# **CHAPTER THREE:** Syllable Units during Oral Reading

The following chapter is under revision as:

Adedeji, V. I., Vasilev, M. R., Kirkby, J. A., & Slattery, T. J. (under revision). *Syllable units during oral reading: Evidence of stability across grades three, four and five*.

## 3.1. Abstract

Children progress from making grapheme-phoneme connections to making grapho-syllabic connections before whole-word connections during reading development (Ehri, 2005b). More is known about the development of grapheme-phoneme connections than is known about grapho-syllabic connections. Therefore, we explored the trajectory of syllable use in English developing readers during oral reading. Fifty-one developing readers in grades three, four and five read aloud sentences with an embedded target word while their eye movements and voices were recorded. Targets contained six-letters and were either one or two syllables. Children in grade five had shorter gaze duration, shorter articulation duration and wider spatial eye-voice span (EVS) than children in grade four. Children in grades three and four did not significantly differ on these measures. A syllable number effect was found for gaze duration but not for articulation duration and spatial EVS. Interestingly, one-syllable words took longer to process compared to two-syllable words suggesting that more syllables may not always signify greater processing difficulty. Importantly, we found no evidence of change in the use of syllable units across the three grades. The findings suggest that syllable representations are relatively stable from grades three to five.

## 3.2. Introduction

Learning to read involves making connections between the sounds of a language and the symbols representing those sounds. Although reading development depends on language development (Nation & Snowling, 2004), the mechanisms underlying both are different. While children develop their language skills by exposure to a language-rich environment (Castles et al., 2018), reading develops primarily through direct instruction (Joo et al., 2021). The English Language National Curriculum is structured so children's word identification processes are trained to be fluent, with small units being introduced before large units (Department for Education, 2013). Word processing times increase with word length in German (Huestegge et al., 2009), Italian (Zoccolotti et al., 2008) and English (Hyona & Olson, 1995; Joseph et al., 2009) languages. Moreover, this effect is larger in younger children compared to older children and adults (Blythe et al., 2011; Huestegge et al., 2009; Joseph et al., 2009), suggesting a transition from the use of small orthographic units such as letters to larger units such as syllables, morphemes, or words as reading skill develops. This transition is highlighted early in children's education with explicit syllable instruction to aid the parsing of longer words (Department for Education, 2013). Thus, allowing children to use pre-existing words of shorter length in reading long words with more than one syllable without resorting to letter-by-letter decoding.

The developmental trajectory of syllable use during word recognition in Spanish, German, Portuguese, and French has been well documented using behavioural data (Alvarez et al., 2016; Campos et al., 2021; Colé et al., 1999; Hasenäcker & Schroeder, 2016). Although, researchers have examined the awareness of syllables and intra-syllabic units such as onsets and rimes in English (Bruck & Treiman, 1992; Duncan & Seymour, 2003; Juel & Holmes, 1981; Treiman, 1986; Treiman et al., 1995; Ziegler & Goswami, 2005), no studies have examined how children develop in using syllable units in English using eye-tracking. Oral reading is the medium through which children learn to read and it shows a stronger relationship with children's reading skills than silent reading (Kim et al., 2019; Rasinski & Hoffman, 2003; Rayner et al., 2013). Additionally, syllables encode information about pronunciation (Hasenäcker & Schroeder, 2016). Given these, it is surprising that studies examining syllable processing have been limited to single word paradigms, and studies in oral sentence reading are rare (but see Juel & Holmes, 1981). Therefore, the current study aimed to address this by investigating syllable representations in children from grades three to five.

### 3.2.1. Reading Development and Syllable Units

Ehri's (2005b, 2017) model of reading development proposes that once children attain alphabetic knowledge- learning letters and their sounds, they progress to making grapho-syllabic connections. Reading instruction encourages children to make a transition from encoding individual letters to encoding larger grain size units, such as affixes, root words, syllables, morphemes, and words. Presentation of words with similar patterns, reinforces these larger units and helps children develop consolidated units of recognition (Ehri, 2005a, 2005b, 2017). These consolidated units reduce memory load and improve fluency, allowing texts to be read more efficiently (Ehri, 2005a). For example, the word *purple* would be encoded as two syllabic units (*pur-ple*) rather than six letter units (*p-u-r-p-l-e*). Whilst familiar words with more than one syllable could be read as whole units, the predominant connections at the consolidated alphabetic phase are grapho-syllabic or grapho-morphemic in nature (Ehri, 2005a, 2005b, 2017). Thus, readers may decode the pronunciation of newly encountered words using syllables.

Skilled readers show evidence of syllable decoding during word recognition. For example, readers take a longer time fixating (Pelczarski et al., 2019), begin pronunciation of (Ferrand, 2000; Ferrand & New, 2003; New et al., 2006; Yap & Balota, 2009) and make lexical decisions (Chetail, 2014; New et al., 2006; Stenneken et al., 2007) on words or non-words with more syllables compared to those with fewer syllables. However, this inhibitory effect (Yap & Balota, 2009), where more syllables prevent faster word recognition compared to fewer syllables, has been flawed by evidence of null effects (Ashby & Clifton, 2005; Drieghe et al., 2019; Fitzsimmons & Drieghe, 2011; Frederiksen & Kroll, 1976) and facilitatory effects where words with fewer syllables may take longer to name (Lee, 2001). Syllable effects may also occur before the word is fixated. That is, readers may be able to extract phonological information in the form of syllables in the parafovea (Ashby, 2010; Ashby & Martin, 2008; Ashby & Rayner, 2004). However, this pre-lexical account is not without controversy (see Drieghe et al., 2019). Interestingly, most of this conflicting evidence has come mainly from the English language, where syllable boundaries are less well defined (Cutler et al., 1986; Kahn, 1976; Seymour et al., 2003). Nevertheless, other experimental manipulations of syllable processing, such as syllable frequency effects, have provided evidence that English readers have access to syllable representations like other languages with simple syllable structures and well-defined syllable boundaries (Spanish: Alvarez et al., 2001; Carreiras et al., 2005; German: Hutzler et al., 2005; English: Macizo & Van Petten, 2006).

Explanations for the use of syllable units have been provided by computational models of polysyllabic word recognition, which account for syllable number effects namely: the Multiple Memory Trace Model (MTMM; Ans et al., 1998) and Connectionist Dual Processing (CDP++; Perry et al., 2010) Model. The MTMM assumes that words are processed via two sequential routes: the global and analytical routes. Words are processed as whole units in the global route. However, processing via the analytical route depends on detecting a vocalic grapheme or nucleus which narrows the processing unit to the largest initial component or syllable recognised as familiar. The MTMM predicts that the naming latency of pseudowords (for skilled adult readers) should increase with an increasing number of syllables, as each new syllable will require a new visual capture. Similarly, the CDP++, which includes two routes to reading aloud, also uses vocalic identification and parsing mechanism in the sublexical route to explain syllable number effects. Critical to the sublexical route of this model is the graphemic parser, which assigns a string of letters into graphemes and categorises them into onset, nucleus, and coda, which are the essential parts of the syllable (Balota et al., 2007; Perry et al., 2010). When two-syllable words are encountered, all graphemes may be assigned erroneously to the first nucleus/vowel, after which a correction is made to assign the second vowel to the second syllable. This process is assumed to make reading aloud of two-syllable words longer than one-syllable words. Although the CDP++ does not make any predictions for developing readers, the MTMM does. It proposes that because developing readers have had fewer encounters with words than skilled readers, they are more likely to use the analytical route for words, thereby activating syllable-level representations (Ans et al., 1998). Several developmental studies using different paradigms, as will be seen below, support this prediction across languages (French: Colé et al., 1999; English: Duncan & Seymour, 2003; Finnish: Häikiö et al., 2016; Häikiö, Hyönä, et al., 2015). Taken together, syllable information is part of a multi-layered structure of phonological information encoded during word recognition in skilled and developing readers (Ashby, 2010; Chateau & Jared, 2003). However, how and when these representations may change during reading development is poorly understood, particularly in English.

Two French studies indicate a difference in syllable processing between children and adults. Colé et al. (1999) assessed syllable processing using the syllable compatibility effect (faster detection of whether a target syllable e.g., *BA*, was present at the beginning of a target word with either the same (*BA.LLON)* or different (*BAL.CON)* initial syllable). After six months of reading instruction (grade one- 6 year olds), developing readers showed no compatibility effects. However, significant compatibility effects were found only for good readers four months later. In contrast, skilled adult readers, showed syllable compatibility effects only for low frequency words (Colé et al., 1999). In the same vein, syllable number predicted French third graders' (8 years) reaction times in a naming and an online identification task compared to fifth graders (10 years) and adults (Bijeljac-Babic et al., 2004). Further evidence for the attenuation of syllable effects with an increase in reading skill has been documented by Hautala et al. (2012). Lexical decisions and naming times of poor Finnish readers in grade two (8 years) increased as the number of syllables in the word and nonword stimuli increased, respectively. However, this effect was absent for typical readers (Hautala et al., 2012). In Spanish, it appears syllables are used similarly across second (7 years) and sixth (11 years) grades. In a word-spotting paradigm, readers were required to spot one syllable words (*FIN*) embedded at the beginning of pseudowords which ended at a syllable boundary (*FIN-LO*) or not (FI-NUS; Alvarez et al., 2016). The results showed that second and sixth graders were equally fast in the syllable boundary condition. Similarly, Hasenäcker and Schroeder (2016) reported that second (7 years) and fourth grade (9 years) German readers were both unaffected by a syllable disruption manipulation where a colon was placed either at the syllable boundary or at a letter following the syllable boundary. However, both groups differed significantly from adults, who did not display reaction time differences based on where the colon was placed.

In English, two studies provide a direct indication that syllables are processed by developing readers. Juel and Holmes (1981) reported that second and fifth graders spent more time reading sentences embedded with two-syllable words compared to one-syllable words during both oral and silent reading. This syllable effect was greater in the oral reading condition. Interestingly, poor readers had a greater syllable effect than good readers during oral reading. However, the syllable effect was similar for both categories of readers in silent reading (see Figure 5 in Juel & Holmes, 1981). Duncan and Seymour (2003) found that 11-year-olds made more errors on three-syllable non-words compared to two-syllable non-words. However, no such effects were present for words. Importantly, it is difficult to draw firm conclusions from the studies by Juel and Holmes (1981) and Duncan and Seymour (2003) as syllable number was confounded by word length. An intervention study by Bhattacharya and Ehri (2004) provided data to suggest that readers in third grade may use syllabic units in reading and spelling new words. Adolescents with an average age of 13 years reading at third, fourth and fifth grade equivalent levels participated in either of two interventions (multisyllabic reading instruction or whole word instruction) or no special instruction (control). The intervention involved practiced reading of 100 multisyllabic words either by analysing them into syllables or reading them as whole words. Their results showed that those receiving multisyllabic instruction performed better than the other groups. In addition, this intervention effect was greater for adolescents reading at third grade compared to fourth and fifth grade level.

Regardless of the experimental paradigm or manipulation, it appears syllables play an important role in word recognition. Sensitivity to syllable units is acquired early after reading instruction for most languages (but see Campos et al., 2021 for diverging evidence in European Portuguese). This pattern may remain relatively stable during primary years for Spanish and German readers (Alvarez et al., 2016; Hasenäcker & Schroeder, 2016). However, for English and French readers, it may change dependent on grade (Bijeljac-Babic et al., 2004) or reading skill (Hautala et al., 2012; Juel & Holmes, 1981). Such cross linguistic differences in the use of large grain size units such as syllables could be attributed to the degree of consistency of grapheme-to-phoneme correspondence between these languages proposed by the Psycholinguistic grain size theory (Ziegler & Goswami, 2005). Whilst German and Spanish have a shallow orthography, French and English have a deep orthography (Seymour et al., 2003). Deep orthographies with inconsistent spelling to sound mapping may result in an earlier transition to whole word reading from syllable decoding compared to transparent orthographies. This explanation may not hold for Finnish which has a transparent orthography and evidence showing differences in syllable processing between poor and typical readers (Hautala et al., 2012). However, one similarity between the English study (Juel & Holmes, 1981) and the Finnish study is the use of a syllable number manipulation which may make proficiency differences more obvious across languages with varying orthographic depth (see Section 1.4.1 in the Introduction). Therefore, developing readers may be more likely to utilise large grain sizes such as syllables early on in reading instruction and start to rely less on syllabic representations as reading skill and experience with print increases (Alvarez et al., 2016; Grainger & Ziegler, 2011; Ziegler & Goswami, 2005).

### 3.2.2. Reading Development and the Eye-voice Span

Sentence reading allows us to examine whether syllable effects generalise beyond word-level paradigms. This approach affords the opportunity to examine children's reading development in its most natural and unobtrusive context. In addition, an oral reading paradigm allows us to measure the coordination of the oculomotor, linguistic and articulatory systems necessary for oral reading comprehension and fluency using the *eye-voice span* (EVS; Kim et al., 2019). The EVS is the distance between the eye and the voice during oral reading (Buswell, 1920; Fairbanks, 1937; Inhoff et al., 2011; Laubrock & Kliegl, 2015; Quantz, 1897).

According to Buswell (1920), a wide spatial EVS allows a reasonable unit of meaning to be recognised and integrated before articulation occurs. Therefore, the eyes travel ahead to allow processing of punctuation, context and meaning, which ensures that reading occurs fluently. Without an EVS, reading would be monotonous, slow, and laden with errors, especially with homographs (words with same spelling but different pronunciation or meaning; Clark, 1995). This suggests that the EVS can be used as an index of reading fluency. For instance, non-fluent readers who are at the very beginning of reading instruction, decode words by sounding out each letter and blending the sounds. These readers may have a significantly narrow EVS. However, as familiarity with words increase, readers' lexical processing efficiency increases, and the span widens (Huestegge et al., 2009; Mancheva et al., 2015; Rayner, 1986; Reichle et al., 2013; Vorstius et al., 2014). For example, Buswell's study on the EVS showed that readers in grade two already have an average EVS of 8-character spaces (Buswell, 1920). However, by the time they reach college, they average an EVS of 15-character spaces. Similarly, Levin and Turner (1966) showed that second graders had smaller EVS (2 words) compared to adults (4 words).

In summary, after about a year or so of reading instruction, beginning readers already show evidence of a span between the eye and the voice, and this span changes as reading experience accumulates. Additionally, it is clear that the EVS represents a useful index to describe reading development. However, whether the change in the use of syllable units, as found in previous single word naming studies, can be measured using the EVS remains an open question.

### 3.3.3. The Present Study

Children in grades three, four and five read aloud sentences embedded with six letter target words of either one or two syllables controlled for number of letters and phonemes. Our study is unique in permitting the use of several dependent measures to understand the time course of syllable processing across reading development. First, gaze duration is the measure of all first pass fixations on the target word before moving away. This measure may be indicative of full lexical access (Reichle et al., 2013; Reichle et al., 1998). Using this measure was deliberate since underlying cognitive processes in children have often been captured by gaze duration due to multiple refixations children make compared to adult readers (Word frequency: Joseph et al., 2013; Word length: Tiffin-Richards & Schroeder, 2015b). Second, articulation duration is the measure of the time between the start and end of target word articulation. This measure has often been examined in relation to word length effects based on the possibility that children often begin articulation before decoding is complete (Gagl et al., 2015; Hautala et al., 2012). However, it's use is far less common in studies examining syllable effects where the time between stimuli presentation to initiation of a vocal response (i.e., naming latency) is used. Finally, an important aspect of this study, is the examination of the spatial EVS, which measures how far the eyes are from the voice when the target word is about to be articulated (see Figure 3.1). We now turn to specific predictions regarding each these measures for syllable number and grade.

Children process syllabic information after sufficient reading instruction (Bhattacharya & Ehri, 2004; Ehri, 2005b, 2017). The limited evidence on the syllable number effect in English developing readers (Duncan & Seymour, 2003; Juel & Holmes, 1981) prevent definite claims from being made about the syllable number effect due to a lack of control for word length in these studies. However, the existing evidence from skilled English readers (Jared & Seidenberg, 1990; but see Lee, 2001; Perry et al., 2010; Yap & Balota, 2009) and developing poor Finnish readers (Hautala et al., 2012) suggest that with more syllables, more time is spent during naming and lexical decisions. With null effects found in adults' eye movements during sentence reading (Ashby & Clifton, 2005; Drieghe et al., 2019), whether inhibitory effects are present for developing English readers is an important question. We propose that findings from naming tasks may extend to eye movement patterns (see Schilling et al., 1998). Additionally, children's eye movements may be more sensitive to a phonological manipulation (Blythe et al., 2015; Jared et al., 2016) and this may be more obvious during reading aloud (Juel & Holmes, 1981). Hence, we predicted that gaze duration should be longer for two syllable words compared to one syllable words because words with more syllables have more vocalic nuclei which make them phonologically more complex in addition to stress assignment and vowel reduction processes (Perry et al., 2010; Yap & Balota, 2009).

It is possible that words with more syllables take longer to articulate due to intra-word pausing at syllable boundaries compared to words with fewer syllables (Proença, 2018). Hence, we predict that articulation duration will be longer for two syllable words compared to one syllable words. To our knowledge, length effects on spatial EVS have been documented once by Halm et al. (2011). Halm et al. (2011) limited the study of length effects to number of letters rather than number of syllables and found longer words resulted in decreased spatial EVS. If the same length effects are applicable, we predict that EVS should be greater for one syllable words compared to two syllable words.

In line with previous eye movement studies (Huestegge et al., 2009; Hyona & Olson, 1995; Johnson et al., 2018; Rayner, 1986), we predict that gaze duration should decrease with higher grades as access to the lexicon becomes faster. With age, spoken word duration decreases and oral reading rate increases (Hasbrouck & Tindal, 2006; Hulme et al., 1984). Therefore, we expect articulation duration to decrease with higher grades. In agreement with prior studies (Buswell, 1920; Levin & Turner, 1966), we also expect the spatial EVS to increase with higher grades. Based on the shift from letter-to-letter decoding to global analysis of words (Ans et al., 1998) and a potential use of larger grain sizes in English readers due to orthographic depth (Ziegler & Goswami, 2005), we predict a reduction in the syllable number effect for all measures as grade increases.

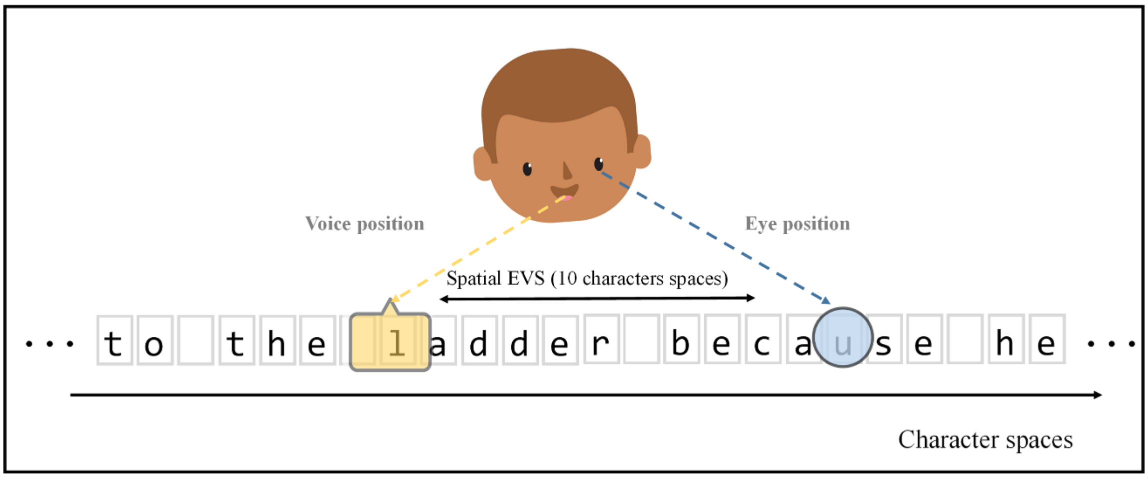


*Figure 3.1. An illustration of the spatial EVS of the target word "ladder"*

### 3.3. Method

#### 3.3.1. Participants

Sixty-four children (35 female) from two primary schools participated after school, parental and child consent were received. All participants reported normal or corrected-to-normal vision and no prior diagnosis of reading disorders and were fluent English speakers and readers. Seven participants spoke at least one other language apart from English. Results did not differ between excluding and including these participants, so they were kept in. Participants were naive as to the purpose of the experiment. One child did not complete the offline measures due to school absence. The eye movement records of two children were discarded: one due to excessive head movements and the other due to at-chance comprehension score. Furthermore, due to technical errors, the voice recordings of 10 participants were lost. This left 51 participants with complete eye movement and voice recordings with an average age of 8.88 years (*SD*=0.93 years;

*range*=7 -10 years). Only the data from these participants were included in all descriptive and inferential statistics. This sample comprised of 13 third graders (6 female, *Mean* = 7.77, *SD*=0.50), 21 fourth graders (13 female, *Mean*= 8.76, *SD*=0.46) and 17 fifth graders (9 female, *Mean*= 9.91, *SD*=0.36).

All children completed standardised measures of reading, spelling, rapid naming speed and intelligence (see Table 3.1 for summary). These were measured with the Test of Word Reading Efficiency 2- Form A (TOWRE; Torgesen et al, 1999), the spelling subtest of the Wechsler Individual Achievement Test II for Teachers (WIAT-II-T; Weschler, 2006), the letters and numbers subtest of the Rapid Automatized Naming and Rapid Alternating Stimulus tests (RAN/RAS; Wolf & Denckla, 2005), and the matrix reasoning and vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence II (WASI-II; Weschler, 2011). IQ scores were within normal range of two standard deviations above and below the mean.

We explored differences in the offline ability measures between the three grades. One way analysis of variance showed that there was no significant difference in intelligence ($F(2,48) = 1.62$, $MSE = 244.8$, $p = .209$), vocabulary ($F(2,48) = 1.84$, $MSE = 102$, $p = .17$), matrix reasoning ($F(2,48) = 0.54$, $MSE = 73.61$, $p = .586$), rapid naming speed – numbers ($F(2,48) = .23$, $MSE = 121.8$, $p = .743$), letters ($F(2,48) = .21$, $MSE = 108.71$, $p = .815$) and spelling ($F(2,48) = 1.24$, $MSE = 153.2$, $p = .299$). However, there was a difference in total word reading efficiency $F(2,48) = 7.10$, $MSE = 73.8$, p=.002, where third graders scored significantly better than fourth graders $t(29.98) = 4.85$, $p < 0.001$, fourth graders were no different from fifth graders $t(34.81) = -1.95$, $p=0.059$, and third graders were not significantly different from fifth grader graders $t(23.89) = -1.98$, $p = 0.059$. In addition, there were significant differences in the sight word efficiency scores, $F(2,48) = 8.95$, $MSE = 67.4$, $p < .001$ and pseudo-word decoding $F(2,48) = 5.27$, $MSE = 105.9$, $p < .001$. These analyses of total word reading efficiency scores indicate that third graders were performing significantly better than the average (111) for their age compared to fourth graders who were on average performing at the average level (100). Furthermore, derived sight word age equivalent scores indicated that third and fourth graders were performing at similar levels, $F(2,48) = 12.49$, $MSE = 1.43$, $p < .001$ (see Table 3.1).

### 3.3.2. Materials and Design

There were 84 experimental passages which comprised of two sentences spanning two lines, with each passage between 70 and 101 characters (*M*= 87.18 characters, *SD*= 8.03). The initial design was a two-year longitudinal study with two time points. As such, the 84 passages were split into two: each with 42 passages. Half the participants were to read one set, and the other half were to read the other set. This was so that each participant could see a different set of passages at time point two. However, after the initial data collection, the Covid-19 pandemic prevented the data collection at timepoint 2. Therefore, the experiment had two independent variables: number of syllables (one vs two) as within-participant and grade (three, four and five) as between-participant variables. A pair of one- and two-syllable target words with similar word frequency counts and number of phonemes (four or five) was embedded in the first sentence of either of two experimental passages (see Figure 3.2). Target words appeared 5.2 words on average into the sentence (*SD*= 1.5, *range*= 3-10).

Table 3.1. Mean and Standard Deviation (In Parenthesis) of Children's Standardized Scores on Off-line Ability Measures of Participants by Grade

| Measure | Third Grade | Fourth Grade | Fifth Grade |
|---|---|---|---|
| WASI-II | 105.77 (17.93) | 106.76 (18.05) | 114.71 (9.28) |
| Vocabulary | 54.54 (12.93) | 55.62 (10.28) | 60.88 (6.96) |
| Matrix Reasoning | 51.23 (7.81) | 51.14 (10.48) | 53.82 (6.13) |
| RAN/RAS-Numbers | 110.62 (11.24) | 108.05 (9.96) | 107.71 (12.11) |
| RAN/RAS-Letters | 103.69 (6.51) | 101.48 (11.84) | 103.00 (10.91) |
| WIAT-II-T-Spelling | 105.23 (11.51) | 101.43 (13.74) | 107.71 (11.14) |
| TOWRE-II-Total Word Reading Efficiency | 111.85(4.41) | 100.62 (9.73) | 106.71 (9.41) |
| TOWRE- Sight Word Efficiency | 112.77 (6.14) | 100.81(8.49) | 103.12 (9.16) |
| TOWRE- Pseudo Word Decoding | 110.00 (6.39) | 100.38 (12.00) | 109.82 (10.34) |
| TOWRE-II Sight Word Age Equivalent | 9.18 (0.90) | 9.16 (0.83) | 11.00 (1.68) |
| Age (years) | 7.77(0.50) | 8.76 (0.46) | 9.91(0.36) |

*Note.* WASI-II IQ scores are shown as standardized scores while mean T scores are shown for the Vocabulary and Matrix reasoning subtests

| Condition | Sentence Frames |
|---|---|
| One syllable | A. He held tightly to the **branch** because he was scared. He had always been afraid of heights. |
| | B. The bird landed on the **branch** and rested there. Ed ran towards it and it flew away. |
| Two syllables | A. He held tightly to the **ladder** because he was scared. He had always been afraid of heights. |
| | B. The bird landed on the **ladder** and rested there. Ed ran towards it and it flew away. |

*Figure 3.2. Example of the experimental sentences*

All target words were six letters (see Appendix A for full list of items) and chosen from the Children's Printed Word Database (CPWD; Stuart et al., 2003). A group of 18 children (*range*=7-10 years) independently rated how well target words fit into two sentence frames (A and B; see Figure 3.2) on a scale of 1(very bad) and 4 (very good). See Table 3.2 for summary of target word characteristics. Sixty-one target words (73%) had an age of acquisition (AoA) norm below 7 years, 20 had an average AoA of 8.51 years (*range*=7.1-11.5) while 3 items did not have an AoA measure. Items were allocated to the different item sets based on the sentence rating and word frequency, such that each item set had similar sentence ratings. The mean word frequency did not differ between the two item sets $F$ (1,80) = .03, $p$ = .86 and two syllable number conditions, $F$(1,80) = .03, $p$ = .88, nor was there an interaction between item set and condition, $F$(1,80) = .003, $p$ = .96. Furthermore, except for phonological neighbourhood size, where there was a marginal difference between one and two syllable conditions, $F$ (1,80) = 3.91, $p$ = .052 and bigram frequency, where there was a marginal difference between item sets, $F$ (1,80) = 3.77, $p$ = .056; all other main or interaction effects of conditions and item sets for the target word characteristics were non-significant. The assignment of conditions to sentence frames and participants to item set was counterbalanced with a full-Latin square design across participants. Items appeared in a pseudo-random order for each participant.

Table 3.2. Mean Target Word Characteristics as a Function of Number of Syllables and Item Set (SD in Parenthesis)

| Measure | Set 1 | | Set 2 | |
|---|---|---|---|---|
| | One-syllable | Two-syllable | One-syllable | Two-syllable |
| Phonemes | 4.52(0.51) | 4.52(0.51) | 4.43(0.51) | 4.43(0.51) |
| CPWD Frequency[a] | 72.14(93.21) | 67.62(82.17) | 67.05(114.81) | 63.82(101.15) |
| Target-word-in-sentence rating | 1.57 (0.34) | 1.56 (0.32) | 1.59 (0.34) | 1.61(0.36) |
| AoA[b] | 6.20(1.47) | 6.01 (1.83) | 6.66 (1.68) | 6.01 (2.00) |
| Orthographic N[c] | 0.76 (0.77) | 0.76 (0.77) | 0.48 (0.68) | 1.0 (1.09) |
| Phonological N[c] | 3.05(2.33) | 2.00(1.67) | 3.10 (2.64) | 2.19 (2.31) |
| Sum. Bigram Frequency[d] | 15030(6588) | 17590(7314) | 18990(6317) | 19630 (8004) |

*Note.* [a]CPWD frequencies given as occurrence per million words. [b] From Brysbaert and Biemiller (2017). [c] N= Neighbourhood, from the CPWD (Stuart et al., 2003). [d] From Balota et al. (2007)

### 3.3.3. Apparatus

Eye movements were recorded with an SR Research EyeLink 1000 Plus desktop-mounted eye-tracker with sampling frequency of 1000Hz. Although viewing was binocular, only the right eye was recorded (except for five participants who had the left eye recorded due to tracking problems). The stimuli were presented on a BenQ XL2410 T LCD monitor with a 1920 x 1080 screen resolution and 60 Hz refresh rate. Voice recordings were taken with a Fifine USB Microphone –K056 Model device with a lag range of 3 to 24 ms. Only the forehead rest was used to allow unhindered articulation while reading. The passage was formatted in a 22-point monospaced Consolas font and appeared as black letters over a white background. The passage was displayed over two lines. The lines were doubled spaced, justified to the left and presented in the middle of the screen vertically and with a 550-pixel offset horizontally. The eye-to-screen was 70cm and each letter subtended ~ $0.34^{o}$ horizontally.

The experiment was programmed in Matlab R2018a (MathWorks, 2014) using the Psychtoolbox v.3.0.11 (Brainard, 1997; Pelli, 1997) and Eyelink (Cornelissen, Peters, & Palmer, 2002) libraries. The experiment was run on a Windows 7 operating system.

### 3.3.4. Procedure

The experiment began after participants gave written consent and verbal instructions were given. Participants were tested in quiet rooms within the school where they completed two sessions: an eye-tracking experiment and a paper-and-pencil offline assessment of reading and cognitive ability. There was no order to completing both sessions as some completed the eye-tracking experiment first and others the offline assessment first. Participants completed a 9- point calibration and validation procedure at the start of the experiment. Both procedures' accuracies were always < 0.40 and recalibration was done whenever the drift check fell below this level and after a 2-minute break scheduled at the middle of the experiment. Participants' fixation on a 50-pixel black gaze box centred at the first letter of the passage triggered the presentation of the passage. Participants were asked to read the passages aloud and to say "done" once they were finished so that the experimenter could terminate the trial. Participants answered TRUE/FALSE comprehension questions that appeared after 14 passages (33%) by pressing one of two buttons on a keyboard. The experiment began with three practice trials to familiarise participants with the instructions. An example comprehension question for the example 1A and 2A in Figure 3.2 is "*He was not afraid of heights. TRUE/FALSE?"*. Both sessions of the experiment lasted for approximately 1 hour with opportunities for break in between sessions.

### 3.3.5. Data Analysis

Eye movement data were manually pre-processed using Eye-doctor v.0.6.5 (Stracuzzi & Kinsey, 2009) to align fixations vertically with preceding and/or successive fixations on the same line. The EMreading R package software was used to extract fixation data for the analysis (Vasilev, 2018). Audio data were pre-processed manually using the PRAAT software (Boersma & Weenink, 2019). The waveforms, spectrogram and formants were used jointly to determine the onset of articulation for each word in the text (see Appendix E for sample PRAAT annotation). Where co-articulation occurred, effort was made to allocate the boundary to a midpoint between the two words. An R script was developed to merge eye fixation data with the audio textgrids and compute the Spatial EVS.

Linear mixed models were used to analyse the data (lme4 package v1.1-21: Bates, Mächler, et al., 2015) in R software v 4.0.3 (R Core Team, 2020). Grade with three levels (third grade, fourth grade, fifth grade) and number of syllables with two levels (one syllable and two syllables) were treated as fixed effects. Successive difference contrasts were used to compare each level of grade with the next. Sum contrasts were used to compare the two levels of syllable number condition against each other (One: 1, Two: -1). Gaze and articulation durations were log-transformed due to skewed distributions (see Appendix B for untransformed analysis) and spatial EVS was untransformed. Initial phoneme properties determine naming onset times and both initial and final phoneme properties determine the articulation duration (Hutzler et al., 2004). Consequently, whether the initial and final phoneme were voiced or voiceless was included into the articulation duration model as a covariate. Similarly, since the spatial EVS measure commences at the onset articulation, the initial phoneme voice characteristic was also included in its model. These were done because the target words were not controlled for this feature experimentally. Additionally, because word frequency and age of acquisition varied within each condition, they were added as continuous covariates in all models. Due to a marginal difference between syllable number condition for phonological neighbourhood size and item sets for bigram frequency (see Materials), phonological neighbourhood size and bigram frequency were included in the models but were removed if they did not significantly improve the model fit using a likelihood test. All continuous covariates were centred to a mean of zero.

Participants and items were treated as crossed random effects. Initially, we adopted a full random structure (Barr et al., 2013), with random intercepts for participants and items; and a random slope for syllable number for participants. If this model did not converge, we removed random slopes from the model. Only the gaze duration model excluded the random slope for syllable number. The results were considered as statistically significant if the $|t|$ values were $\geq 1.96$. Cohen's d effect sizes are reported for the significant effects.

### 3.4. Results

All participants scored above chance level (50%) on the comprehension questions. (mean= 84.9%; SD: 12.3%; range: 64.3%-100%). Six trials were removed due to tracking loss. Fixations less than 80ms within one-character of a temporally adjacent fixation were merged while other fixations that were less than 80ms were discarded (11.68%). Two data files were created, one for the gaze and articulation duration analysis and the other for the spatial EVS analysis. The following exclusions were made from both data sets: trials in which participants made errors on target words (5.09% and 4.71%), found target words difficult to pronounce (3.05 % and 3.26%), read the trial silently before reading aloud and trials with long pauses and several errors in the trial (1.42% and 1.17%). Additionally, words with blinks (7.07%) were removed from the gaze and articulation duration analysis leaving a total of 1686 data points. Likewise, due to blink artefacts on fixations while target words were being articulated, 4.01% of target words were removed leaving a total of 1898 data points for the spatial EVS analysis. There was no significant difference in the number of errors made for each condition (*b* = 0.05, *SE* = 0.18, *z* =0.29, *p*=0.77) or set (*b* = 0.35, *SE* = 0.21, *z* = 1.66, *p*=0.10)[1], neither was there an interaction (*b* = 0.10, *SE* = 0.18, *z* = 0.56, *p*=0.58). Descriptive statistics for all dependent measures by grade and syllable number are shown in Table 3.3.

The linear mixed model (LMM) results for gaze duration, articulation duration and spatial EVS are shown in Table 3.4, 3.5 and 3.6 respectively and jointly illustrated in Figure 3.3. The gaze duration result revealed that there was a main effect of syllable number (*d*= 0.15). Interestingly, gaze duration was longer when processing one syllable words compared to two syllable words which was opposite to our prediction. Additionally, children in grade four had longer gaze duration on target words than children in grade five (*d*=0.51). However, the gaze durations of children in grades three and four did not differ. Frequency and Age of Acquisition had significant effects on gaze duration of target words in the usual directions. Specifically, gaze duration decreased with greater frequency and increased with greater age of acquisition. The interaction between syllable number and grade was not significant[2].

Table 3.3. Mean and Standard Deviations (in Parenthesis) for Dependent Measures across Grade and Number of Syllables

| Syllable | Third grade | | Fourth grade | | Fifth grade | |
|---|---|---|---|---|---|---|
| | One | Two | One | Two | One | Two |
| Gaze duration (ms) | 587 (458) | 520 (390) | 604 (509) | 530 (440) | 397 (215) | 358 (218) |
| Articulation duration (ms) | 632 (236) | 573 (159) | 623 (271) | 566 (179) | 518 (171) | 490 (133) |
| Spatial EVS* | 8.6 (4.5) | 9.0 (4.0) | 9.0 (4.8) | 9.4 (4.1) | 11.4 (4.7) | 11.4 (4.9) |
| Spatial EVS in words | 1.1 (1.0) | 1.3 (0.9) | 1.2 (1.1) | 1.3 (1.0) | 1.7 (1.1) | 1.7 (1.1) |

*Note*. Spatial EVS was measured as number of character spaces between initial letter of the target word when it was articulated and the fixation location in the passage.

[1] A marginal difference between the two item sets were found, therefore this factor was included in the models if it improved the model fit.

[2] The offline measures in Table 1 were examined in relation to the syllable effect. However, only sight word efficiency subtest of the TOWRE led to a significant interaction (see Appendix B).

Table 3.4. LMM Analyses Showing Gaze duration as a Function of Grade and Number of syllables

| Fixed effects | *b* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 2.607 | 0.017 | 154.505 | <0.01 |
| Syllable number | 0.022 | 0.01 | 2.302 | **0.024** |
| Grade 3-4 | 0.013 | 0.037 | 0.339 | 0.736 |
| Grade 4-5 | -0.147 | 0.035 | -4.233 | **<0.01** |
| Word Frequency | -0.03 | 0.01 | -3.116 | **0.003** |
| AoA | 0.044 | 0.01 | 4.581 | **<0.01** |
| Syllable number * Grade 3-4 | 0.006 | 0.015 | 0.388 | 0.698 |
| Syllable number * Grade 4-5 | 0.002 | 0.014 | 0.141 | 0.888 |
| Random effects | Var. | SD | Corr. | |
| Item intercept | 0.004 | 0.066 | | |
| Subject intercept | 0.009 | 0.097 | - | |
| Residual | 0.054 | 0.234 | | |

*Note*. Statistically and marginally significant *p* values are formatted in bold and italics. Phonological N did not improve the model fit for GD.

Fifth grade children had significantly lower articulation duration than those in the fourth grade (*d*= 0.46). The effect of syllable number and its interaction with grade was not significant. However, age of acquisition influenced articulation duration where early acquired words had a shorter duration compared to late acquired words. Additionally, initial phoneme voicing contributed to the variance in the model where words beginning with voiced phonemes were faster to articulate than words with voiceless initial phonemes.

Children in fifth grade had a larger spatial EVS than children in fourth grade (*d*= -0.47). However, children in third and fourth grade did not differ significantly in the spatial EVS measure. The main effect of syllable number and its interaction with grade was not significant. Furthermore, as age of acquisition of words increased, spatial EVS decreased significantly.

Table 3.5. LMM Analyses Showing Articulation duration as a Function of Grade and Number of syllables

| Fixed effects | *b* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 2.734 | 0.016 | 175.258 | <0.01 |
| Syllable number | 0.004 | 0.008 | 0.515 | 0.608 |
| Grade 3-4 | 0.002 | 0.022 | 0.11 | 0.913 |
| Grade 4-5 | -0.074 | 0.02 | -3.663 | **0.001** |
| Word Frequency | -0.007 | 0.006 | -1.037 | 0.303 |
| AoA | 0.035 | 0.006 | 5.472 | **<0.01** |
| Phonological Neighbourhood size | *-0.011* | *0.006* | *-1.695* | *0.094* |
| Initial phoneme voicing | 0.022 | 0.007 | 3.076 | **0.003** |
| Final phoneme voicing | -0.006 | 0.016 | -0.379 | 0.706 |
| Syllable number * Grade 3-4 | <0.01 | 0.008 | -0.052 | 0.959 |
| Syllable number * Grade 4-5 | -0.005 | 0.007 | -0.762 | 0.45 |
| Random effects | Var. | SD | Corr. | |
| Item intercept | 0.002 | 0.050 | | |
| Subject intercept | 0.006 | 0.060 | | |
| Syllable number slope (subject) | 0.000 | 0.012 | 0.63 | |
| Residual | 0.009 | 0.010 | | |

*Note*. Statistically and marginally significant *p* values are formatted in bold and italics.

Table 3.6. LMM Analyses Showing Spatial EVS as a Function of Grade and Number of syllables

| Fixed effects | *b* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 0.051 | 0.156 | 37.81 | <0.01 |
| Syllable number | 0.191 | 0.594 | 0.324 | 0.747 |
| Grade 3-4 | 2.342 | 0.548 | 0.321 | 0.750 |
| Grade 4-5 | 0.072 | 0.149 | 4.273 | **<0.01** |
| Word Frequency | -0.882 | 0.149 | 0.486 | 0.628 |
| AoA | -0.281 | 0.159 | -5.925 | **<0.01** |
| Initial phoneme voicing | *-0.014* | *0.268* | *-1.77* | *0.081* |
| Syllable number * Grade 3-4 | 0.189 | 0.246 | -0.051 | 0.96 |
| Syllable number * Grade 4-5 | 0.051 | 0.156 | 0.769 | 0.445 |
| Random effects | Var. | SD | Corr. | |
| Item intercept | 0.720 | 0.849 | | |
| Subject intercept | 2.305 | 1.518 | | |
| Syllable number slope (subject) | 0.052 | 0.229 | 0.53 | |
| Residual | 17.230 | 4.151 | | |

*Note*. Statistically and marginally significant *p* values are formatted in bold and italics.

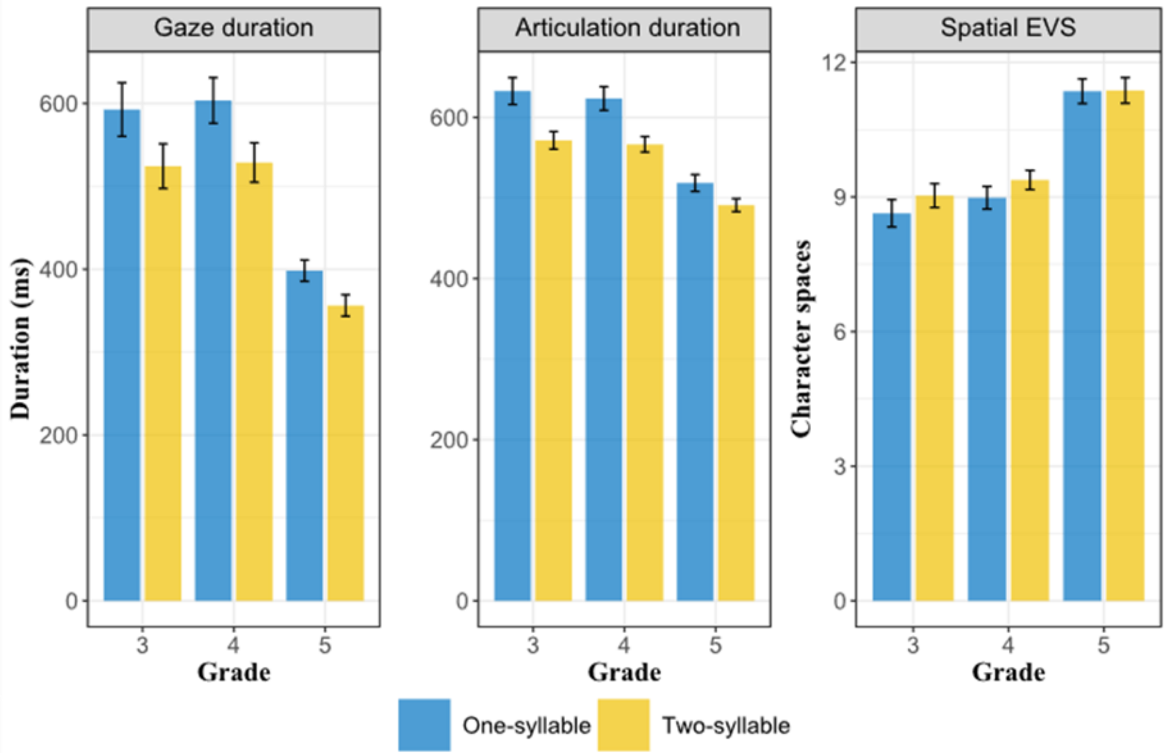


*Figure 3.3. Bar plots showing means and ± 1 SE for gaze duration, articulation duration and spatial EVS on target words by syllable number and grade*

### 3.5. Discussion

The present study investigated changes in syllable use among third, fourth, and fifth grade children by simultaneous eye movement and voice recording. Contrary to our expectations, across all three grades, gaze duration was longer for one compared to two-syllable words suggesting that children spent less time processing two syllables compared to one syllable words. In line with our hypotheses, gaze duration and articulation duration decreased with higher grades but only significantly so for the grade four and five contrast. Similarly, grade five children had a wider spatial EVS compared to grade four children. Additionally, we found no evidence of an interaction between the syllable number manipulation and grade level for any of our measures.

Surprisingly, our results showed that gaze durations were longer for one-syllable words than for two-syllable words. The direction of this effect is in the opposite direction to majority of the previous experimental and corpus analyses of syllable number. In English, words with more syllables taking longer to process than words with fewer syllables is considered typical and inhibitory (Stenneken et al., 2007; Yap & Balota, 2009). Furthermore, this pattern has been documented mostly in adults (Jared & Seidenberg, 1990; but see Lee, 2001; New et al., 2006; Pelczarski et al., 2019; Perry et al., 2010; Yap & Balota, 2009) and scarcely in children (Duncan & Seymour, 2003; Juel & Holmes, 1981). Precisely why our results diverge from this pattern is not immediately obvious. The argument that methodological differences could account for the discrepancy is countered by evidence from Pelczarski et al. (2019) who found an inhibitory trend using eye-tracking, although for naming non-words. Thus, it could be that with less familiar words, inhibitory effects are more likely. This proposition is consistent with data from Jared and Seidenberg (1990) where it took longer to name low frequency words that were two syllables than one syllable. Similarly, Yap and Balota (2009) found larger effects on low frequency words. The use of a sublexical (Coltheart, 2006; Coltheart et al., 2001) or analytic (Ans et al., 1998) route while reading less familiar words may lead to syllabic ac tivation compared to using a lexical route for high frequency words. Although, this explanation seems less likely given the inhibitory pattern found in an analysis of the English Lexicon Project for both high and low frequency words (New et al., 2006; Yap & Balota, 2009) after controlling for a number of variables. Therefore, stimuli characteristics, experimental method and participant sample may interact to account for the observed differences between these studies and the current study (see below for further discussion).

In contrast to these inhibitory effects, Lee (2001) found that skilled adults took longer to name one syllable words compared to two syllable words. The opposite effect was found in a four-letter condition with one and two syllable words. Lee (2001) attributed these findings to the spelling-sound mappings where each phoneme in the two-syllable words mapped on to a single grapheme, whereas some phonemes in one-syllable words mapped on to multi-letter graphemes. Such multi-grapheme mapping was more common in the six-letter condition than the four-letter condition. However, this explanation cannot hold for the current study as the number of phonemes and graphemes were controlled for across the two conditions.

There are other plausible explanations for a facilitatory effect. First, the occurrence of consonant clusters which are more likely for one syllable words (Frederiksen & Kroll, 1976; McLeod & Arciuli, 2009) may slow processing (Bruck & Treiman, 1990; Frederiksen & Kroll, 1976; Saha et al., 2021; Treiman, 1985, 1991). While the CDP++ and MTMM models attribute syllabic effects to presence of more vowel nuclei, it seems that more vowels that are spaced out make reading easier. Gagl et al. (2015) showed that one-syllable German words with consonant clusters (*Herbst*) yielded longer naming latencies than two-syllable words with the same number of letters and phonemes but without clustering (*Mantel*) in developing readers (7 to 9 years). However, this facilitatory effect was driven primarily by six-letter words rather than four-letter words, where a trend of inhibition was apparent (Appendix C; Gagl et al., 2015). This suggest that the direction of the syllable number effect is dependent on word length. The presence of consonant clusters in one-syllable words becomes more likely than two syllable words as word length increases. Therefore, the longer the word, the more likely inhibitory consonant cluster effects will be found. In the current study, 93% of one-syllable words had consonant clusters at the onset, while 2% of two syllable words had consonant cluster. Although Jared and Seidenberg (1990) found inhibitory effects for six letter words which likely had consonant clusters for the one-syllable condition, the words were low frequency and number of phonemes was not controlled for. Two syllable words were more likely to have more phonemes than one syllable words (see Table 5 in Jared and Seidenberg, 1990). Taken together, these studies suggest that the syllable number effect is likely dependent on properties of the experimental stimuli. Therefore, future studies should seek to control for a range of variables that may confound the syllable effect. Another consideration may need to be given to the initial phoneme which impacts naming latencies. However, because ensuring control on a range of factors (see Table 3.2) and holding the sentence frames constant (see Figure 3.2) across the conditions was central to our design, initial phoneme control proved difficult to achieve. This was exacerbated as our items were limited to grade two words available in the CPWD. Thus, future studies may consider the inclusion of tightly controlled experimental stimuli of non-words with a learning phase.

A second potential account is that initial syllables may serve as crucial activating units during visual lexical access. The basic orthographic syllable of a word activates lexical representation of words (see Taft, 1986). Similarly, first syllable frequency influences lexical access (Alvarez et al., 2001; Carreiras et al., 2005; Croot et al., 2017; Hutzler et al., 2005; Macizo & Van Petten, 2006). Therefore, if first syllables are “access codes” used during multisyllabic word reading (Taft, 1986), then two syllable words whose first syllable is likely four letters or less may have been activated much quickly than one syllable words whose first syllable would always be the whole word (i.e., six letters in the current study). In addition, the recognition of the first syllable of a two-syllable word would potentially activate more lexical candidates; for example, *pur-* may activate *purple*, *purpose*, *pursuit* etc., leading to facilitation in English (Macizo & Van Petten, 2006) and is consistent with our findings. Studies by Milledge et al. (2021) and

Pagan et al. (2016) which show significant costs to substituting and transposing the first three letters of a two-syllable word in the parafovea may provide support that indeed beginning letters are crucial for lexical access. However, the extent to which such syllable information can be processed parafoveally needs to be empirically tested in children (see Ashby & Rayner, 2004 for evidence of parafoveal processing of syllables in adults).

It is perhaps interesting to consider why our findings differ from previous eye-tracking evidence on syllable number in skilled adult readers during sentence reading. Previous studies have found null effects on fixation duration measures (skilled adult readers: Ashby & Clifton, 2005; Drieghe et al., 2019). This pattern could be due to the silent reading modality employed in such studies and the possibility that for adults, syllable number may not impact upon the decisions to trigger a saccade to the next word. Therefore, syllable effects could be more obvious during reading aloud with children (Juel & Holmes, 1981), because they have a greater sensitivity to syllabic representations than adults (Colé et al., 1999; Hasenäcker & Schroeder, 2016).

We failed to find syllable number effects on articulation duration Similar to Gagl et al. (2015), this contradicts the idea of intra-word syllabic pausing in typically developing readers (Proença, 2018). Critically, we failed to find evidence that syllable effects influence stages beyond word identification as spatial EVS did not differ significantly between our syllable conditions. This null effect is perhaps consistent with findings on delayed naming where syllable effects are absent (Juphard et al., 2006; Klapp et al., 1973).

The reduction of gaze duration with grade likely results from an increase in lexical processing efficiency (Mancheva et al., 2015; Reichle et al., 2013). This finding agrees with previous eye movement research (Blythe & Joseph, 2011; Kim et al., 2019; Vorstius et al., 2014). Analysis of the sight word reading age equivalent scores across the three grades (see Table 3.1) revealed no difference between grades three and four, but a significant difference between grades four and five similar to our eye-movement reading measures. This corroborates the finding that reading proficiency rather than age-related changes explain changes in oculomotor control during reading (Blythe & Joseph, 2011). A similar grade effect was found for articulation duration where this time was shorter for children in grade five than grade four but not significantly different between grades three and four. Therefore, other stages beyond phonological encoding, word recognition and semantic integration such as articulatory planning and production reflect lexical skills rather than just speech maturation that comes with age. While evidence points to an age dependent decrease in speech and articulation rates (Hulme et al., 1984; Logan et al., 2011; Sturm & Seery, 2007) resulting from maturation of speech motor organs (Kent, 1976), our findings support the view that cognitive skills in the form of reading expertise also shape articulatory processes. Consistent with this view, Popescu and Noiray (2021) show that better readers show less intersegmental coarticulation (a measure of speech fluency) while repeating non-words. Additionally, adolescents aged 12 to 15 years with reading difficulties make more speech production errors during picture naming and repetition of multisyllabic words and phrases (Catts, 1986). Such a proposal is consistent with the idea that the language domain influences speech production processes (German & Newman, 2007; Saletta et al., 2015).

For the first time in a little over a century, we document grade effects on the spatial EVS using modern eye-tracking equipment. The results replicate previous research showing that the EVS increases with an increase in grade ((Buswell, 1920; Holgerson, 1977; Levin & Turner, 1966). More importantly, it indicates that reading proficiency may be more vital than grade as no difference in the EVS was found for grades three and four. At the onset of target word articulation, the eyes of grade five children were two character spaces ahead than grade four children. Further analysis using the EVS in words revealed that approximately one word intervened between the eye and the voice for grade four and two words for grade five (see Table 3.3). This suggests that the spatial EVS reflects ease of word recognition and oral reading fluency. The eyes travelling further ahead of the voice in grade five readers compared to grade four readers also reflects increased capacity to hold information in working memory for later articulation (Laubrock & Kliegl, 2015). Combined with historical evidence, our findings support Buswell’s finding that between grade three and five (US grades two, three and four), the average EVS is about 10 characters (Table 1 & 2; Buswell, 1920). However, our findings are at odds with an EVS estimate of three words for grade two children measured using the lights-off paradigm (Levin & Turner, 1966) and suggest that this paradigm likely overestimates the EVS even for a developmental sample (see Laubrock & Kliegl, 2015 for evidence in adults).

Other landmark effects in word recognition and reading research exerted strong influences on one or more of our dependent variables. Word frequency and AoA significantly predicted gaze duration replicating prior research (Coltheart et al., 1988; Dirix & Duyck, 2017; Joseph et al., 2013; Juhasz & Rayner, 2003; Juhasz & Rayner, 2006). However, only AoA significantly predicted articulation duration and spatial EVS of the target words. It has been suggested that early acquired words may serve as hubs within semantic networks (Steyvers & Tenenbaum, 2005), which provides them a benefit, above and beyond word frequency, in word recognition tasks. Therefore, if early AoA words hold a special place within semantic networks, they may also lead to benefits in semantic integration with the other words of the sentence resulting in benefits beyond gaze duration. This may provide some basis for the differentiation of AoA effects above frequency effects for developing readers (Juhasz & Sheridan, 2020; Zevin & Seidenberg, 2004).

The present study found no evidence that the use of syllable information as indexed by syllable number, changes across grades three, four and five. English is considered a deep orthography with inconsistent spelling-to-sound mappings where the same spelling may yield different pronunciations and the same pronunciations may yield different spellings. Therefore, readers may prefer to use large grain size earlier in reading development. If so, then readers in higher grades may respond to one and two syllable words as whole units compared to readers in lower grades. The absence of an interaction between syllable number and grade may indicate that the use of syllable units is relatively stable from grades three to five, similar to what has been found in Spanish (Alvarez et al., 2016) and German (Hasenäcker & Schroeder, 2016). Further exploratory analysis of the offline ability measures showed that only the sight word efficiency subtest of the TOWRE interacted significantly with syllable number within our sample (see Appendix B, Figure B1). Previous evidence examining reading skill differences in the syllable number effect have yielded inconsistent results in adults (Butler & Hains, 1979). Butler and Hains (1979) reported an interaction between syllable number and vocabulary score on naming latency only when the main effect of syllable number was excluded from the model. Therefore, more research is needed using a larger sample. Additionally, future studies may investigate changes in the use of syllable units outside these three grades i.e., grade two and six and compare to a group of skilled adult readers. This will substantiate whether the syllable effect has reached a point of stability after grade three or it is still developing after grade five.

### 3.5.1. Conclusion

This study was first to examine syllable number effects in English developing readers using eye movement and voice recording. Critically, we found a stable effect of syllable number among children in grades three to five where it took longer to process six letter one-syllable words than corresponding six letter two syllable words. These findings further our understanding of word recognition during oral sentence reading and suggest that children are sensitive to the complexity of sub-syllabic units (i.e., consonant clusters). Despite advancement in reading instructional methods over the last century and our use of more sophisticated eye-tracking equipment, the average EVS estimate obtained for our sample agrees closely with Buswell (1920). Such reliability over time justifies the call for more research to disentangle specific aspects of cognition and written language processes that the EVS reflects.

### 3.6. Chapter Overview

The syllable effect on gaze duration found in this Chapter provided some evidence that syllabic or intra-syllabic representations are activated during word recognition processes. However, one important finding from this Chapter, was the grade-to-grade change in the EVS that was largely dependent on reading skills. Grades three and four did not differ in any of the dependent measures or word reading skill measure. However, differences were found between grades four and five on all the dependent measures and word reading skill. This study revealed that the EVS estimate within the current sample was very reliable, agreeing closely with research conducted over 100 years ago. With such reliable estimates, the next Chapter focused on how the EVS modulates eye movement behaviour between lines compared to within lines in order to address some of the hypothesis generated in the Discussion section of Chapter 2.