale: Groot Woordenboek der Nederlandse Taal: 3 Dl. (11th edition). Utrecht: Van Dale Lexicografie.

Gelder de, B., Vroomen, J. & Bertolson, P. (1993). The effects of alphabetic-reading competence on language representation in bilingual Chinese subjects. *Psychological research,* 55, pp. 315-321.

Gibson, E. & Levin, H. (1976). *The psychology of reading*. Cambridge, Massachusetts and London: The MIT Press.

Gill JR, J.T. (1989). The relationship between word recognition and spelling in the primary grades, *Reading Psychology*, 10*(2),* pp. 117-136.

Hamilton, M.E. & Barton, D. (1983). Adults' definitions of `Word': The effects of literacy and development. *Journal of Pragmatics*, 7*(5)*, pp. 581-594.

Hussien, A.M. (2014). The Effect of Learning English (L2) on Learning of Arabic Literacy (L1) in the Primary School, *International Education Studies*, 7*(3)*, pp. 88-98.

Johnston, F.R. (2000). Exploring classroom teachers' spelling practices and beliefs. *Literacy Research and Instruction*, 40*(2),* pp. 143-155.

Kenstowicz, M. (1986). Notes on Syllable Structure in Three Arabic Dialects. *Revue québécoise de linguistique*, 16*(1)*, pp. 101–127.

Kurvers, J. (2002). *Met ongeletterde ogen. Kennis van taal en schrift van analfabeten.* Amsterdam: Amsterdam University Press.

Kurvers, J. & Zouw van der, K. (1990). *In de ban van het schrift: over analfabetisme en alfabetisering in een tweede taal.* Amsterdam, Lisse: Swets & Zeitlinger.

Kurvers, J., Hout van, R. & Vallen, T. (2007) *Literacy and word boundaries. Low educated second language and literacy acquisition: Research, policy and practice. Proceedings of the second annual forum,* pp. 44-64. In N. Faux & H. Major (Eds.). Richmond, VA: The Literacy institute.

https://www.cbs.nl/maatwerk/2017/31/migratie-naar-achtergrond-1e-helft-2016-en-2017

Chaney, C. (1989). I pledge a legiance to the flag: Three studies in word segmentation. *Applied Linguistics,* 10, pp. 261-281.

Daniels, P. T. (1996). The study of writing systems. In P.T. Daniels & W. Bright (Eds.), *The World's Writing Systems*, pp. 579-586. New York: Oxford University Press.

Edelsky, C. (1982). Writing in A Bilingual Program: The Relation of L1 and L2 Texts. *TESOL Quarterly*, 16*(2)*, pp. 211-228.

Ehri, L.C., Nunes, S.R., Willows, D.M., Schuster, B.V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 36, pp. 250-287.

Geerts, G. & Heestermans, H. (1984). Van Dale: Groot Woordenboek der Nederlandse Taal: 3 Dl. (11th edition). Utrecht: Van Dale Lexicografie.

Gelder de, B., Vroomen, J. & Bertolson, P. (1993). The effects of alphabetic-reading competence on language representation in bilingual Chinese subjects. *Psychological research,* 55, pp. 315-321.

Gibson, E. & Levin, H. (1976). *The psychology of reading*. Cambridge, Massachusetts and London: The MIT Press.

Gill JR, J.T. (1989). The relationship between word recognition and spelling in the primary grades, *Reading Psychology*, 10*(2),* pp. 117-136.

Hamilton, M.E. & Barton, D. (1983). Adults' definitions of `Word': The effects of literacy and development. *Journal of Pragmatics*, 7*(5)*, pp. 581-594.

Hussien, A.M. (2014). The Effect of Learning English (L2) on Learning of Arabic Literacy (L1) in the Primary School, *International Education Studies*, 7*(3)*, pp. 88-98.

Johnston, F.R. (2000). Exploring classroom teachers' spelling practices and beliefs. *Literacy Research and Instruction*, 40*(2),* pp. 143-155.

Kenstowicz, M. (1986). Notes on Syllable Structure in Three Arabic Dialects. *Revue québécoise de linguistique*, 16*(1)*, pp. 101–127.

Kurvers, J. (2002). *Met ongeletterde ogen. Kennis van taal en schrift van analfabeten.* Amsterdam: Amsterdam University Press.

Kurvers, J. & Zouw van der, K. (1990). *In de ban van het schrift: over analfabetisme en alfabetisering in een tweede taal.* Amsterdam, Lisse: Swets & Zeitlinger.

Kurvers, J., Hout van, R. & Vallen, T. (2007) *Literacy and word boundaries. Low educated second language and literacy acquisition: Research, policy and practice. Proceedings of the second annual forum,* pp. 44-64. In N. Faux & H. Major (Eds.). Richmond, VA: The Literacy institute.

McCarthy, J, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, 12*(3)*, pp. 373-418. Massachusetts: The MIT Press; In F. Katamba (Eds.) (2004). *Morphology. Critical Concepts in Linguistics. Vol III: Morphology: its Relation to Phonology*, pp. 230-278. London: Routledge.

Miller, C. (2003). Linguistic Policies and Language Issues in the Middle East. In A. Usuki & H. Kato (Eds.), *Islam in the Middle Eastern Studies: Muslims and Minorities,* pp. 149 174. Osaka, Japan: JCAS Symposium Series 7.

Mishra, R. & Stainthorp, R. (2007). The relationship between phonological awareness and word reading accuracy in Oriya and English: A study of Oriya-speaking fifth-graders. *Journal of Research in Reading*, 30*(1),* pp. 23-37.

Morais, J., Cary, L., Allegria, J. & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition,* 24, pp. 45-64.

Morais, J. (1991). Constraints on the development of phonemic awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy,* pp. 5-27. Hillsdale, NJ: Lawrence Erlbaum Associates.

Morais J. (1993). Phoenmic awareness, language and literacity. In R.M. Joshi and C.K. Leong (Eds.), *Reading disabilities: Diagnosis and component processes*, pp. 175-184. Dordrecht: Kluwer Academic Publishers.

Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading*, 30*(1)*, pp. 7–22.

Nelson, L. (1989). Something borrowed, something new: Teaching implications of developmental spelling research. *Reading Psychology*, 10, pp. 255-274.

Peterson, K.M., Reis, A., & Ingvar, M. (2001). Cognitive processing in literate and illiterate subjects: A review of some recent behavioral and functional neuroimaging data. *Scandinavian Journal of Pyschology*, 42, pp. 251-267.

Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. (2013). What do spelling errors tell us? Classification and analysis of errors made by Greek schoolchildren with and without dyslexia. *Reading and Writing: An Interdisciplinary Journal*, 26*(5)*, pp. 615-646.

Read, C., Yun-Fei, Z., Hong-Yin, N. & Bao-Qing, D. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition,* 24, pp. 31-44.

Reddy, P.P. & Koda, K. (2013) Orthographic constraints on phonological awareness in biliteracy development, *Writing Systems Research*, 5*(1)*, pp. 110-130.

Reis, A. & Castro-Caldas, A. (1997). Illiteracy: a cause for biased cognitive development. *Journal of the International Neuropsychological Society*, 3*(5)*, pp. 444-450.

Saville-Troike, M. (2012). *Introducing Second Language Acquisition.* Cambridge: Cambridge University Press.

Scribner, S. & Cole, M. (1981). *The psychology of literacy.* Cambridge: Harvard University Press.

SoNaR (2013). Text corpus of the Institute for Dutch Lexicology (INL) [Database]. Retrieved from: http://opensonar.inl.nl/

United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics. (2016). *Literacy rate, adult total* [Dataset]. Retrieved from: https://data.worldbank.org/indicator/SE.ADT.LITR.ZS

Veldhuis, D. & Kurvers, J. (2012). Offline segmentation and online processing units. The influence of literacy. *Written language & literacy*, 15*(2)*, pp. 165-184.

Verhoeven, L. & Aarnoutse, C. (1999). Interactief taalonderwijs: balans tussen constructief en instructief leren. *Spiegel*, 17/18*(3/4)*, pp. 9-29.

Verhoeven, L. & Aarnoutse, C. & Leeuwe van, J. (2000). Ontwikkeling van beginnende geletterdheid. *Pedagogische Studiën*, 77*(5/6),* pp. 307-325.

Versteegh, K. (2001). *The Arabic Language.* Edinburgh: Edinburgh University Press.

Viise, N.M. (1996). A study of the spelling development of adult literacy learners compared with that of classroom children. *Journal of Literacy Research*, 28, pp. 561-587.

Walter, S.L. & Davis, P.M. (2005). *The Eritrea National Reading Survey.* Dallas: TX: SIL International.

WebCelex (1995). [Database]. Retrieved from: www.celex.mpi.nl

Vries de, M., Winkel te, L.A. (1954). Woordenlijst voor de Spelling der Nederlandsche Taal. Den Haag, Leiden, Arnhem: Martinus Nijhof.

Wright, M.W. (2002). *Ideologies and methodologies in language and literacy instruction in postcolonial Eritrea.* Unpublished Doctoral Dissertation. Philadelphia: University of Pennsylvania.

Wydell, T. N. & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition*, 70, pp. 273–305.

Ziegler, J. C. & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, pp. 3–29.

Appendices

I. Test scheme with words per category

Words that are marked with an asterix (*) are the words that are the less frequent words.

Category 1: Vowels				Item Count
Rating on two aspects	3. vowel written (1) / not written (0) 4. correct vowel (1) / incorrect vowel (0) a double vowel is graded as an incorrect vowel ➔ for each category, only the mentioned syllable gets graded. e.g. category 'first syllable stressed open', only the first syllable is graded on these two aspects			Total number of items in this category: 63 (126 points possible) Per subcategory min. 6 high frequency words. If no 6 high freq available, add up with low freq. Strive for 10 words, max. 10 per category. (high freq. = frequency in CELEX>100 per 1,000,000 words). (less freq. = frequency in CELEX>30 and <60 per 1,000,000).
Item restrictions	- no double vowels, only a, e, i, o, u - the graded syllables are only syllables that have a CV (open) or CVC (closed) pattern - try to keep ungraded syllables simple (try to avoid diphthongs) - no verbs - no proper names, place names, etc.			
Subcategories	**Syllables**	**Stress**	**Syllable type**	**Count**
	Monosyllabic words	Stressed	Open	- (removed for lack of relevance and to shorten the test)
			Closed	b**e**d d**i**t g**a**f h**u**n k**o**p n**e**t r**o**l r**u**g v**e**r w**e**g
	Bisyllabic words	1st syllable stressed	Open	b**e**ter d**a**gen j**a**ren

73

				kamer leven lichaam nodig samen vader water
			Closed	binnen dokter landen lippen middel minder rustig verder volgens zonder
		1st syllable unstressed	Open	begrip bekend belang direct gebied gedrag manier moment muziek
			Closed	nogal partij persoon terwijl vandaag verband
		2nd syllable stressed	Open	- (removed for lack of data)
			Closed	geluk geval begin besef* bevel* gemak*
		2nd syllable unstressed	Open	hoge welke beste derde onze nota* (little in corpus high freq)
			Closed	bezig basis

					later le**d**en ta**f**el ze**k**er

Category 2: Long sound vowels					Item Count
Rating on two aspects		3. correct number of vowels (1) / incorrect number of vowels (0) 4. correct vowel(s) (1) / incorrect vowel(s) (0) a double vowel where a single vowel is expected gets graded as an incorrect vowel and vice-versa. For each subcategory, only the mentioned syllable gets graded. (e.g. for the category 'single vowel 1st syllable stressed open', only the vowel in the first syllable of the word is graded on the two aspects.)			Total number of items in this category: 54 (108 points possible) Per subcategory min. 6 high frequency words. If no 6 high freq available, add up with low freq. Strive for 10 words, max. 10 per category (high freq. = frequency in CELEX>100 per 1,000,000 words). (less freq. = frequency in CELEX>30 and <60 per 1,000,000).
Item restrictions		- the graded syllables are only syllables that have a CV or CVV (open) or CVC or CVVC (closed) pattern. Try to avoid double consonants. - try to vary vowels as much as possible - try to keep ungraded syllables simple - no verbs - no proper names, place names, etc.			
Subcategories	**Number of vowels**	**Syllables**	**Stress**	**Syllable type**	**Count**
	Single vowel (e.g. a)	Monosyllabic	stressed	Open	- (removed for lack of relevance and to shorten the test)
				Closed	- (always double vowel)
		Bisyllabic	1st syllable stressed	Open	b**a**sis b**e**ter b**e**zig d**a**gen h**e**le h**o**ge j**a**ren k**a**mer l**a**ter l**e**den
				Closed	-

					(always double vowel)
			1st syllable unstressed	Open	moment zodat zowel muziek helaas* lawaai*
				Closed	- (always double vowel)
			2nd syllable stressed	Open	- (removed for lack of data)
				Closed	- (always double vowel)
			2nd syllable unstressed	Open	- (removed for lack of data)
				Closed	(always double vowel)
	Double vowel (e.g. aa)	Monosyllabic	stressed	Open	mee twee zee
				Closed	deel hoog jaar meer taal voor week zaak zoon
		Bisyllabic	1st syllable stressed	Open	- (only in compositions)
				Closed	beetje daarom hoogte laatste maanden voordat woorden
			1st syllable unstressed	Open	-
				Closed	waarbij waarmee daaruit*
			2nd syllable stressed	Open	- (removed for lack of data)
				Closed	alleen bestaan elkaar gewoon meteen

					omh**oo**g
					pers**oo**n
					vand**aa**g
					verh**aa**l
					min**uu**t*
		2nd syllable unstressed	Open	-
			Closed	lich**aa**m
				afspr**aa**k*
				uitspr**aa**k*
				noodz**aa**k*
				voord**ee**l*
				zom**aa**r*

Category 3: End consonants				Item Count
Rating on two aspects	3. correct number of end-consonants (1) / incorrect number of end-consonants (1) 4. correct end-consonants (1) / incorrect end-consonants (0)			Total number of items in this category: 60 (120 points possible) Per subcategory min. 6 high frequency words. If no 6 high freq available, add up with low freq. Strive for 10 words, max. 15 per category (high freq. = frequency in CELEX>100 per 1,000,000 words). Words in blue are less frequent words (less freq. = frequency in CELEX>30 and <60 per 1,000,000).
Item restrictions	- try to vary the consonants as much as possible - no diphthongs as graded consonants - try to keep ungraded syllables simple, try to avoid diphthongs - no verbs - no proper names, place names, etc.			
Subcategories	**Number of consonants**	**Syllables**	**Next syllable starts with consonant?**	**Count**
	Single end consonant	Monosyllabic	-	brie**f** boe**k** haa**r** ka**n** ko**p** pa**s** raa**m** ru**g** ve**r**

				wet
		Bisyllabic (first syllable)	No	daarom vanaf daaruit* veelal*
			One C	beste derde dokter hoogte minder rustig terwijl verhaal volgens welke zonder
			Two C	afstand eerste indruk invloed laatste diensten* minstens* verstand* voorstel* centrum*
	Double end consonant	Monosyllabic	-	half hart hulp kans kant niets punt volk vorm werk
		Bisyllabic (first syllable)	No	- (removed for lack of data)
			One C	antwoord standpunt* godsdienst*
			Two C	hoofdstuk ontstaan* grondslag*

II. Words with their CELEX frequencies

1	*afspraak*	36		70	*maanden*	155	
2	*afstand*	117		71	*manier*	333	
3	*alleen*	554		72	*mee*	641	
4	*antwoord*	177		73	*meer*	1457	
5	*basis*	102		74	*meteen*	211	
6	*bed*	284		75	*middel*	130	
7	*beetje*	297		76	*minstens*	42	
8	*begin*	127		77	*minder*	213	
9	*begrip*	123		78	*minuut*	35	
10	*bekend*	171		79	*moment*	270	
11	*belang*	207		80	*muziek*	115	
12	*besef*	43		81	*net*	421	
13	*bestaan*	143		82	*niets*	794	
14	*beste*	215		83	*nodig*	393	
15	*beter*	120		84	*nogal*	134	
16	*bevel*	41		85	*noodzaak*	39	
17	*bezig*	175		86	*nota*	33	
18	*binnen*	386		87	*omhoog*	123	
19	*boek*	250		88	*ontstaan*	48	
20	*centrum*	46		89	*onze*	727	
21	*brief*	114		90	*partij*	102	
22	*daarom*	325		91	*pas*	235	
23	*daaruit*	42		92	*persoon*	108	
24	*dagen*	325		93	*punt*	124	
25	*deel*	312		94	*raam*	112	
26	*diensten*	39		95	*rol*	183	
27	*derde*	160		96	*rug*	173	
28	*direct*	126		97	*rustig*	125	
29	*dit*	2233		98	*samen*	308	
30	*dokter*	130		99	*standpunt*	46	
31	*eerste*	880		100	*taal*	130	
32	*elkaar*	811		101	*tafel*	189	
33	*gaf*	318		102	*terwijl*	622	
34	*gebied*	183		103	*twee*	996	
35	*gedrag*	143		104	*uitspraak*	49	
36	*geluk*	105		105	*vader*	547	
37	*gemak*	41		106	*vanaf*	137	
38	*geval*	411		107	*vandaag*	118	
39	*gewoon*	221		108	*veelal*	46	
40	*godsdienst*	40		109	*ver*	221	
41	*grondslag*	44		110	*verband*	177	
42	*haar*	4531		111	*verder*	267	

43	half	183	112	verstand	47
44	hart	183	113	verhaal	161
45	helaas	42	114	volgens	292
46	hele	292	115	volk	122
47	hoge	304	116	voor	139
48	hoog	159	117	voordat	138
49	hoofdstuk	119	118	voordeel	45
50	hoogte	310	119	voorstel	44
51	hulp	115	120	vorm	242
52	hun	2126	121	waarbij	213
53	indruk	143	122	waarmee	180
54	invloed	173	123	water	353
55	jaar	734	124	week	155
56	jaren	406	125	weg	453
57	kamer	318	126	welke	326
58	kan	548	127	werk	496
59	kans	171	128	wet	146
60	kant	235	129	woorden	310
61	kop	117	130	zaak	204
62	laatste	260	131	zee	138
63	landen	114	132	zeker	389
64	later	258	133	zodat	302
65	lawaai	31	134	zomaar	42
66	leden	116	135	zonder	878
67	leven	222	136	zoon	151
68	lichaam	264	137	zowel	206
69	lippen	102			

III.   Used words in the test with their pseudo words

Letters in bold are the parts of the words that got graded.

		category 1: vowels	category 2: long vowels	category 3: consonant endings	Pseudo-word
1	*afspraak*		afspr**aa**k		onspraak
2	*afstand*			a**f**stand	afstik
3	*alleen*		all**ee**n		olleen
4	*antwoord*			a**nt**woord	antwolg
5	*basis*	bas**is**	b**a**sis		gasis
6	*bed*	b**e**d			det
7	*beetje*		b**ee**tje		leetje
8	*begin*	beg**i**n			bevin
9	*begrip*	b**e**grip			bedrip
10	*bekend*	b**e**kend			bekest
11	*belang*	b**e**lang			belaf
12	*besef*	bes**e**f			besek
13	*bestaan*		best**aa**n		beskaan
14	*beste*	best**e**		be**s**te	keste
15	*beter*	b**e**ter	b**e**ter		geter
16	*bevel*	bev**e**l			revel
17	*bezig*	bez**i**g	b**e**zig		hezig
18	*binnen*	b**i**nnen			rinnen
19	*boek*			boe**k**	voek
20	*centrum*			ce**n**trum	pentrum
21	*brief*			brie**f**	krief
22	*daarom*		d**aa**rom	da**a**rom	vaarom
23	*daaruit*		d**aa**ruit	da**a**ruit	vaaruit
24	*dagen*	d**a**gen	d**a**gen		tagen
25	*deel*		d**ee**l		weel
26	*diensten*			die**n**sten	wiensten
27	*derde*	derd**e**		de**r**de	serde
28	*direct*	d**i**rect			lirect
29	*dit*	d**i**t			rit
30	*dokter*	d**o**kter		do**k**ter	wokter
31	*eerste*			ee**r**ste	veerste
32	*elkaar*		elk**aa**r		arkaar
33	*gaf*	g**a**f			taf
34	*gebied*	g**e**bied			gebies
35	*gedrag*	g**e**drag			getrag
36	*geluk*	gel**u**k			reluk
37	*gemak*	gem**a**k			gevak
38	*geval*	gev**a**l			gezal
39	*gewoon*		gew**oo**n		gezoon
40	*godsdienst*			go**ds**dienst	godsdiert
41	*grondslag*			gro**nd**slag	rondslak
42	*haar*			haa**r**	zaar

43	*half*			half	talf
44	*hart*			hart	wart
45	*helaas*		helaas		pelaas
46	*hele*		hele		zele
47	*hoge*	hoge	hoge		roge
48	*hoog*		hoog		voog
49	*hoofdstuk*			hoofdstuk	hoofdstak
50	*hoogte*		hoogte	hoogte	roogte
51	*hulp*			hulp	mulp
52	*hun*	hun			lun
53	*indruk*			indruk	indras
54	*invloed*			invloed	invloer
55	*jaar*		jaar		kaar
56	*jaren*	jaren	jaren		karen
57	*kamer*	kamer	kamer		bamer
58	*kan*			kan	tan
59	*kans*			kans	rans
60	*kant*			kant	dant
61	*kop*	kop		kop	wop
62	*laatste*		laatste	laatste	waatste
63	*landen*	landen			janden
64	*later*	later	later		jater
65	*lawaai*		lawaai		jawaai
66	*leden*	leden	leden		jeden
67	*leven*	leven			jeven
68	*lichaam*	lichaam	lichaam		lachaam
69	*lippen*	lippen			jippen
70	*maanden*		maanden		zaanden
71	*manier*	manier			wanier
72	*mee*		mee		lee
73	*meer*		meer		seer
74	*meteen*		meteen		seteen
75	*middel*	middel			widdel
76	*minstens*			minstens	kinstens
77	*minder*	minder		minder	linder
78	*minuut*		minuut		binuut
79	*moment*	moment	moment		doment
80	*muziek*	muziek	muziek		buziek
81	*net*	net			ket
82	*niets*			niets	riets
83	*nodig*	nodig			kodig
84	*nogal*	nogal			wogal
85	*noodzaak*		noodzaak		roodzaak
86	*nota*	nota			rota
87	*omhoog*		omhoog		zahoog
88	*ontstaan*			ontstaan	ontsteen
89	*onze*	onze			inze
90	*partij*	partij			kartij

#					
91	*pas*			pa**s**	vas
92	*persoon*	p**e**rsoon	pers**oo**n		tersoon
93	*punt*			pu**nt**	vunt
94	*raam*			raa**m**	laam
95	*rol*	r**ol**			nol
96	*rug*	r**u**g		ru**g**	tug
97	*rustig*	r**u**stig		ru**s**tig	bustig
98	*samen*	s**a**men			samer
99	*standpunt*			st**and**punt	standmuns
100	*taal*		t**aa**l		gaal
101	*tafel*	taf**e**l			lafel
102	*terwijl*	t**e**rwijl		te**r**wijl	terwijk
103	*twee*		tw**ee**		pree
104	*uitspraak*		uitspr**aa**k		uitstaak
105	*vader*	v**a**der			zader
106	*vanaf*			va**n**af	danaf
107	*vandaag*	v**a**ndaag	vand**aa**g		vandaak
108	*veelal*			vee**l**al	deelal
109	*ver*	ve**r**		ve**r**	ber
110	*verband*	v**e**rband		ve**r**haal	verhaak
111	*verder*	v**e**rder			berder
112	*verstand*			ve**r**stand	kerstand
113	*verhaal*		verh**aa**l	ve**r**haal	derhaal
114	*volgens*	v**o**lgens		vo**l**gens	polgens
115	*volk*			vo**lk**	solk
116	*voor*		v**oo**r		loor
117	*voordat*		v**oo**rdat		woordat
118	*voordeel*		voord**ee**l		loordeel
119	*voorstel*			voo**r**stel	doorstel
120	*vorm*			vo**r**m	dorm
121	*waarbij*		w**aa**rbij		laarbij
122	*waarmee*		w**aa**rmee		laarmee
123	*water*	w**a**ter			waper
124	*week*		w**ee**k		neek
125	*weg*	w**e**g			neg
126	*welke*	welk**e**		we**l**ke	velke
127	*werk*			we**rk**	verk
128	*wet*			we**t**	det
129	*woorden*		w**oo**rden		voorten
130	*zaak*		z**aa**k		baak
131	*zee*		z**ee**		lee
132	*zeker*	zek**e**r			leker
133	*zodat*		z**o**dat		zodak
134	*zomaar*		zom**aa**r		womaar
135	*zonder*	zond**e**r		zo**n**der	tonder
136	*zoon*		z**oo**n		zook
137	*zowel*		z**o**wel		zowek