Reading through the lifespan: Individual differences in psycholinguistic effects

Rob Davies*

Lancaster University

Ruth Arnell

Exeter University

Julia M.H. Birchenough

Oxford Brookes University

Debbie Grimmond

Reading University

Sam Houlson

Bristol University

Running head: Reading through the lifespan

Author Note: Rob Davies, Department of Psychology, Lancaster University; Ruth Arnell, Department of Education, Exeter University; Julia Birchenough, Department of Psychology, Social Work and Public Health, Oxford Brookes University; Debbie Grimmond, Institute of Education, Reading University; Sam Houlson, Norah Fry Research Centre, Bristol University. The research reported in this article was supported by Early Career Research and Central Research fund awards to Rob

Davies, at Oxford Brookes University. Preliminary findings on the pronunciation data were reported at the meeting of the Society for the Scientific Study of Reading in Berlin, July 2010. Some lexical decision observations were reported in dissertations submitted in partial fulfilment of MSc degree requirements at Oxford Brookes University by Debbie Grimmond and Sam Houlson. Sub-sets of the pronunciation observations were collected by Rebecca Baxter, Emma Findlay, Hajyra Ilyas, Claire Phelan and Rosanna Sparks, under the supervision of the corresponding author, as part of undergraduate research projects at Oxford Brookes University.

Correspondence concerning this article should be sent to Rob Davies, Department of Psychology, Lancaster University, Bailrigg, LA1 4YF, UK. E-mail: r.davies1@lancaster.ac.uk.

Reading through the lifespan: Individual differences in psycholinguistic effects

Abstract

The effects of psycholinguistic variables are critical to the evaluation of theories about the cognitive reading system. However, reading research has tended to focus on the impact of key variables on average performance. We report the first investigation examining variation in psycholinguistic effects across the life-span, from childhood into old age. We analysed the performance of a sample of 535 readers, aged 8-83 years in lexical decision and pronunciation tasks. Our findings show that the effects on reading of two key variables, frequency and AoA, decrease in size with increasing age over the life-span. We observed the systematic modulation by age and reading ability of these and other psycholinguistic effects alongside a global U-shaped effect of age. Diffusion model analyses suggest that developmental speed-up in decision responses can be attributed to the increasing quality of evidence accumulation in reaction to words, while the ageing-related slowing can be attributed to decreasing efficiency of stimulus encoding or response execution processes. An analysis of spoken response durations furnishes a consistent picture in which the slowing of pronunciation responses with age can be attributed to slowing articulatory processes. We think our findings can be explained by theoretical accounts that incorporate learning as the basis for the development of structure in the reading system. However, an adequate theory shall have to include assumptions about both developmental learning and later ageing. Our results warrant a life-span theory of reading.

Keywords: reading; individual differences; development; aging; word recognition

Reading through the lifespan: Individual differences in psycholinguistic effects

Children can read most words presented in isolation by about nine years of age. What happens then? The adult reader surely knows many more words, can read them more quickly and can extract more information about them. Are there qualitative differences between the child, the young adult and the older adult reader? Or can we assume that development terminates in a mature reading system that thereafter varies only quantitatively?

We conducted a study to map variation in the cognitive reading system. We tracked the state of the system by estimating item effects on reading performance due to critical word properties, focusing on effects of word frequency and Age-of-Acquisition (AoA). Our study is the first to examine variation in psycholinguistic effects from childhood into old age. To anticipate our results, we found that the frequency and AoA effects were smaller for older readers. We argue that the modulation of frequency and AoA effects by age can be explained only in models of the cognitive reading system that assume that these effects reflect the impact of learning on the structure of reading processes or representations. However, a complete empirical account of reading over the lifespan must incorporate the fact that the psycholinguistic effects, and the interactions reflecting their modulation by individual differences, are observed in the context of a large, global, U-shaped effect of age on response latencies. Latencies first decreased from childhood into adulthood, and then gradually increased in association with increasing age. Diffusion model analyses show that the U-shaped effect of age on lexical decision latencies can be explained by age-related variation in the quality of information extracted given the stimulus, and in perceptual stimulus encoding or response execution speed. Analyses of spoken response durations show that age-related slowing in pronunciation can be attributed to differences in response execution efficiency. We discuss later how a theoretical account of reading could embed or integrate an explanation of systematic individual differences in an explanation of cognitive development and ageing over the life-span.

Psycholinguistic effects and the emerging, skilled and aging reading system

Experimental research has employed simple tasks like word naming or lexical decision to uncover the properties of the reading system. In these tasks, words are presented in isolation so that the demands on the reading system are narrowed to probe the most basic functions: visual recognition and the encoding of pronunciations. Experimental evidence has accumulated to show that the average healthy young adult is faster to respond to words that occur more frequently or in more linguistic contexts (e.g. Adelman, Brown, & Quesada, 2006; Balota & Chumbley, 1984; Brysbaert & New, 2009), to words that are shorter and look similar to more other words (e.g. Andrews, 1989, 1992; Yarkoni, Balota, & Yap, 2008), that have referents that are easier to imagine (e.g. Strain, Patterson, & Seidenberg, 1995), that were learnt earlier in life (e.g. Cortese & Khanna, 2007; J. Monaghan & Ellis, 2002), and that have pronunciations that obey the rules for the spelling-sound mappings of constituent graphemes (e.g. Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) or that are consistent with the pronunciation of similar-looking words (e.g. Andrews, 1982; Glushko, 1979; Taraban & McClelland, 1987).

Knowing what item attributes affect reading performance has motivated and constrained the development of theories about how cognitive reading processes function (e.g. Coltheart, 1978; Coltheart et al., 2001; Glushko, 1979; Harm & Seidenberg, 2004; McClelland & Rumelhart, 1981; Morton, 1969; P. Monaghan &

Ellis, 2010; Murray & Forster, 2004; Norris, 2006, 2009; Perry, Ziegler, & Zorzi, 2007; Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989). In the last 30 years, the assumptions of theoretical accounts have been tested in computational simulations of psycholinguistic effects but, with some exceptions (Adelman, Sabatos-DeVito, Marquis, & Estes, 2014; Dilkina, McClelland, & Plaut, 2008; Ziegler, Castel, Pech-Georgel, Alario, & Perry, 2008) discussed later, the evaluation of models and thus of theories has concerned their capacity to simulate the influence of psycholinguistic effects on the *average* performance of *adult* readers. Current theories, and their implementations, may account for the development of skilled performance (e.g. Seidenberg & McClelland, 1989) but have nothing to say about healthy ageing. In contrast, there are theoretical accounts of ageing (general accounts, e.g. Hale, & Myerson, 1996; Salthouse, 1996; accounts concerned with reading, e.g. Spieler & Balota, 2000) whose remit begins with skilled performance (though see e.g. Cerella & Hale, 1994). At present, there are no theoretical accounts of the development of the reading system from childhood to old age. Should there be such an account, and what form should it take?

We argue that a lifespan account of the cognitive reading system would be warranted by observations indicating that psycholinguistic effects are substantially modulated by age. We focus on the effects of frequency and AoA because empirical accounts of reading phenomena mandate the investigation of both factors together (e.g. Cortese & Khanna, 2007; Morrison & Ellis, 1995) given their inter-relation as measures, and because theoretical accounts place one or both factors among the key influences on the functioning of the cognitive reading system (e.g. Coltheart et al., 2001; P. Monaghan & Ellis, 2010; Seidenberg & McClelland, 1989). We set out the theoretical context for our investigation by discussing the predictions that can be derived from existing assumptions. Differences between the two most influential accounts of the reading system, the dual route model (Coltheart, 1978; Coltheart & Rastle, 1994; Coltheart et al., 1993; Coltheart et al., 2001) and the connectionist 'triangle' model (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989) imply diverging predictions concerning how item effects should vary as a result of individual differences.

Theoretical accounts of the cognitive reading system

The dual route model (Coltheart et al., 2001) assumes that the reading system operates over symbolic representations of knowledge about letters, words and grapheme-phoneme correspondences (GPCs). In a dual route system, in reading aloud, lexical and non-lexical routes both feed activation to the phoneme level but responses to non-words are assembled from the serial activation of phonemes through GPCs. Responses to words require the lexical route, in which word knowledge is represented with phonological and orthographic units. A lexical unit's activation is a function of inputs plus a frequency constant dependent on the corresponding word's estimated frequency of occurrence, multiplied by a scaling parameter. The contribution of the constant explains the frequency effect in the dual route account. Critically, for our concerns, the connections in the lexical route (in DRC and CDP+) are pre-specified, not learned. The verbal theory of the dual route account of reading (Coltheart et al., 2001; see also Castles & Coltheart, 1993) assumes that development involves the acquisition of the lexical and non-lexical routes but the computational model implementing the theory is explicitly not adaptive, at least, not in the lexical route (cf. CDP+ learning in the non-lexical route, Perry et al., 2007). We may, therefore, derive predictions about variation in psycholinguistic effects but such predictions are conjectures without simulation studies to test them.

We contend that as the frequency effect in the DRC results from the contribution of a frequency scaling variable to lexical unit activation, it can be predicted that the frequency effect should grow in size over development, as the performance of the reader grows to depend on a lexical reading route. Once the reading system has matured, the assumptions of the theory do not warrant the expectation of further change through adulthood into old age. However, Balota and colleagues (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Spieler & Balota, 2000) suggested that continued exposure to known words, along with exposure to new words, could be associated with a process of unitization in which the representations supporting reading are gradually compiled into unitary representations rather than sets of sublexical representations. Such a process would be associated with continued growth in the importance of lexical effects, like the frequency effect, through adulthood into old age. The dual route account of reading (Coltheart et al., 2001) does not address the AoA effect but the effect can be explained in the account by assuming that lexical activation is scaled by AoA and frequency. Granted that extension, we can predict that both AoA and frequency effects grow in size through development but then either plateau in adulthood or continue to grow into old age.

In contrast to existing dual route accounts (with the partial exception of CDP+, Perry et al., 2007), connectionist models are designed to learn from experience. Though connectionist models vary between implementations, the predictions that can be derived concerning individual differences in psycholinguistic effects flow from the principles of nonlinearity, adaptivity and distributed representations governing the approach (Harm & Seidenberg, 2004; Plaut et al., 1996; Seidenberg & McClelland, 1989). Connectionist models implement the assumption that the reading system operates over networks of sub-symbolic representations of word orthography, phonology and semantics (semantic representations are implemented in Harm & Seidenberg, 2004). Exposure to a word will cause changes to the weights on network connections, with these adaptations reducing error in output because the weight changes result in increased input to output units that should be active and decreased input to output units that should be inactive. A word that is experienced more frequently will have more opportunities to drive helpful weight changes. However, the nonlinearity of the function linking input to output activation means that output activation will tend to asymptote towards 0 or 1 as input activation increases, ensuring that increased input activation will translate into progressively smaller decrements in error (Plaut et al., 1996). This predicts, in general, that the effects of those item attributes that influence the efficacy of mappings in the connectionist reading system should decrease in size as experience accumulates.

Reading performance is shaped not just frequency but also by AoA (see, for reviews: Johnston & Barry, 2006; Juhasz, 2005), and the effects of both variables may be predicted to reduce with increasing age. It has been argued that the AoA effect can be taken to reflect the influence of those variables that determine the relative ease with which words can be learnt (Zevin & Seidenberg, 2002, 2004; see, also, Bonin, Barry, Méot, & Chalard, 2004; Bonin, Méot, Mermillod, Ferrand, & Barry, 2009; Mermillod, Bonin, Méot, Ferrand, & Paindavoine, 2012). However, computational simulations reported by P. Monaghan & Ellis (2010) demonstrate that the AoA effect is observed, independent of potentially confounding factors, where a network is trained with words encountered in an order corresponding to their age-graded frequency of occurrence. P. Monaghan and Ellis (2010) showed that a word’s point of entry to a network's training regime influenced the network's performance over and above the impact of the word's continuing frequency of encounter. This order effect results, as experience accumulates, from a reduction in the plasticity of network connections and from the adaptation of the network to support responses to earlier-acquired items. It has been argued, however, that the AoA effect would diminish over increasing age if it reflects the impact of readers' cumulative experience of words (Barry, Johnston, & Wood, 2006; Ghyselinck, Lewis, & Brysbaert, 2004; Morrison, Hirsh, Chappell, & Ellis, 2002). Readers who learn some words earlier in life will tend also to accumulate more experience of those words. Morrison et al. (2002) and Barry et al. (2006) noted that absolute differences in AoA may be the same for older and younger adults but that differences in accumulated frequency would be smaller for the older adults, in proportion to their greater lexical experience overall. Thus, if AoA effects owe something to differences in cumulative frequency then the AoA effect should be smaller for older readers.

While we have drawn out predictions from dual route or connectionist assumptions that the frequency and AoA effects may change over the lifespan, Murray and Forster's (2004) account of lexical access supposes that the frequency effect does not change. Murray and Forster (2004) proposed that the frequency effect reflects the ordering of sets of candidates for recognition by the relative frequency of words. Lexical access occurs at the successful termination of a search through lexical entries, where those entries are ordered by relative frequency, and the search begins with the highest frequency words. Murray and Forster (2004) asserted that search based theories predict the frequency effect should be independent of absolute frequency. It should not change with increasing overall experience.

Observations concerning individual differences in the effects of frequency and AoA

Current theoretical accounts sustain a variety of predictions concerning individual differences in the frequency and AoA effects. That variety is matched by a remarkable inconsistency among empirical observations. Previous studies have reported similar frequency effects on reading in younger and older children (reading in Italian, Burani, Marcolini & Stella, 2002), and in younger and older adults (reading in English, Allen, Madden, & Crozier, 1991; Allen, Madden, Weber, & Groth, 1993, expt. 2; Bowles & Poon, 1981; Cohen-Shikora & Balota, 2016; Tainturier, Tremblay & Roch Lecours, 1989). In these studies, the frequency effect is not significantly modulated by age while in other studies the frequency effect has been found not to interact either with differences in adult vocabulary size (Butler & Hains, 1979) or variation in adult print exposure (Lewellen, Goldinger, Pisoni, & Greene, 1993).

Balota and colleagues have reported larger frequency effects in older compared to younger adults' reading (Balota & Ferraro, 1993, 1996; Balota et al., 2004; Spieler & Balota, 2000; but see Cohen-Shikora & Balota, 2016). Thus, Balota et al. (2004), found that the frequency effect is larger in older adults' lexical decision and pronunciation latencies. However, Yap, Balota, Sibley, and Ratcliff (2012) observed that the frequency effect was smaller in word naming but (slightly) larger in lexical decision in young adults with higher vocabulary levels. In a number of studies, the frequency effect on reading has been observed to be smaller over increasing age, reading skill or print exposure. Allen and colleagues found a smaller frequency effect for older compared to younger adults in visual word recognition (Allen et al., 1993, expt. 3), Tainturier, Tremblay and Roch Lecours (1992) found a smaller frequency effect in lexical decision for more educated adult readers, Morrison et al. (2002) observed a frequency effect in young but not older adults' word pronunciation, and Sears, Siakaluk, Chow, and Buchanan (2008; also, Chateau and Jared, 2000), found a smaller frequency effect on lexical decision among adults with higher print exposure.

This variation in observations means it is unclear how the frequency effect varies among readers. This lack of clarity is important because, as we have seen, different accounts of the cognitive reading system predict different trends. There have been only two investigations of variation in the AoA effect. These have indicated that the AoA effect is similar in younger and older adults for word naming (Barry et al., 2006; Morrison et al., 2002) and lexical decision (Barry et al., 2006). However, no previous study has examined if the AoA effect changes from childhood to old age.

Inconsistencies among previous observations may result from the limitations inherent in comparisons between group-level average effect estimates. Abstracting continuous variation to group differences will reduce the sensitivity of analyses not just because more participants must be tested to detect effects (Cohen, 1983) but also because effects may be missed where the influence of individual differences is curvilinear (Cohen, Cohen, Aiken, & West, 2003). This is relevant because a review of the effect of age on cognitive processing over the lifespan (Cerella & Hale, 1994) showed that age-related variation in response speed is best described by a U-shaped curve, as we discuss next. However, inconsistencies among previous observations may result, also, from limitations in the range of ages or reading abilities sampled in previous studies. If age-related changes are confined to specific phases of development or ageing, then the age ranges in which reading is tested may have a critical influence on the character of the item effects observed. Our study addressed both limitations by examining the effect of age as a continuous variable and sampling readers from childhood to old age.

Examining individual differences among readers

Any variation in critical psycholinguistic effects over differences in age must be understood in the broader context of large-scale age-related changes in cognition. Hartshorne and Germaine (2015) reported that performance on working memory tasks peaked at around 30 years, later than performance on processing speed (digit symbol coding, peak around 20 years) and earlier than performance on vocabulary (peak at 65 years). Such observations are broadly consistent with those, in previous reports based on cross-sectional and longitudinal comparisons (e.g. Horn, 1982; Rabbitt, Diggle, Holland, & McInnes, 2004; Rabbitt et al., 2004; Salthouse, 2004, 2014), indicating long term stability or a slow rise in vocabulary knowledge up to ages in the 60s. In comparison, it has been generally observed that ageing is associated with slower response speed (Cerella, 1985; Salthouse, 2004) in simple cognitive tasks.

It has been argued that age-related cognitive changes are dominated by a single slowing factor (Salthouse, 1996; Verhaegen & Salthouse, 1997), given evidence for associations or commonalities among cognitive speed measures, and between cognitive speed and other ability measures. However, analyses reported by Hale, Myerson and colleagues suggest distinct rates of age-related slowing in lexical and non-lexical processing domains (Hale & Myerson, 1996; Lima, Hale, & Myerson, 1991; cf. Cerella, 1985), with greater slowing for responses in non-lexical tasks like line discrimination than in lexical tasks like lexical decision. This claim appears to be consistent with evidence for age-related slowing in stimulus encoding and response output processing but similar activation of lexical or semantic information in younger compared to older adults (Allen, Bucur, Grabbe, Work, & Madden, 2011; Allen et al., 1991; Allen et al., 1993; Bowles & Poon, 1985; Madden, 1992). The marked slowing reported in stimulus encoding or response execution may be linked to observations of age-related changes in visual-sensory (Faubert, 2002; Schneider & Pichora-Fuller,

2000) or orthographic processes (Allen et al., 1993; Allen et al., 2011; Madden, 1992) and in response execution (Allen et al., 1993; Stelmach, Goggin, & Amrhein, 1988).

Characterizing age-related changes in terms of substantial differences in more peripheral, stimulus encoding or response production, processes and limited differences in more central, lexical, processes resembles a diffusion model account of age-related changes in lexical decision (Ratcliff, Thapar, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2010). The diffusion model (Ratcliff, Gomez, & McKoon, 2004) assumes that the mechanism underlying a binary choice like lexical decision is the accumulation of noisy information from a stimulus over time. In a trial, information accumulates towards the word or the non-word response criteria at a rate, the drift rate, determined by the quality of information produced from processing the stimulus. When a criterion is reached, the response is initiated. Ratcliff et al. (2004) showed that the effect of word frequency on lexical decision performance, across a series of experiments, could be simulated by the diffusion model given only variation in drift rate values, with higher drift rates associated with higher compared to lower frequency words. Critically, the diffusion model furnishes an explanation for the changing effect of age on cognitive response speed over the lifespan.

Studies of the effect of age in development have shown a curvilinear increase in response speed, age-related speeding, comparing children with young adults (Hale, 1990; Kail, 1986, 1991). Analyses of data from multiple studies on ageing have shown a curvilinear decrease in response speed, age-related slowing, comparing younger with older adults (e.g. Hale, Myerson, & Wagstaff, 1987). Very few researchers have examined cognitive speed from childhood to old age, within the same study, using the same task, but Cerella and Hale (1994) noted that a U-shaped effect of age was evident in the speed of response in a simple choice task (Noble, Baker, & Jones, 1964), in the Stroop task (Comalli, 1965), and in simple matching tasks (Hale, 1990; Hale, Lima, & Myerson, 1991). This U-shaped effect can be taken as the sum of two age functions operating over the lifespan. Cerella and Hale (1994) argued that the age-related change in response speed reflects quantitative changes in processing rate rather than qualitative differences in process structure. In this account, the developmental increase in response speed corresponds to the development of stimulus encoding and response execution processes, as well as more central stimulus evaluation and response selection processes. The slowing of processing speed, comparing younger to older adults, corresponds to a decline in central response selection processes. Broadly, this account can be related to a diffusion model theory.

In a series of studies, Ratcliff and colleagues showed that the U-shaped effect of age on lexical decision times can be explained by variation between individuals in drift rate parameters. Importantly, their findings demonstrated that the slower decision speeds of children and older adults, compared to young adults, are explained by distinct sets of factors. Ratcliff, Love, Thompson, and Opfer (2012) showed that children were slower than adults because they extracted lower quality information from stimuli (as indicated by lower drift rates), were more conservative in making decisions (they had wider boundary separations), and were slower at stimulus encoding and response output (longer non-decision times). For lexical decision data observed using a different stimulus set, Ratcliff et al. (2010; Ratcliff, Thapar et al., 2004) showed that older adults' decisions were slower than those of younger adults because, while the quality of information extracted from stimuli was similar among individuals in the different age groups, the older adults preferred more conservative decision-making, and were slower in stimulus encoding or response execution.

The aim of our investigation was to better understand the lifespan development of the reading system, from childhood to old age. Our primary focus was on age-related variation in the frequency and AoA effects but previous findings mandate that any observed variation in item effects should be examined in the context of large-scale age-related differences in reading performance. Therefore, we studied reading in children, younger adults and older adults using the oral reading (word pronunciation) task and the lexical decision task. As far as we know, no previous single study has examined visual word recognition across the lifespan. Taken together, the findings from the diffusion model analyses indicated that the developmental speed-up and ageing-related slow-down in decision latencies result from different sources of variation in the decision-making process (Ratcliff et al., 2012; Ratcliff et al., 2010). This warranted an analysis of the effects of individual differences on the parameters of the diffusion model of lexical decision. Given previous observations, we supposed that variation in diffusion model parameters could explain both an expected U-shaped effect of age on decisions in our study, and the modulation of the frequency and AoA effects by age.

However, the diffusion model is fitted to binary choice tasks and cannot directly inform our understanding of age-related changes in word pronunciation. Connectionist simulations (Plaut et al., 1996; Seidenberg & McClelland, 1989; Zevin & Seidenberg, 2002) show that the accuracy of output activation improves towards asymptote as training experience accumulates. We would predict, then, that reading performance should improve with increasing age in lexical decision and in pronunciation. A diffusion model account can incorporate this developmental trend in terms of growth in the quality of information extracted from word stimuli. And it can explain the later age-related slowing of decisions in terms of variation in boundary separation or non-decision (stimulus encoding and response execution) time. What about potential age-related slowing in pronunciation? If we located the impact of ageing on pronunciation in the activation of representations in the word recognition system, a connectionist account of that system would predict only decreasing pronunciation latencies with increasing age. We investigated the possibility that age-related differences in response execution could help to explain slowing in pronunciation as well as in lexical decision. We did this by analyzing not just the latency but also the duration of spoken word responses in the pronunciation task.

We reasoned that if variation in response duration reflected not just overt articulation speed but also the efficiency of articulatory coding, and if some of the effects observed to influence pronunciation latencies reflected not just lexical access but also response execution processes, then those effects should influence response duration also. Evidence from speech production corpora suggests that word durations are shorter for high frequency (Bell et al., 2003; Gahl, 2008; Gahl & Strand, 2016; Moers, Meyer, & Janse, 2016; but see Seyfarth, 2014) or high neighbourhood density words (Gahl, Yao, & Johnson, 2012; Gahl & Strand, 2016), though recent studies of single word production have indicated a limited influence on durations due to lexical frequency (Mousikou & Rastle, 2015). Previous research is limited on how individual differences influence durations. Critically for our concerns, Moers et al. (2016; see, also, Balota & Duchek, 1988; Huttenlocher, 1984) reported that, in a corpus of recordings of continuous text reading, older children produced shorter pronunciation durations than younger children, while older adults produced longer durations than younger adults. These findings predict a U-shaped effect of age on pronunciation durations, along with lexical neighbourhood and perhaps frequency effects.

The present study

We conducted our study to address the question: Are there qualitative or quantitative differences between the child, the young adult and the older adult reader? We examined variation in the cognitive reading system by estimating effects of critical word properties on reading performance in lexical decision and pronunciation, focusing on the influence of frequency and AoA. We tested variation in psycholinguistic effects over a broad age range, with a large sample of participants, to examine if the moderation of critical item effects by age was limited to a specific interval. We tested linear and curvilinear effects of variation in age and ability so that we could, firstly, accurately estimate the effects of individual differences and, secondly, examine if the moderation of psycholinguistic effects by those individual differences was different in different intervals of age or ability. Critically, we estimated interactions between the effects of word properties and the effects of reader attributes to test the modulation of psycholinguistic effects by age. Our analyses primarily concerned the factors that influenced reading response latencies but to gain insight into the impact of age on visual word recognition, we examined the factors that influenced diffusion model parameters calculated from lexical decision data. To gain insight into the effect of age on pronunciation, we examined the factors affecting both the latency and the duration of spoken responses. Differences in age can be associated with differences in reading skill but while the accumulation of experience may drive the performance of the reading system towards maximal efficiency, the accumulation of age can also result in slowing response speed. This warrants an examination of individual differences simultaneously across age and skill. In the following, we report such an investigation.

Method

Participants

We tested 609 participants, 207 on a word naming task and 402 on a lexical decision task. Of these, we excluded the data for 71 participants before analysis: 30 whose first language was not English and 41 who had been diagnosed with dyslexia. Data for one participant were lost through experimenter error. Data for a further two participants were excluded before analysis due to the presence of missing values for those participants on the test of phonological awareness. The results we report were yielded by analysis of the data for the 535 participants remaining after these exclusions. (We repeated our analyses with all participants for whom we had complete data, including dyslexic readers and speakers of English as a second language, and found no differences in critical results; see Supplementary Materials.)

Participant recruitment and testing procedures were approved by the Oxford Brookes University Research Ethics Committee. Adult participants were recruited from the local Oxford population, from the Oxford Brookes University community, and from colleges and businesses in South-eastern England. Children were recruited from schools in Oxfordshire, Berkshire and South Eastern Ireland. Permission was obtained from schools to conduct the research with their pupils and consent was obtained from children's parents prior to testing as well as from the children themselves at the start of test sessions. All participants had normal or corrected-to-normal vision.

All participants were tested individually in a quiet room, and completed three tasks in random order: the experimental reading test, either lexical decision or pronunciation (reading aloud); standardized tests of word naming and non-word naming reading skill (TOWRE Sight Word Efficiency and Phonemic Coding sub-tests, Torgessen, Wagner, & Rashotte, 1999); and a test of phonological awareness skill (the Spoonerisms sub-test, Phonological Assessment Battery, Frederickson, Frith, & Reason, 1997). The Spoonerisms task assesses phonological awareness by requiring participants to replace the first sound of a word with a new sound or to swap the initial sounds in two words to produce two new words. Participants were scored on accuracy for a set of 20 items that were presented orally. In each TOWRE sub-test, participants were instructed to name printed lists of stimuli as quickly and accurately as possible; there were 104 words and 63 non-words. Performance was scored on accuracy, and reading times were recorded if the test was completed within 45s. We note that 76 participants completed the TOWRE sight word test in less than 45s, 96 completed the phonemic coding sub-test in less than 45s. We calculated a skill measure for performance in each TOWRE sub-test by dividing accuracy scores by naming times. Participant scores are summarized in Table 1.

(Table 1 about here)

Summary statistics do not furnish a clear picture of the distribution of ages and ability scores across our participant sample. Figure 1 shows that: 1. more children and young adults than older adults were tested; 2. most phonological awareness scores were at or near ceiling; 3. the distribution of word or non-word pronunciation scores was reasonably symmetric; and 4. the distributions of mean RTs in the two tasks, calculated per person, were also symmetric. In our analyses, we addressed the potential impact of the distribution of individual differences variables on our estimates of their effects, and of their interaction with psycholinguistic effects. To anticipate, convergent evidence from multiple analytic methods indicated that our findings were robust to imbalances in the distribution of individual differences scores.

(Figure 1, about here)

A number of participants completed one or more additional tasks during a test session, interleaved with the critical tests in random order. These included: tests of print exposure (the UK version of the Author Recognition Test, Masterson & Hayes, 2007; the Children's Author Recognition Test, Stainthorp, 1997); the Wechsler Adult Intelligence Scale blocks sub-test (Wechsler, 2008); and an orthographic choice test (based on Olson, Kliegl, Davidson, & Foltz, 1985). Resource constraints meant these tasks could not be administered to all participants, resulting in missing data for large proportions of the participant sample. Thus, performance measures for these tasks were not included in our analyses.

Materials

We selected 160 words for use in both pronunciation and lexical decision tasks. We selected monomorphemic items that were likely to be known by all participants, having an estimated AoA (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2013) of six years, on average. Though effects of spelling-sound regularity or consistency were not the focus of our analyses, we nevertheless controlled for these variables, ensuring a balance of regular and irregular words in our stimulus set. The 160 items consisted of 80 pairs of words: one with a regular pronunciation, another with an irregular exceptional pronunciation (with GPC regularity determined according to the N-watch database; Davis, 2003). Regular and irregular words were matched on: frequency (Log10 SUBTLEX-UK Contextual Diversity, van Heuven, Mandera, Keeulers, & Brysbaert, 2014); rated imageability (ratings collected in our laboratory); rated AoA (Kuperman et al., 2013); length in letters, as well as orthographic neighbourhood size, Orthographic Levenshtein Distance (OLD), summed bigram frequency, mean bigram frequency and bigram frequency by position (English Lexicon Project norms, Balota et al., 2007); all comparisons, independent samples t-tests, 2-tailed p-values > 0.1. We report summary values for words in Table 2, and explain the variables in the following sub-section.

(Table 2 about here)

For the lexical decision task, we selected word-like pronounceable non-words from those made available in the ELP database (Balota et al., 2007). Non-words were selected as pair-wise matches to the words on: initial phoneme; length in letters; orthographic neighbourhood size; summed, mean and position specific bigram frequency; all comparisons, independent samples t-tests, 2-tailed p-values > 0.2.

Psycholinguistic Variables

For each word, we collated values for critical psycholinguistic variables that would be used to analyze reading behaviours. This included orthographic neighbourhood size (Coltheart et al., 1977), the number of words of the same length that can be created from the item by a single letter substitution, as well as the average Orthographic Levenshtein Distance (Yarkoni et al., 2008), the mean distance from a word to its 20 nearest orthographic neighbours, calculated as the number of letter substitutions, deletions or additions required to transform one word into another. Following Yap and Balota (2009), we calculated the orthography-to-phonological

(OLD/PLD) consistency, the ratio of its OLD to PLD, for each word. In addition, we collated the summed, mean, and position-specific bigram frequency. These measures were accessed from the English Lexicon Project (Balota et al., 2007).

Analyses of lexical decision latencies from the ELP, reported by Brysbaert and New (2009) and Adelman et al. (2006), have shown that estimates of word frequency calculated in terms of Contextual Diversity perform better in explaining variance than frequency calculated in terms of frequency of occurrence. Contextual Diversity (CD) refers to the number of documents in a corpus that include a word. van Heuven et al. (2014) have further shown that frequency estimates derived from a UK corpus (SUBTLEX-UK) performed better at explaining variance in the lexical decision latencies of British participants (the British Lexicon Project, Keuleers, Lacey, Rastle, & Brysbaert, 2012) than did estimates derived from a US corpus (SUBTLEX, Brysbaert & New, 2009). Therefore, we used the CD measure derived from the SUBTLEX-UK corpus to estimate the frequency effect in our analyses.

We used estimates of AoA provided by Kuperman et al. (2013). In their survey, Kuperman et al. asked participants to enter the age at which they thought they had learnt target words. We used word imageability ratings collected in our laboratory following a procedure employed in previous studies (e.g. Cortese & Fuggett, 2004). We asked 29 participants to rate how easily each word aroused a mental image, on a 7-point scale from 1 (very difficult) to 7 (very easy), and calculated mean rated imageability per word.

In addition to the predictors of theoretical interest, our analyses included variables coding for the phonetic characteristics of word initials. We used a commonly employed scheme representing the presence or absence of 10 phonetic features: voiced, nasal, fricative, liquid, bilabial, labiodental, alveolar, palatal, velar, and glottal. This was done to capture variance due to word initial phonemes' phonetic characteristics (Kessler, Treiman, & Mullenix, 2002; Spieler & Balota, 1997).

Experimental task procedure

In both reading tasks, participants were tested using Windows XP or Windows 7 laptops. Stimuli were presented in black 32-point Times New Roman font on a grey field. Participants were seated 40cm from displays, and words subtended 2.5 degrees of visual angle on average (mean = 2.5, SD = .5). Stimulus presentation and response recording were administered using the DMDX application (Forster & Forster, 2003). The sequence of events in a trial was as follows: blank screen for 500ms; fixation point (*) shown at centre of screen for 500ms; stimulus presented for 2,000ms (response interval). In both tasks, words were presented in random order in blocks, and blocks were presented in random order per session for each participant. Participants were invited to take breaks between blocks.

In the pronunciation task, participants were asked to read aloud the words shown on screen as quickly and as accurately as possible. The 160 critical stimulus words were randomly assigned to five blocks of 32 words, with 16 regular and 16 irregular words presented in each block. Participants wore microphone headsets and their vocal responses were recorded directly to hard disk. Response latencies and durations were extracted from recordings using the CheckVocal application (Protopapas, 2007). The test began with 10 practice items selected to match the critical items.

The lexical decision task consisted of 320 trials: 160 word and 160 non-word trials. Participants were asked to decide whether a stimulus was a word or non-word, indicating their decision by keypress using a USB gamepad as quickly and as accurately as possible. Trials were split into three sections: a practice section (20 trials); section A (160 trials); and section B (160 trials). Trials in sections A and B were presented in 4 blocks of 40. Sections A and B were completed in random order.

Results

Our investigation focused on the effects of participant age, word frequency and AoA on the latency of responses in the pronunciation and lexical decision tasks. Our review of previous experimental observations, and current theoretical accounts, indicated that the psycholinguistic effects could be hypothesized to be qualitatively the same but, depending on the theory, (1.) remain the same size (2.) increase in size over increasing age or (3.) decrease in size over increasing age. The critical test of the hypotheses lay in the estimation of interaction effects due to the modulation of the frequency and AoA effects by the age effect.

To aid the interpretation of the critical effects estimated in our main analyses, we then conducted two sets of further analyses. In the first, we examined diffusion model parameters, estimating the drift rate, boundary separation and non-decision time parameter values, for each participant, from the speed and accuracy of their lexical decisions. In the second, we analyzed the effects of item attributes and individual differences on the duration of spoken responses.

Our models included not just terms to capture the effects of frequency, AoA, age and the interactions between age and frequency, and age and AoA, but but also predictor variables capturing other individual differences or item attributes. Estimating the critical effects alongside effects due to these other variables allowed us to control for potential confounds due to differences between participants in reading or phonological awareness skill, or of variation among words in imageability, spelling-sound regularity, word length, orthographic similarity, or bigram frequency. We checked whether the estimates of critical effects depended on the measures of orthographic similarity (e.g. Coltheart's N or OLD) entered as predictor variables. They did not. We report effects associated with the non-critical variables in summary form though we discuss in some detail the modulation of the frequency effect by reading skill. We report the results of check analyses in Supplementary Materials.

We were concerned to estimate the effects of the individual differences variables allowing for the fact that the relationship between age and RT could be curvilinear. Two common methods of estimating curvilinear effects include the use of polynomial terms (e.g. quadratic or cubic effects) or restricted cubic splines. We report curvilinear effects estimated using splines but checked whether critical results differed if polynomial terms were used instead. They did not.

Linear mixed-effects (LMEs) models of response latencies are mandated by the structure inherent in our data, as a consequence of the use of a repeated measures design for our experimental tasks (Baayen, Davidson, & Bates, 2008). Observations obtained using such models are now common in the empirical literature. However, in previous research, questions like those we investigated have been examined using two alternate methods that can be understood as approximations to LMEs (Gelman & Hill, 2006; Snijders & Bosker, 2012): subject-level (slopes-as-outcomes) and item-level (averaged outcomes) analyses.

Researchers have examined individual differences in item effects by firstly estimating the item effects separately for each participant then testing the impact on the per-subject item effect coefficients of individual differences (e.g. Balota et al., 2004; Lorch & Myers, 1990), a method often referred to as slopes-as-outcomes analyses. This approach has been widely used, and depicting the modulation of item effects by individual differences in terms of variation in subject-level coefficients, allowed ready comparison with previous results. As will be seen, the depiction and analysis of variation in subject-level coefficients clearly demonstrates the modulation of frequency and AoA effects by age. However, this approach suffers the limitation that it does not distinguish the true variability of the second-level regression coefficients from the sampling variability of the subject-level coefficients (Snijders & Bosker, 2012). In addition, slopes-as-outcomes analyses ignore the fact that, for repeated measures designs, there is similarity between the data for different participants because they are all responding to the same stimuli. What is required is an analytic approach that appropriately takes into account the error variance structure in the data, including random variance due to participant sampling and random variance due to item sampling. This is the approach we use in our main analyses, Linear Mixed-effects (LME) models with crossed random effects.

Researchers have also examined variation in item effects by examining, separately for different participant groups, the effects of item attributes on the mean by-items latency of correct responses. This approach is advantageous because it allows the reliable estimation of item effects, though it is somewhat less powerful than an LME analysis (Baayen, Davidson, & Bates, 2008). In addition, it permits the presentation of item effects estimated for different, theoretically interesting, groups. Critically, using the approach allowed us to examine if differences in the sampling of participants from different parts of the age range biased our findings. Our observations showed that they did not. However, where the target for investigation, as here, are the potential interactions by which item effects may be modulated by individual differences, the item-level analysis method suffers from serious limitations, as we discuss later.

In the following, we describe the treatment of data used in our analyses, the summary features of reading performance, the inter-correlation of predictor variables, and the analytic approach, before presenting the results of the statistical models.

The treatment of data used in analyses and summary features of performance

We collected 113,920 lexical decision observations, including correct and incorrect word and non-word classifications, excluding responses made by dyslexic readers or speakers of English as a second language. We further excluded 640 responses made by two participants with missing scores on the phonological awareness task. We estimated diffusion model parameters from the latency and accuracy of the 113,280 observations remaining.

We analyzed responses to words to estimate the effects of word and participant attributes. Out of a total of 56,640 lexical decision responses to words, we were able to examine the influences on the latencies of 53,290 correct responses. We collected 28,800 responses in the word pronunciation task, excluding responses made by dyslexic readers or speakers of English as a second language. Excluding 710 observations that corresponded to errors and 49 that corresponded to response latencies < 200ms, we were able to examine the influences on the latencies of 28,041 correct pronunciation responses.

Levels of accuracy in the reading tasks were very high: the mean percentage of correct responses to words was 94% for lexical decision (SD = 8), and 97% for word pronunciation (SD = 6). The distribution of percentage accuracy values and mean RTs, calculated per person by task, is presented in Figure 1. It can be seen that, except for a small number of outliers, most participants performed at a level of accuracy greater than 90%. The most appropriate method for analyzing response accuracy, for repeated measures data, is the use of Generalized Linear Mixed-effects Models (GLMMs). In practice, with such high levels of response accuracy, it was not possible to reliably estimate critical effects. We found that GLMMs would not converge. We therefore report only analyses of response latencies, though accuracy is incorporated in the calculation of diffusion model parameters.

The average of mean RTs calculated per participant for each experimental task was faster in pronunciation (M = 592ms, SD = 114) than in lexical decision (M = 681ms, SD = 155). The distribution of RTs of responses made to each word by each participant were highly skewed, as is usually observed in experimental reading research. The skew in the raw RT distribution is usually ameliorated, for statistical analysis in psycholinguistic research, either by log transforming RTs (log10 or $\log_e$ of RTs) or by taking the reciprocal of RTs (-1/RT or -1000/RT). In practice, the log or reciprocal transformations are similar in effectively rendering the latency distribution more normal. The assumption, in linear and LME models, that residuals are normally distributed, warranted the analysis of logRT or -1/RTs. We report the results of logRT models in the present article but also estimated effects on raw RT and -1/RT, reporting the models in Supplementary Materials. We note here that the pattern of results was essentially the same irrespective of the outcome measure, though we discuss later how some critical effects were not reliably detected in models of RT.

<u>Correlations between RT and participant or item attribute variables</u>

The correlations between variables are a key consideration in examining the impact of participant or item attributes. In Table 3, we report for each task the

(Pearson's r) correlations between RT and the critical participant and word attributes. The large number of observations means that most correlations are significant, therefore we discuss only their sizes.

(Table 3, about here)

The same pattern of correlations was observed in both tasks. TOWRE word and non-word pronunciation skill measures were highly correlated with each other (r >= .8), as were the orthographic neighbourhood size and OLD measures of orthographic similarity (r = .95). Likewise, the three bigram frequency measures were also highly correlated (BG-Sum with BG-Mean, r ~ .9; BG Frequency-by-position with BG-Sum or BG-Mean, r ~ .6). These high correlations reflect the relationship between the variables as, largely, alternate measures of the same underlying dimensions: pronunciation skill; orthographic similarity; and bigram frequency. All other correlations were small or moderate though we observed TOWRE skills measures correlated with phonological awareness skills (spoonerisms, r ~.55), that length correlated with orthographic neighbourhood size and OLD measures (r ~ .65), and that frequency correlated with AoA (r ~ .55). These correlations replicate relationships observed in previous research. It is interesting, because it will be relevant to later discussion of analysis results, that the correlations between RT and measures of standardized pronunciation skill were larger for pronunciation (r ~ .4) than for lexical decision (r ~ .3).

The very high correlations between some item or participant attributes warranted precautions against the problems associated with multicollinearity. High intercorrelations among predictors make it difficult to distinguish the unique contribution of each to outcome variance. We standardized numeric predictor variables to remove multicollinearity due to scaling, to eliminate multicollinearity between linear and curvilinear terms or main effects and interactions (Cohen et al., 2003), as well as to improve the interpretability of coefficients. In the reported analyses we did not enter as predictors all available item attribute variables, selecting one measure of orthographic similarity (neighbourhood size), and one measure of bigram frequency (BG-mean). In addition, given the very high correlation between word and non-word pronunciation ability variables, we created a single ability measure by averaging together the TOWRE word and non-word pronunciation skill scores for each participant. The results of alternative approaches, using a different measure of orthographic similarity (OLD), or both, using all measures of bigram frequency together, or using both separate measures of pronunciation skill as predictors showed that the findings on the critical effects were robust, remaining the same across all alternates. (The alternate analyses, and their results, are presented in Supplementary Materials.)

It is well known that bare correlations between pairs of variables may not reflect the critical features of the relationships between the variables. It is useful then to examine scatterplots showing the relationships between RT and the critical participant or word attribute variables for each task, presented in figures 2 and 3, with points representing raw response latencies and the LOESS smoothers representing the bivariate relationship between RT and each variable. The scatterplots showing the relationship between RT and age reveal a U-shaped effect that is occluded by the small negative correlation coefficients ($r = -.15$) calculated between this pair of variables for each dataset. Latencies first decreased, from childhood into adulthood, then increased, in adulthood, over increasing age. We can see, also, that the relationship between RT and the word or non-word pronunciation skill measures are also curvilinear, with the impact of increased skill appearing to diminish for the highest skill levels. Lastly, the relationships between RT and participant attributes were clearly larger (r ~ .3) than those between RT and word attributes (r ~ .1).

(figures 2 and 3, about here)

Slopes-as-outcomes subject-level analyses

As a first step towards answering the question, if frequency and AoA effects are modulated by age or other individual differences, we performed slopes-as-outcomes analyses. These analyses were completed in two steps. In the first, analyzing each participant separately, we estimated the effects on response latencies of: word initial phoneme coding variables; length; orthographic neighbourhood size; mean bigram frequency; word frequency (log 10 SUBTLEX-UK CD); AoA; and imageability. The resulting subject-level coefficient estimates for the psycholinguistic effects were collated as the outcome variable (hence, slopes-as-outcomes) for the second step, in which we analyzed the effects of variation in participant attributes (age, reading skill, phonological awareness skill) on subject-level coefficients.

In Figure 4, we present scatterplots showing, for each task, variation in subject-level frequency or AoA effects estimates in relation to individual differences in age, reading skill and phonological awareness. In these plots, each point represents the estimated effect coefficient for a participant, while the lines represent the associated standard error.

(Figure 4, about here)

The plots clearly show that the size of the frequency and AoA effects decrease substantially over increasing age, reading skill and phonological awareness skill. The AoA effect appears to decrease with increasing age according to a curvilinear function, plateauing at a smaller size than first seen, at around 250 months (21 years). The AoA effect also appears to be smaller for participants with higher phonological awareness scores, at least, in pronunciation. The frequency effect decreases over increasing age, reading skill or phonological awareness skills according to more nearly monotonic functions, though the confidence intervals about the smoothers indicate that there is clearly uncertainty, corresponding to sparser sampling, at the extremes of the age or skills ranges.

We analyzed the per-subject coefficient estimates in linear models, in slopes-as-outcomes analyses (model summaries are presented in Table 4). Our analyses showed that the AoA and frequency effects were significantly modulated by the curvilinear effect of age in both lexical decisions and pronunciation. With increasing age, the AoA effect tended to decrease (decisions, age effect on AoA, coefficient estimate = -.008; pronunciation, age coefficient = -.006) but further age increases were associated with a deceleration or weakening in the impact of age on the AoA effect (decisions, age' effect on AoA, coefficient = .010; pronunciation, age' coefficient = .008). Note that the influence of age on the AoA effect for pronunciations was near significant at the .05 level while all other mentioned effects were significant at that level. Estimates of the coefficients of the age effect on subject-level frequency effects suggested that the frequency effect was slightly amplified over increasing age (decisions, age effect on frequency, coefficient estimate = -.003; pronunciation, age coefficient = -.016) but the age effect was curvilinear and estimates of the coefficients of the curvilinear component, the age' effect, showed that over the long term in our sample, the impact of increasing age was substantially to decrease the frequency effect (decisions, age' effect on frequency, coefficient = .010; pronunciation, age' coefficient = .028).

In pronunciation, the subject-level AoA effect was significantly modulated, also, by phonological awareness skill (effect of phonological awareness on AoA, coefficient = -.004; curvilinear effect of phonological awareness' on AoA, coefficient = .004) while the frequency effect was modulated by differences in reading skill (effect of reading skill on frequency, coefficient = .021; curvilinear effect of skill' on frequency, coefficient = -.012).

(Table 4, about here)

In summary, the slopes-as-outcomes analyses reveal the modulation of frequency and AoA effects by age. Check analyses (see Supplementary Materials) show that these age effects were robustly observed across variation in outcome transformation or model specification but while the effects were consistently detected on log10(RT) and -1/RT, they were detected less consistently in RT analyses. There were indications, also, that individual differences in psycholinguistic effects were significantly influenced by task. It is striking that the modulation of the frequency and AoA effects by age, certainly in the decision data, appears to occur year-on-year among, especially, children and young adults. It is equally striking, however, that the uncertainty about subject-level coefficient estimates varied considerably between individuals. This variation is not taken into account in slopes-as-outcomes analyses but is taken into account in LMEs.

Item-level analyses for separate age groups

The analyses of subject-level estimates indicated that the frequency and AoA effects were smaller for older, more skilled, readers. However, Figure 1 made clear that more participants were sampled from the younger than from the older part of the age range. In the present section, we report analyses that examine the modulation of the frequency and AoA effects by individual differences in an approach that addressed this imbalance in participant sampling.

Typically, in previous psycholinguistic studies in which data collection does not conform to a factorial design, the analytic approach has been to estimate the effects of word attributes on by-items mean RTs in multiple regression analyses, separately for each age group if participants are sampled across an age range. The analysis of by-items data for different age groups is useful because the mean RTs can be calculated for roughly equal subsets of observations. This allowed us to check if the modulation of the AoA and frequency effects by age were somehow an artefact of the greater sampling of participants from the younger participants in the age range. We split the data into three subsets ordered by age group. Children were defined as participants with ages < 216 months (18 years), young adults as participants with ages >= 216 months but < 360 months (30 years), and older adults as participants with ages >= 360 months. (These boundaries are, of course, arbitrary.) We calculated the by-items mean RTs for lexical decision and pronunciation data, separately for each age group. The by-items mean RTs were the average latency of every correct response to each stimulus word, calculated over the responses made by the participants in an age group. We then estimated the effects of word attributes on by-items mean RTs. For these analyses, there were equal numbers of by-items RTs per task for each age group. Table 5 presents, a side-by-side comparison of the coefficients estimates for the lexical decision and pronunciation models for each age group.

The coefficients for both the effects of frequency and AoA decreased with increasing age, in lexical decision and pronunciation. The frequency effect decreased comparing by-items estimates across age groups, slightly in decisions (for children frequency effect coefficient = -.026, for young adults coefficient = -.026, for older adults coefficient = -.021) but substantially in pronunciation (for children frequency effect coefficient = -.019, for young adults coefficient = -.009, for old adults coefficient = -.008). The AoA effect decreased substantially across age groups in both decisions (for children AoA effect coefficient = .011, for young adults coefficient = .006, for old adults coefficient = .006) and pronunciation (for children frequency effects coefficient = .006, for young adults coefficient = .002, for old adults coefficient = .001).

(Table 5, about here)

The results of the by-items analyses clearly replicate those of the subject-level analyses. The frequency and AoA effects are smaller for young or older adults than for children. However, while we were able to examine the variation in item effects between age groups, we could not do so while simultaneously taking into account differences in reading or phonological awareness skill. In addition, aggregating data to by-items group means reducing the sensitivity of any analysis of the effect of an individual differences variable (as in differences between age groups) when the underlying dimension is known to vary continuously (Cohen, 1983). Also, scatterplots (Figure 3) showed that the relationship between response latencies and age is probably curvilinear, so that estimating the age-related differences has to take this into account whereas comparisons between age groups cannot do so. Finally, whereas it is common to compare psycholinguistic effects coefficients for different groups qualitatively, recent analyses show that such comparisons carry the risk of detecting spurious interactions (Nieuwenhuis, Forstmann, & Wagenmakers, 2011). To address these limitations, we directly tested the modulation of the effects of psycholinguistic variables by individual differences using Linear Mixed-Effects models.

Linear mixed-effects models

The results of the subject-level and item-level analyses indicated that the frequency and AoA effects decreased in size from childhood into old age. The methods of analysis correspond to approaches that have been widely used in previous research. However, we have noted that both methods suffer from critical limitations.

To address our research questions, we examined response latencies using Linear Mixed-Effects (LME) models. LMEs do not suffer the limitations of subject- or item-level analyses because they incorporate the estimation of replicable 'fixed' effects on performance, for example, the effects of word frequency or participant age, while taking into account random effects due to variation between subjects or between items in average response latency (random intercepts) or in the average slopes of the fixed effects (Baayen et al., 2008). In addition, LMEs permit new insights. By using LMEs, we were able to estimate the effects of individual differences simultaneously with the effects of item attributes, while taking into account random error variance due to differences between participants or stimuli. In doing so, we accounted for variation in average latencies that might otherwise confound the estimation of interactions between the effects of item attributes and the effects of individual differences (Faust et al., 1999). This meant that we could directly estimate the modulation of the effects of frequency and AoA by the effect of age, rather than estimating individual differences separately for each variable (as in the subject-level analysis). It meant that we could estimate curvilinear effects of age and other individual differences. And it meant that we could, for the first time, report estimates of the amount of variance in reading latencies explained by both word properties and individual differences. Most psycholinguistic researchers know that the effects of individual differences tend to be larger than the effects of word properties. Our report is the first to say by how much.

Models were fit using the maximum likelihood procedure with the lme4 package (version 1.1-12, Bates et al., 2016) in R (version 3.3.1, R Development Core Team, 2016). In model summaries, we report estimated coefficients (and SEs) of hypothesized effects along with p-values obtained using the Satterthwaite approximation for degrees of freedom (lmerTest version 2.0-32, Kuznetsova, Brockhoff, & Christensen, 2016). We report marginal R^2_m, the variance explained by the fixed effects as a proportion of the sum of all the variance components (Johnson, 2014; Nakagawa & Schielzeth, 2013; calculated using the MuMIn package, version 1.15.6, Bartoń, 2016), i.e. the ratio of the variance attributable to the fixed effects compared to the sum of the variance due to the fixed effects, the random effects and the residuals.

We report estimates for models in which we included random effects (Barr, Levy, Scheepers, & Tily, 2013; but see Bates, Kliegl, Vasishth, & Baayen, 2015) corresponding to: (1.) unexplained differences between average latency of responses made by different participants, or to different word stimuli; and (2.) unexplained i.e. random differences between participants in the slope of the frequency and AoA effects, and random differences between words in the slope of the age effect. Reported models were first fitted to the complete data-set for typically developing readers, then refitted following the removal of observations associated with large residuals (standardized residuals > 2.5, following Baayen, 2008), to mitigate the influence of outliers.

In addition to the predictors of theoretical interest, our analyses included variables coding for trial order, and for the phonetic characteristics of word initials. We entered trial order in analyses to capture variance associated potentially with order effects, for example, due to participants tiring over test sessions (cf. Baayen, 2008). We coded initials to capture variance due to word initial phonemes' phonetic characteristics (Kessler et al., 2002; Spieler & Balota, 1997).

In the following, we first report the results of models of the lexical decision and pronunciation data. We then report the results of a model of the cross-task dataset, including latencies of responses in both tasks. In the text, we distinguish significant critical effects where $t > 2$, $p < .05$.

Lexical decision results

We found that lexical decision latencies were significantly influenced by the curvilinear effects of age and reading skill. A model summary is presented in Table 6. The U-shaped effect of age apparent in Figure 2 was captured by the model. With increasing age, latencies tended to decrease. The age coefficient estimate of -.139 indicates that, on average, logRT decreased by .139 for unit increase in age. But the model estimates show how, for older participants, the trend reversed so that latencies grew longer with increasing age. The age' coefficient estimate of .239, for the curvilinear component of the effect, indicates a countervailing trend such that logRT would tend to increase not decrease with increasing age for older participants. In short, because the age effect was curvilinear, as captured by the significant splines terms, the age effect varied, its size and direction depending on the specific age value.

Increasing reading skill was also associated with shorter latencies. The skill coefficient estimate of -.041 indicating that, on average, logRT decreased by .041 for unit increase in aggregate skill score. But the rate of decrease in latencies was substantially reduced for the highest skill levels, as seen in Figure 2. The skill' coefficient estimate of .021, for the curvilinear component of the effect, indicates that the average rate of decrease tended to decelerate by .021 over increasing skill for the most skilled readers.

Decision latencies were significantly affected by critical psycholinguistic effects. Latencies were shorter for longer words (the coefficient estimate of -.013 indicating logRT decreased by .013 for unit increase in length), more frequent words (the coefficient estimate of -.028 indicating logRT decreased by .028 for unit increase in log CD), that were easier to image (coefficient = -.007 indicating logRT decreased by .007 for unit increase in imageability), and earlier acquired (coefficient = .007 indicating logRT increased by .007 for unit increase in AoA). (Recall that numeric predictors were standardized so that the estimates indicate the effect of increasing the standardized variable by one (raw score SD) unit relative to 0, and that a value of 0 for the standardized variable corresponds to the mean value for the original variable.)

Most importantly, the effects of individual differences modulated the effects of word attributes. Critically, we observed a significant interaction between the age and AoA effects such that, with increasing age, the size of the AoA effect tended to decrease. The age x AoA effect coefficient of -.007 shows how the size of the AoA effect decreased by .007 for unit increase in age, on average. The AoA effect was different for different participant ages. Model estimates showed that, for further increase in age, the diminution in the AoA effect decelerated. The age' x AoA effect coefficient estimate of .008 indicates how the trend for the AoA effect to decrease would itself weaken, with increasing age. This interaction captures both the general trend for the AoA effect to decrease with increasing age, and the fact that that trend grows more shallow for older participants, as seen in Figure 4. The imageability effect also diminished with increasing age (age x imageability effect = .005) though again the diminution in the imageability effect over increasing age decelerated for older participants (age' x imageability effect = -.006). The interaction between the frequency and age effects was not significant in this analysis though the frequency effect tended to diminish over increasing age, at least, for some of the age range (age' x frequency effect = .006). Notably, the age by frequency interaction was significant in alternate analyses, of -1/RT and RT, as we discuss later.

The effect of orthographic neighbourhood size decreased with increasing reading skill. The impact of differences in phonological awareness skill also significantly modulated the effects of length, neighbourhood size and imageability. For higher levels of phonological awareness skill, each psycholinguistic effect significantly decreased in size.

We found that a model including just the effects of word properties was associated with an R^2_m of .041. Adding the effects of age, reading skill and phonological awareness skill increased R^2_m to .154. Adding the curvilinear components of the effects of age, reading and phonological awareness skill increased R^2_m to .229. Adding pairwise interactions between effects of individual differences and effects of psycholinguistic variables further increased R^2_m to .233. Likelihood Ratio Tests (LRTs, Pinheiro & Bates, 2000; see Supplementary Materials) showed that each increment in model complexity was significantly associated with improved model fit to data (all LRTs, $p < .001$).

(Table 6, about here)

Pronunciation results

There were substantial similarities between the pattern of effects observed for pronunciation as for lexical decision latencies (see Table 6). We found that pronunciation latencies were also significantly influenced by the curvilinear effects of age and skill, according to similar functions. Latencies tended to decrease with increasing age (age effect, coefficient = -.039) but with further increases in age the trend reversed so that, for older participants, latencies increased with increasing age (age' effect, coefficient = .097). Latencies also tended to decrease with increasing skill (reading skill effect, coefficient = -.061), but the rate of decrease in latencies diminished for higher skill levels (reading skill' effect, coefficient = .023). In addition, the latencies of pronunciation responses were significantly shorter for more frequent (coefficient = -.011), easier to image (coefficient = -.004) words, just as in lexical decision. In contrast to the lexical decision data, longer words elicited longer latencies in the pronunciation task (coefficient = .006).

Critically, significant interactions showed that increasing age was associated with decreases in the sizes of the effects of frequency (age x frequency effect, coefficient = -.010; age' x frequency effect, coefficient = .019), AoA (age x AoA effect, coefficient = -.008; age' x AoA effect, coefficient = .012) and imageability (age x imageability effect, coefficient = -.005; age' x imageability effect, coefficient = .008). The significant interactions capture the trends evident in the Figure 4 plots showing the decrease in subject-level estimates of the frequency and AoA effects over increasing age. In addition, the effects of word length and regularity were also significantly modulated by differences in age such that the length effect grew larger (age x length effect, coefficient = .009; age' x length effect, coefficient = -.013) while the regularity effect grew less marked with increasing age (age x regularity effect, coefficient = .013; age' x regularity effect, coefficient = -.020).

Differences in reading skill significantly interacted with effects due to orthographic neighbourhood size, bigram frequency, regularity, word frequency and imageability. For increasing reading skill, the effects of neighbourhood size, frequency and imageability decreased. However, the small effects of regularity and bigram frequency appeared to increase very slightly. The significant interaction between effects of phonological awareness skill and neighbourhood size indicated that the neighbourhood effect was slightly larger for participants with stronger phonological skills.

A model including just the effects of word properties was associated with R^2_m of .055. Adding the effects of age, reading skill and phonological awareness skill increased R^2_m to .306. Adding the curvilinear effects of age, reading skill and phonological awareness skill increased R^2_m to .326. Adding all pairwise interactions between effects of individual differences and effects of psycholinguistic variables further increased R^2_m to .339. Likelihood ratio tests (Supplementary Materials) showed that each increment in model complexity improved model fit to data (all LRTs, p <.001).

Cross-task analysis results

We examined differences between tasks in a cross-task analysis combining both lexical decision and pronunciation response data, and including task as a fixed effect (Table 7). Overall, we found a significant U-shaped effect of age. Latencies tended to decrease with increasing age (age effect, coefficient = -.146) but with further increases in age the trend reversed so that, for older participants, latencies increased with increasing age (age' effect, coefficient = .275). We found a curvilinear effect of reading skill such that readers with stronger skills produced faster responses (reading skill effect, coefficient = -.039) but the rate of decrease in latencies diminished at the higher skill levels (reading skill' effect, coefficient = .021). Overall, longer (coefficient = -.011), more frequent (coefficient = -.028), easier to image (coefficient = -.006), earlier-acquired (coefficient = .006) words elicited responses with shorter latencies.

Critically, significant interactions showed that increasing age was associated with smaller AoA (age x AoA effect, coefficient = -.008; age' x AoA effect, coefficient = .010) and imageability effects (age x imageability effect, coefficient = .005; age x imageability effect, coefficient = -.008). As in the lexical decision analysis, there was a non-significant trend for the frequency effect to be smaller with increasing age (age' x frequency effect, coefficient = .006).

The neighbourhood size effect was smaller for more skilled readers. In contrast, overall, the word length effect appeared to be larger for more skilled readers. In addition, there were significant interactions between the effects of variation in phonological awareness skill and the effects of length, neighbourhood size and imageability. The neighbourhood and length effects were found to be larger for participants with higher awareness skills. However, the imageability effect appeared to be smaller for those with better awareness skills.

The cross-task analysis indicated significant interactions between the effect of task and the effects of word and participant attributes. Critically, the curvilinearity in the age effect was less marked in pronunciation than in lexical decision (as can be seen in figures 2 and 3). There was a general tendency for latencies to decrease with increasing age but this age effect was smaller in pronunciation (task x age effect, coefficient = .108). The age effect, as noted, was curvilinear. However, the second-order (curvilinear term) increase in latencies, with increasing age, was also found to be smaller in pronunciation (task x age' effect, coefficient = -.177). A significant task by length interaction (task x length effect, coefficient = .015) reflected the fact that in lexical decision, longer words elicited faster responses but in pronunciation they elicited slower responses. The interactions between the task and frequency (task x Log10(CD) effect, coefficient = .017) and the task and AoA effects (task x AoA effect, coefficient = -.007) reflected the smaller size of these effects in pronunciation.

We found significant three-way interactions indicating that the modulation of psycholinguistic effects by individual differences varied between tasks. As seen in the plots of the variation in subject-level estimates of the frequency effect over age (Figure 4), the decline in the frequency effect with increasing age was sharper, more dramatic, for pronunciation than for lexical decision (task x age x Log10(CD) effect, coefficient = -.011; task x age' x Log10(CD) effect, coefficient = .015). This was also true for the modulation of the imageability effect by age (task x age x imageability effect, coefficient = -.009; task x age' x imageability effect, coefficient = .014).

We found a significant interaction between the effects of task, reading skill and frequency (task x reading skill x Log10(CD) effect, coefficient = .019; task x reading skill' x Log10(CD) effect, coefficient = -.013) corresponding to the pattern (seen in Figure 4) in which the diminution in the frequency effect with greater reading skill was stronger in pronunciation than in lexical decision. We also found significant interactions between task, reading skill and regularity as well as task, reading skill and imageability reflecting the fact that the modulation of these effects by reading skill was also stronger in pronunciation. Significant three-way interactions between task, phonological awareness and length, and between task, awareness and AoA, suggested that the length effect tended to be facilitatory with increasing awareness skill, in lexical decision but not in pronunciation, while the AoA effect tended to be larger with increasing awareness skill in pronunciation than in lexical decision.

(Table 7, about here)

A model including just the effects of word properties was associated with an R^2_m of .033. Adding the effects of age, reading skill and phonological awareness skill increased R^2_m to .188. Adding the curvilinear components of the effects of age, reading and phonological awareness skill increased R^2_m to .237. Adding pairwise interactions between effects of individual differences and effects of psycholinguistic variables further increased R^2_m to .241. Adding interactions between the effect of task and the effects of word or participant attributes increased R^2_m to .284. Likelihood ratio tests (see Supplementary Materials) showed that each increment in model complexity was significantly associated with improved model fit to data (all LRTs, $p < .001$).

Summary

The effects of frequency and AoA were smaller with increasing participant age. The decrease in the size of the AoA effect was significant, substantial, and robustly observed across reading tasks, and across variation in analysis approach. The decrease in the size of the frequency effect for older readers was also significant but was more robustly observed in pronunciation than in lexical decision. The frequency effect was also smaller for more skilled readers. In addition, we found that the imageability effect decreased with increasing age in both lexical decision and in pronunciation, with increasing phonological awareness skill in lexical decision, and with increasing reading skill in pronunciation. The neighbourhood effect decreased with increasing reading skill in both lexical decision and pronunciation. In contrast, it increased with increasing phonological awareness skill in both tasks. Lexical effects were observed to be stronger in lexical decision than in pronunciation. Longer words elicited shorter latencies in lexical decision but longer latencies in pronunciation. The interactions between the psycholinguistic effects and the effects of individual differences were observed in the context of large curvilinear effects of age, reading and phonological awareness skill. The impact of individual differences accounted for greater portions of variance than did the item effects.

Robustness checks

We established the robustness of our findings by examining whether the pattern of results varied in relation to differences in analysis method. Our checks showed that the U-shaped effect of age, and the modulation of the frequency and AoA effects by differences in age were consistently observed across all analyses (see Supplementary Materials), irrespective of variation in: (1.) outcome transformation (RT, log10(RT), -1/RT); (2.) method used to estimate curvilinearity in individual differences effects (using polynomial terms or splines); (3.) inclusion or exclusion of participants who were not typically developing monolingual speakers of English; (4.) selection of control variables, models including neighbourhood size or Levenshtein Distance measures or both, models including just mean bigram frequency or summed, mean and position specific bigram frequency, models including a single aggregate measure of reading skill or separate measures of word or non-word reading skill; (5.)

inclusion of random effects, fitting models with random effects of subjects or items on just intercepts, or models with random effects on intercepts and on the slopes of the critical effects; and (6.) exclusion of outlier observations, fitting models with or without the exclusion of large (>2.5) standardized residuals.

The modulation of the AoA effect by age was observed as an interaction that was reliably detected without exception in all analyses. The interaction between the age and frequency effects was found to affect pronunciation in an interaction that was reliable detected, also, across all analyses. The age by frequency interaction was not found in our analysis of lexical decision logRTs when random effects of participants or items on slopes were included in the model. However, it was found in LME models of logRTs when random slopes were not included. It was also found in all LME models when -1/RT was the outcome variable, and in models when RT was the outcome variable even when random slopes were included. We think the fairest account of the effect on lexical decision of the interaction between age and frequency is that there is strong evidence that it is present but that the interaction is not reliably detected in all analyses, suggesting it may be confounded, in part, with the presence of random between-subjects deviations in the average effect.

Diffusion model analysis

We have seen that the frequency and AoA effects on reading latencies decrease with increasing age, from childhood into old age. The modulation of psycholinguistic effects by age is observed in our study alongside a U-shaped age effect, and a reading skill effect that decelerated for higher skill levels. We conducted a diffusion model analysis of lexical decision performance, to gain insight into the processing differences that could underlie these trends.

Previous work by Ratcliff and colleagues (Ratcliff et al., 2010, 2012) has shown that, in lexical decision, the U-shaped effect of age on response latencies can be explained by age-related variation in the quality of information extracted from stimuli (drift rate), relative conservatism over decision-making (boundary separation), and non-decision time (time taken by stimulus encoding and response execution). However, the data comparing children to adults and older adults to younger adults were from different studies employing different stimulus sets. This means it is possible that stimulus differences confounded age group differences in diffusion model parameters. The examination of the age effect on lexical decision over the lifespan would benefit, therefore, from analyzing diffusion model parameters calculated using response data collected from different age groups given the same stimuli. This is the analysis we report in the current section.

In addition, recent work (McKoon & Ratcliff, 2016; Zegeurs et al., 2011) has indicated that differences in reading ability affect drift rates and boundary separation, such that readers with lower levels of literacy skill tend to have lower drift rates but more conservative decision criteria. No previous study has examined the effects of both age and reading ability on decision making processing components. In the analyses we report, we examined the relationship between diffusion model parameters and age, as well as reading ability and phonological awareness skill.

Granted that age-related variation in decision making components could help to explain the influence of age and other individual differences on lexical, we investigated if the same diffusion model analysis would help in understanding the interactions between the effects of frequency or AoA and the effect of age. Ratcliff et al. (2004) observed an effect of word frequency on drift rate in lexical decision. Diffusion model parameter differences between words of differing AoA have never previously been analyzed. No previous study has examined if decision making components are shaped by interactions between item and participant attributes.

To calculate diffusion model parameters, we used the EZ-diffusion function (Wagenmakers, Van Der Maas, & Grasman, 2007). We examined the responses for each participant per word type: (1.) early vs. late acquired words; (2.) high frequency vs. low frequency words. We sub-divided the lexical decision data, classifying words with AoA greater than the median for the sample as late acquired, words with AoA < median as early acquired, words with frequency values > median as high-frequency, and words with frequencies < median as low-frequency. We began our analysis by plotting the bivariate relationships between diffusion model parameters and individual differences variables. Figure 5 shows the diffusion model parameters calculated for each participant, for early or late acquired words. Figure 6 shows the parameters calculated for high or low frequency words. These plots show similar patterns.

(Figures 5 and 6, about here)

As expected, drift rate increased over increasing age, from childhood into adulthood, but then plateaued at a high level through adulthood into old age. Boundary separation decreased over increasing age, from childhood into adulthood, but then increased gradually into old age. Strikingly, non-decision time sharply decreased from childhood into adulthood, and then steadily increased again, through adulthood into old age. In comparison, drift rate increased over increased reading or phonological awareness skill, with the rate of increase getting smaller for higher skill levels. Boundary separation tended to decrease over increasing reading or phonological awareness skill but the rate of decline was small. Non-decision time also tended to decrease over increasing skill levels but the rate of change plateaued for higher skill levels. We found that differences between word types were associated only with differences in drift rate. High frequency or early acquired words had higher drift rates than low frequency or late acquired words. However, the difference in drift rate associated with differences in frequency or AoA entirely converged for older readers.

We tested the effects of individual differences and word type on the diffusion model parameters, drift rate, boundary separation and non-decision time in separate LME models. The outcome variables consisted of parameter values calculated for each participant from their lexical decision performance, with the response data sub-divided into data about responses to (1.) early vs. late acquired words (AoA models); or (2.) low vs. high frequency words (frequency models). Our models used restricted cubic splines to estimate curvilinear effects due to individual differences and estimated the effects of participant attributes, item type (low or high frequency, early or late acquired words), and interactions between attributes and types. A summary of the models is presented in Table 8.

(Table 8, about here)

We found a curvilinear effect of age on drift rate. Drift rate was higher with increasing age in both the AoA and the frequency models of diffusion parameters (AoA model, age effect, coefficient = .102; frequency model, age effect, coefficient = .125). However, the significant curvilinear age' term indicated that the age effect different for different ages: the rate of age-associated increase in drift rate decelerated for older participants, as seen in Figures 5 and 6 (AoA model, age' effect, coefficient

= -.142; frequency model, age' effect, coefficient = -.153). In comparison, drift rate increased with higher reading skill levels at a steady rate (AoA model, reading skill effect, coefficient = .025; frequency model, reading skill effect, coefficient = .031; the skill' effect was not significant in either analysis). In addition, drift rate was higher for early acquired than for late acquired words (AoA model, AoA word-type effect, coefficient = -.058), and for high frequency compared to low frequency words (frequency model, frequency word-type effect, coefficient = .072). The effect of frequency did not significantly interact with the effects of age or of skill. However, a near-significant interaction at p =.1 (AoA model, age x AoA effect, coefficient = .022) suggested the effect of AoA on drift rate was smaller for older readers.

We found that boundary separation was only influenced by reading skill, decreasing with increasing skill (AoA model, reading skill effect, coefficient = .002; reading skill' effect, coefficient = -.009; frequency model, reading skill effect, coefficient = .003; reading skill' effect, coefficient = -.007).

Non-decision time tended to decrease over increasing age (AoA model, age effect, coefficient = -.134; frequency model, age effect, coefficient = -.146) but the age effect on non-decision time was significantly curvilinear; its slope was different for different ages. The higher-order age' effect coefficient indicated that as age increased the trend in the age effect shifted from decreasing non-decision time to increasing non-decision time, as seen in Figures 5 and 6 (AoA model, age' effect, coefficient = .246; frequency model, age' effect, coefficient = .258). Non-decision time also decreased over increasing levels of reading skill (AoA model, reading skill effect, coefficient = -.042; frequency model, reading skill effect, coefficient = -.056) but the rate of change (see Figures 5 and 6), decreased for the highest skill levels (AoA model, reading skill' effect, coefficient = .029; frequency model, reading skill' effect, coefficient = .031). Interestingly, non-decision time was longer for late compared to early acquired words (AoA model, AoA word type effect, coefficient = .025). There was no effect of frequency on non-decision time but a significant interaction suggested that the decrease in non-decision time with increasing reading skill was less for responses to high compared to low frequency words (frequency model, reading skill x frequency effect, coefficient = .018; the reading skill' x frequency effect was not significant).

In summary, the results of our diffusion model analyses indicate that the pronounced U-shaped effect of age on decision latencies can be explained by supposing that age-related changes occur in drift rate and in non-decision time. These changes are detected while taking into account differences in reading skill. Differences in drift rate are linked to differences in frequency and AoA. Our results suggested that the AoA effect on drift rate was modulated by age while the frequency effect on non-decision time was modulated by reading skill.

Spoken response duration analysis

Diffusion model analyses are applied only to binary choice task data. However, we wanted to examine, also, how much of the effects of individual differences or item attributes on pronunciation could be said to correspond to variation in the efficiency of lexical access or response execution processes. We supposed that variation in durations should not only reflect overt articulation speed but also the efficiency of articulatory coding processes. We reasoned that if some of the effects observed to influence pronunciation latencies were due to variation in the efficiency of coding processes (in diffusion model terms, response execution processes) then those effects should influence response duration also. If they did not influence duration then they could be taken to influence only variation in the quality of lexical access (drift rate in diffusion model terms). Given previous findings, we expected to observe a U-shaped effect of age on pronunciation durations, encompassing a decrease in durations from childhood to adulthood but an increase in durations through adulthood into old age. In addition, we expected to find effects due to word frequency and neighbourhood size and, potentially, interactions indicating the modulation of the item effects by the effects of individual differences.

We extracted the durations of spoken responses in the pronunciation task by subtracting response onsets from response offsets (both transcribed using CheckVocal). Unfortunately, we were unable to transcribe the responses of 12 participants due to the corruption of DMDX trial information (.azk) files. Our analyses were otherwise of the same observations as were analyzed in the response latency models. Plots of the bivariate relationships between response duration and critical variables (Figure 8) show that age had a U-shaped effect on durations, as expected, but that there were effects, also, due to individual differences in reading skill. Figure 7 shows, in addition, that response durations were related to differences in word length and neighbourhood size.

(Figure 7, about here)

We fitted an LME model with log10(durations) as the dependent variable and, as predictor variables (see Table 9 for a summary): participant age, reading skill and phonological awareness skill; along with item initial phoneme features, word length, neighbourhood size, mean bigram frequency, frequency, AoA, and imageability; and pairwise interactions between the effects of the individual differences variables and the effects of the psycholinguistic variables.

(Table 9, about here)

The model indicated a U-shaped effect of age, there was a tendency for durations to decrease with increasing age (age effect, coefficient = -.136) but with further age increases the trend reversed so that greater age was associated with longer durations (age' effect, coefficient = .248). This age effect was observed alongside an effect of reading skill. More skilled readers produced responses with shorter pronunciation durations (reading skill effect, coefficient = .030; reading skill' effect, coefficient = -.037). In addition, longer words elicited longer responses (length effect, coefficient = .020) while words with larger neighbourhood elicited shorter durations (coefficient = -.010). A significant interaction indicated that increasing age modulated the neighbourhood effect (age x neighbourhood effect, coefficient = -.013; age' x neighbourhood effect, coefficient = .021). Inspection of a scatterplot relating age differences to subject-level estimates of the neighbourhood effect on durations (Figure 8) indicates that the neighbourhood effect deepened from childhood to adulthood but that the rate of change decreased or reversed in direction with further ageing (sparse data at older ages makes the later form of the curve less certain). A significant interaction also indicated that differences in phonological awareness skill modulated the AoA effect on durations (awareness skill x AoA effect, coefficient = .002; awareness skill' x AoA effect, coefficient = -.003). However, the plot relating subject-level estimates of the AoA effect to awareness skill suggests that the interaction depends on the presence of a small number of participants with exceptionally low levels of awareness skill.

(Figure 8, about here)

In summary, the analysis of pronunciation durations suggests that increasing age may be associated with first a decrease in articulation duration (response execution) time then, through adulthood into old age, an increase in articulation time. The U-shaped effect of age on durations was observed independent of an effect of reading skill and effects of word length or neighbourhood size. Interestingly, the neighbourhood effect was modulated by differences in age.

Discussion

How does reading vary over the lifespan, from childhood into old age? How do the effects of psycholinguistic variables like word frequency or AoA interact with the effects of age, and other individual differences? Our results show that differences in age, but also differences in reading ability and phonological awareness systematically modulated critical psycholinguistic effects. We analyzed responses to the same words, in two tasks, lexical decision and pronunciation, in a large sample of readers, including children, as well as younger and older adults. We analyzed reading performance using a comprehensive array of methods, but the findings from all approaches converged on the same result: the frequency and AoA effects decrease over increasing age. As readers grew older, their performance was less affected by how common the words are in the language, or by when in life they learnt the words.

This observation of the modulation of the frequency and AoA effects by age is new but fits into a larger pattern of age-related change that replicates and extends previous findings. We observed a U-shaped effect of age over the life-span. Reading response latencies decrease over increasing age, from childhood into adulthood, but then increase over increasing age, through adulthood into old age. Our data adds to a theoretically significant body of evidence from a very small number of studies in which performance has been compared across the life-span in the same task, for the same stimuli.

Our observation of interactions between the effect of age and the effects of word frequency and AoA have important implications for theoretical accounts of the cognitive reading system. Our results require that explanations of psycholinguistic effects must assume that the structure of the reading system develops through learning. As we discuss in the following, this is inconsistent with existing theoretical accounts of frequency and AoA effects, and with at least one class of general theoretical account of the reading system. However, we argue that our results additionally show that learning-based theories of the development of the reading system cannot satisfactorily explain the life-long performance of that system without explaining the U-shaped effect of age. The findings we report from diffusion model and pronunciation duration analyses offer insights into how existing reading models can be extended to explain the large-scale impact of individual differences. We begin our discussion by examining current inconsistencies in the observation of individual differences in the frequency and AoA effects, especially where our results differ from the results of previous studies. We then discuss the theoretical implications of our findings.

Inconsistencies in the empirical account of individual differences in the frequency and AoA effects

We observed the same interactions, between the frequency and age effects, and between the AoA and age effects, across a comprehensive examination of variants in data analysis methods. Historically, experimental psycholinguistic data have been analyzed using a variety of methods and, even if employing the same strategy, have often adopted different tactics. Recent studies have clarified how variation in decision-making in analysis can have decisive effects on the outcomes of analyses (Silberzahn & Uhlman, 2015; Simmons, Nelson, & Simonsohn, 2011). We examined if common variants in general analysis approach or, within an approach, in method, would modify the critical effects of interest in our own data, the interactions between the effect of age and the effects of frequency or AoA. They did not. Overall, we found the same pattern: the frequency and AoA effects were smaller for older readers. Given the robustness of our observations, we contend that differences between our results and previous findings should be explained in terms of the impact of differences in stimulus or participant sampling, rather than of differences in analytic method.

Few previous studies have examined if the AoA effect varies across individual differences. In the two previous studies of which we are aware, Morrison and colleagues (2002) and Barry et al. (2006) observed significant AoA and age effects in reading but did not find that the AoA effect was different for different age groups. The contrast between our results and these previous observations of a null interaction may be the result of differences in statistical sensitivity related to differences in sample size and differences in the treatment of continuous variables. Morrison et al. (2002) compared AoA and frequency effects on word pronunciation of 24 early and

24 late acquired words in 28 younger and 32 older participants. Barry et al. (2006) compared the effect of AoA in younger and older adults using two tasks, lexical decision (9 younger and 10 older readers) and pronunciation (10 younger and 10 older readers), analyzing responses to 24 early and 24 late acquired words. In both studies, age and AoA effects, and the potential interactions between these effects, were examined in analyses of variance, testing the effects of differences between the average latency of responses to early vs. late-acquired words produced by younger vs. older reader age groups.

There are two salient differences between our study and the previous studies. Firstly, we tested many more participants and, in each task, recorded their responses to many more stimuli. In comparison, samples of the size used in previous studies may not have had sufficient power to detect the interaction. In addition, we estimated the effects of AoA and age as the effects of continuous variables. Cohen (1983) demonstrated that estimating continuous effects in terms of dichotomous variables, as here, differences between responses to sub-groups of words made by sub-groups of participants, substantially reduces the power to detect the effects if they are present. Together, these differences made it likely, we suggest, that the critical interaction that we found could not be detected in the previous studies. The same account may explain differences between our observations, and previous findings, concerning the interaction between age and frequency effects.

Several previous studies have examined if the frequency effect varies in association with age and other individual differences but, as noted previously, they have yielded inconsistent results. Some studies have indicated similar frequency effects in younger compared to older readers in studies of children (Burani et al., 2002) and adults (Allen et al., 1991; Allen et al., 1993; Cohen-Shikora & Balota, 2016; Tainturier et al., 1989). Other studies have shown larger frequency effects in older compared to younger adult readers (Balota & Ferraro, 1993, 1996; Balota et al., 2004; Spieler & Balota, 2000). Further studies have found, as we did, that the frequency effect is smaller for older compared to younger adult readers (Morrison et al., 2002; see, also, Allen & Madden, 1989, who reported a frequency by age interaction in a study of the effect of word frequency on letter detection) or for readers with higher estimated levels of print exposure (Chateau & Jared, 2000; Sears et al., 2008; but see Lewellen et al., 1993).

We think that where previously researchers have not detected a reliable interaction between the frequency and age effects, differences in stimulus or participant sampling, as well as differences in analytic methods, can explain the null results. In each of the previous studies cited, fewer participants were tested than in the present study (Allen et al., 1991, tested 24 younger and 24 older adults on 84 words; Allen et al., 1993, tested 20 younger and 20 older participants on 216 words in experiment 1, groups of 20 adults on 480 words in experiment 2; Burani et al., 2002, tested 90 children on 80 words; Tainturier et al., 1989, tested 20 younger and 20 older adults on 90 words). In addition, in each of these studies, the analytic approach compared response to different word types from different participant age groups. As we have argued, these differences in sample size or analytic method would reduce the probability that a frequency by age interaction could be detected.

It is more difficult to explain the contrast between our observation of the frequency by age interaction, and that reported by Balota and colleagues over a series of studies (Balota & Ferraro, 1993, 1996; Balota et al., 2004; Spieler & Balota, 2000). That both sets of observations have yielded an age by frequency interaction seems clear, why they go in different directions is not. There are methodological differences between the studies that are salient. The first lies in the sampling of stimuli and participants. Compared to the studies reported by Balota and colleagues, we tested many more participants but (at least, compared to Balota et al., 2004; Spieler and Balota, 2000) using many fewer words. It is unclear which differences were decisive.

In addition, we observed the frequency by age interaction in models that accounted, also, for effects due to individual differences in reading ability or phonological awareness skill. However, we did not test the vocabulary, print exposure, or orthographic knowledge of our participants though recent research has indicated that these dimensions are likely to be critical in shaping cognitive reading processes (Andrews & Hersch, 2010; Chateau & Jared, 2000; Sears et al., 2008; Lewellen et al., 2000; Yap et al., 2012). Yap and colleagues (2012) reported that, in pronunciation, participants with higher vocabulary scores were less sensitive to the effect of a Principal Component related to word frequency, while in lexical decision higher vocabulary scores were associated with greater sensitivity to frequency. In contrast, Chateau and Jared (2000; Sears et al., 2008, but see, Lewellen et al., 2000) observed a smaller frequency effect on lexical decision in readers with higher print exposure. To these observations, we can now add our observation of interactions between the frequency effect and the effect of reading skill. In unpublished work in our laboratory, measures of reading skill correlate very highly with measures of vocabulary, print exposure, and spelling knowledge, suggesting that differences in these variables do not explain the variation in observations of frequency by age interaction in the studies reviewed. Further, we replicated the interaction in both a sample of typically developing readers, and a larger sample including also dyslexic readers and speakers of English as a second language, suggesting the specific composition of our participant sample did not bias the results. Clearly, further research is required to examine variation in psycholinguistic effects over a comprehensive range of individual differences measures.

However, in a recent further study, in which 148 adults were tested on 1200 words, Cohen-Shikora and Balota (2016) reported that subject-level estimates of the frequency effect were correlated with age differences in pronunciation but not in lexical decision. In other words, the frequency effect was not moderated by age in lexical decision. The frequency effect on pronunciation was found to be smaller for older readers, but this relationship disappeared when individual differences in vocabulary were partialled out. Cohen-Shikora and Balota (2016) concluded that the frequency effect is stable across the lifespan. This finding is inconsistent with previous reports from the same group. It is inconsistent, also, with our observations. The contradiction renders more salient a potentially critical difference between our study and previous studies.

We estimated variation in the frequency effect over age, sampling participants across the range from childhood to old age. In all previous studies, comparisons have either been among children or between younger and older adults. Our observations suggest that this is important because the principle phase in which the frequency effect changes, with age, in the lifespan may lie in the transition from childhood into adulthood. In our item-level analysis, the frequency effect appeared to be larger in children's latencies than in younger or older adults' latencies, while it appeared to be similar in size for the latter groups. In our subject-level analysis, the per-subject estimates of the frequency effect coefficient varied in relation to age, but the age effect on the frequency coefficients was curvilinear; it appeared to be stronger for younger ages. The most appropriate approach to repeated measures data is the use of LME analyses, and in our models a significant interaction between the frequency and age effect was detected in pronunciation but not in lexical decision, with a near-significant interaction detected in the cross-task data analysis. In fact, the modulation in lexical decision of the frequency effect by age was reliably detected in some variants of our LME analyses but not in others. We think then that the best supported conclusion is that there is a tendency for the frequency effect to decrease over increasing age, but that that interaction is more prominent in pronunciation, perhaps especially in the transition from childhood into adulthood. This finding is entirely consistent with the observations reported by Cohen-Shikora and Balota (2016) but it restricts the overall conclusion to stability in the frequency effect in later adulthood.

Implications for theoretical accounts of the frequency and AoA effects, and of the cognitive reading system

The theoretical implications of our results are clear. We found that the frequency and AoA effects decrease over increasing age. These interactions are incompatible with theoretical accounts of lexical access that assume that the effects of word attributes do not vary with age. We think they require the assumption of learning mechanisms as the basis for development of structure in the reading system.

Theoretical accounts of the frequency and AoA effects

Our observation of the decrease in the frequency effect with increasing age is inconsistent with the proposal that lexical access is delivered by a search process ordered by relative frequency independent of absolute frequency (Murray & Forster, 2004). Murray and Forster (2004) asserted that search based theories predict that the frequency effect should not change with increasing overall experience, in part, because in previous research at that time, the frequency effect did not decrease with increasing age. The similarity of the effect in younger and older adults (e.g. Tainturier et al., 1989) was held to be inconsistent with an explanation of the frequency effect based on learning because the assumption of learning predicts the diminution of the frequency effect over increasing age in the lifespan. That is precisely the interaction we observed but it is worth considering if that means a search based theory of lexical access cannot account for our observations. Critically, in an extension of their model, Murray and Forster (2004) proposed that a search based lexical access account can explain the AoA effect if the search process is ordered by cumulative frequency. Lexical access would be determined by the rank ordering of search candidates by their relative frequency but a AoA effect would reflect changes in relative frequency over the lifespan. We think that if search based theories of lexical access can admit the ordering of search candidates by cumulative frequency then they could, in principle, account for the frequency by age interaction. This is because cumulative frequency accounts of the AoA effect are based on the expectation that the frequency effect should be smaller for older readers.

In early work on AoA, researchers examined the possibility that differences in cumulative frequency could explain the AoA effect (Caroll & White, 1973; Lewis, Gerhand, & Ellis, 2001; Zevin & Seidenberg, 2002), motivated by the observation that words that vary in AoA also vary in the total number of times they have been encountered. The results of experimental and simulation studies appeared, for a time, to support a rejection of the cumulative frequency account of the AoA effect. Critically, a cumulative frequency account predicts that AoA effects should decrease over increasing age but this prediction was not supported in experimental data (Barry et al., 2006; Morrison et al., 2002). Given the finding that AoA effects on reading appeared to be similar in younger and older adults (see also Ghyselinck et al., 2004), researchers concluded that the impact of AoA is fixed when words are learned (Morrison et al., 2002).

Computational simulations of reading development have suggested, however, that reading performance is likely to be affected by AoA as well as by cumulative frequency, and frequency trajectory. While a word's cumulative frequency is its total frequency of occurrence, a word's frequency trajectory is given by the distribution of its frequency of occurrence varying over time. The results of computational simulations of reading development reported by Zevin and Seidenberg (2002), at first appeared to demonstrate that, given realistic input and output representations based on English orthography and phonology, differences in performance associated with differences in AoA (operationalized as frequency trajectory) disappeared if cumulative frequency was equated. However, later simulations reported by P. Monaghan and Ellis (2010) demonstrated that if models with realistic input and output representations are trained with a realistic distribution over time of the frequency of encounter of words then an AoA effect is observed, independent of the effect of cumulative frequency. This independent AoA effect results from an association between the point at which words enter the vocabulary (the age of acquisition) and the effect on performance of the relative plasticity of network connections. P. Monaghan and Ellis (2010) found that an effect of frequency trajectory could also be detected, in addition to the effect of the point-of-entry, but that the trajectory effect accounted for a relatively small amount of variance in network performance.

Our observation that the AoA effect decreases with increasing age shows that it was premature to conclude that the AoA effect remains fixed through the lifespan. The age by AoA interaction is clearly consistent with the cumulative frequency account. However, there is evidence that it would be premature now to identify the AoA effect with the impact of cumulative frequency. The simulations reported by P. Monaghan and Ellis (2010) distinguished an AoA effect on reading independent of a cumulative frequency effect. Our observation of distinct frequency, AoA, and age by AoA interaction effects implies that a component of the AoA effect can be linked to cumulative frequency but also that there must be an independent AoA effect.

Two additional observations are germane. Firstly, we found that the AoA effect was larger in lexical decision than word pronunciation. Secondly, we found that the interaction between the age and AoA effects is apparent, in the subject-level analysis plots, as a steeper decline over age in the size of the AoA effect in pronunciation than in lexical decision. These observations are relevant because the AoA effect can be argued to have components linked to both orthography-to-phonology mappings, as required for reading aloud, and to semantic processing, as required, arguably, for lexical decision. The P. Monaghan and Ellis (2010) simulations demonstrated that an AoA effect could be observed in a network implementing only the orthography-to-phonology mapping. However, a range of evidence suggests that AoA effects are more prominent where task performance draws on semantics. A review of multi-task investigations reported by Brysbaert and Ghyselinck (2004) indicated that the AoA effect has been found to be larger in tasks, like object naming, that rely on semantics. Computational simulations have shown that the AoA effect is stronger where there is an arbitrary relation between the input and output patterns that must be learned (Ellis & Lambon Ralph, 2000; Lambon Ralph & Ehsan, 2006; Mermillod et al., 2012; Zevin & Seidenberg, 2002), and input-output mappings involving semantics can be characterized as arbitrary. Finally, Cortese and Khanna (2007) observed that the AoA effect is larger in lexical decision than in word naming, supporting the interpretation that the lexical decision task emphasizes semantics (Chumbley & Balota, 1984).

We propose that the modulation of the AoA effect by age may indicate that a component of the AoA effect reflects the impact of the order of learning on the efficient of orthography-to-phonology mappings. This is the component revealed in the age by AoA interaction observed in the pronunciation data. The AoA effect is reduced over increasing age, consistent with a cumulative frequency account, but it is not altogether eliminated over increasing age, consistent with P. Monaghan and Ellis' (2010) observation of independent AoA and cumulative frequency effects. In comparison, we observed that the decline of the AoA effect progresses steadily with increasing age, in lexical decision. We propose that this may be explained by supposing that the AoA effect reflects the involvement of semantic knowledge in lexical decisions, and that mappings to or from semantics approach asymptote more slowly as experience accumulates.

Theoretical accounts of the reading system

Given our observations, what must general theoretical accounts of the cognitive reading system explain? Theories about the reading system have been evaluated on their capacity to account for benchmark effects (e.g. as listed by Coltheart et al., 2001). We suggest that reading theories must now explain two additional findings: (1.) the frequency and AoA effects decrease over increasing age; and (2.) the modulation of the item effects takes place in the context of a broader shift such that reading performance is influenced by a U-shaped effect of age on latencies. How far can existing theories explain, or can be extended to explain, these results?

No prediction of decline in the frequency effect can be derived from verbal descriptions of the dual route account (e.g. Castles & Coltheart, 1993) or from current implementations of the dual route account as a computational model (Coltheart et al., 2001; Perry et al., 2007). In an implementations of a dual route type theory, the CDP+ model (Perry et al., 2007), the non-lexical route develops in a two-layer associative (TLA) network that learns to map orthographic to phonological representations. This non-lexical route is sensitive to the statistical distribution of orthographic to phonological mappings but cannot learn whole word associations while the lexical route, in both the DRC and CDP+ implementations, does not and cannot learn.

However, a dual route reading system, if that is what underlies skilled reading behaviour, must emerge from development. Granted development, we should expect behavioural phenomena that reflect the functioning of reading routes to grow stronger in association with the increasing development of those routes. We think that this predicts larger lexical effects over increasing age from the beginning of development to the emergence of the mature reading system. We note that the current theory has nothing to say about the AoA effect but we suppose that the AoA effect could arise in a dual route reading system by influencing lexical unit activation, similar to the impact of frequency, or, given current claims about the locus of the AoA effect (Ghyselinck & Brysbaert, 2006), by modulating the activation of semantic units in the lexical semantic route, or the strength of links to or from semantics in that route. The critical questions that future simulation work must address are, firstly, whether this is, in fact, how a dual route system behaves in development and, secondly, whether prolonged exposure to word stimuli and other age-related effects are associated with changes in psycholinguistic effects of the kind we report.

The potential for modifications of a dual route model to simulate the critical effects we have reported are delimited by the results of recent simulation studies of individual differences in reading in adults (Adelman et al., 2014) and children

(Ziegler et al., 2008). Adelman et al. (2014) examined the correlations between subject-level estimates of psycholinguistic effects with reference to the capacity of computational models to simulate those correlations. Their aim was to evaluate the adequacy of theoretical assumptions, implemented in the models, because if the effects of psycholinguistic variables are assumed to have common loci then they should be vulnerable to the same sources of variation between individuals and, as a result, participants showing a strong effect of one variable should show a strong effect of the co-located variable, yielding a correlation. Adelman et al. (2014) observed the effects of critical psycholinguistic variables on adult pronunciation (100 participants, aged 17-55 years). They calculated the effects of psycholinguistic variables for each participant, and then calculated the correlations between the subject-level estimates of the different psycholinguistic effects.

Most relevant to our discussion, Adelman et al. (2014) conducted a comprehensive examination of the performance of 250,000 implementations of the DRC and CDP+ computational models, each implementation representing a different combination of parameter values. Their analyses showed that several hundred parameter sets were capable of simulating observed item effects, and that among these models many represented parameter sets capable of simulating observed subject-level estimates of item effects. Critically, the simulations were found to be capable of recovering the observed correlations between psycholinguistic effects. Moreover, Adelman et al. (2014) showed that the DRC and CDP+ simulations were less successful than a version of the DRC which word frequency did not bias the input to orthographic lexical units (as it does in the DRC, Coltheart et al., 2001) but did modulate the weights on connections from orthographic to phonological units. They thus demonstrated that some assumptions about reading system architectures, but not others, can furnish the scope for variation in parameters that captures individual differences in psycholinguistic effects.

The Adelman et al. (2014) observations are theoretically important but in evaluating the potential for modifications of dual route model parameters to fit our observations, we think that future studies shall have to address theoretical accounts of the main drivers for the interactions between the frequency or AoA effects and age, and for the overarching impact of individual differences, the U-shaped effect of age and the curvilinear effect of reading ability. The Adelman et al. (2014) simulations show that a variety of implementations of the dual route architecture can capture individual differences in psycholinguistic effects. Simulation studies reported by Ziegler et al. (2008) further demonstrate that dual route implementations can simulate observed individual differences, granted adaptations informed by measures of ability. However, without further simulations, it is unclear if instances of the dual route architecture could simulate developmental, skilled and aged reading. More generally, it is unclear if a dual route account could explain why psycholinguistic effects should vary in association with individual differences in age and ability in the form observed.

In comparison, our findings appear to have ready explanation in a connectionist account. The interactions between frequency or AoA and age or reading skill are consistent with the gradual ceiling effect predicted to result from the assumption, in connectionist systems, of asymptotic learning based on distributed representations and a nonlinear input-output function (e.g. Plaut et al., 1996; Van Orden, Pennington, & Stone, 1990). More broadly, the predictions of a connectionist account would explain our observations of decreases in the effects of imageability with increasing age, and of orthographic neighbourhood size with increasing reading skill. These interactions are consistent with previous observations that the neighbourhood effect is smaller for responses to more frequent words (Andrews, 1989, 1992; Sears, Hino & Lupker, 1995). They are also consistent with observations that while the neighbourhood effect has been found to influence reading in children (Laxon, V. Coltheart, & Keating, 1988; Laxon, Gallagher, & Masterson, 2002), it appears to be smaller for more skilled emergent readers (Laxon et al., 1988). Likewise, the interactions we report are consistent with previous observations that the imageability effect appears to be more robustly observed for low frequency words (Strain et al., 1995; but see J. Monaghan & Ellis, 2002). In general, the principle features of connectionist reading models, asymptotic learning based on distributed representations and a nonlinear input-output function, explain the reduction in the effects of word properties as a function of the increasing approach of the system towards maximal efficiency as experience accumulates and skill develops.

Recent computational studies demonstrate that connectionist models of the reading system may indeed be capable of simulating individual differences in psycholinguistic effects of the kind we observed. Dilkina, McClelland, and Plaut (2008; see also Plaut, 1997) have argued that variation in patterns of preserved or impaired reading ability seen in different brain-injured patients may be explained by individual differences in biology or experience, while assuming commonality in reading system structure. Dilkina et al. (2008) found that variation in the preservation of semantic and lexical abilities, observed in patients, could be simulated using variants of the same network structure differing in training regime, orthography-to-phonology pathway size, or the extent or location of damage to the network. Training regime was varied by manipulating the number of presentations of input patterns to the orthographic layer, while the orthography-to-phonology pathway size was varied by manipulating the number of hidden units connecting the orthographic to the phonological layer. Remarkably, the data for most patients could be fitted by the same network given only variation in training regime.

We predict that the observed decrease in the frequency and AoA effects with increasing age should reflect the diminution in the impact of experience as network efficiency approaches asymptote, in a connectionist model given prolonged training. While there is variation in structure among different studies, this should be true in any connectionist model, given the principle that the network unit activation function takes a nonlinear sigmoidal form. Our observation of the modulation of frequency and imageability item effects by differences in reading skill could be explained by the impact of an association between differences in reading skill and variation in the size of the orthography-to-phonology pathway. As the pathway is strengthened, there will be less room for the impact of word semantics in pronunciation (Plaut et al., 1996; Strain et al., 1995). We would, equally, expect the modulation of frequency by skill (distinct from the frequency by age interaction) whether the frequency effect is located in orthography-to-phonology connections (P. Monaghan & Ellis, 2010; Seidenberg & McClelland, 1989) or in other aspects of the reading system. These predictions are consistent with the results of Dilkina et al.'s (2008) observation that more training and a larger OP pathway are associated with high levels of performance and that, at that high level, item effects tend to converge.

As we have noted, existing simulations present a curvilinear trajectory for network performance, as it improves towards asymptote (e.g. Zevin & Seidenberg, 2002). The curvilinear function matches that observed in the developmental phase of our data. No current connectionist simulations address the impact of ageing but we think it is possible that the introduction of "neural noise" (Li, Lindenberger, & Sikström, 2001) may slow network response output, in line with the age-related slowing we and others have observed. Such a manipulation may help to explain the age-related slowing we observed in pronunciation but it may not be sufficient to account for the age effect in lexical decision. Li et al. (2001) link age-related changes in the dopaminergic system to age-related declines in cognitive performance via a reduction in the distinctiveness of neural representations. In their account, attenuation in the dopaminergic modulation of synaptic transmission means that network unit activation in response to input signals is reduced, more variable, that is, noisier, relative to background levels of random activation variation. In our analysis, age-related slowing appears to be related to an increase in non-decision time. This account is supported by the results of our analysis of pronunciation durations which suggest that the slowing we observed in pronunciation is associated with an increase in articulatory coding (response execution) processing efficiency. We think, then, that the assumption of age-related increase in neural noise would equip a connectionist account to explain the U-shaped effect of age on reading latencies if that neural noise influenced the resolution of network activation on output representations.

<u>Implications for theoretical accounts of cognitive development and aging</u>

Researchers usually focus on the effects of word properties in item-level analyses or on the effects of individual differences. The benefit of a multilevel analysis of reading is that it afforded two new insights. The first is that psycholinguistic effects systematically vary in relation to individual differences in age and reading ability. The second is that this variation happens against a background of large, over-arching, effects on performance due to individual differences. Our mixed-effects models showed that the effects of word properties, and their modulation by individual differences is significant, but that the dominant source of variance in reading performance are those individual differences.

We found that the variance explained by the (fixed) psycholinguistic effects in our mixed-effects models was about 5% (the marginal R^2_m, Nakagawa & Schielzeth, 2013), the variance explained by the fixed effects due to individual differences including age was about 20-25%, and that explained by the interactions between the psycholinguistic effects and the individual differences effects accounted for an additional 1%. We are the first, we believe, to report variance explained at the item-level and at the person-level on the basis of the same model of the same trial-level data. Traditional item-level analyses have yielded estimates of variance explained of around 50% (e.g. Balota et al., 2004), thus delimiting the work that must be done by models of reading (Spieler & Balota, 1997). Our results show that if researchers do not average response data to by-items mean latencies, 'washing out' subject-level variability, an adequate account of reading must attend both to differences between participants and to differences between words in fully accounting for systematic variance in reading behaviours. This conclusion simply mirrors the conclusions drawn previously by Seidenberg and Plaut (1997) in response to earlier item-level estimates of how much variance must be explained by models of reading (Spieler & Balota, 1997). What is new is our conclusion that any account of reading must explain the systematic variation in psycholinguistic effects due to differences in age and reading ability, appearing in the context of a global U-shaped effect of age.

Our diffusion model analyses indicated that we can account for the U-shaped effect of age on lexical decision latencies in terms of age-related changes in, especially, drift rate, the quality of information accumulated from stimuli, and non-decision time, the time taken by and thus the relative efficiency of stimulus encoding and response execution processes. Our findings extend those previously reported by Ratcliff and colleagues (2004, 2010, 2012) though we did not observe, as they did, the large effect of age on boundary separation in our analyses. This may be because the models we fitted of diffusion model parameters included reading ability as well as age, and it was differences in ability that accounted for variation in boundary separation. The results of our diffusion analyses suggest that the speed-up in response latencies from childhood to adulthood can be explained by an age-related increase in the quality of information extracted from stimuli. This is a trend that decelerates in adulthood, with drift rate remaining at a high level into old age. In addition, there is an age-related decrease in non-decision time, reflecting an increase in the efficiency of stimulus encoding and response execution processes into adulthood. These results appear consistent with accounts of development in which the efficiency of lexical access processes (e.g. P. Monaghan & Ellis, 2010; Seidenberg & McClelland, 1989) or the quality of lexical representations (e.g. Perfetti, 2007; Perfetti & Hart, 2001, 2002) improves through childhood into adulthood.

The later slow down in response speed can be explained, in our analysis, by an increase in non-decision time. This is consistent with previously observed age-related changes in visual-sensory (Faubert, 2002; Schneider & Pichora-Fuller, 2000), orthographic encoding (Allen et al., 1993; Allen et al., 2011; Madden, 1992) and response output execution processes (Allen et al., 1993; Stelmach et al., 1988). In our own analysis of pronunciation latencies, the increase in the length effect with increasing age may reflect a decrease in the efficiency of response encoding processes. That would fit with the observation of a U-shaped effect of age on pronunciation durations, implying a speed-up from childhood into adulthood, and a slow-down through adulthood into old age, of response execution processes.

Could we, then, account for the results we observed, the interaction between frequency or AoA effects, along with the U-shaped effect of age overall, simply by embedding a developmental model of the reading system within a broader framework in which lifespan development and ageing is most important to stimulus encoding and response execution, that is, more peripheral cognitive processes? Such an account would resemble a two-factor slowing model of ageing (e.g. Hale & Myerson, 1996) in which age-related cognitive slowing is greater among non-lexical than among lexical processes. It is possible to envisage a theory that explains our results assuming the features of a connectionist lexical access system embedded within a diffusion model decision process. However, Norris (2009; see also Norris, 2006) identifies a logical problem with such an account.

In the diffusion model, the lexical access process outputs a wordness value that determines the relative drift rate. But the critical assumption in this account is that differences in response speed to low versus high frequency words result from differences in drift rate. The lexical access system is assumed to terminate in an output wordness value at the same time irrespective of differences in word properties Norris (2009). It is difficult to reconcile this limitation with the substantial, systematic, and broad age-related changes in frequency, AoA, and imageability we observed in both lexical decision and pronunciation. Perhaps all such changes are associated with a global developmental change in drift rate. That is a possibility that could be tested in future research. However, a promising approach is presented in the Bayesian reader model, which subsumes the diffusion model as a special case (Norris, 2006, 2009). In the Bayesian reader model, noisy input information is integrated with prior knowledge of the likelihood of occurrence of a word to evaluate the probable identity of a stimulus. In this account, lexical access and decision processes operate simultaneously and are integrated. A testable possibility is that prior knowledge of the probability of occurrence of a word in a reading context could be informed by frequency (context distinctiveness), AoA, and other factors and that that information could shift, in line with our observations, as readers accumulate more experience and skill over the life-span. However, even if a Bayesian reader or a diffusion model could successfully simulate the pattern of change we observed, the pattern of effects on pronunciation that we report would be outside its scope. Future simulation work shall have to examine if extensions of a Bayesian reader model, or of a connectionist model, are sufficient to account for the results observed.

Conclusions

The effects of psycholinguistic variables are critical to the evaluation of theories about the cognitive reading system. Our findings show that the effects on reading of two key variables, frequency and AoA, decrease in size with increasing age over the life-span. In answer to the question with which we began: the reading system does change, that change is seen in the reduction in psycholinguistic effects over increasing age. But the systematic modulation of psycholinguistic effects was observed in the context of substantial over-arching effects due to age and individual differences in reading ability. From childhood to adulthood, reading responses speed up, but through adulthood into old age, responses slow down, that is, a marked U-shaped effect of age. We think our findings can be explained by theoretical accounts that incorporate learning as the basis for the development of structure in the reading system. However, an adequate theory shall have to include assumptions about both developmental learning and later ageing: a life-span theory of reading.

References

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity not frequency determines word naming and lexical decision times. *Psychological Science, 17* 814-823.

Adelman, J. S., Sabatos-DeVito, M. G., Marquis, S. J., & Estes, Z. (2014). Individual differences in reading aloud: A mega-study, item effects, and some models. *Cognitive Psychology, 68*, 113-160.

Allen, P. A., Bucur, B., Grabbe, J., Work, T., & Madden, D. J. (2011). Influence of encoding difficulty, word frequency, and phonological regularity on age differences in word naming. *Experimental Aging Research, 37*, 261-292.

Allen, P. A., Madden, D. J., & Crozier, L. C. (1991). Adult age differences in letter-level and word-level processing. *Psychology and Aging, 6*, 261-271.

Allen, P. A., Madden, D. J., Weber, T. A., & Groth, K. E. (1993). Influence of age and processing stage on visual word recognition. *Psychology and Aging, 8*, 274-282.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory and Cognition, 10*, 565-575.

Andrews, S. (1989). Frequency and neighbourhod effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 802-814.

Andrews, S. (1992). Frequency and neighbourhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 234-254.

Andrews, S., & Hersch, J. (2010). Lexical Precision in Skilled Readers: Individual Differences in Masked Neighbor Priming. *Journal of Experimental Psychology-General, 139*, 299-318.

Baayen, R. H. (2008). *Analyzing linguistic data*: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Balota, D. A., & Chumbley, J. L. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 340-357.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283-316.

Balota, D. A., & Duchek, J. M. (1988). Age-related differences in lexical access, spreading activation, and simple pronunciation. *Psychology and Aging, 3*, 84-93.

Balota, D. A., & Ferraro, F. R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer type. *Journal of Memory and Language, 32*, 573-592.

Balota, D. A., & Ferraro, F. R. (1996). Lexical, sublexical, and implicit memory processes in healthy young and healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology, 10*, 82-95.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftus, B., . . . Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445-459.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Barry, C., Johnston, R. A., & Wood, R. F. (2006). Effects of age of acquisition, age, and repetition priming on object naming. *Visual Cognition, 13*, 911-927.

Bartoń, K. (2016). MuMIn: multi-model inference. *R package version*, *1.15.6*. URL: https://cran.r-project.org/web/packages/MuMIn/index.html

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., Grothendieck, G., & Green, P. (2016). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, *1.1-12*. URL: https://github.com/lme4/lme4/

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America, 113*, 1001-1024.

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never-ending story? Journal of Memory and Language, 50, 456–476.

Bonin, P., Méot, A., Mermillod, M., Ferrand, L., & Barry, C. (2009). The effects of age of acquisition and frequency trajectory on object naming: Comments on

Pérez (2007). The Quarterly Journal of Experimental Psychology, 62, 1132–1140.

Bowles, N. L., & Poon, L. W. (1981). The effect of age on speed of lexical access. *Experimental Aging Research, 7*, 417-425.

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition*, *13*, 992-1011.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved frequency measure for American English. *Behavior Research Methods, 41*, 977-990.

Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpho-lexical reading develop in readers of a shallow orthography? *Brain and Language, 81*, 568-586.

Butler, B., & Hains, S. (1979). Individual differences in word recognition latency. *Memory and Cognition, 7*, 68-76.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *The Quarterly Journal of Experimental Psychology*, *25*, 85-95.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 148-180.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 98*, 67-83.

Cerella, J., & Hale, S. (1994). The rise and fall in information-processing rates over the life span. *Acta Psychologica, 86*, 109-197.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory and Cognition, 28*, 143-153.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory and Cognition, 12*, 590-606.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*, 249-253.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences (3rd. edition)*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen-Shikora, E. R., & Balota, D. A. (2016). Visual word recognition across the adult lifespan. *Psychology and Aging*, *31*, 488-502.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies in information processing* (pp. 151-216). London: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed processing approaches. *Psychological Review, 100*, 589-608.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535-555). Hillsdale, NJ: LEA.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1197-1211.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204-256.

Comalli, P. E. (1965). Cognitive functioning in a group of 80-90-year-old men. *Journal of Gerontology, 20*, 14-17.

Cortese, M. J., & Fugett, A. (2004). Imageability Ratings for 3,000 Monosyllabic Words. *Behavior Methods and Research, Instrumentation, & Computers, 36*, 384-387.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *The Quarterly Journal of Experimental Psychology, 60*, 1072-1082.

Davis, C. J. (2005). N-watch: A program for deriving neighbourhood size and other psycholinguistic statistics. *Behavior Research Methods, 3*, 65-70.

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2008). A single-system account of semantic and lexical deficits in five semantic dementia patients. *Cognitive Neuropsychology*, *25*, 136-164.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1103-1123.

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 56*, 164-176.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124.

Frederickson, N., Frith, U., & Reason, R. (1997). *Phonological Assessment Battery [PhAB]: Manual and test materials*. London: nfer Nelson Publishing Company Ltd.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language, 84*, 474–496.

Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language, 66*, 789–806.

Gahl, S., & Strand, J. F. (2016). Many neighborhoods: Phonological and perceptual neighborhood density in lexical production and perception. *Journal of Memory and Language, 89*, 162-178.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*, 43-67.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology, 5*, 674-691.

Hale, S., 1990. A global developmental trend in cognitive processing speed. *Child Development, 61*, 653-663.

Hale, S., Lima, S. D., & Myerson, J. (1991). General cognitive slowing in the nonlexical domain: An experimental validation. *Psychology and Aging*, *6*, 512-521.

Hale, S., & Myerson, J. (1996). Experimental evidence for differential slowing in the lexical and nonlexical domains. *Aging Neuropsychology and Cognition, 3*, 154-165.

Hale, S., J. Myerson and D. Wagstaff, 1987. General slowing of nonverbal information processing: Evidence for a power law. Journal of Gerontology 42, 131-136.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662-720.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science, 26*, 433-443.

Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), Aging and cognitive processes (pp. 237–278). Boston: Plenum Press.

Huttenlocher, J. (1984). Word recognition and word production in children. In H. Bouma & D. G. Bouwhuis (Eds.), Attention and performance X (pp. 447-457). Hillsdale, NJ: Erlbaum.

Johnson, P. C. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*, 944-946.

Johnston, R., & Barry, C. (2006). Age of acquisition and lexical processing. Visual Cognition, 13, 789-845.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*, 684-712.

Kail. R. 1986. Sources of age differences in speed of processing. *Child Development, 57*, 969-987.

Kail, R. 1991. Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin, 109*, 490-501.

Keeulers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decisiond ara for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44*, 287-304.

Kessler, B., Treiman, R., & Mullenix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language, 47*, 145-171.

Kuperman, V., Stadthagen-Gonzales, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods, 44,* 978-990.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2016). lmerTest: Tests in Linear Mixed Effects Models. *R package version*, *2.0-32*. URL: https://cran.r-project.org/web/packages/lmerTest/index.html

Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition, 13*, 928-948.

Laxon, V., Gallagher, A., & Masterson, J. (2002). The effects of familiarity, orthographic neighbourhood density, letter-length and graphemic complexity on children's reading accuracy. *British Journal of Psychology, 93*, 269-287.

Laxon, V. J., Coltheart, V., & Keating, C. (1988). Children find friendly words friendy too: Words with many orthographic neighbours are easier to read and spell. *British Journal of Educational Psychology, 58*, 103-119.

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. G. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General, 122*, 316-330.

Lewis, M. B., Gerhand, S., & Ellis, H. D. (2001). Re-evaluating age-of-acquisition effects: are they simply cumulative-frequency effects?. *Cognition*, *78*, 189-205.

Li, S. C., Lindenberger, U., & Sikström, S. (2001). Aging cognition: from neuromodulation to representation. *Trends in Cognitive Sciences*, *5*, 479-486.

Lima, S. D., Hale, S., & Myerson, J. (1991). How general is general slowing? Evidence from the lexical domain. *Psychology and Aging, 6*, 416-425.

Lorch, R. F., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 149-157.

Madden, D. J. (1992). Four to ten milliseconds per year: Age-related slowing of visual word identification. *Journal of Gerontology, 47*, 59-68.

Masterson, J., & Hayes, M. (2007). Development and data for UK versions of an author and title recognition test for adults. *Journal of Research in Reading*, *30*, 212-219.

McClelland, J., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375-407.

McKoon, G., & Ratcliff, R. (2016). Adults with poor reading skills: how lexical knowledge interacts with scores on standardized reading comprehension tests. *Cognition*, *146*, 453-469.

Mermillod, M., Bonin, P., Méot, A., Ferrand, L., & Paindavoine, M. (2012). Computational evidence that frequency trajectory theory does not oppose but emerges from Age-of-Acquisition theory. *Cognitive Science, 36*, 1499–1531.

Moers, C., Meyer, A., & Janse, E. (2016). Effects of Word Frequency and Transitional Probability on Word Reading Durations of Younger and Older Speakers. *Language and Speech*, doi: 10.1177/0023830916649215

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*, 183-206.

Monaghan, P., & Ellis, A. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language, 63*, 506-525.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 116-133.

Morrison, C. M., Hirsh, K. W., Chappell, T. D., & Ellis, A. W. (2002). Age and age of acquisition: An evalution of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology, 14*, 435-459.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76*, 165-178.

Mousikou, P., & Rastle, K. (2015). Lexical frequency effects on articulation: a comparison of picture naming and reading aloud. Frontiers in Psychology, DOI: 10.3389/fpsyg.2015.01571

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: the rank hypothesis. *Psychological Review*, *111*, 721-756.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133-142.

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: A problem of significance. *Nature Neuroscience, 14*, 1105-1107.

Noble. C., Baker, B., & Jones, T. N. (1964). Age and sex parameters in psychomotor learning. *Perceptual and Motor Skills, 19*, 935-945.

Norris, D. (2006). The Bayesian reader: explaining word recognition as an optimal Bayesian decision process. *Psychological review*, *113*, 327-357.

Norris, D. (2009). Putting it all together: A unified account of word recognition and reaction-time distributions. *Psychological Review*, *116*, 207-219.

Olson, R.K., Kliegl, R., Davidson, B.J., & Foltz, G. (1985). Individual and development differences in reading disability. In G.E. MacKinnon & T.G. Waller (Eds.), *Reading research: Advances in theory and practice* (pp. 1-64). Orlando: FL: Academic Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357-383.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 67-86). Washington, DC: American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). Amsterdam: John Benjamins.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*, 273-315.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-plus (Statistics and Computing)*. New York: Springer.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12*, 765-805.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56-115.

Protopapas, A. (2007). CheckVocal: A program to facilitate checking accuracy and response time of vocal rwesponses from DMDX. *Behavior Research Methods, 39,* 859-862.

Rabbitt, P., Diggle, P., Holland, F., & McInnes, L. (2004). Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 59*, 84-97.

Rabbitt, P. M. A., McInnes, L., Diggle, P., Holland, F., Bent, N., Abson, V., ... & Horan, M. (2004). The University of Manchester longitudinal study of cognition in normal healthy old age, 1983 through 2003. *Aging Neuropsychology and Cognition, 11*, 245-279.

R Core Team (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, *111*, 159-182.

Ratcliff, R., Love, J., Thompson, C. A., & Opfer, J. E. (2012). Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses. *Child Development, 83*, 367-381.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical decision task. *Psychology and Aging, 19*, 278-289.

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology, 60*, 127-157.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428.

Salthouse, T. A. (2004). What and when of cognitive aging. Current directions in psychological science, 13, 140-144.

Salthouse, T. A. (2014). Quantity and structure of word knowledge across adulthood. Intelligence, 46, 122-130.

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implication of perceptual deterioation for cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds.), The handbook of aging and cognition (2nd ed., pp. 155–219). Mahwah, NJ: Erlbaum.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighbourhood size and neighbourhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 876-900.

Sears, C. R., Siakaluk, P. D., Chow, V. C., & Buchanan, L. (2008). IS there an effect of print exposure on the word frequency effect and the neighbourhood size effect? *Journal of Psycholinguistic Research, 37*, 269-291.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Seyfarth, S. (2014). Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition, 133*, 140-155.

Silberzahn, R., & Uhlmann, E. L. (2015). Crowdsourced research: Many hands make tight work. *Nature*, *526*, 189-191.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359-1366.

Snijders, T.A., & Bosker, R.J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling (2nd Edition)*. Springer Berlin Heidelberg.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8*, 411-416.

Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging, 15*, 225-231.

Stainthorp, R. (1997). A children's author recognition test: A useful tool in reading research. *Journal of Research in Reading*, *20*, 148-158.

Stelmach, G. E., Goggin, N. L., & Amrhein, P. C. (1988). Aging and the restructuring of precued movements. *Psychology and Aging, 3*, 151-157.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 1140-1154.

Tainturier, M. J., Tremblay, M., & Lecours, A. R. (1989). Aging and the word frequency effect: A lexical decision investigation. *Neuropsychologia, 27*, 1197-1203.

Tainturier, M. J., Tremblay, M., & Roch Lecours, A. (1992). Educational level and the word frequency effect: A lexical decision investigation. *Brain and Language, 43*, 460-474.

Taraban, R., & McClelland, J. L. (1987). Conspiracy effects in word pronunciation. *Journal of Memory and Language, 26*, 608-631.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *TOWRE Test of word reading efficiency*. Austin,TX: Pro-ed.

van Heuven, W. J. B., Mandera, P., Keeulers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*, 1176-1190.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97*, 488-522.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age–cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. Psychological bulletin, 122, 231-249.

Wagenmakers, E. J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14*, 3-22.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60*, 502-529.

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual Differences in Visual Word Recognition: Insights From the English Lexicon Project. *Journal of Experimental Psychology-Human Perception and Performance, 38*, 53-79.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin and Review, 15*, 971-979.

Zeguers, M. H., Snellings, P., Tijms, J., Weeda, W. D., Tamboer, P., Bexkens, A., & Huizenga, H. M. (2011). Specifying theories of developmental dyslexia: a diffusion model analysis of word recognition. *Developmental Science*, *14*, 1340-1354.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*, 1-29.

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of- Acquisition efects in reading aloud: Tests of cumulat- ive frequency and frequency trajectory. *Memory and Cognition, 32*, 31–38.

Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F. X., & Perry, C. (2008). Developmental dyslexia and the dual route model of reading: Simulating individual differences and subtypes. *Cognition, 107*, 151-178.

Table 1. Summary of participant attributes

	N	Mean	SD	Minimum	Maximum
Task: decisions					
Age (months)	357	337.4	210.26	96	996
Age (years)	357	28.12	17.52	8	83
TOWRE words accuracy	357	85.11	14	26	104
TOWRE words time	357	44.55	1.55	36	45
TOWRE words skill	357	1.92	0.36	0.58	2.89
TOWRE nonwords accuracy	357	50.73	10.86	10	63
TOWRE nonwords time	357	43.93	3.22	26	45
TOWRE nonwords skill	357	1.17	0.31	0.22	2.38
Phonological awareness (Spoonerisms)	355	25.16	4.39	7	30
Task: naming					
Age (months)	181	299.55	191.84	116.04	912
Age (years)	181	24.96	15.99	9.67	76
TOWRE words accuracy	181	89.48	14.1	29	104
TOWRE words time	181	44.23	1.78	37	45
TOWRE words skill	181	2.03	0.37	0.64	2.81
TOWRE nonwords accuracy	181	52.89	10.53	12	63
TOWRE nonwords time	181	43.56	3.15	30	45
TOWRE nonwords skill	181	1.23	0.3	0.27	2.1
Phonological awareness (Spoonerisms)	181	25.41	4.55	3	30

Note that phonological awareness scores were not recorded for two participants who performed the lexical decision task, and experimental word naming data were not recorded for one participant who performed the naming task. Scores for these participants are included in the summary but their data were excluded before analysis of the experimental reading data.

Table 2. Summary of word attributes

	Mean	SD	Min	Max
Length (letters)	4.3	0.7	3.0	6.0
OLD	1.5	0.3	1.0	2.1
PLD	1.3	0.3	1.0	2.0
Orthographic neighbourhood size	7.1	5.0	0.0	24.0
BG-Sum	5566.7	2890.3	418.0	13656.0
BG-Mean	1675.9	813.9	168.3	4149.7
BG-Frequency by position	1202.8	593.9	70.0	2799.0
Log10 (SUBTLEX-UK Context Distinctiveness)	3.8	0.7	1.6	4.6
SUBTLEX-UK word form frequency per million	411.4	1004.1	0.2	7903.6
Rated Imageability	4.4	1.3	1.7	6.8
Rated Age-of-Acquisition	5.8	2.1	2.8	12.0

OLD = Orthographic Levenshtein Distance; PLD = Phonological Levenshtein Distance; BG-Sum = summed bigram frequency; BG-Mean = mean bigram frequency; BG-Frequency by position = bigram frequency by position; SUBTLEX-UK word form frequency per million = total frequency of occurrence of word, per million, in SUBTLEX-UK corpus; Log10 (SUBTLEX-UK Context Distinctiveness) = log base 10 of context distinctiveness count, SUBTLEX-UK corpus.

Table 3. Summary of the bivariate correlations between RT and the critical participant and word attribute variables in each task

Lexical decisions data	RT (ms)	Age	Word skill	NW skill	PhA	Length	N-size	BG-Sum	BG-Mean	BG-Freqpos	OLD	Log10CD	IMG	AoA	Regularity
Age (months)	-0.15***														
TOWRE words skill	-0.30***	0.32***													
TOWRE nonwords skill	-0.28***	0.32***	0.80***												
Phonological awareness (Spoonerisms)	-0.25***	0.28***	0.53***	0.61***											
Length (letters)	0.03***	0.01	0.01**	0.01**	0.01*										
Orthographic neighbourhood size	-0.04***	-0.01*	-0.01**	-0.01**	-0.01*	-0.61***									
BG-Sum	0.01*	0.00	0.00	0.01	0.00	0.46***	-0.05***								
BG-Mean	0.00	0.00	0.00	0.00	0.00	0.07***	0.21***	0.89***							
BG-Frequency by position	0.02***	0.00	0.00	0.00	0.00	0.36***	0.10***	0.66***	0.57***						
OLD	0.04***	0.01*	0.01**	0.01**	0.01*	0.67***	-0.95***	0.05***	-0.25***	-0.06***					
Log10 (SUBTLEX-UK Context Distinctiveness)	-0.15***	-0.03***	-0.03***	-0.03***	-0.02***	-0.34***	0.33***	-0.08***	0.05***	-0.05***	-0.37***				
Rated Imageability	-0.01**	-0.01	-0.01	-0.01	0.00	-0.05***	-0.03***	-0.07***	-0.05***	-0.06***	0.03***	-0.31***			
Rated Age-of-Acquisition	0.12***	0.03***	0.02***	0.02***	0.02***	0.29***	-0.17***	0.09***	-0.01	0.09***	0.22***	-0.53***	-0.15***		
Regularity	0.01*	0.00	0.00	0.00	0.00	-0.09***	0.06***	-0.01**	0.05***	-0.11***	-0.06***	-0.03***	0.02***	0.09***	
OP consistency	0.02***	0.00	0.00	0.00	0.00	0.14***	-0.31***	-0.03***	-0.09***	-0.04***	0.33***	-0.12***	0.06***	0.09***	0.14***
Pronunciation data															
Age (months)	-0.15***														
TOWRE words skill	-0.44***	0.47***													
TOWRE nonwords skill	-0.42***	0.45***	0.87***												
Phonological awareness (Spoonerisms)	-0.34***	0.40***	0.56***	0.58***											
Length (letters)	0.05***	0.01	0.02**	0.02**	0.02**										
Orthographic neighbourhood size	-0.06***	-0.01	-0.01	-0.01	-0.01*	-0.60***									
BG-Sum	0.02***	0.00	0.01	0.01	0.01	0.47***	-0.05***								
BG-Mean	0.01	0.00	0.00	0.00	0.00	0.07***	0.21***	0.89***							
BG-Frequency by position	0.04***	0.00	0.00	0.01	0.01	0.36***	0.10***	0.66***	0.57***						
OLD	0.06***	0.01	0.01	0.01	0.01*	0.66***	-0.95***	0.05***	-0.25***	-0.07***					
Log10 (SUBTLEX-UK Context Distinctiveness)	-0.11***	-0.01*	-0.03***	-0.03***	-0.03***	-0.32***	0.33***	-0.08***	0.05***	-0.04***	-0.37***				
Rated Imageability	-0.01	0.00	0.00	0.00	0.00	-0.03***	-0.03***	-0.07***	-0.06***	-0.05***	0.03***	-0.30***			
Rated Age-of-Acquisition	0.09***	0.01*	0.03***	0.03***	0.03***	0.27***	-0.16***	0.08***	-0.01	0.09***	0.21***	-0.54***	-0.17***		
Regularity	-0.02***	-0.01	-0.01	-0.01	0.00	-0.08***	0.06***	-0.01	0.06***	-0.10***	-0.07***	-0.03***	0.03***	0.09***	
OP consistency	0.02**	0.00	0.00	0.00	0.00	0.12***	-0.31***	-0.03***	-0.08***	-0.04***	0.33***	-0.12***	0.04***	0.09***	0.12***

Age = Age (months); Word skill = TOWRE words skill; NW skill = TOWRE nonwords skill; PhA = Phonological awareness (Spoonerisms); Length = Length (letters); N-size = Orthographic neighbourhood size; BG-Freqpos = BG-Frequency by position; Log10CD = Log10 (SUBTLEX-UK Context Distinctiveness); IMG = Rated Imageability;AoA = Rated Age-of-Acquisition; *** if p < .001; ** if p < .01; * if p < .05

Table 4. Summary of slopes-as-outcomes models of the effects of age, reading and phonological awareness skill on the subject-level coefficients of the frequency and AoA effects on lexical decision and pronunciation latencies.

	AoA effect on decisions latencies					Frequency effect on decisions latencies				
Effect	Estimate	SE	t	p		Estimate	SE	t	p	
Intercept	0.006	0.002	3.06	0.002	**	-0.029	0.003	-11.131	< 0.001	***
Age	-0.008	0.002	-3.361	0.001	***	-0.003	0.003	-1.052	0.293	
Age'	0.010	0.004	2.602	0.010	**	0.010	0.005	2.154	0.032	*
Reading skill	0.001	0.002	0.771	0.441		0.002	0.002	0.803	0.422	
Reading skill'	-0.001	0.002	-0.757	0.450		0.001	0.002	0.328	0.743	
Phonological awareness	0.002	0.001	1.325	0.186		0.001	0.002	0.404	0.687	
Phonological awareness skill'	-0.003	0.002	-1.596	0.111		0.002	0.002	0.992	0.322	
	F (6, 348 df) = 4.4, p < .001; Adjusted R^2 = .05					F (6, 348 df) = 7.0, p < .001; Adjusted R^2 = .09				
	AoA effect on pronunciation latencies					Frequency effect on pronunciation latencies				
Effect	Estimate	SE	t	p		Estimate	SE	t	p	
Intercept	-0.001	0.002	-0.658	0.512		-0.010	0.003	-3.82	< 0.001	***
Age	-0.006	0.003	-1.835	0.068	.	-0.016	0.004	-3.702	< 0.001	***
Age'	0.008	0.006	1.314	0.191		0.028	0.008	3.628	< 0.001	***
Reading skill	0.001	0.002	0.547	0.585		0.021	0.003	7.194	< 0.001	***
Reading skill'	-0.001	0.002	-0.462	0.645		-0.012	0.002	-5.188	< 0.001	***
Phonological awareness	-0.004	0.001	-2.591	0.010	*	0.002	0.002	1.076	0.284	
Phonological awareness skill'	0.004	0.002	2.358	0.020	*	-0.001	0.002	-0.641	0.522	
	F (6, 173 df) = 4.8, p < .001; Adjusted R^2 = .11					F (6, 173 df) = 30.1, p < .001; Adjusted R^2 = .49				

*** if p <= .001; ** if p < .01

Table 5. Summary of linear models of effects of word attributes on by-items mean RTs, estimated separately for each task and age group.

Lexical decisions	Children Coefficients	SE	t	p		Young adults Coefficients	SE	t	p		Old adults Coefficients	SE	t	p	
Intercept	2.897	0.007	392.7	< .001	***	2.772	0.008	333.1	< .001	***	2.818	0.007	393.1	< .001	***
Voice	-0.008	0.004	-1.9	0.064	.	-0.005	0.005	-1.1	0.283		-0.006	0.004	-1.3	0.185	
Nasal	0.005	0.007	0.7	0.477		0.014	0.008	1.7	0.092	.	0.013	0.007	1.9	0.064	.
Fricative	0.000	0.007	0.1	0.957		0.014	0.008	1.7	0.084	.	-0.001	0.007	-0.1	0.900	
Liquid	0.003	0.007	0.4	0.664		0.013	0.008	1.6	0.118		-0.003	0.007	-0.5	0.620	
Bilabial	-0.003	0.007	-0.4	0.712		-0.013	0.008	-1.5	0.125		-0.022	0.007	-3.0	0.003	**
Labiodental	-0.005	0.010	-0.5	0.626		-0.025	0.011	-2.3	0.022	*	-0.015	0.009	-1.7	0.101	
Alveolar	-0.001	0.008	-0.1	0.933		-0.006	0.009	-0.7	0.498		-0.008	0.008	-1.0	0.311	
Palatal	-0.007	0.009	-0.8	0.454		-0.014	0.010	-1.4	0.154		-0.009	0.009	-1.1	0.280	
Velar	0.007	0.007	1.0	0.339		0.001	0.008	0.2	0.865		-0.011	0.007	-1.5	0.125	
Glottal	0.005	0.009	0.5	0.585		0.003	0.010	0.3	0.761		-0.009	0.009	-1.0	0.313	
Length (letters)	-0.006	0.002	-2.6	0.011	*	-0.011	0.002	-4.6	0.000	***	-0.011	0.002	-5.2	< .001	***
Orthographic neighbourhood size	-0.005	0.002	-2.5	0.014	*	-0.003	0.002	-1.2	0.251		-0.005	0.002	-2.4	0.017	*
BG-Mean	0.002	0.002	1.1	0.267		0.002	0.002	0.9	0.366		0.003	0.002	1.6	0.110	
Regularity	-0.003	0.003	-0.9	0.379		-0.002	0.003	-0.7	0.471		0.000	0.003	0.0	0.989	
Log10 (SUBTLEX-UK CD)	-0.026	0.002	-11.8	< .001	***	-0.026	0.002	-10.5	< .001	***	-0.021	0.002	-9.9	< .001	***
Rated Imageability	-0.012	0.002	-6.7	< .001	***	-0.009	0.002	-4.7	< .001	***	-0.009	0.002	-5.1	< .001	***
Rated Age-of-Acquisition	0.011	0.002	5.6	< .001	***	0.006	0.002	2.8	0.006	**	0.006	0.002	3.0	0.003	**
	F (17, 142 df) = 30.8, p < .001; Adjusted R^2 = .76					F (17, 142 df) = 18.4, p < .001; Adjusted R^2 = .65					F (17, 142 df) = 17.3, p < .001; Adjusted R^2 = .64				

Pronunciation	Children Coefficients	SE	t	p		Young adults Coefficients	SE	t	p		Old adults Coefficients	SE	t	p	
Intercept	2.862	0.011	257.6	< .001	***	2.748	0.008	343.1	< .001	***	2.781	0.007	390.4	< .001	***
Voice	0.008	0.007	1.2	0.224		0.004	0.005	0.8	0.417		0.004	0.004	0.9	0.367	
Nasal	-0.014	0.011	-1.3	0.201		-0.024	0.008	-3.0	0.003	**	-0.017	0.007	-2.5	0.014	*
Fricative	-0.038	0.011	-3.5	0.001	***	-0.040	0.008	-5.1	< .001	***	-0.043	0.007	-6.2	< .001	***
Liquid	-0.008	0.011	-0.7	0.480		-0.011	0.008	-1.4	0.163		-0.007	0.007	-1.1	0.295	
Bilabial	-0.009	0.011	-0.8	0.409		-0.007	0.008	-0.9	0.369		-0.002	0.007	-0.3	0.751	
Labiodental	-0.016	0.014	-1.1	0.267		-0.002	0.010	-0.2	0.825		0.006	0.009	0.7	0.511	
Alveolar	-0.027	0.012	-2.2	0.030	*	-0.018	0.009	-2.0	0.046	*	-0.011	0.008	-1.4	0.173	
Palatal	-0.005	0.013	-0.4	0.703		0.001	0.010	0.1	0.943		0.014	0.008	1.6	0.109	
Velar	-0.015	0.011	-1.3	0.181		-0.008	0.008	-1.1	0.279		-0.002	0.007	-0.3	0.790	
Glottal	-0.036	0.013	-2.7	0.008	**	-0.041	0.010	-4.3	< .001	***	-0.038	0.009	-4.4	< .001	***
Length (letters)	0.001	0.003	0.2	0.842		0.003	0.002	1.1	0.273		0.002	0.002	0.8	0.424	
Orthographic neighbourhood size	-0.006	0.003	-2.1	0.041	*	-0.002	0.002	-1.0	0.321		-0.004	0.002	-2.0	0.049	*
BG-Mean	0.004	0.003	1.4	0.160		0.003	0.002	1.6	0.115		0.004	0.002	2.5	0.012	*
Regularity	-0.008	0.004	-1.9	0.061	.	-0.007	0.003	-2.3	0.025	*	-0.005	0.003	-1.9	0.060	.
Log10 (SUBTLEX-UK CD)	-0.019	0.003	-5.6	< .001	***	-0.009	0.002	-3.8	< .001	***	-0.008	0.002	-3.7	< .001	***
Rated Imageability	-0.006	0.003	-2.2	0.026	*	-0.003	0.002	-1.5	0.129		-0.004	0.002	-2.4	0.020	*
Rated Age-of-Acquisition	0.006	0.003	2.1	0.037	*	0.002	0.002	0.7	0.468		0.001	0.002	0.3	0.794	
	F (17, 142 df) = 12.7, p < .001; Adjusted R^2 = .56					F (17, 142 df) = 11.3, p < .001; Adjusted R^2 = .52					F (17, 142 df) = 14.1, p < .001; Adjusted R^2 = .58				

*** if p < .001; ** if p < .01; * if p <.05; ~ if p < .1; Log10 (SUBTLEX-UK CD) = Log10 (SUBTLEX-UK Context Distinctiveness)

Table 6. Summary of the lexical decision and pronunciation data models of log10(RT)

	Lexical Decision					Pronunciation				
	Estimate	SE	t	p		Estimate	SE	t	p	
Intercept	2.700	0.014	191.46	< 0.001	***	2.727	0.013	211.02	< 0.001	***
Trial order	< 0.001	< 0.001	3.97	< 0.001	***	< 0.001	< 0.001	7.02	< 0.001	***
Voice	-0.007	0.003	-2.14	0.034	*	0.008	0.004	2.24	0.026	*
Nasal	0.014	0.006	2.45	0.016	*	-0.021	0.006	-3.81	< 0.001	***
Fricative	0.003	0.006	0.58	0.563		-0.040	0.006	-6.95	< 0.001	***
Liquid	0.005	0.006	0.92	0.358		-0.007	0.006	-1.32	0.190	
Bilabial	-0.013	0.006	-2.32	0.022	*	-0.008	0.006	-1.39	0.166	
Labiodental	-0.013	0.007	-1.81	0.072	~	-0.007	0.007	-0.88	0.383	
Alveolar	-0.006	0.006	-0.95	0.342		-0.020	0.006	-3.11	0.002	**
Palatal	-0.011	0.007	-1.66	0.099	~	0.001	0.007	0.17	0.863	
Velar	-0.001	0.006	-0.27	0.790		-0.010	0.006	-1.80	0.073	~
Glottal	-0.001	0.007	-0.18	0.860		-0.042	0.007	-6.03	< 0.001	***
Age	-0.139	0.014	-9.98	< 0.001	***	-0.039	0.020	-1.99	0.049	*
Age'	0.239	0.023	10.27	< 0.001	***	0.097	0.035	2.77	0.006	**
Reading skill	-0.041	0.010	-4.24	< 0.001	***	-0.061	0.013	-4.86	< 0.001	***
Reading skill'	0.021	0.009	2.44	0.015	*	0.023	0.010	2.27	0.025	*
Phonological awareness	-0.008	0.007	-1.16	0.248		-0.008	0.007	-1.06	0.293	
Phonological awareness'	0.009	0.009	0.92	0.360		< 0.001	0.010	-0.04	0.969	
Length (letters)	-0.013	0.002	-5.48	< 0.001	***	0.006	0.002	2.85	0.005	**
Orthographic neighbourhood size	-0.002	0.002	-0.78	0.438		-0.001	0.002	-0.42	0.675	
BG-Mean	0.001	0.002	0.39	0.700		< 0.001	0.002	-0.22	0.830	
Regularity	-0.002	0.003	-0.71	0.477		-0.002	0.003	-0.81	0.417	
Log10 (SUBTLEX-UK CD)	-0.028	0.003	-10.44	< 0.001	***	-0.011	0.002	-4.41	< 0.001	***
Rated Imageability	-0.007	0.002	-3.63	< 0.001	***	-0.004	0.002	-2.24	0.025	*
Rated Age-of-Acquisition	0.007	0.002	3.13	0.002	**	-0.002	0.002	-1.14	0.254	
Age x Length	< 0.001	0.002	0.08	0.938		0.009	0.002	3.54	< 0.001	***
Age' x Length	< 0.001	0.003	-0.03	0.977		-0.013	0.004	-3.09	0.002	**
Age x Orth N-size	0.003	0.002	1.65	0.099	~	< 0.001	0.002	0.08	0.936	
Age' x Orth N-size	-0.006	0.003	-1.94	0.052	~	0.001	0.004	0.15	0.879	
Age x BG-Mean	-0.001	0.001	-1.02	0.306		-0.002	0.002	-0.90	0.366	
Age' x BG-Mean	0.002	0.002	0.98	0.328		0.004	0.003	1.09	0.275	
Age x Regularity	0.005	0.003	1.63	0.104		0.013	0.004	3.40	0.001	***
Age' x Regularity	-0.003	0.004	-0.61	0.540		-0.020	0.006	-3.06	0.002	**
Age x Log10 (SUBTLEX-UK CD)	< 0.001	0.002	0.05	0.964		-0.010	0.003	-3.15	0.002	**
Age' x Log10 (SUBTLEX-UK CD)	0.006	0.004	1.46	0.146		0.019	0.006	3.33	0.001	***
Age x Imageability	0.005	0.002	3.09	0.002	**	-0.005	0.002	-2.31	0.021	*
Age' x Imageability	-0.006	0.003	-2.46	0.014	*	0.008	0.004	2.15	0.032	*
Age x AoA	-0.007	0.002	-3.51	< 0.001	***	-0.008	0.003	-3.34	0.001	***
Age' x AoA	0.008	0.003	2.48	0.014	*	0.012	0.004	2.77	0.006	**
Reading skill x Length	-0.003	0.001	-1.92	0.054	~	-0.002	0.002	-1.35	0.177	
Reading skill' x Length	0.001	0.001	0.97	0.334		0.001	0.001	0.43	0.666	
Reading skill x Orth N-size	0.005	0.001	3.74	< 0.001	***	0.006	0.002	3.65	< 0.001	***
Reading skill' x Orth N-size	-0.004	0.001	-3.32	0.001	***	-0.003	0.001	-2.71	0.007	**
Reading skill x BG-Mean	-0.002	0.001	-1.55	0.122		-0.003	0.001	-2.50	0.012	*
Reading skill' x BG-Mean	0.001	0.001	1.55	0.122		0.003	0.001	2.81	0.005	**
Reading skill x Regularity	0.001	0.002	0.58	0.564		-0.006	0.002	-2.37	0.018	*
Reading skill' x Regularity	< 0.001	0.002	-0.27	0.791		0.004	0.002	1.97	0.049	*
Reading skill x Log10 (SUBTLEX-UK CD)	-0.001	0.002	-0.32	0.752		0.014	0.002	6.70	< 0.001	***
Reading skill' x Log10 (SUBTLEX-UK CD)	0.002	0.002	1.46	0.145		-0.008	0.002	-4.58	< 0.001	***
Reading skill x Imageability	-0.001	0.001	-0.92	0.359		0.006	0.001	4.15	< 0.001	***
Reading skill' x Imageability	0.001	0.001	0.73	0.464		-0.003	0.001	-2.66	0.008	**
Reading skill x AoA	< 0.001	0.001	0.21	0.834		< 0.001	0.002	0.27	0.786	
Reading skill' x AoA	< 0.001	0.001	-0.36	0.720		< 0.001	0.001	-0.02	0.983	
Phonological awareness x Length	-0.003	0.001	-3.32	0.001	***	-0.001	0.001	-0.57	0.570	
Phonological awareness' x Length	0.004	0.001	2.95	0.003	**	-0.001	0.001	-0.90	0.371	
Phonological awareness x Orth N-size	-0.002	0.001	-2.17	0.030	*	-0.002	0.001	-2.19	0.028	*
Phonological awareness' x Orth N-size	0.003	0.001	2.09	0.037	*	0.001	0.001	0.40	0.689	
Phonological awareness x BG-Mean	0.001	0.001	1.01	0.314		0.001	0.001	1.34	0.180	
Phonological awareness' x BG-Mean	-0.001	0.001	-0.84	0.403		0.000	0.001	-0.38	0.707	
Phonological awareness x Regularity	-0.002	0.001	-1.54	0.125		0.001	0.001	0.83	0.409	
Phonological awareness' x Regularity	0.001	0.002	0.67	0.504		0.001	0.002	0.41	0.684	
Phonological awareness x Log10 (SUBTLEX-UK CD)	0.001	0.001	0.78	0.437		0.001	0.001	0.60	0.547	
Phonological awareness' x Log10 (SUBTLEX-UK CD)	0.001	0.002	0.72	0.470		0.001	0.002	0.35	0.728	
Phonological awareness x Imageability	0.002	0.001	2.48	0.013	*	0.000	0.001	0.03	0.976	
Phonological awareness' x Imageability	-0.001	0.001	-1.12	0.264		0.001	0.001	0.79	0.428	
Phonological awareness x AoA	0.001	0.001	1.28	0.199		-0.002	0.001	-1.64	0.103	
Phonological awareness' x AoA	-0.002	0.001	-1.59	0.113		0.002	0.001	1.85	0.066	~

*** if p < .001; ** if p < .01; * if p < .01; ~ if p < .1

Table 7. Summary of the cross-task (lexical decision and pronunciation data) model

Cross-task data: decision and pronunciation latencies	Estimate	SE	t	p	
Intercept	2.702	0.012	216.83	< 0.001	***
Trial order	< 0.001	< 0.001	6.64	< 0.001	***
Voice	-0.002	0.003	-0.61	0.544	
Nasal	-0.001	0.005	-0.29	0.772	
Fricative	-0.012	0.005	-2.64	0.009	**
Liquid	0.001	0.005	0.16	0.872	
Bilabial	-0.011	0.005	-2.22	0.028	*
Labiodental	-0.011	0.006	-1.76	0.081	~
Alveolar	-0.011	0.005	-2.14	0.034	*
Palatal	-0.007	0.006	-1.19	0.235	
Velar	-0.004	0.005	-0.86	0.390	
Glottal	-0.015	0.006	-2.64	0.009	**
Task	0.023	0.020	1.10	0.270	
Age	-0.146	0.013	-11.16	< 0.001	***
Age'	0.275	0.024	11.48	< 0.001	***
Reading skill	-0.039	0.008	-4.60	< 0.001	***
Reading skill'	0.021	0.008	2.70	0.007	**
Phonological awareness	-0.008	0.006	-1.29	0.197	
Phonological awareness'	0.009	0.009	1.07	0.287	
Length (letters)	-0.011	0.002	-5.25	< 0.001	***
Orthographic neighbourhood size	-0.001	0.002	-0.38	0.706	
BG-Mean	0.001	0.002	0.61	0.543	
Regularity	-0.002	0.003	-0.63	0.531	
Log10 (SUBTLEX-UK CD)	-0.028	0.002	-11.87	< 0.001	***
Rated Imageability	-0.006	0.002	-3.75	< 0.001	***
Rated Age-of-Acquisition	0.006	0.002	2.94	0.003	**
Task x Age	0.108	0.028	3.91	< 0.001	***
Task x Age'	-0.177	0.051	-3.47	0.001	***
Task x Reading skill	-0.028	0.019	-1.49	0.137	
Task x Reading skill'	0.004	0.016	0.26	0.797	
Task x Task x Phonological awareness	0.000	0.012	0.01	0.991	
Task x Phonological awareness'	-0.009	0.015	-0.64	0.520	
Task x Length (letters)	0.015	0.003	5.79	< 0.001	***
Task x Orthographic neighbourhood size	-0.002	0.003	-0.65	0.515	
Task x BG-Mean	-0.002	0.002	-0.76	0.448	
Task x Regularity	< 0.001	0.004	-0.11	0.914	
Task x Log10 (SUBTLEX-UK CD)	0.017	0.003	4.85	< 0.001	***
Task x Imageability	0.001	0.002	0.58	0.563	
Task x AoA	-0.007	0.003	-2.48	0.013	*
Age x Length	-0.001	0.002	-0.35	0.729	
Age' x Length	0.001	0.003	0.26	0.799	
Age x Orth N-size	0.003	0.002	1.50	0.135	
Age' x Orth N-size	-0.006	0.003	-1.88	0.060	~
Age x BG-Mean	-0.002	0.001	-1.53	0.126	
Age' x BG-Mean	0.004	0.002	1.47	0.142	
Age x Regularity	0.004	0.003	1.57	0.117	
Age' x Regularity	-0.003	0.005	-0.57	0.569	
Age x Log10 (SUBTLEX-UK CD)	< 0.001	0.002	-0.01	0.990	
Age' x Log10 (SUBTLEX-UK CD)	0.006	0.004	1.52	0.130	
Age x Imageability	0.005	0.002	3.60	< 0.001	***
Age' x Imageability	-0.008	0.003	-3.04	0.002	**
Age x AoA	-0.008	0.002	-4.11	< 0.001	***
Age' x AoA	0.010	0.003	3.03	0.003	**
Reading skill x Length	-0.003	0.001	-2.44	0.015	*
Reading skill' x Length	0.002	0.001	1.42	0.156	
Reading skill x Orth N-size	0.005	0.001	3.97	< 0.001	***
Reading skill' x Orth N-size	-0.004	0.001	-3.60	< 0.001	***
Reading skill x BG-Mean	-0.001	0.001	-1.25	0.210	
Reading skill' x BG-Mean	0.001	0.001	1.26	0.209	
Reading skill x Regularity	0.001	0.002	0.45	0.651	
Reading skill' x Regularity	< 0.001	0.002	-0.14	0.889	
Reading skill x Log10 (SUBTLEX-UK CD)	-0.001	0.002	-0.57	0.566	
Reading skill' x Log10 (SUBTLEX-UK CD)	0.003	0.001	1.90	0.057	~
Reading skill x Imageability	-0.001	0.001	-1.07	0.286	
Reading skill' x Imageability	0.001	0.001	0.78	0.433	
Reading skill x AoA	< 0.001	0.001	0.18	0.858	
Reading skill' x AoA	< 0.001	0.001	-0.38	0.708	

*** if p < .001; ** if p < .01; * if p < .01; ~ if p < .1

Table 7. Summary of the cross-task (lexical decision and pronunciation data) model (continued)

Cross-task data: decision and pronunciation latencies	Estimate	SE	t	p	
Phonological awareness x Length	-0.003	0.001	-3.64	< 0.001	***
Phonological awareness' x Length	0.004	0.001	3.21	0.001	**
Phonological awareness x Orth N-size	-0.002	0.001	-2.57	0.010	*
Phonological awareness' x Orth N-size	0.003	0.001	2.52	0.012	*
Phonological awareness x BG-Mean	0.001	0.001	1.41	0.157	
Phonological awareness' x BG-Mean	-0.001	0.001	-1.18	0.238	
Phonological awareness x Regularity	-0.002	0.001	-1.68	0.093	~
Phonological awareness' x Regularity	0.001	0.002	0.63	0.531	
Phonological awareness x Log10 (SUBTLEX-UK CD)	0.002	0.001	1.47	0.141	
Phonological awareness' x Log10 (SUBTLEX-UK CD)	< 0.001	0.001	0.34	0.738	
Phonological awareness x Imageability	0.002	0.001	2.72	0.007	**
Phonological awareness' x Imageability	-0.001	0.001	-1.34	0.180	
Phonological awareness x AoA	0.001	0.001	1.58	0.114	
Phonological awareness' x AoA	-0.002	0.001	-1.84	0.066	~
Task x Age x Length	0.012	0.004	3.30	0.001	***
Task x Age' x Length	-0.018	0.006	-2.86	0.004	**
Task x Age x Orth N-size	0.000	0.004	-0.12	0.906	
Task x Age' x Orth N-size	0.003	0.006	0.46	0.646	
Task x Age x BG-Mean	-0.001	0.003	-0.21	0.834	
Task x Age' x BG-Mean	0.002	0.005	0.39	0.698	
Task x Age x Regularity	0.010	0.005	1.96	0.050	~
Task x Age' x Regularity	-0.020	0.010	-2.08	0.037	*
Task x Age x Log10 (SUBTLEX-UK CD)	-0.011	0.005	-2.32	0.020	*
Task x Age' x Log10 (SUBTLEX-UK CD)	0.015	0.009	1.68	0.094	~
Task x Age x Imageability	-0.009	0.003	-3.01	0.003	**
Task x Age' x Imageability	0.014	0.006	2.43	0.015	*
Task x Age x AoA	-0.002	0.004	-0.52	0.604	
Task x Age' x AoA	0.004	0.007	0.55	0.584	
Task x Reading skill x Length	-0.002	0.002	-0.68	0.494	
Task x Reading skill' x Length	0.001	0.002	0.44	0.663	
Task x Reading skill x Orth N-size	< 0.001	0.002	0.04	0.966	
Task x Reading skill' x Orth N-size	0.001	0.002	0.65	0.516	
Task x Reading skill x BG-Mean	-0.001	0.002	-0.67	0.506	
Task x Reading skill' x BG-Mean	0.001	0.002	0.75	0.456	
Task x Reading skill x Regularity	-0.010	0.004	-2.63	0.009	**
Task x Reading skill' x Regularity	0.006	0.003	2.08	0.037	*
Task x Reading skill x Log10 (SUBTLEX-UK CD)	0.019	0.003	5.55	< 0.001	***
Task x Reading skill' x Log10 (SUBTLEX-UK CD)	-0.013	0.003	-4.62	< 0.001	***
Task x Reading skill x Imageability	0.007	0.002	3.00	0.003	**
Task x Reading skill' x Imageability	-0.003	0.002	-1.79	0.074	~
Task x Reading skill x AoA	0.002	0.003	0.75	0.453	
Task x Reading skill' x AoA	-0.001	0.002	-0.37	0.710	
Task x Phonological awareness x Length	0.003	0.002	1.94	0.052	~
Task x Phonological awareness' x Length	-0.005	0.002	-2.42	0.016	*
Task x Phonological awareness x Orth N-size	0.000	0.002	0.29	0.772	
Task x Phonological awareness' x Orth N-size	-0.002	0.002	-1.31	0.190	
Task x Phonological awareness x BG-Mean	< 0.001	0.001	-0.20	0.840	
Task x Phonological awareness' x BG-Mean	0.001	0.001	0.49	0.623	
Task x Phonological awareness x Regularity	0.005	0.002	1.94	0.052	~
Task x Phonological awareness' x Regularity	-0.001	0.003	-0.31	0.758	
Task x Phonological awareness x Log10 (SUBTLEX-UK CD)	-0.001	0.002	-0.49	0.628	
Task x Phonological awareness' x Log10 (SUBTLEX-UK CD)	< 0.001	0.003	0.03	0.977	
Task x Phonological awareness x Imageability	-0.002	0.001	-1.62	0.106	
Task x Phonological awareness' x Imageability	0.003	0.002	1.66	0.097	~
Task x Phonological awareness x AoA	-0.003	0.002	-1.92	0.055	~
Task x Phonological awareness' x AoA	0.005	0.002	2.25	0.025	*

*** if p < .001; ** if p < .01; * if p < .01; ~ if p < .1

Table 8. Summary of models of relationships between diffusion model parameter values and individual differences.

Word types: early vs. late acquired	Drift					Separation					Non-decision time				
	Estimate	SE	t	p		Estimate	SE	t	p		Estimate	SE	t	p	
Intercept	0.353	0.015	23.74	< 0.001	***	0.132	0.005	24.81	< 0.001	***	0.306	0.017	18.49	< 0.001	***
Word type	-0.058	0.012	-4.80	< 0.001	***	-0.007	0.005	-1.50	0.135		0.025	0.012	2.07	0.040	*
Age	0.102	0.016	6.27	< 0.001	***	-0.005	0.006	-0.93	0.351		-0.134	0.018	-7.35	< 0.001	***
Age'	-0.142	0.028	-5.14	< 0.001	***	0.003	0.010	0.26	0.796		0.246	0.031	7.96	< 0.001	***
Reading skill	0.025	0.011	2.33	0.020	*	0.002	0.004	0.41	0.683		-0.042	0.012	-3.43	0.001	***
Reading skill'	-0.010	0.010	-0.97	0.333		-0.009	0.004	-2.40	0.017	*	0.029	0.012	2.46	0.014	*
Phonological awareness	0.013	0.008	1.58	0.115		-0.002	0.003	-0.58	0.560		-0.006	0.009	-0.67	0.506	
Phonological awareness'	-0.006	0.011	-0.54	0.587		0.002	0.004	0.57	0.572		0.015	0.013	1.23	0.218	
Type x Age	0.022	0.013	1.66	0.099	~	0.004	0.005	0.82	0.414		-0.018	0.013	-1.36	0.174	
Type x Age'	-0.009	0.023	-0.42	0.675		0.002	0.009	0.20	0.839		0.018	0.022	0.82	0.413	
Type x Reading skill	0.005	0.008	0.60	0.552		0.000	0.004	0.02	0.984		-0.010	0.009	-1.12	0.265	
Type x Reading skill'	-0.003	0.008	-0.39	0.694		0.003	0.003	0.80	0.427		0.000	0.008	0.04	0.972	
Type x Phonological awareness	0.002	0.006	0.31	0.754		0.000	0.003	-0.18	0.858		-0.002	0.006	-0.25	0.805	
Type x Phonological awareness'	0.007	0.009	0.82	0.415		-0.002	0.004	-0.53	0.598		0.002	0.009	0.22	0.827	
Word types: high vs. low frequency	**Drift**					**Separation**					**Non-decision time**				
	Estimate	SE	t	p		Estimate	SE	t	p		Estimate	SE	t	p	
Intercept	0.297	0.014	21.25	< 0.001	***	0.125	0.005	27.53	<2e-16	***	0.334	0.015	21.75	< 0.001	***
Word type	0.072	0.013	5.38	< 0.001	***	0.001	0.005	0.13	0.898		-0.015	0.013	-1.17	0.245	
Age	0.125	0.016	8.03	< 0.001	***	-0.002	0.005	-0.42	0.676		-0.146	0.017	-8.57	< 0.001	***
Age'	-0.153	0.026	-5.85	< 0.001	***	0.005	0.009	0.59	0.556		0.258	0.029	8.97	< 0.001	***
Reading skill	0.031	0.011	2.86	0.004	**	0.003	0.004	0.87	0.385		-0.056	0.012	-4.63	< 0.001	***
Reading skill'	-0.014	0.010	-1.32	0.188		-0.007	0.003	-2.05	0.041	*	0.031	0.011	2.73	0.007	**
Phonological awareness	0.014	0.008	1.76	0.080	~	-0.003	0.003	-1.09	0.277		-0.007	0.009	-0.75	0.453	
Phonological awareness'	-0.002	0.011	-0.16	0.876		0.001	0.004	0.30	0.765		0.017	0.012	1.45	0.149	
Type x Age	-0.009	0.015	-0.59	0.556		-0.009	0.005	-1.74	0.084	~	0.021	0.014	1.50	0.134	
Type x Age'	-0.019	0.025	-0.75	0.452		0.007	0.009	0.72	0.475		-0.025	0.024	-1.02	0.307	
Type x Reading skill	-0.007	0.009	-0.82	0.415		-0.003	0.003	-1.00	0.319		0.018	0.009	2.11	0.036	*
Type x Reading skill'	0.008	0.009	0.91	0.364		0.001	0.003	0.43	0.667		-0.008	0.008	-1.00	0.317	
Type x Phonological awareness	0.001	0.006	0.15	0.883		0.001	0.002	0.54	0.593		0.001	0.006	0.11	0.916	
Type x Phonological awareness'	-0.008	0.010	-0.86	0.392		-0.001	0.004	-0.31	0.760		-0.001	0.009	-0.16	0.876	

*** if p < .001; ** if p < .01; * if p < .01; ~ if p < .1

Table 9. Summary of linear mixed-effects model of pronunciation durations

	Estimate	SE	t	p	
Intercept	2.652	0.022	121.29	< 0.001	***
Trial order	< 0.001	< 0.001	15.57	< 0.001	***
Voice	-0.003	0.010	-0.32	0.750	
Nasal	0.022	0.016	1.40	0.165	
Fricative	0.047	0.016	2.95	0.004	**
Liquid	0.031	0.016	1.96	0.052	~
Bilabial	-0.011	0.016	-0.71	0.480	
Labiodental	-0.011	0.021	-0.53	0.595	
Alveolar	-0.009	0.018	-0.50	0.619	
Palatal	-0.040	0.019	-2.07	0.040	*
Velar	-0.020	0.016	-1.25	0.214	
Glottal	0.002	0.020	0.09	0.929	
Age	-0.136	0.026	-5.31	< 0.001	***
Age'	0.248	0.046	5.41	< 0.001	***
Reading skill	0.030	0.016	1.83	0.069	~
Reading skill'	-0.037	0.013	-2.84	0.005	**
Phonological awareness	-0.016	0.010	-1.68	0.095	~
Phonological awareness'	0.007	0.013	0.49	0.624	
Length (letters)	0.020	0.005	4.05	< 0.001	***
Orthographic neighbourhood size	-0.010	0.005	-2.03	0.043	*
BG-Mean	0.004	0.004	1.05	0.294	
Regularity	0.001	0.007	0.09	0.928	
Log10 (SUBTLEX-UK CD)	0.005	0.005	0.98	0.328	
Rated Imageability	0.007	0.004	1.72	0.088	~
Rated Age-of-Acquisition	0.006	0.005	1.28	0.201	
Age x Length	-0.001	0.003	-0.53	0.597	
Age' x Length	0.004	0.005	0.89	0.376	
Age x Orth N-size	-0.013	0.003	-4.82	< 0.001	***
Age' x Orth N-size	0.021	0.005	4.61	< 0.001	***
Age x BG-Mean	0.000	0.002	0.19	0.847	
Age' x BG-Mean	-0.003	0.004	-0.82	0.415	
Age x Regularity	0.004	0.004	0.90	0.371	
Age' x Regularity	-0.002	0.007	-0.34	0.737	
Age x Log10 (SUBTLEX-UK CD)	0.001	0.003	0.52	0.607	
Age' x Log10 (SUBTLEX-UK CD)	-0.005	0.005	-1.05	0.294	
Age x Imageability	0.001	0.002	0.24	0.811	
Age' x Imageability	-0.001	0.004	-0.32	0.752	
Age x AoA	< 0.001	0.003	-0.06	0.954	
Age' x AoA	0.001	0.005	0.18	0.856	
Reading skill x Length	-0.001	0.002	-0.58	0.563	
Reading skill' x Length	-0.001	0.001	-0.48	0.631	
Reading skill x Orth N-size	-0.001	0.002	-0.40	0.690	
Reading skill' x Orth N-size	< 0.001	0.001	-0.01	0.996	
Reading skill x BG-Mean	< 0.001	0.001	0.14	0.886	
Reading skill' x BG-Mean	-0.001	0.001	-0.83	0.408	
Reading skill x Regularity	0.004	0.003	1.38	0.167	
Reading skill' x Regularity	-0.001	0.002	-0.49	0.623	
Reading skill x Log10 (SUBTLEX-UK CD)	0.002	0.002	0.87	0.382	
Reading skill' x Log10 (SUBTLEX-UK CD)	-0.002	0.001	-1.07	0.285	
Reading skill x Imageability	0.002	0.001	1.17	0.241	
Reading skill' x Imageability	-0.002	0.001	-1.82	0.069	~
Reading skill x AoA	-0.001	0.002	-0.66	0.507	
Reading skill' x AoA	0.001	0.001	0.37	0.711	
Phonological awareness x Length	< 0.001	0.001	0.30	0.761	
Phonological awareness' x Length	0.001	0.001	0.55	0.580	
Phonological awareness x Orth N-size	0.001	0.001	0.81	0.416	
Phonological awareness' x Orth N-size	-0.002	0.001	-1.48	0.139	
Phonological awareness x BG-Mean	< 0.001	0.001	0.17	0.862	
Phonological awareness' x BG-Mean	0.001	0.001	0.58	0.559	
Phonological awareness x Regularity	-0.002	0.001	-1.16	0.247	
Phonological awareness' x Regularity	< 0.001	0.002	0.24	0.813	
Phonological awareness x Log10 (SUBTLEX-UK CD)	0.002	0.001	1.57	0.118	
Phonological awareness' x Log10 (SUBTLEX-UK CD)	-0.002	0.001	-1.17	0.242	
Phonological awareness x Imageability	0.001	0.001	0.78	0.437	
Phonological awareness' x Imageability	-0.001	0.001	-0.91	0.364	
Phonological awareness x AoA	0.002	0.001	1.88	0.060	~
Phonological awareness' x AoA	-0.003	0.001	-2.07	0.039	*

*** if p < .001; ** if p < .01; * if p < .01; ~ if p < .1

Figure 1. Histograms showing the distribution of participant ages, phonological awareness skill (Spoonerisms score /30), word pronunciation skill (TOWRE words accuracy /103 divided by time), nonword pronunciation skill (TOWRE nonwords accuracy /63 divided by time), and percentage correct as well as mean RT (average RT of correct responses to words) in lexical decision and pronunciation tasks.

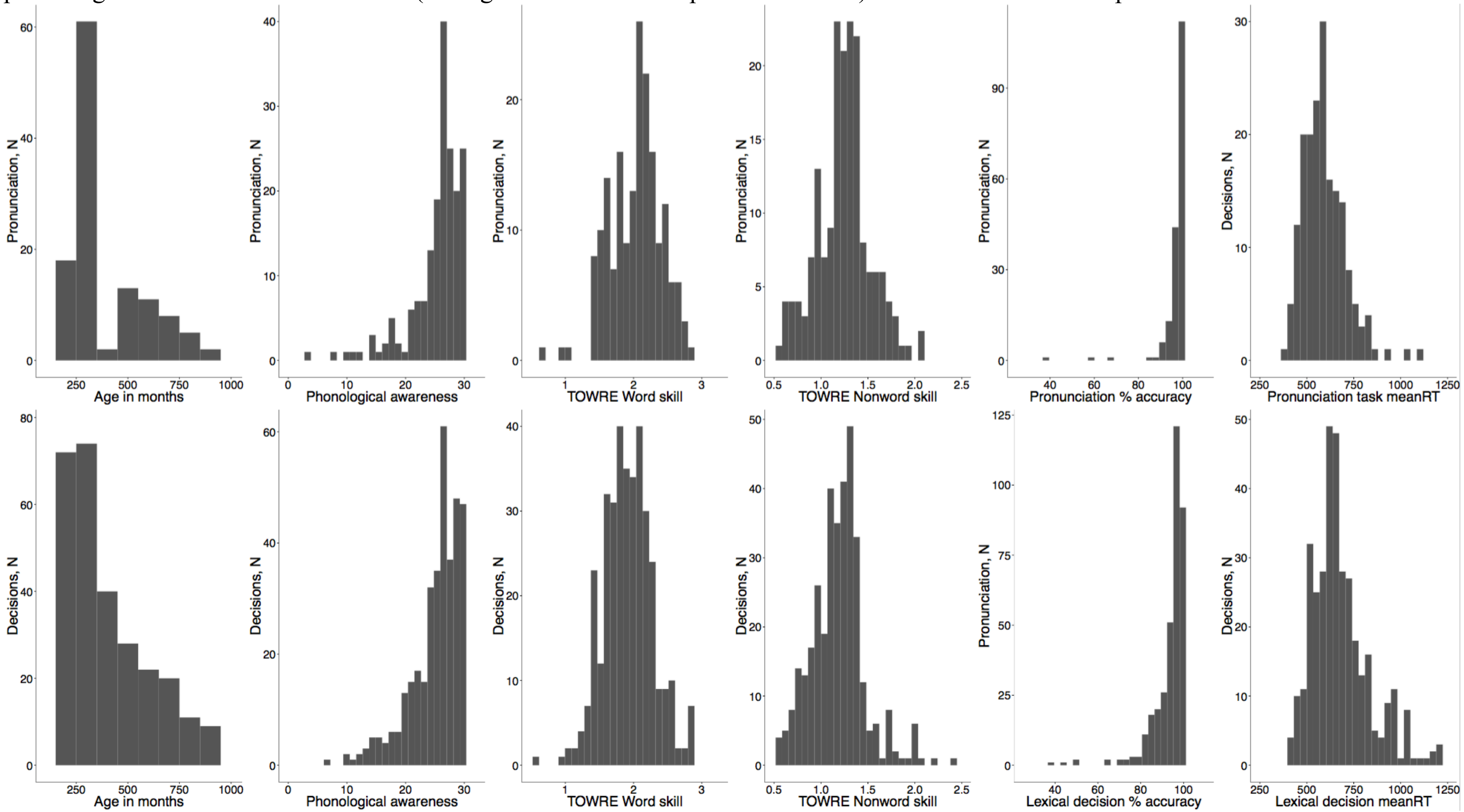


Figure 2. Scatterplots showing the relationship between lexical decision latencies and critical participant or word attributes. Each grey point represents the response made by a participant to a word. Each black line shows a LOESS smoother representing the locally weighted polynomial regression relating latencies to attribute values.

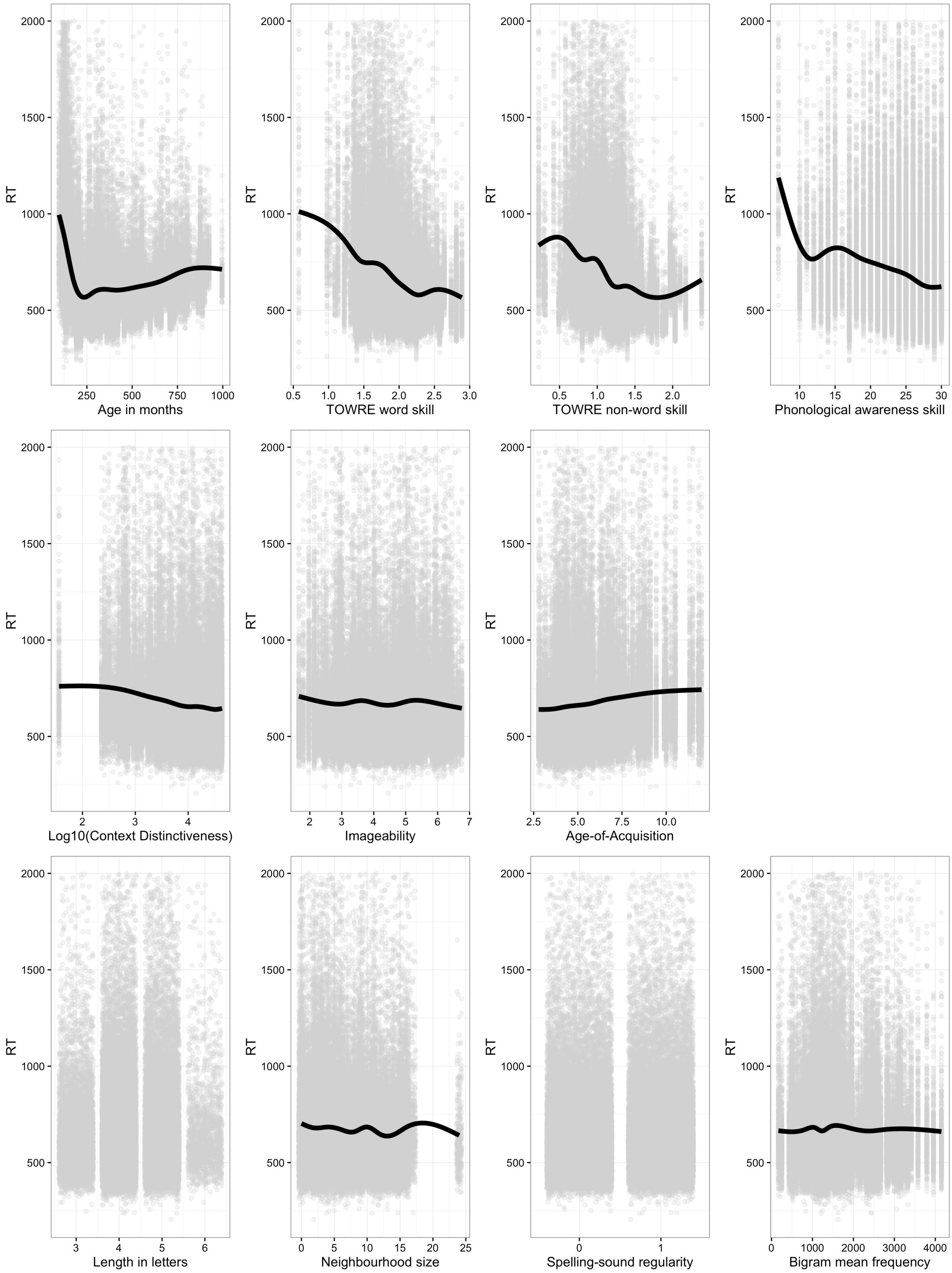


Figure 3. Scatterplots showing the relationship between pronunciation latencies and critical participant or word attributes. Each grey point represents the response made by a participant to a word. Each black line shows a LOESS smoother representing the locally weighted polynomial regression relating latencies to attribute values.

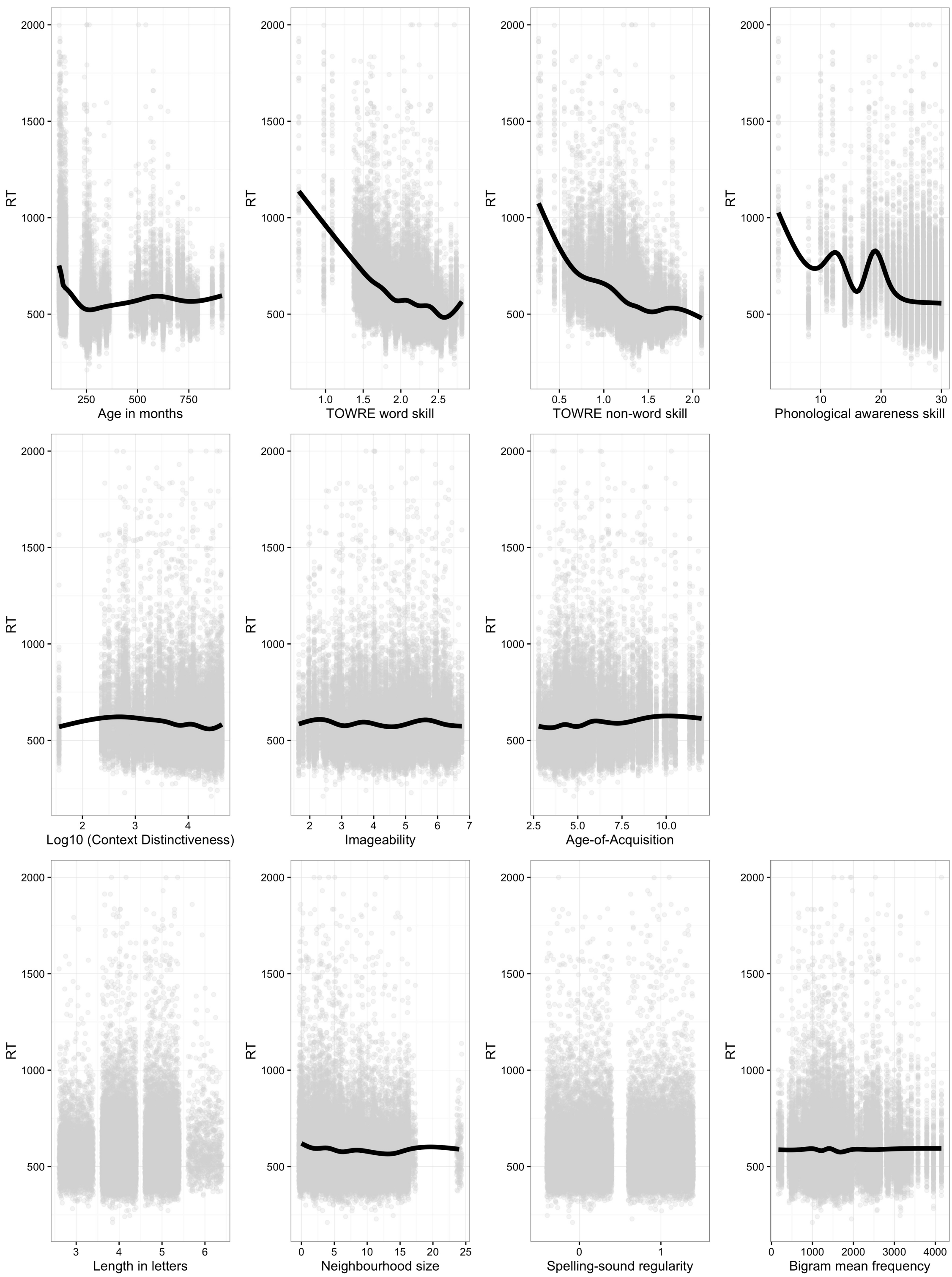


Figure 4. Scatterplots showing the subject-level estimates of frequency or AoA effects on lexical decision and pronunciation RTs. Points represent individual participant coefficient estimates. Line ranges represent associated standard errors. Smoothers represent the bivariate relationship (LOESS estimates) between the subject-level coefficient estimates and individual differences in age, TOWRE word or non-word reading skill or phonological awareness skill. Note “Reading skill” is the aggregate measure combining word and non-word ability scores.

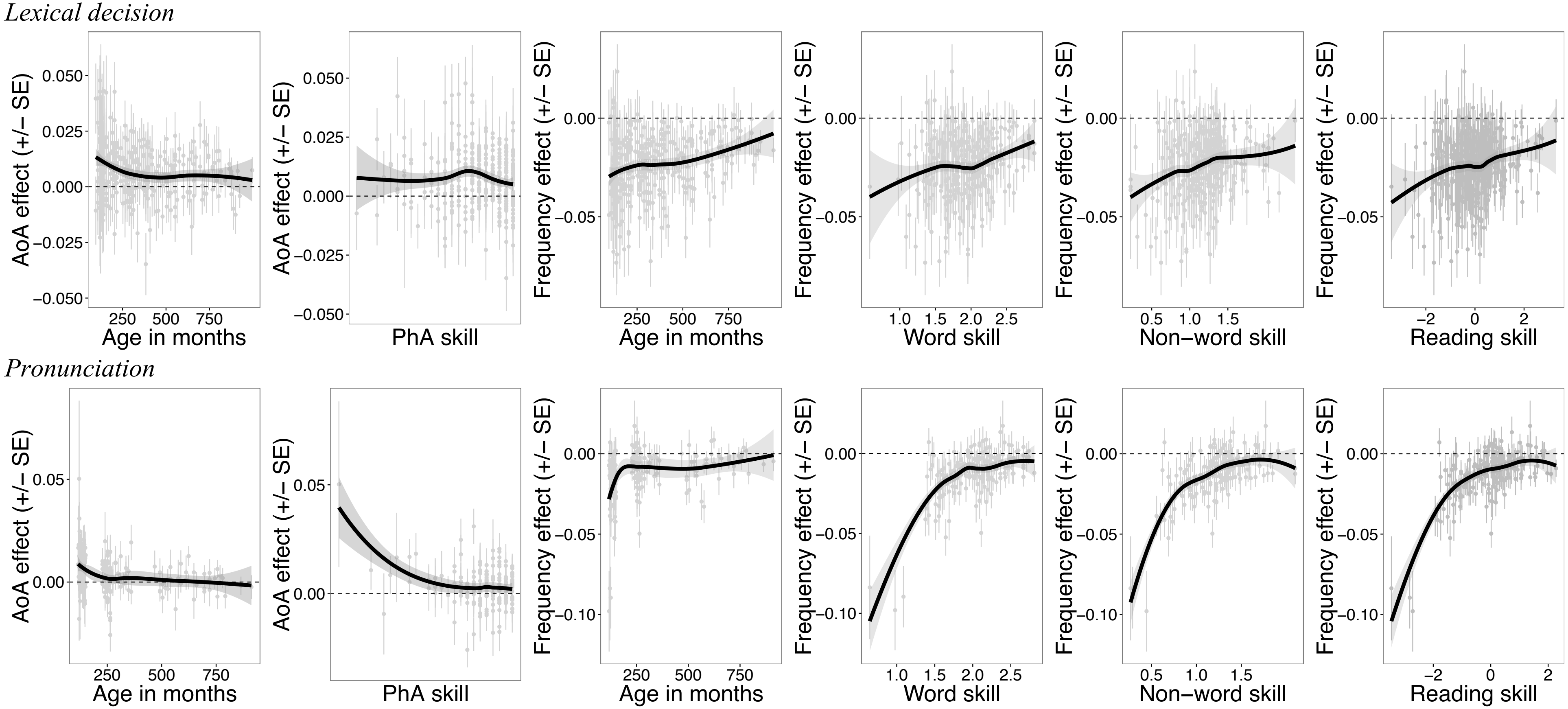


Figure 5. Scatterplot showing variation in diffusion model parameter values in relation to individual differences in participant age, reading skill, and phonological awareness skill. Diffusion model parameter values are calculated separately for each participant, on the basis of data about the accuracy and speed of their responses to words, here: early vs. late acquired words.

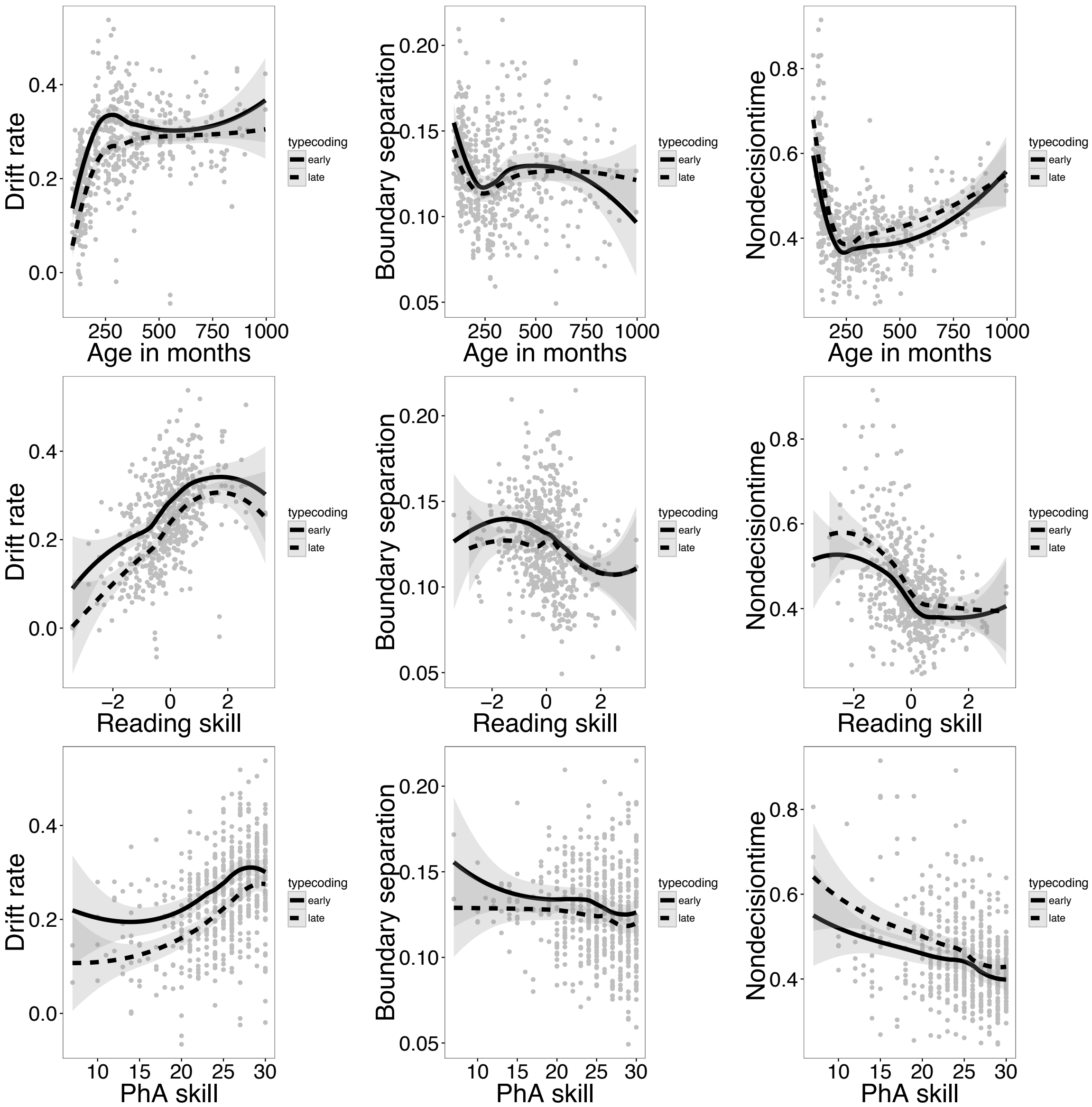


Figure 6. Scatterplot showing variation in diffusion model parameter values in relation to individual differences in participant age, reading skill, and phonological awareness skill. Diffusion model parameter values are calculated separately for each participant, on the basis of data about the accuracy and speed of their responses to low vs. high frequency words.

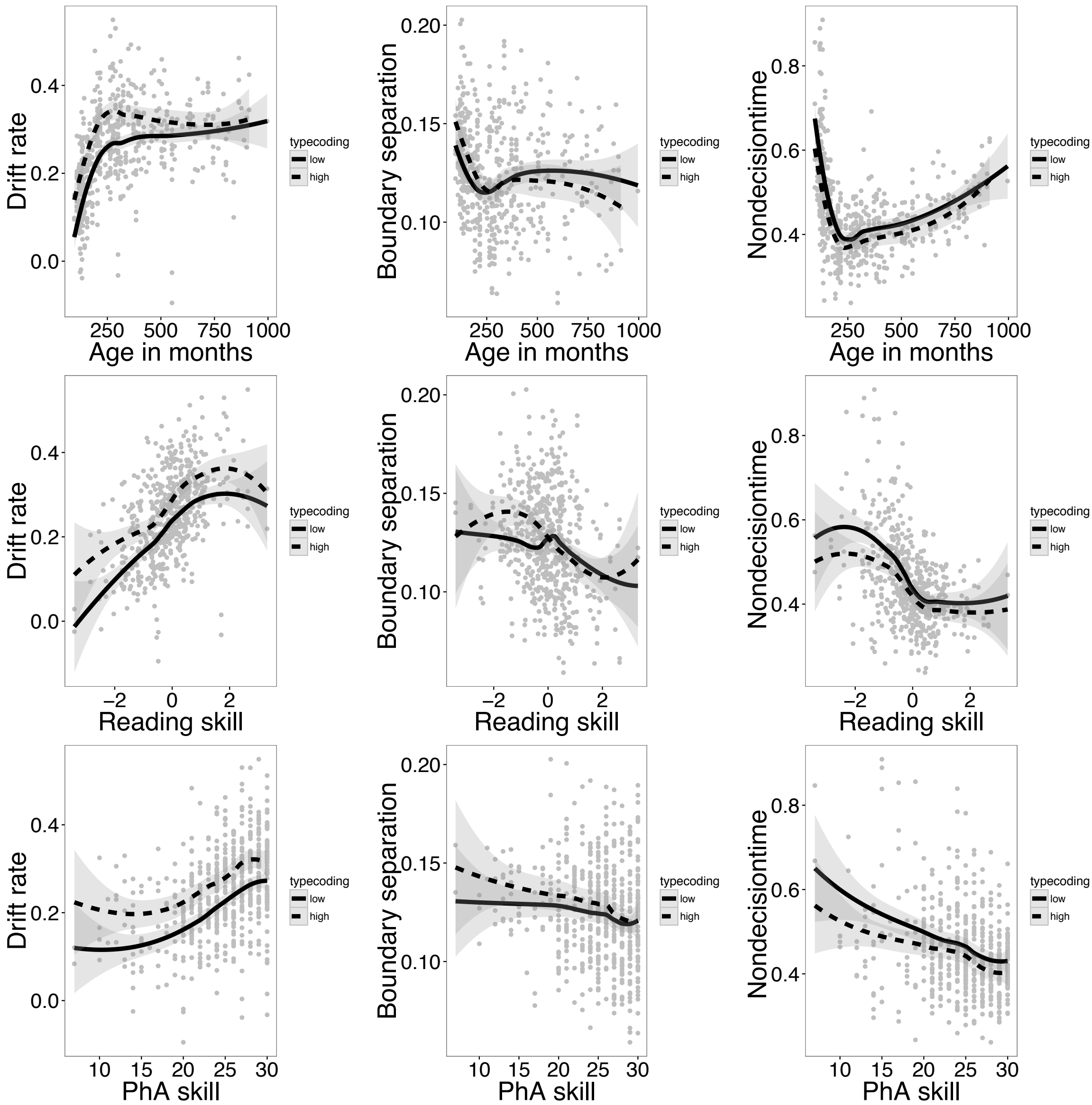


Figure 7. Scatterplot showing the bivariate relationship between pronunciation spoken response duration and critical variables, including measures of individual differences and of item attributes.

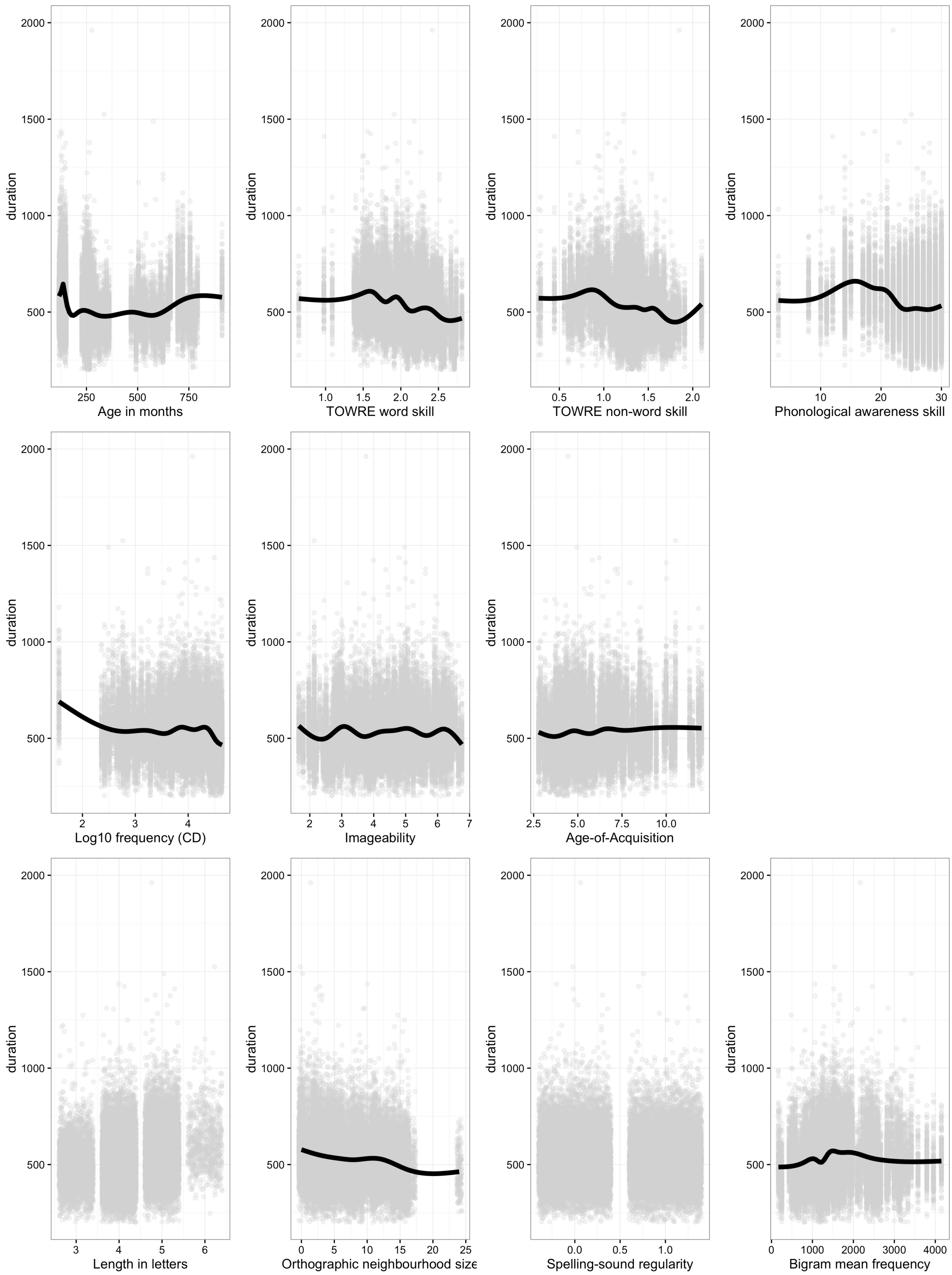
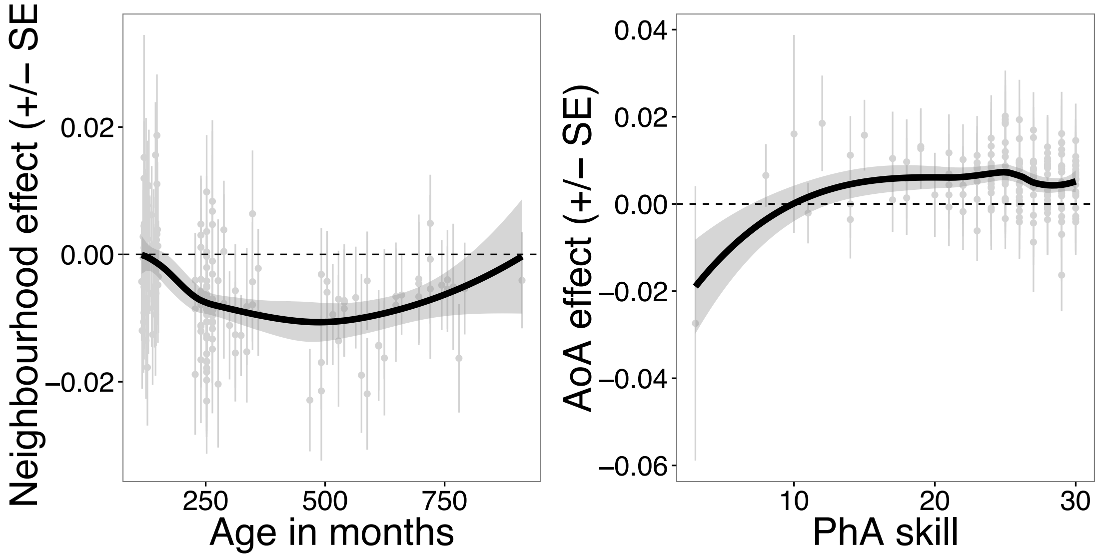


Figure 8. Scatterplots showing the subject-level estimates of AoA and neighbourhood effects on pronunciation durations. Points represent individual participant coefficient estimates. Line ranges represent associated standard errors. Smoothers represent the bivariate relationship (LOESS estimates) between the subject-level coefficient estimates and individual differences in age or phonological awareness skill.



Appendix. The words presented in the lexical decision and pronunciation tasks, together with by-items mean RTs (averaged over correct response participant RTs) for each task, and critical psycholinguistic attributes

word	decisions mean RT (ms)	pronunciation mean RT (ms)	OLD	PLD	Regularity	length (letters)	Orthographic N-size	BG_Sum	BG_Mean	BG Freq_by_Pos	Imageability	AoA	SUBTLEX-UK Log10(CD)
act	681	595	1.5	1.0	1	3	5	2603	1301.5	258	4.0	6.4	4.0
ask	648	585	1.5	1.6	0	3	4	1678	839	207	3.3	2.9	4.4
both	681	624	1.7	1.3	0	4	4	2992	997.333	860	2.4	4.8	4.4
box	636	596	1.4	1.1	1	3	10	860	430	380	6.0	4.3	4.1
broad	708	653	1.7	1.4	0	5	2	4189	1047.25	1348	2.9	10.5	3.4
bronze	705	644	2.1	1.9	1	6	0	7848	1569.6	1397	5.3	10.0	3.2
calf	749	629	1.9	1.1	0	4	3	4968	1656	1239	5.7	6.6	2.8
can	651	576	1.0	1.0	1	3	15	5286	2643	1371	5.2	4.3	4.6
care	638	588	1.0	1.0	0	4	24	9310	3103.333	2195	4.2	5.7	4.2
carve	799	654	1.7	1.6	1	5	3	6886	1721.5	2072	4.0	6.4	2.8
chance	649	593	1.8	1.7	0	6	2	9175	1835	1687	2.8	7.8	4.4
cheat	675	601	1.7	1.5	1	5	4	9262	2315.5	1502	3.3	5.1	3.0
cheer	667	580	1.8	1.2	0	5	3	10298	2574.5	2421	4.9	5.0	3.5
chew	719	590	1.8	1.5	1	4	4	3575	1191.667	850	5.1	4.1	3.0
child	632	567	1.9	1.9	0	5	3	4364	1091	1170	6.1	5.1	4.1
choice	663	586	2.0	2.0	1	6	0	6985	1397	1328	2.8	5.2	4.1
chop	733	624	1.5	1.4	1	4	7	3462	1154	1000	5.2	6.7	3.3
chose	733	703	1.6	1.8	0	5	6	5587	1396.75	1360	2.3	4.7	3.6
class	628	596	1.5	1.2	0	5	8	5621	1405.25	1293	5.2	5.0	4.0
clean	630	587	1.7	1.5	1	5	3	9293	2323.25	1302	5.4	3.9	4.0
clerk	736	688	1.9	1.8	0	5	1	10427	2606.75	1097	4.6	6.7	2.7
club	661	603	1.9	1.8	1	4	1	1537	512.333	571	5.5	5.9	4.0
coast	675	630	1.7	1.2	1	5	3	7910	1977.5	2601	5.9	6.4	3.8
cod	804	634	1.0	1.0	1	3	16	3008	1504	1887	5.6	11.5	2.8
cold	661	587	1.2	1.0	0	4	12	4100	1366.667	2295	5.0	4.0	4.1
comb	732	632	1.6	1.0	0	4	5	4120	1373.333	2479	6.2	5.5	2.7
cook	650	593	1.3	1.2	0	4	11	3628	1209.333	2268	6.1	4.2	3.7
cope	742	609	1.0	1.0	1	4	17	4887	1629	2233	2.9	9.4	3.7
cost	719	611	1.4	1.3	1	4	9	6883	2294.333	2731	3.6	5.7	4.1
cough	680	631	1.7	1.4	0	5	6	4940	1235	2485	5.9	4.3	2.9
could	646	636	1.9	1.4	0	5	2	5512	1378	2434	2.1	4.3	4.6
deal	637	633	1.1	1.0	1	4	16	7480	2493.333	1923	3.4	6.0	4.3
death	640	645	1.9	1.5	0	5	1	9920	2480	2308	5.4	5.4	4.1
down	653	606	1.6	1.3	1	4	4	1439	479.667	485	4.6	4.9	4.6
draw	674	607	1.6	1.7	0	4	5	3732	1244	1081	5.1	4.1	3.9
ease	741	652	1.4	1.0	0	4	6	5298	1766	777	2.0	9.1	3.5
eat	636	600	1.0	1.0	1	3	11	5996	2998	651	5.9	2.8	4.1
eye	633	604	1.5	1.0	0	3	8	418	209	70	6.8	3.8	4.2
face	617	539	1.2	1.0	1	4	11	3524	1174.667	824	6.2	3.8	4.4
fact	636	566	1.6	1.3	1	4	5	3031	1010.333	735	3.0	6.5	4.4
fade	741	541	1.4	1.0	1	4	11	4080	1360	803	3.8	7.0	3.0
faith	672	574	2.0	1.4	1	5	0	4847	1211.75	1118	3.2	7.6	3.7
fall	653	539	1.3	1.0	0	4	12	5109	1703	1365	4.7	4.7	4.1
false	694	570	2.0	1.8	0	5	0	5938	1484.5	1245	3.0	6.7	3.5
far	686	544	1.0	1.1	0	3	16	3393	1696.5	1261	3.4	4.9	4.5
farm	646	554	1.6	1.6	1	4	6	3976	1325.333	1478	5.9	3.9	3.7
fast	613	546	1.3	1.1	0	4	11	5445	1815	1437	4.9	3.7	4.1
fault	696	552	1.8	1.4	0	5	2	2302	575.5	628	3.0	6.9	3.8
fear	646	552	1.3	1.0	0	4	12	5351	1783.667	1368	4.3	4.8	3.9
feel	632	517	1.4	1.0	1	4	9	3175	1058.333	551	3.9	5.1	4.5
from	658	562	1.9	1.7	1	4	2	3738	1246	1177	2.1	4.4	4.6
front	665	576	1.8	1.8	0	5	2	10145	2536.25	1569	3.8	5.2	4.3
gain	723	650	1.4	1.0	1	4	9	8428	2809.333	939	2.8	7.1	3.5
get	636	585	1.2	1.1	0	3	14	2565	1282.5	520	2.4	3.2	4.6
give	631	608	1.5	1.7	0	4	7	3352	1117.333	586	3.8	4.3	4.6
glass	652	623	1.8	1.5	0	5	3	5478	1369.5	1112	6.2	4.5	3.9
gloom	716	635	1.8	1.5	1	5	2	4048	1012	799	4.0	9.0	2.9
good	617	601	1.4	1.5	0	4	8	1822	607.333	677	3.9	3.6	4.6
grace	692	607	1.4	1.0	1	5	8	6954	1738.5	1773	2.5	7.3	3.3
grant	713	610	1.7	1.6	0	5	4	10359	2589.75	2000	3.0	12.0	3.5
guard	687	649	2.0	1.3	1	5	0	4407	1101.75	801	5.6	6.3	3.5
guide	668	624	1.8	1.0	1	5	3	4260	1065	761	4.4	7.1	3.6
halt	773	592	1.4	1.3	0	4	9	4580	1526.667	1078	3.7	8.1	3.2
hang	703	544	1.4	1.0	1	4	10	8547	2849	1439	5.2	6.7	4.1
hate	655	541	1.0	1.0	1	4	16	9721	3240.333	1602	3.7	5.5	3.9
have	641	542	1.2	1.1	0	4	12	3429	1143	891	2.9	3.7	4.6
head	621	569	1.2	1.0	0	4	13	4615	1538.333	1307	5.8	3.4	4.4
heat	650	561	1.2	1.0	1	4	11	7745	2581.667	1416	5.0	5.3	3.8
help	630	526	1.6	1.5	1	4	7	3566	1188.667	816	4.4	3.7	4.5
her	633	575	1.3	1.3	0	3	6	7911	3955.5	985	4.1	5.1	4.5
his	670	568	1.4	1.0	0	3	6	3845	1922.5	976	4.0	3.7	4.6
hit	663	527	1.0	1.0	1	3	15	3263	1631.5	397	5.8	4.8	4.3
house	629	532	1.6	1.7	0	5	5	6368	1592	1449	6.6	3.2	4.4
hurt	663	556	1.7	1.2	1	4	5	2647	882.333	922	5.0	4.0	3.9
ice	659	585	1.5	1.0	1	3	4	3939	1969.5	73	6.0	3.9	3.8
job	641	585	1.2	1.5	0	3	14	503	251.5	220	4.7	5.4	4.4
join	649	588	1.8	1.6	1	4	3	7389	2463	642	4.0	5.8	4.2
keep	617	560	1.5	1.0	1	4	9	2442	814	395	2.7	4.4	4.5
land	655	575	1.2	1.0	1	4	11	7264	2421.333	1468	5.5	5.2	4.1
large	636	577	1.9	1.8	1	5	2	6501	1625.25	1665	4.4	5.7	4.1
lark	791	606	1.4	1.4	1	4	9	5254	1751.333	1406	3.6	11.8	2.7

word	decisions mean RT (ms)	pronunciation mean RT (ms)	OLD	PLD	regularity	length (letters)	Orthographic N-size	BG_Sum	BG_Mean	BG Freq_by_Pos	Imageability	AoA	SUBTLEX-UK Log10(CD)
last	648	566	1.2	1.0	0	4	14	7036	2345.333	1462	3.4	4.3	4.6
laugh	639	595	2.0	1.1	0	5	0	3279	819.75	709	6.0	3.8	3.9
learn	663	596	1.8	1.5	0	5	1	8669	2167.25	1512	3.8	4.4	4.1
less	658	567	1.6	1.2	1	4	8	9476	3158.667	1000	2.8	4.0	4.3
long	627	564	1.5	1.2	1	4	7	9994	3331.333	1705	4.5	4.2	4.5
look	613	568	1.4	1.1	0	4	11	2624	874.667	807	4.5	4.1	4.6
love	604	583	1.2	1.4	0	4	12	3911	1303.667	758	5.0	5.2	4.5
man	619	549	1.0	1.0	1	3	16	5163	2581.5	1372	6.3	3.1	4.5
match	644	573	1.6	1.3	1	5	8	7617	1904.25	1681	4.9	5.7	4.0
mend	767	589	1.3	1.2	1	4	11	7319	2439.667	1364	4.9	9.1	2.9
monk	700	625	1.7	1.4	0	4	5	5856	1952	1644	5.5	10.3	2.6
month	679	619	1.8	1.9	0	5	2	9911	2477.75	2268	3.3	5.8	4.1
mood	661	580	1.4	1.0	1	4	9	2425	808.333	971	3.7	6.6	3.7
mould	750	668	1.8	1.0	0	5	3	4003	1000.75	1141	5.1	11.3	3.1
move	619	553	1.4	1.4	0	4	9	3406	1135.333	926	4.7	4.6	4.4
noise	662	588	1.7	1.7	0	5	4	5939	1484.75	950	4.2	4.5	3.8
nor	907	627	1.6	1.0	0	3	5	3553	1776.5	971	2.0	8.6	3.6
now	639	575	1.1	1.2	1	3	15	1485	742.5	431	3.3	5.3	4.6
nurse	655	601	1.8	1.7	1	5	3	5125	1281.25	1133	6.2	5.8	3.5
part	664	572	1.0	1.1	1	4	14	5066	1688.667	1908	3.0	5.1	4.5
pink	649	582	1.2	1.2	1	4	14	8002	2667.333	901	6.7	3.8	3.7
pint	721	690	1.2	1.7	0	4	10	10734	3578	1222	6.0	8.4	3.3
plant	645	591	1.6	1.6	0	5	4	9297	2324.25	1590	6.2	4.0	3.7
plus	661	569	1.8	1.7	1	4	2	2826	942	621	3.4	6.4	3.9
posh	699	595	1.6	1.4	1	4	9	3215	1071.667	822	4.6	11.9	3.4
pure	663	586	1.4	1.8	0	4	9	6292	2097.333	1204	3.2	8.0	3.5
put	650	585	1.1	1.1	0	3	15	1338	669	574	2.6	3.7	4.6
range	685	601	1.7	1.7	0	5	0	11723	2930.75	1621	2.9	7.1	3.9
rash	760	624	1.3	1.0	1	4	11	5857	1952.333	881	5.2	5.8	2.8
reach	668	593	1.5	1.3	1	5	8	9338	2334.5	2799	4.0	4.9	4.0
real	639	583	1.2	1.9	0	4	14	9421	3140.333	2432	2.5	5.0	4.4
rear	745	661	1.1	1.0	0	4	16	9322	3107.333	2630	4.0	6.9	3.3
rent	707	590	1.2	1.0	1	4	15	11376	3792	2427	3.9	8.9	3.4
rinse	761	651	2.0	1.6	1	5	0	13656	3414	1454	5.0	5.0	2.5
roast	669	612	1.8	1.1	0	5	4	7580	1895	1123	5.3	8.5	3.1
saw	687	529	1.1	1.0	0	3	15	962	481	409	4.8	5.4	4.4
scarce	824	636	2.0	1.9	0	6	0	7709	1541.8	1136	2.1	10.5	2.8
scold	770	566	1.9	1.7	0	5	1	4977	1244.25	677	5.1	8.5	1.6
sense	688	543	1.8	1.2	1	5	2	9734	2433.5	1753	2.9	7.3	4.2
shall	668	537	1.7	1.3	1	5	6	7121	1780.25	1218	1.7	8.4	4.2
share	637	532	1.1	1.0	0	5	15	9947	2486.75	1798	3.7	4.8	4.0
she	661	539	1.6	1.0	1	3	4	3044	1522	754	4.0	3.6	4.5
short	651	540	1.5	1.5	0	5	7	6246	1561.5	1227	4.6	4.3	4.2
shout	661	552	1.6	1.5	1	5	6	5106	1276.5	1021	5.2	4.7	3.6
show	643	538	1.3	1.0	0	4	10	3175	1058.333	936	4.8	6.2	4.5
skate	735	558	1.7	1.4	1	5	2	8916	2229	1126	5.9	5.4	2.6
skill	664	529	1.6	1.2	1	5	5	3701	925.25	597	3.0	6.8	3.5
staff	651	516	1.9	1.5	0	5	2	6077	1519.25	1145	5.0	10.0	3.9
stay	651	532	1.4	1.1	1	4	8	5865	1955	1069	3.0	4.2	4.3
stood	671	527	1.9	1.7	0	5	2	6392	1598	1114	4.2	4.4	3.8
stop	645	499	1.5	1.3	1	4	5	5682	1894	1091	5.0	2.9	4.4
store	659	510	1.2	1.1	0	5	13	12227	3056.75	1715	4.2	4.8	3.7
stuff	651	524	1.8	1.4	1	5	4	4832	1208	941	4.7	5.0	4.3
swan	708	551	1.4	1.6	0	4	10	4343	1447.667	753	6.3	6.3	3.0
sweat	681	530	1.7	1.6	0	5	3	6676	1669	931	5.6	7.3	3.2
swell	739	538	1.6	1.4	1	5	5	4031	1007.75	545	4.7	7.4	2.8
swim	654	531	1.7	1.5	1	4	6	1550	516.667	396	6.4	4.2	3.4
taste	652	602	1.6	1.0	0	5	6	11289	2822.25	2087	4.4	4.3	3.9
test	642	588	1.3	1.0	1	4	13	12449	4149.667	1575	5.1	6.3	4.1
thaw	810	650	1.8	1.4	0	4	4	2694	898	733	3.5	8.1	2.5
thump	767	643	1.8	1.5	1	5	3	3121	780.25	631	4.8	7.6	2.4
train	667	589	1.6	1.1	1	5	6	12612	3153	2034	6.3	4.0	3.9
truth	655	595	1.9	1.7	0	5	0	4617	1154.25	1148	2.9	4.4	4.0
want	632	572	1.3	1.6	0	4	12	7165	2388.333	1437	3.4	4.2	4.6
wasp	708	613	1.7	2.0	0	4	5	2945	981.667	899	5.9	5.6	2.4
wealth	690	619	2.0	1.6	0	6	1	7057	1411.4	1355	4.8	8.8	3.4
weave	747	630	1.8	1.2	1	5	3	4583	1145.75	1318	5.1	9.9	2.8
went	623	580	1.2	1.1	1	4	15	7318	2439.333	1146	2.5	3.4	4.5
whisk	743	625	2.0	1.9	1	5	0	4303	1075.75	754	5.7	8.7	2.8
wide	662	607	1.3	1.0	1	4	13	3915	1305	775	4.1	5.8	3.9
width	757	721	1.9	1.8	1	5	0	2653	663.25	701	3.6	8.8	2.8
wife	647	585	1.6	1.5	1	4	9	1542	514	465	5.5	5.7	4.1
will	646	592	1.0	1.0	1	4	16	3517	1172.333	1137	2.7	7.5	4.6
with	685	594	1.7	1.5	0	4	3	3968	1322.667	734	2.2	4.4	4.6
wolf	654	595	1.9	1.9	0	4	2	1737	579	691	6.4	4.5	3.0
worm	702	608	1.5	1.5	0	4	9	3675	1225	1103	6.2	3.9	2.9
worse	676	616	1.7	1.3	0	5	4	6592	1648	1501	2.0	2.8	4.1
yawn	718	586	1.7	1.3	0	4	7	505	168.333	156	6.4	5.3	2.4
yet	676	576	1.3	1.3	1	3	13	1536	768	405	1.8	7.0	4.4