The Influence of Orthographic Consistency on Eye Movement Behaviour in Word and Sentence Processing

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Dipl. Psych. Anne Katrin Rau
aus Tübingen

Tübingen
2014

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	17.3.2015
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatterin:	Prof. Dr. Karin Landerl
2. Berichterstatterin:	Prof. Dr. Claudia Friedrich

Danke

Danke an Prof. Dr. Karin Landerl – liebe Karin, danke für ein spannendes Thema, für eine spannende Zeit in Deiner Tübinger Arbeitsgruppe, und vor allem danke dafür, dass Du mich auch auf die Distanz noch so lange und so gut betreut hast.

Danke an Prof. Dr. Claudia Friedrich – liebe Claudia, danke dass Du mich als Doktorandin und Mitarbeiterin übernommen hast und mir eine neue Aufgabe in Deiner Arbeitsgruppe gegeben hast, danke natürlich auch dafür, dass Du meine Dissertation liest und bewertest.

Danke an Prof. Dr. Korbinian Möller – danke Korbinian für die Hilfe bei der Überführung der eye tracking Daten in Excel und SPSS, für das gefühlte zwanzig-malige Durchlesen meines ersten papers und für Dein wertvolles feedback.

Danke an Dr. Kristina Moll und an Prof. Dr. Maggie Snowling für die unkomplizierte und ausgesprochen angenehme Kooperation in York, und für die schöne Zeit in York mit Dir, Kristina. Danke auch Euch beiden für die wertvollen Kommentare zum JECP und zum eingereichten dritten paper.

Danke den alten und neuen Kolleginnen und Kollegen Patrick Schleifer, Dr. Beate Kajda, Andy Kramer, Larissa Schellenberger, Maren Reinl, Franziska Schleger, Judith Janschewski, Prof. Dr. Birgit Träuble, Prof. Dr. Mechthild Kiegelmann, Dr. Ulrike Schild, und Anne Bauch.

Danke allen teilnehmenden Kindern, ihren Eltern, Lehrern und Schulen. Spezielles Danke an Juri “Sucho” Suchowerski.

Danke den Hilfskräften in York, Rachel Kirmond und Calvin Wong.

Danke meiner Familie, vor allem dem Christoph.

Contents

Summary

It has long been established that the English orthography is exceptionally inconsistent with regards to the relationship of letters and sounds, and that this inconsistency causes a substantial delay in initial reading acquisition of English-speaking children compared to children learning to read a more consistent orthography.

Perhaps it is little surprising that readers appear to adapt their reading behaviour to the demands of the orthography they are reading. In this context, psycholinguistic grain size theory (Ziegler & Goswami, 2005) has suggested that small-unit bottom-up processing should be more prevalent in readers of consistent orthographies, and large-unit top-down processing should be more prevalent in readers of inconsistent orthographies.

The research project presented here sought to test whether this is so. Three studies investigated word and sentence processing in the consistent German and the inconsistent English orthography by studying eye movement behaviour in developing and adult readers.

Study 1 investigated the transition in predominant reading strategy from serial sublexical decoding to more parallel lexical processing as a function of word frequency in the consistent German orthography. The least experienced readers of grade 2 were found to apply serial sublexical decoding as a default reading strategy to most items, largely independent of word frequency, whereas more experienced readers of grades 3 and 4, and adults, were increasingly relying on direct lexical access.

In a direct cross-linguistic comparison, studies 2 and 3 investigated the influence of orthographic consistency on the time-course of word and sentence processing. The two studies found evidence for more small-unit bottom-up processing on the part of the German, and more large-unit top-down processing on the part of the English readers, for both the local word, and the global sentence level. The overall processing pattern differed between orthographies: while German readers showed a plodder-like reading style with more diligent first-pass reading and less re-reading, English readers showed an explorer-like reading style with more word skippings and more regressive eye movements. In sum, orthographic consistency impacts upon both local word recognition and global sentence processing, in both developing and skilled readers.

It would be interesting for future studies to establish whether German-English bilinguals process nearly identical sentences in a similar way in both languages, or whether they adapt their eye movement behaviour to the demands of the orthography they are reading.

Part I

Theoretical Background and Scientific Context

1 Reading and Reading Development in Alphabetic Orthographies

I take it you already know
Of tough and bough and cough and dough
Others may stumble, but not you
On hiccough, thorough, laugh, and through.
And cork and work and card and ward
And font and front and word and sword
Well done! And now if you wish, perhaps
To learn of less familiar traps.
Beware of heard, a dreadful word
That looks like beard and sounds like bird.
And dead: it's said like bed and not like bead -
For goodness sakes don't call it deed.
Watch out for meat and great and threat,
They rhyme with suite and straight and debt.
A moth is not a moth in mothers,
Nor both in bother, broth in brother.
And here is not a match for there,
And dear and fear for bear and pear.
And then there's dose and rose and lose -
Just look them up--and goose and choose,
And do and go, then thwart and cart.
Come, come, I've hardly made a start!
A dreadful language? Man alive!
I'd mastered it when I was five.

(Brush up on your English, https://www.wordsmith.org, n.d.)

1.1 Introduction: English – an outlier orthography

Until rather recently, the vast majority of empirical and theoretical work in reading research was done with English speakers reading in their native language, and for a long time the universal validity of these findings went unquestioned. In 2008, David Share published an influential article which quite explicitly addressed the problems arising from "overreliance on an outlier orthography" in reading research. By reviewing studies carried out in orthographies other than English which had reported findings quite different to previously reported English-based findings, Share pinpointed very clearly the nuisance of the anglocentricities of ongoing reading research and practice that many reading researchers in other languages had been criticising for a number of years.

But what is it that makes English an "outlier" orthography? Well, an impression of the particularities of English orthography is provided by the poem cited at the outset of this chapter. The fundamental principle of alphabetic writing systems is that letters or letter clusters (i.e., graphemes) represent speech sounds (i.e., phonemes) (Landerl, 2006). Shallow orthographies represent the linguistically shallow level of phonology, and are characterised by a highly consistent relationship between graphemes and phonemes. An example of a perfectly shallow, or consistent alphabetic orthography is Finnish, as each letter represents exactly one phoneme, and each phoneme is represented by exactly one letter. Deep orthographies represent the linguistically deeper level of morphology, and are characterised by a highly variable and inconsistent relationship between graphemes and phonemes. English is considered *the* most inconsistent alphabetic orthography of all (Borgwaldt, Hellwig, & de Groot, 2005; Frost, 2012a; Landerl, 2006; Share, 2008). Thus, while the majority of alphabetic orthographies is less consistent than Finnish, none is quite as inconsistent as English. Indeed, English is exceptionally inconsistent on both the level of grapheme-phoneme-correspondences, which is relevant for reading, and the level of phoneme-grapheme-correspondences, which is relevant for spelling. These inconsistencies are responsible for the fact that some words are spelled very similarly, but are pronounced differently (such as *heard* and *beard* in the poem cited above), while at the same time, other words are pronounced the same, but are not spelled the same (such as *heard* and *bird*).

According to Share (2008), the long standing focus on English-based findings has led to a research agenda addressing theoretical and applied issues with limited relevance to a universal science of reading. The issues he lists which have received disproportionate attention in the past include reading accuracy, phonological awareness, early reading instruction, the architecture of stage models of reading development, the definition and remediation of reading disability, and the role of lexical-semantic and supralexical information in word recognition.

Recently, Share (2014) has extended his critique to what he terms “alphabetism”. This bias is reflected in a tendency to interpret alphabetic writing systems as inherently superior to non-alphabetic writing systems (i.e., logographic scripts with symbols representing words or word parts (morphemes), such as Chinese; syllabaries with symbols representing syllables, such as Japanese kana; or consonantal scripts with symbols representing the consonants, but not the vowels of the language, such as Arabic or Hebrew). Share is right in stating that the majority of current reading research is done in alphabetic orthographies, even though the majority of the world’s readers read in non-alphabetic scripts. The current piece of work is concerned with reading and reading development in a consistent and an inconsistent alphabetic orthography, and in this sense “eurocentristic”. Nevertheless, findings from cross-linguistic comparisons between alphabetic orthographies of differing orthographic depth such as the ones presented here are just as important as cross-script comparisons when it comes to creating an empirical basis for the development of more universal theories of reading and reading development.

1.2 Models of reading and reading development

1.2.1 Stage models of reading development

Early theories of reading development have suggested that reading development progresses in qualitatively different stages. One of the most influential models at the time was that of Frith (1985) positing an initial logographic stage, followed by an alphabetic, and finally an orthographic stage. In the logographic stage, a child is proposed to be able to recognise a limited number of familiar words based on salient graphic features (e.g., recognition of the word “dog” on the basis of the little tail). At this stage, a child is said unable to pay much attention to letter order, let alone phonological factors. At the alphabetic stage, a child has learned the grapheme-phoneme-correspondence rules and applies them in a systematic manner. This strategy allows for the pronunciation of entirely novel words, or nonwords. The final orthographic stage refers to the instant analysis of words into orthographic units without phonological conversion. Frith based her model on a number of previously published findings, and there have equally been subsequent studies supporting the existence of these stages (Seymour & Elder, 1986; Stuart & Coltheart, 1988). However, all of the supportive evidence was based on English-speaking children learning to read in English.

The applicability of the model to more consistent orthographies has consequently been challenged. Thus, Wimmer and Hummer (1990) found little evidence for early logographic reading in German-speaking first graders.

1.2.2 Dual route theory

One of the most influential models of skilled reading, which has equally been applied to reading acquisition, is the dual-route model of reading aloud (Coltheart, 1978, 1985, 2006; Coltheart, Curtis, Atkins, & Haller, 1993). According to this model, there are two parallel reading routines, a nonlexical and a lexical route. Nonlexical word recognition works serially via application of grapheme-phoneme-correspondence rules; a process also termed phonological decoding (Olson, Wise, Ring, & Johnson, 1997), or phonological recoding (Share, 1995). Lexical word recognition is described as an automatic, direct, parallel process, and is basically a very fast and efficient retrieval of a word's orthographic representation from the mental lexicon. It appears that the main inspiration for the proposal of the two reading procedures in dual route theory has been the distinction between words that conform to the rules of English grapheme-phoneme-correspondences (so-called regular words and nonwords) and words that do not conform to these rules (so-called irregular words, or exception words). Thus, a fundamental claim of dual-route theory is that while regular words (e.g., *hit*) can be read via either route, nonwords (e.g., *gilp*) can only be read via the nonlexical route, and exception words (e.g., *pint*) can only be read correctly via the lexical route.

This focus on the nonword/exception word dimension has been object to criticism because the majority of more consistent orthographies do not have a lot of, if any, exception words. For many of these highly transparent orthographies, the nonlexical route alone would in theory be sufficient for successful word recognition. Nevertheless, it has often been shown that readers in consistent orthographies equally use the lexical route (German: Rau, Moeller, & Landerl, 2014; Italian: Pagliuca, Arduino, Barca, & Burani, 2008; Serbo-Croatian: Katz & Feldman, 1983), and that they start doing so from an early point in reading development (Dutch: Wesseling & Reitsma, 2000; German: Rau et al., 2014; Italian: Burani, Marcolini, & Stella, 2002). In this context, Share (2008) noted that the central dualism inherent in all skill learning is a transition from slow, deliberate, step-by-step, unskilled performance to rapid, highly automatised, one-step, or "holistic" performance. He has therefore proposed that a universally applicable "unfamiliar-to-familiar" dimension ought to replace the nonword-exception word dimension mainly applicable to English. Importantly, unlike the between-item nonword/exception word dualism of original dual route theory, the proposed unfamiliar-to-familiar dualism is a within-item developmental transition. From a developmental perspective, it is easily conceivable that every word is visually unfamiliar at first. Every word is thus functionally a nonword at first which needs to be phonologically decoded by applying grapheme-phoneme-correspondence rules. Conversely, a nonword can become a familiar, well-known letter string if it is encountered often enough (e.g., the *gruffalo* is a familiar letter string to readers of the storybook by Julia Donaldson and Axel Scheffler).

In fact, the idea that the process of phonological recoding serves as an item-based mechanism for the build-up of word-specific representations (on which subsequent rapid orthographic word recognition relies) exactly corresponds to the predictions of Share's self-teaching hypothesis (Share, 1995; for a recent computational implementation, see Grainger, Lété, Bertand, Dufau, & Ziegler, 2012). Following the idea that the transition from effortful, step-by-step recoding to automatic, holistic word recognition happens in an item-based, rather than in a stage-based way, self-teaching hypothesis proposes that each successful decoding encounter with an unfamiliar word provides an opportunity to acquire word-specific orthographic information. It appears that a relatively small number of successful decoding encounters are sufficient for acquiring orthographic representations (e.g., de Jong & Share, 2007; Reichle & Perfetti, 2003). While it is much less important in skilled word recognition, phonological recoding is considered *the* principal means by which the learner attains word recognition proficiency (this applies especially, but not exclusively, to consistent orthographies).

An important element of self-teaching hypothesis is the claim that the process of word recognition will depend primarily on the actual frequency to which a child has been exposed to a particular word. Since orthographic information is acquired rapidly, high-frequency words are likely to be recognised visually with minimal phonological processing from a very early point in reading acquisition. In contrast to this, novel, less frequent words for which the child has yet to acquire orthographic representations will be more dependent on phonological processing.

1.2.3 Connectionist models of word recognition

Connectionist models of reading aloud are typically computational models. The concept in computational modelling is that the underlying theory should provide a solid explanation for empirical phenomena while the computational model itself should be able to simulate the data as precisely as possible. Early computational models of reading were only able to simulate a very narrow range of phenomena, and had a range of fundamental limitations, such as only being able to handle four-letter words. More recent models have produced much more accurate simulations of an increasingly wide range of empirical data (for a comprehensive review of computational models of reading, see Norris, 2013).

Connectionist models of word recognition challenge the notion that there are different routes for different kinds of words (e.g., Harm & Seidenberg, 2004; Plaut, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996). Instead, they postulate the operation of a single interactive mechanism for reading both words and nonwords. Three sets of processing units are proposed to interact in word recognition – a set of grapheme units representing orthography, a set of phoneme units representing phonology, and a set of semantic units. Within this framework, words (and nonwords) correspond to distributed patterns of activity across each set of units. The underlying triangle model encompasses two pathways between written and spoken words: a first pathway mapping from orthography to phonology directly, and a second pathway, mapping from orthography to phonology through semantics.

In spite of the recent advances in the domain of computational models of reading (for instance, computational models combining connectionist with dual route features: Ans, Carbonnel, & Valdois, 1998; Perry, Ziegler, & Zorzi, 2014), individual models tend to focus on a single domain of behaviour, such as reading aloud, eye movements, or lexical decision (Norris, 2013). There are two further issues which have largely been ignored by connectionist models. First, connectionist models have typically focused on skilled word recognition, and are therefore of limited applicability when it comes to reading *development* (Frost, 2012b). Second, most connectionist models do not provide a comprehensive framework for explaining cross-linguistic differences in word recognition (Frost, 2012a; but see Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004; and Perry et al., 2014 for attempts to apply such models cross-linguistically).

Thus, in the absence of an integrated connectionist computational model of reading which adequately explains reading development in different orthographies; more traditional theories are more readily applicable in the context of the present research endeavour.

1.2.4 Models considering the influence of orthographic consistency

Provided that the self-teaching aspect of the with-item unfamiliar-to-familiar transition replaces the between-item nonword-versus-exception word distinction in dual route theory, the latter remains a useful model in the field of reading development and skilled reading which can also adequately describe reading development in orthographies of differing consistency (cf. Ziegler, Perry, & Coltheart, 2000).

But even well before Share proposed the universally applicable unfamiliar-to-familiar dimension to replace the anglo-specific nonword-versus-exception word distinction in 2008, it has been suggested that the consistency of an orthography should impact upon the speed of the respective reading procedures. If one thinks of the two routes as two horses engaging on a word recognition race, Paap and Noel (1991) predicted that the lexical horse should be equally fast across orthographies, while the nonlexical horse should be faster in shallow orthographies.

Indeed, there have been numerous studies showing initial reading acquisition in general, and the acquisition of phonological recoding skills in particular to progress faster in consistent than in inconsistent orthographies (e.g., Aro & Wimmer, 2003; Ellis & Hooper, 2001; Frith, Wimmer, & Landerl, 1998; Landerl, Wimmer, & Frith, 1997; Thorstad, 1991; Wimmer & Goswami, 1994). Importantly, while factors such as age at onset of formal reading instruction, or the actual method of reading instruction, are also likely to play a role in explaining the different rate of reading acquisition in consistent and inconsistent orthographies, the central factor appear to be differences in orthographic consistency (Landerl, 2000; Seymour, Aro, & Erskine, 2003).

1.2.4.1 Orthographic depth hypothesis

An early and very influential account considering the influence of orthographic consistency on the reading process was the orthographic depth hypothesis, the weak version of which stated that "the degree to which the prelexical process is active is a function of the depth of the orthography; prelexical analytic processes are more functional in shallow orthographies" (Frost, 1994, p. 117; Frost, Katz, & Bentin, 1987). Put more simply, this means that readers of shallow orthographies rely more strongly on phonological recoding processes than do readers of deep orthographies.

With regards to reading acquisition, the orthographic depth hypothesis implies several further hypotheses (Aro, 2006). Accordingly, the high consistency of grapheme-phoneme-correspondences in shallow orthographies should promote a rapid development of phonological recoding skills. In contrast, the complexity and inconsistency of grapheme-phoneme-correspondences in deep orthographies should force beginning readers to supplement phonological recoding strategies with reading strategies aiming at the level of the rime, or the level of whole words.

1.2.4.2 Psycholinguistic grain size theory

These hypotheses which naturally follow from the orthographic depth hypothesis were further specified and given a complete theoretical framework by Ziegler and Goswami (2005) in the form of psycholinguistic grain size theory. The considerably lower rate at which reading development progresses in inconsistent orthographies is thus explained by the need to develop reading strategies targeting psycholinguistic units at a variety of grain sizes.

Specifically, readers of English not only need to develop a strategy for whole-word recognition in order to be able to read the fairly high number of irregular words such as *choir* or *yacht*, they also need to develop a strategy for recognising rhyme analogies to enable the reading of irregular words sharing a consistently pronounced letter cluster (e.g., words such as *flight*, *night*, *bright*, *light*). Importantly, they further need to develop a strategy for the conversion of graphemes to phonemes in order to be able to read unknown words. In contrast to this, for children learning to read a consistent orthography, the strategy of systematic grapheme-to-phoneme conversion is perfectly sufficient in the early phase of reading development. It therefore appears that reading acquisition for children learning to read an inconsistent orthography is delayed for two reasons: first, by the need to develop a number of reading strategies targeting different grain sizes and second, by the fact that both the acquisition and the successful application of grapheme-phoneme-correspondence rules are harmed by their inherent inconsistent nature.

Hence, grain size theory claims that readers of consistent orthographies rely more on small-unit processing, and readers of less consistent orthographies rely more on large-unit processing. It is important to note at this point that such differences in the size of preferred processing units are predicted between both beginning and skilled readers of different orthographies. Conversely, some researchers have argued that the development of reading in languages of differing orthographic depth converges as lexical influences become increasingly dominant in the course of reading development in consistent orthographies, too (e.g., de Jong, 2006).

1.3 Empirical findings on the influence of orthographic consistency on reading

1.3.1 Empirical findings from consistent orthographies

Consistent with the idea that differences in word recognition processes between shallow or deep orthographies are to be mainly expected early in reading development, most studies have investigated such processes in primary school-aged children rather than in more skilled readers.

There is an ample body of research supportive of the notion that early reading in consistent orthographies is mainly based on small-unit grapheme-phoneme decoding from studies investigating the development of word recognition in a single orthography (e.g., Finnish: Holopainen, Ahonen, & Lyytinen, 2002; German: Rau et al., 2014; Greek: Porpodas, 2006; Italian: Zoccolotti, de Luca, di Pace, Gasperini, Judica, & Spinelli, 2005; Zoccolotti, de Luca, di Filippo, Judica, & Martelli, 2009; Spanish: Goswami, Gombert, & Fraca de Barrera, 1998; Signorini, 1997). Many studies have taken the word length effect (i.e., the increase in processing time with increasing number of letters) as indicative of serial phonological recoding (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Martens & de Jong, 2006; Rau et al., 2014; Weekes, 1997; Zoccolotti et al., 2005, 2009). The size of the word length effect typically decreases with increasing reading skill, a finding which has been interpreted to reflect a gradual shift from serial, letter-by-letter reading towards more parallel processing of increasingly larger parts of words (e.g. van den Boer, de Jong, Haentjens-van Meeteren, 2013, Zoccolotti et al., 2005, 2009).

1.3.2 Empirical findings from inconsistent orthographies

The claim that early reading in less transparent orthographies is more reliant on units larger than the single grapheme is based on evidence from a number of older studies with children learning to read in English. While it is widely accepted that the English orthography is exceptionally inconsistent with regards to small-unit grapheme-phoneme correspondences, it has been shown that larger grain sizes such as rimes are not quite as inconsistent (e.g., *link*, *sink*, *wink*, *drink*, *blink*) (Treiman, Mullenix, Bijeljac-Babic, & Richmond-Welty, 1995). A number of studies have shown that from an early phase of reading development beginning readers of inconsistent orthographies make use of analogies to known words when asked to read low-frequency words and nonwords (e.g., Brown & Deavers, 1999; Goswami, 1986; Laxon, Coltheart, & Keating, 1988; Laxon, Masterson, & Moran, 1994; Treiman, Goswami, & Bruck, 1990). Likewise, some studies have found evidence for early reading at even larger grain sizes than the rime. Error analyses on those words that children failed to read correctly often revealed a tendency to attempt to read words as a whole. Thus, young children would often read an existing, familiar word instead of the less familiar word or nonword (e.g., read *policeman* instead of *postman*, or *children* instead of *chocolate*) (Laxon et al., 1994; Seymour et al., 2003; Seymour & Elder, 1986).

It is crucial to note that even early word recognition in consistent orthographies is neither *entirely* reliant on small grain sizes (e.g., Burani, Marcolini, & Stella, 2002; Davies, Cuetos, & Glez-Seijas, 2007), nor is it *entirely* reliant on larger grain sizes in less consistent orthographies (e.g., Duncan, Seymour, & Hill, 1997; Goswami, Ziegler, Dalton, & Schneider, 2001). Indeed, this is not what grain size theory posits. Its central claim is rather that reading in consistent orthographies is *more* dependent on small-unit processing, and reading in less consistent orthographies is *more* dependent on large-unit processing. In order to appropriately test such a comparative prediction, direct cross-linguistic studies are the method of choice.

1.3.3 Empirical findings from cross-linguistic studies in developing readers

Since the 1990s, a growing number of studies have directly compared the development of reading ability in languages of differing orthographic depth. From these studies, different types of evidence have emerged supportive of the claim that readers of consistent orthographies rely more on small-unit processing while readers of less consistent orthographies rely more on large-unit processing.

Firstly, nonword reading accuracy has consistently been reported to be much lower in children learning to read more inconsistent orthographies than in children learning to read more consistent orthographies (e.g., Aro & Wimmer, 2003; Frith et al., 1998; Landerl, 2000; Seymour et al., 2003; Thorstad, 1991; Wimmer & Goswami, 1994), a finding reflecting the fact that the process of serially decoding small-unit grapheme-phoneme-correspondences is more readily available to beginning readers of consistent orthographies than to beginning readers of inconsistent orthographies. It should be noted, though, that in many of the studies children learning to read the consistent orthography were considerably older than the children learning to read the inconsistent English orthography. Even though age does not appear to be a main factor in explaining cross-linguistic differences in word processing, the procedure of phoneme blending does require some memory capacity which may well be partly dependent on age.

Secondly, the kind of errors children learning to read orthographies of differing consistency tend to make when reading also point to the use of differing word recognition strategies. Thus, a high percentage of reading errors in beginning reader of less consistent orthographies consists of refusals or null attempts and word substitutions for both words and nonwords (e.g., Ellis & Hooper, 2001; Frith et al., 1998; Seymour et al., 2003). The refusal to try and read an unfamiliar word or nonword can be interpreted as reflecting the inability to apply small-unit grapheme-phoneme-correspondences while the tendency to reply with a familiar word when asked to read another word or a nonword can be interpreted as an attempt at whole-word recognition. In contrast, beginning readers of more consistent orthographies have been reported to produce mainly nonword errors reflecting a reliance on a small-unit strategy of alphabetical decoding (e.g., Ellis & Hooper, 2001; Seymour et al., 2003).

Thirdly, a few such direct comparisons have shown stronger word length effects in children learning to read a consistent orthography than in children learning to read English. Thus, Goswami and colleagues (1998) found children learning to read the consistent Spanish orthography to read monosyllabic words much better than bisyllabic words in terms of both accuracy and speed while children learning to read the less consistent English and French orthographies showed much less of a difference between words of different length. Similarly, children learning to read the consistent Welsh orthography were found to produce an almost perfectly linear relationship between word length and naming latency while children learning to read English showed a non-linear increase in naming latency up to nine letters and an actual decrease in naming latency for words longer than this (Ellis & Hooper, 2001).

Fourthly, a number of studies have shown stronger evidence for the use of lexical analogies at the rime level in English children than in children learning to read more consistent orthographies. In a study using English-German cognates with either few or many body neighbours (i.e., words that share the same orthographic rime (body), such as *street*, *meet*, *feet*), the facilitatory effect of neighbourhood size was found to be larger in English than in German children, a finding which was interpreted as indicating more large-unit processing on the part of the English children (Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Körne, 2003). In addition to this, length effects were found to be reduced in high-N words as compared to low-N words for English, but not for German children, a finding interpreted to suggest that German children continued to process smaller units even when large-unit information was present. There are, however, important methodological limitations that apply to this study. First, Ziegler and colleagues used only monosyllabic words, which may be appropriate for readers of English, but certainly not for readers of German. Second, a substantial proportion of the words they termed cognates were not truly comparable either semantically (e.g., *See* and *sea*; *Fan* and *fan*; *Kleid* and *cloth*; *Biest* and *beast*; *Mahl* and *meal*), phonologically and orthographically (e.g., *Zeh* and *toe*; *Tag* and *day*; *Stuhl* and *chair*), or were probably much more frequent in English than in German (*Akt* and *act*; *Pfund* and *pound*). Third, and importantly, for both small and large body neighbourhoods, the English targets had more neighbours on average than the German words (small neighbourhood: 3.3 vs 4.3 body neighbours for German and English, respectively; large neighbourhood: 9.8 vs 13.2

body neighbours for German and English, respectively). Given that the words with a large body neighbourhood had considerably more neighbours in English than in German, the larger facilitatory effect in English as opposed to German children cannot unequivocally be attributed to more large-unit processing on their part.

Most other studies showing stronger large-unit effects in English children than in children learning to read more consistent orthographies have used different types of nonwords as stimuli. Typically, processing times for young readers of more consistent orthographies did not differ much whether nonwords shared an orthographically familiar letter sequence with existing words or not. In contrast, English children (and in one study, French children) showed a clear benefit for nonwords that could be read by analogy to existing words (e.g., *bicket*; real-word analogue: *ticket*) as compared to phonologically identical nonwords for which no such analogy was evident (e.g., *bikket*); a benefit which was either less pronounced or absent in beginning readers of more consistent orthographies (Goswami et al., 1998; Goswami, Porpodas, & Wheelwright, 1997; Goswami, Ziegler, Dalton, & Schneider, 2003). However, a methodological shortcoming of these studies is that the reading material was not the same for English and the more consistent orthographies (which is admittedly hard to achieve in languages as distant as English and Greek).

Beginning readers of the inconsistent English orthography show stronger effects of frequency and lexicality than beginning readers of more consistent orthographies. The frequency effect denotes the processing advantage of high-frequency words over low-frequency words in terms of reading accuracy and/or speed. Similarly, the lexicality effect denotes the processing advantage of existing words over nonwords. Importantly, the larger the processing difference between familiar words and less familiar words (i.e., low-frequency words or nonwords), the more pronounced the dissociation of reading procedures between words of the two categories. Thus, stronger effects of frequency and lexicality are to be expected for readers who predominantly process larger units than for readers who predominantly process smaller units (e.g., Paap & Noel, 1991). Consistent with this notion, Wimmer and Goswami (1994) found the processing difference between words and nonwords in terms of mean reading time per item was at least double in English children than that in German children for three different age groups. Similarly, Frith and colleagues (1998) reported much stronger effects of both frequency and lexicality in young readers of English than in young readers of German. Landerl and colleagues (1997) reported no effect of word frequency in normally-reading German children selected as age-level controls for children with reading difficulties for either reading accuracy or reading speed. In contrast, English children showed higher error rates on low-frequency words than on high-frequency words even though they were one year older than their German counterparts and had received about two more years of reading instruction.

1.3.4 Empirical findings from cross-linguistic studies in skilled readers

The studies mentioned above have contrasted the development of word recognition in primary-school aged children learning to read a consistent orthography with those learning to read an inconsistent orthography. There are only a few studies reporting cross-linguistic differences between skilled adult readers.

An older study comparing naming and lexical decision between adult readers of the deep English and the shallow Serbo-Croatian orthography found an influence of orthographic consistency on both tasks (Katz & Feldman, 1983). The naming process seemed to be lexically mediated in English adults since their naming latencies were faster when the target was preceded by a semantically related word. While the same effect was also found in Serbo-Croatian adults, it was less pronounced, suggesting less lexical involvement in word pronunciation in the more consistent orthography. Further, naming and lexical decision latencies were correlated for English, suggesting a lexical influence on both processes, but mostly uncorrelated for Serbo-Croatian, suggesting different processes to operate.

More recently, Paulesu and colleagues (2000) found naming latencies in both word and nonword reading to be lower in Italian than in English university students. Importantly, while both groups were faster at naming words than at naming nonwords, this lexicality effect was more pronounced for readers of English. In two subsequent experiments using positron emission tomography (PET), the authors found a different pattern of brain activation in an explicit and an implicit reading task. While the Italian readers showed greater activation in brain regions associated with sublexical phonological processing, the English readers showed greater activation in brain regions associated with lexical and semantic processing.

The other cross-linguistic study reporting differences in (non)word processing between skilled adult readers found the influence of word length to be stronger in German than in English-speaking university students. In contrast, while the English-speaking readers did show a significant facilitatory effect of body neighbourhood size on (non)word naming latencies, the German readers did not (Ziegler, Perry, Jacobs, & Braun, 2001). Importantly however, similar to the study by the same first author reported above (Ziegler et al., 2003), this study included only monosyllabic items, and body neighbourhood size was larger for English than for German items, for both words and nonwords, and for both small and large body neighbourhoods. Again, the possibility cannot be ruled out that the finding of a facilitatory effect of body neighbourhood size on English, but not German readers was in part due to the fact that the English stimuli had more neighbours than the German stimuli.

1.3.5 Methodological constraints in cross-linguistic research

While there is little doubt that direct cross-linguistic studies are the preferred way to investigate word processing in different alphabetic orthographies, there are a number of methodological constraints that apply to such studies. Major methodological difficulties faced by cross-linguistic studies of reading development relate to the matching of participants on the one hand, and the matching of reading material on the other hand. These points will be discussed in the following, and chapter 2 will explain how the research project presented here has dealt with them.

1.3.5.1 Age at school entry

Cross-linguistic comparisons of reading development typically involve children from different countries, and consequently from different educational systems. An important factor influencing reading development is the age of school entry which actually differs quite substantially between some countries. For instance, in the ambitious cross-linguistic comparison of 13 European orthographies by Seymour and colleagues (2003), the Scottish first graders were about two years younger than the first graders from Finland, Norway, Sweden, Denmark, Germany, and Austria, and still one year younger than the first graders from all other countries. When cross-linguistically comparing children of the same grade level, children are therefore not necessarily comparable in terms of age, and those children entering school at a younger age may well be less advanced with regards to general cognitive and linguistic development (Landerl, 2000).

The confound introduced by differing age can of course be overcome by comparing children of the same age, but this will in turn mean that those children entering school at a younger age will have received more reading instruction, which is a confound with reading experience. Even though it is surely possible to compare a group of eight year-old second grade children from one orthography with a group of eight year-old third grade children *and* a group of seven year-old second grade children from another orthography, there remains the trouble that children of the two orthographies will always differ in either age or reading experience (if age at school entry differs). In one study comparing phonological recoding in German and English developing readers, children were selected from corresponding grades, and then further matched on “reading age” according to their performance in a standardised word reading test (Goswami et al., 2001). However, this approach to matching children on reading age is questionable since the two standardised reading tests were not readily comparable between orthographies, and were not designed for this purpose either. It should be noted though that age at the onset of formal reading education is unlikely to play a central role in explaining cross-linguistic differences in reading acquisition. Thus, in the study by Seymour and colleagues, the Scottish children were the youngest out of all other groups, and performed much worse in both word and nonword reading than all other groups, but the Danish children were among the oldest to enter school, and also performed much worse than most other groups in word and nonword reading.

1.3.5.2 Method of reading instruction

Besides age of school entry, another factor which may differ between educational systems of different countries is the dominant method of reading instruction. The methodological limitation which applies to many studies in this context is that the approach to reading instruction typically reflects a language’s orthographic consistency.

The most extensively studied orthography is English; and in accordance with the exceptionally inconsistent orthography of the English language, English-speaking countries have been exceptional with regards to the method of reading instruction. In contrast to other alphabetical orthographies, which typically teach reading through phonics (Aro & Wimmer, 2003), the English-speaking children in many of the studies published to date have mostly been taught reading by more eclectic approaches utilising both whole-word recognition and grapheme-phoneme-conversion elements (e.g., Frith et al., 1998; Goswami et al., 1997, 2001; Wimmer & Goswami, 1994).

Two notable exceptions have compared word reading proficiency between English and a more consistent orthography (German/Welsh) in groups of children being taught to read with comparable methods: Landerl (2000) compared word and nonword reading in German and English children of grades 2, 3, and 4 with the groups of one English school having received the standard eclectic approach of reading instruction, and the groups of a different English school having received a straight forward phonics approach similar to that of the German children. Ellis and Hooper (2001) compared word reading in Welsh children learning to read in English with Welsh children learning to read in Welsh, both of whom received the mixed approach of phonics and whole-word teaching typically prevalent in English-speaking countries. Importantly, the results of these two studies using matched approaches to reading instruction were highly consistent with studies using orthography-specific approaches in showing a superior performance of the children learning to read in a consistent orthography.

Consistent with the notion that the use of both whole-word and phonics elements in teaching to read the English orthography is perhaps an adequate response of dealing with the exceptional inconsistency of English orthography, the methodological limitation regarding different instructional methods to teaching reading which apply to the vast majority of cross-linguistic studies is perhaps inevitable.

1.3.5.3 Cultural background

A last factor on the participant side which may influence reading development is cultural background. For cross-linguistic studies comparing reading across different countries, this factor is nearly impossible to control for. Some studies simply refrain from reporting what cultural background their participants had (e.g., Goswami et al., 1998, 2003), some note that their participants came from a variety of socioeconomic backgrounds (e.g., Patel, Snowling, & de Jong, 2004), some report that the schools participants were recruited from were not located in an underprivileged area (e.g., Aro & Wimmer, 2003; Seymour et al., 2003), while others state that they recruited their participants from comparable surroundings in the respective countries, i.e., from suburbs of a larger city, or from a metropolitan area in both countries (e.g., Caravolas et al., 2012; Frith et al., 1998; Ziegler et al., 2003). The study with Welsh children mentioned above (Ellis & Hooper, 2001) is exceptional in that it provided a natural control for cultural background: the children participating in their study all came from the same area and were thus highly similar comparable in terms of cultural background.

However, the fact that word reading ability was much higher in the children learning to read the consistent Welsh orthography than in children learning to read the inconsistent English orthography speaks against a large influence of cultural factors in cross-linguistic studies.

1.3.5.4 Reading material

Quite naturally, *what* is being read will directly influence *how well* it will be read. Hence, in cross-linguistic research it is crucial that the stimulus material is comparable for participants of the different orthographies. The matching of reading material poses a particular difficulty when the languages to be studied are very different. Thus, English and Finnish would be an ideal comparison from the point of differing orthographic consistency, but the two languages are quite unrelated from a linguistic point of view. To give an example, Finnish words tend to be polysyllabic and are thus typically much longer than English words, while at the same time, unlike for English, Finnish syllables are simple and rarely include complex consonant clusters (Aro, 2006). When comparing reading development between orthographies with differing language structure, the construction of comparable word material is less of a problem than the construction of comparable nonword material. Words of comparable length and frequency can thus be considered to be a sufficiently good match (e.g., Ellis & Hooper, 2001; Seymour et al., 2003), even though it would of course be preferable to use identical or close-to identical stimuli (i.e., cognates) across orthographies as many comparative studies of German and English have done (e.g., Frith et al., 1998; Landerl et al., 1997; Landerl & Wimmer, 2000; Wimmer & Goswami, 1994; Ziegler et al., 2003). However, nonwords are more complicated to match between rather distant languages. Thus, Landerl (2006) argues that the use of identical nonwords in languages as different as English and Turkish will be disadvantageous to one or the other group. The alternative to using identical nonwords for different orthographies is to construct different nonwords for different orthographies. Quite obviously, this approach is equally, if not more problematic. For instance, the study of Seymour and colleagues used mono- and bisyllabic nonwords consisting of dominant and consistent grapheme-phoneme-correspondences in each language, all with simple syllable structure. This unusually simple syllable structure, together with the fact that all nonwords were printed in lower case, resulted in a number of the German nonwords, such as *uki*, *mipu*, or *jefi*, to appear rather “un-wordlike”.

1.4 Additional background on orthography and reading acquisition in English and German

1.4.1 English orthography and reading acquisition in English

As mentioned at the outset of this chapter, alphabetic orthographies differ in the consistency of their letter-sound relationships, and English is exceptionally inconsistent in terms of both grapheme-phoneme-correspondences, and in terms of phoneme-grapheme-correspondences. The reasons for this are numerous and are partly to do with the large number of loan words which have kept their original spelling, but are mainly to do with the fact that generally, modern English spelling does not reflect the sound changes that have occurred since the late 15th century (English orthography, http://en.wikipedia.org, n.d.). The inconsistency of grapheme-phoneme correspondences of English causes a substantial delay in initial reading acquisition, which is taken to result from two sources: first, the difficulty in acquiring and applying notoriously unreliable grapheme-phoneme-correspondence rules, and second, the need to develop and apply different reading strategies targeting processing units of differing sizes.

In the United Kingdom, children enter primary school at the age of five and are thus a year or over a year younger than children in other European countries when they start formal reading education (Seymour et al., 2003). However, formal reading education in England starts even before primary school, in reception year at the age of four (Blythe & Joseph, 2011; Caravolas et al., 2012).

There has been a long standing debate about which method of teaching children to read may be most appropriate to cater for the needs of the inconsistent English orthography, and no generally accepted conclusion has been reached as to yet (e.g., Wyse & Goswami, 2008). Providing detailed information on the different methods of teaching reading would go beyond the scope of this section. It may be sufficient to briefly outline the two major competing methods of reading instruction, whole language and phonics. The whole language approach trains children to focus on words, sentences, and paragraphs as a whole rather than letters. A subtype of whole language is the look-and-say method, or sight word method, which involves children to learn by heart a sight vocabulary of 50-100 words, to recognise the words they have memorised, and to guess unknown words from context or initial letter clues. In contrast, phonics is a teaching approach emphasising grapheme-phoneme-correspondences, specific rules and their use in reading and spelling. The goal is to enable beginning readers to decode words by sounding them out (Learning to read, http://en.wikipedia.org, n.d.; Phonics, http://en.wikipedia.org, n.d.). While reading instruction in the UK has long included a mixed approach combining whole-word and phonics methods (e.g., Frith et al., 1998; Landerl, 2000), the USA’s National Reading Panel regime (National Institute of Child Health and Human Development [NICHD], 2000), the Australian (Australian Government, Department of Education Science and Training, 2005), and the UK government (the National Literacy Strategy; Department for Education & Employment, 1998) have all recently recommended that a systematic phonics instruction be used in primary school education. Since 2006, the more specific recommendation in the UK is to use synthetic phonics (Rose, 2006).

The controversy about the ideal teaching method is closely linked to the debate on the question as to whether early reading in English is based primarily on small units, such as phonemes, or on larger units, such as onsets and rimes, or whole words (Aro, 2006; Seymour & Duncan, 1997). Importantly, as outlined above, it seems that the inconsistent English orthography forces readers to apply both small and large unit decoding strategies from the very start of reading development.

Perhaps because initial reading acquisition is no easy task in the inconsistent English orthography, there has traditionally been a great interest in factors predictive of basic word reading ability. There are three important skills which have consistently been shown to predict progress in reading development in English, namely, letter knowledge, awareness of phonemes in spoken words, and rapid automatised naming of visual stimuli (Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013). Since the research presented here focuses on actual word and sentence processing rather than their precursors or predictors, this field of research will not be discussed in more detail.

1.4.2 German orthography and reading acquisition in German

In contrast to English orthography, German orthography is characterised by highly consistent grapheme-phoneme-correspondences. In consequence, knowledge of all grapheme-phoneme-correspondence rules generally permits the correct pronunciation of the majority of German words (exception words are mostly loan words of foreign origin such as *Computer*, imported from English, or *Restaurant*, imported from French). In the reading direction, German can therefore be considered as a shallow orthography. Since German, just like English, adheres to the principle of morpheme consistency, German is less consistent with regards to phoneme-grapheme-correspondences. However, in contrast to English orthography, the morphological principle never overrides the phonological principle in German orthography. Further, the umlaut graphemes help to retain both morpheme and phoneme consistency at the same time (e.g., the plural of *Wald* is spelled *Wälder*, not *Welder*) (Landerl, in press). The fact that most phonemes have two or more grapheme correspondences renders spelling more complex than reading. Particularly inconsistent is the orthographic marking of vowel length – for instance, a long vowel can be marked by vowel doubling (e.g., *See*), by inverting a “silent h” after the vowel (e.g., *Zeh*), or it may not be orthographically marked at all (e.g., *Gel*) (Landerl, 2006). In the spelling direction, German is therefore considered more of a deep orthography (not as deep as English, though).

Formal reading instruction in German-speaking countries mostly only starts from school entry, which is typically at six years of age. Reading instruction at most schools is phonics based, and after the first one or two years in primary school, the instruction focus is mainly on spelling (Landerl, in press). In contrast to the long lasting debate on which method of reading instruction is optimal in English-speaking countries, phonics is the established and generally accepted approach to teaching reading in German-speaking countries. Given the highly consistent grapheme-phoneme-correspondences in German, this approach is indeed the most suitable (e.g., Landerl, 2000).

The high consistency of German orthography on the level of grapheme-phoneme-correspondences facilitates insight into the alphabetic principle. Systematic grapheme-phoneme based decoding is a highly reliable default reading strategy in the consistent German orthography, and it also supports the build-up of word-specific orthographic representations (Share, 1995). These orthographic representations in an internal lexicon are important for two reasons – first, as a basis for the fast and direct word recognition which is characteristic of skilled reading and second, as a basis for orthographically correct spelling. The combination of a phonologically transparent orthography and a systematic phonics teaching approach provide a solid basis for successful initial reading acquisition in German.

Thus, word decoding ability develops in the course of the first few months of reading instruction with reading accuracy reaching a high level even for unfamiliar words at the end of first grade (Landerl & Wimmer, 2008; Seymour et al., 2003; Wimmer & Hummer, 1990). The main challenge in learning to read German results from its rather complex grammatical language structure. Due to inflection, most words comprise more than one syllable and a considerable number of words comprise complex consonant clusters. Since beginning readers rely mainly on systematic sublexical decoding, it is the processing efficiency for long and morphologically complex words which differentiates between skilled and less skilled readers (Landerl, in press).

There has been some debate about whether the three core abilities known to predict word reading ability in English are equally important in more consistent alphabetic orthographies (e.g., Share, 2008). Thus, phonological awareness appears to predict spelling rather than reading ability in German (e.g., Landerl & Wimmer, 2008; Moll, Fussenegger, Willburger, & Landerl, 2009), and direct cross-linguistic studies have shown phonological awareness to be less important in predicting reading development for German than for English (e.g., Mann & Wimmer, 2002). The most relevant predictor of word and nonword reading ability in German seems to be rapid automatised naming (Moll et al., 2009). More recently however, a longitudinal study comparing predictors of reading development between English and the consistent orthographies of Spanish and Czech has shown phoneme awareness, letter-sound knowledge, and rapid automatised naming to be of comparable importance as predictors of variations in reading development across orthographies of differing depth (Caravolas et al., 2013).

1.5 Eye tracking in reading research

1.5.1 Advantages of eye tracking in reading research

In the past, much has been learned about word recognition processes and their development in studies using naming or lexical decision tasks. However, such tasks present single words in isolation, and the generalizability of such results to more natural reading situations has rightly been questioned (e.g., Juhasz & Rayner, 2003). An attractive alternative to collecting naming or response latencies in naming or lexical decision tasks is to present word lists, sentences, or complete texts and record eye movements while participants read in a quite natural and undisturbed way. Other advantages of eye tracking are that data is collected online, and that no secondary tasks need to be employed (Rayner, Sereno, Morris, Schmauder, & Clifton Jr., 1989). It is furthermore possible to study the processing of specified target words and of meaningful texts or sentences at the same time (it should be noted though that the analysis of specific target words presented in meaningful context bears a couple of drawbacks. First, the context is likely to support word recognition, and second, properties of the words immediately preceding or following the target word may influence word recognition of the target). One of the most central benefits of eye tracking is that unlike tasks such as naming or lexical decision that provide only a single reaction time measure, eye movement data give the researcher a wealth of different measures which can be used to see how a variable's impact unfolds over time (Cunnings & Clahsen, 2007; Juhasz & Rayner, 2003).

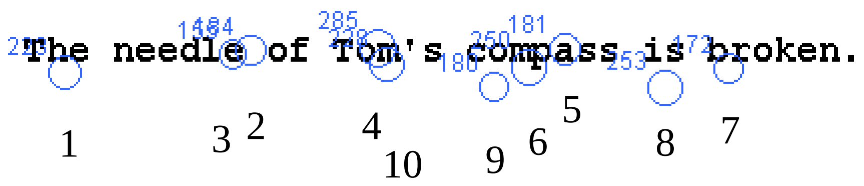
The history of eye tracking research dates back until the beginning of the last century where many basic facts about eye movements were discovered (Rayner, 1998). It is however only since the 1970s that researchers have actually used eye movement behaviour to try and infer underlying cognitive processes. Direction and sequence of eye movements are generally taken to indicate which word or word part is currently being processed, and the fixation durations are taken to indicate the mental effort devoted to word recognition (e.g., Radach & Kennedy, 2004). It is not the case, though, that processing strictly coincides with the current fixation at any point in time. Thus, many studies have shown that there may be a substantial amount of pre-processing of word *n+1* during the current fixation of word *n* (parafoveal-on-foveal, or parafoveal preview effects, e.g., Kennedy & Pynte, 2005; Kliegl, Nuthmann, & Engbert, 2006). Similarly, processing may spill over from one word to the next, and a difficult word can thus increase the fixation time of the following word (spillover effects, e.g., Kliegl et al., 2006; Rayner & Duffy, 1986). Importantly, the relationship between fixation positions and fixation durations and actual local processing is nevertheless strong enough to produce reliable effects when pooled over participants and items. In this sense, eye movement measures are considered to provide an extremely sensitive index of local processing load, and have thus contributed a great deal in developing and testing psycholinguistic hypotheses about the processing of written language (Radach & Kennedy, 2004).

1.5.2 Basic characteristics of eye movements in reading

Despite our subjective perception of a continuous reading process, reading is actually a succession of very fast, fairly well coordinated forward movements of the eyes (saccades), and longer periods of relative stability (fixations). Visual information intake only happens during fixations, and the saccades serve to bring into foveal view the word or word part we are planning to process next (Rayner, 1998). The majority of saccades are progressive saccades which are placed in the reading direction, and have amplitudes of a maximum of 20 letters (Radach, Günther, & Huestegge, 2012 report a mean amplitude of about six to eight letters for skilled readers of German; Rayner, 1998 reports a mean amplitude of seven to nine letters for skilled readers of English). There are also saccades going against the reading direction (so-called regressive saccades, or regressions), and their amplitudes are typically only about half that of progressive saccades. Such rather short (mostly intraword-) regressions either occur when the previous saccade was misplaced and landed too far ahead in the text, or when the reader has problems processing the currently fixated word (Rayner, 1998). Longer interword regressions occur when the reader is having trouble comprehending the text. As mentioned above, fixation durations are used as an online indicator of processing load, and typical fixation durations are reported to last about 200-250 ms (for skilled readers of English; Rayner, 1998), or 220-250 ms (for skilled readers of German; Radach et al., 2012). While the majority of the words in a text are fixated at least once during reading, a certain amount of words is skipped. The probability of fixating a word clearly increases with increasing word length, and word skipping therefore mostly occurs for very short, high-frequency function words, such as articles and personal pronouns.

1.5.3 Eye movement measures

As mentioned before, there is a wealth of different measures which can be extracted from eye movement data, and it depends on the research question which measures are most appropriate. Hyönä, Lorch, and Rinck (2003, Table 16.4) provide a comprehensive overview over commonly used eye movement measures on both the word- and the region-, or sentence level. An important concept which helps the understanding of the different measures is that of a “pass”. The first pass is the first encounter with a word or otherwise defined region of text before the eyes move away from it to either a previous or a subsequent word or region, while the second pass entails all reinspections occurring during the second encounter (Hyönä & Olson, 1995). Since there may be more than two passes for very difficult words or passages of text, measures of second- and third-pass processing are often combined in number of re-reading fixations and summed duration of re-reading time. Measures of first-pass processing are typically regarded as reflecting early word recognition processes, such as initial visual recognition and lexical access (e.g., first fixation duration), or word identification (e.g., gaze duration), and reinspection measures are regarded as reflecting late word recognition processes, such as the integration of a word into the sentence context (Kuperman & van Dyke, 2011). A special case is the total word reading time measure, or total viewing time, which combines both early and later processing stages, and which is taken to provide a general index of processing load (Cunnings & Clahsen, 2007), or general comprehension (Boston, Hale, Kliegl, Patil, & Vasishth, 2008). A comprehensive overview of first-pass and later measures of eye movements and the stages in parsing processes they are assumed to represent is provided by Boston and colleagues (2008; Tables 3 and 4). Figure 1 provides an illustration of early and late measures of single word recognition for the target word “compass” featuring in one the sentences from the current research project. The fixation pattern stems from one of the English children within the sample. The numbers below the sentence correspond to the order in which fixations occurred (in the case of the target word “compass”, these are fixations 5, 6, and 9).



Early measures:
First fixation duration: fixation 5
Gaze duration: fixations 5 + 6
Late measures:
Re-reading time: fixation 9
Total reading time: fixations 5 + 6 + 9

Figure 1. Early and late measures of processing time for the target word “compass”.

1.5.4 Eye movements and individual differences in reading skill

Importantly, eye movement behaviour is dependent on a number of variables, some relating to task demands, to text characteristics, and to characteristics of the reader. Thus, saccade amplitudes and fixation durations depend on whether reading mode is silent or oral with longer saccades and shorter fixation durations in silent reading (Rayner, 1998). While eye movement behaviour also differs as a function of text difficulty, the perception of whether a word or a text is difficult or not is obviously largely dependent on the level of individual reading skill. Such individual differences in reading ability translate into eye movement measures in a fairly straightforward fashion. There are two ways in which the influence of reading skill has typically been studied in this context.

One approach is to compare reading eye movements of dyslexic children with those of typically developing children of the same age or reading level. There is a large body of research on the differences in eye movements of dyslexic and age-matched control children when reading in different orthographies (e.g. English: Hyönä & Olson, 1995; Rayner, 1985; German: Hutzler & Wimmer, 2004; Trauzettel-Klosinski, Koitzsch, Dürrwächter, Sokolov, Reinhard, & Klosinski, 2010; Greek: Hatzidaki, Gianneli, Petrakis, Makaronas, & Aslanides, 2011; Italian: de Luca, Borrelli, Judica, Spinelli, & Zoccolotti, 2002; de Luca, di Pace, Judica, Spinelli, & Zoccolotti, 1999). Across orthographies, dyslexic children show a higher number of fixations, slightly longer fixation durations, and often also a higher number of regressions than their typically reading peers.

The other approach to study the influence of different reading skill on eye movement behaviour is to compare typically developing children of different ages cross-sectionally, or to compare the same group of typically developing children at different ages longitudinally. Findings have been largely comparable for English and more consistent orthographies in showing skill-related developmental differences (English: Blythe, Liversedge, Joseph, White, & Rayner, 2009; Buswell, 1922; McConkie et al., 1991; Rayner, 1985; Taylor, 1965; Finnish: Blythe, Häikiö, Bertram, Liversedge, & Hyönä, 2011; German: Huestegge, Radach, Corbic, & Huestegge, 2009).

Thus, as reading skill increases, fixation durations and sentence reading times decrease, as does the number of fixations. While the probability for making another fixation on a previously fixated word also decreases, the probability for skipping a word in first pass increases with increasing reading skill. With regards to number, or percentage of regressions[1], findings have been seemingly less consistent with some studies showing a slight *de*crease with increasing reading expertise (English: Buswell, 1922; Rayner, 1985), some showing a more constant level of regressions among children of different reading experience (English: Blythe et al, 2009; McConkie et al., 1991; Taylor, 1965; German: Huestegge et al., 2009), and one very recent study showing an *in*crease in the proportion of interword regressions from grade 1 to 5 (English: Vorstius, Radach, & Lonigan, 2014).

The reason for these inconsistent findings is probably that those studies showing constant regressions rates have included both intra- and interword regressions in their measures, and the study showing an increase has used only interword regressions. Radach and colleagues (2012) argue that the constant regression rate in many former studies is likely to result from an interesting trade-off involving a decrease in the number of intraword regressions, and a simultaneous increase in the number of interword regressions over the first few years of reading development. The number of intraword regressions is typically high in beginning readers because early reading is more careful than more advanced reading, and the number of interword regressions tends to increase in the course of reading development as more advanced reading also becomes more risky.

[1] The percentage of regressions denotes the percentage of regressive eye movements in relation to all eye movements (i.e., saccades and regressions).

1.5.5 A different pattern of eye movements in different types of dyslexics

As established in the previous section, many interesting studies have used eye tracking to investigate individual differences in reading skill by comparing groups of differing reading skill. A different approach was taken by Olson, Kliegl, Davidson, and Foltz in 1985 who analysed differences in reading style within a group of English-speaking dyslexic readers. Based on two main eye movement parameters, the respective percentage of interword regressions and word-skipping progressive saccades, they identified two distinct subgroups of dyslexic readers. The so-called plodders showed relatively few word skippings and few between-word regressions, and were described as moving forward along lines of text in a steady and continuous fashion with a high number of short forward saccades. In contrast, the eye movement pattern of the so-called explorers was characterised by a relatively higher number of word skippings, a higher number of regressions, and less word-to-word progressive movements.

This classification has later been suggested to fit the distinction between developmental surface and developmental phonological dyslexia as proposed by Castles and Coltheart (1993). The characteristic deficit in surface dyslexia being an impeded access to the orthographic lexicon, and in phonological dyslexia a difficulty in applying serial phonological recoding, surface dyslexia has been claimed to correspond to the plodders', and phonological dyslexia to the explorer's pattern of eye movements (de Luca et al., 1999). Rayner (1998) has suggested that skilled readers can be divided into similar styles. What makes this classical plodder/explorer distinction particularly interesting in the context of the present research project is the circumstance that the plodder reading style is in fact a small-unit approach, and the explorer reading style a large-unit approach to text processing.

2 Project Findings on Word and Sentence Processing in German and English Readers

2.1 Rationale for and design outline of the present research project

The aim of the research project presented here was to study the influence of orthographic consistency on eye movement behaviour in word and sentence processing. To achieve this, a sentence-reading task was devised and administered to German and English children and adults in an eye tracking experiment. This multi-dimensional experiment forms the basis for the entire research project and the three studies reported below.

As established in the first chapter, the cross-linguistic comparison of reading and reading development bears important methodological problems relating to the difficulty of matching children across countries with differing educational systems, matching reading material across orthographies, and to choosing an experimental task reflecting the natural process of reading while maintaining sufficient experimental control over the presented stimulus material.

2.1.1 Choice of orthographies

The choice of German as a consistent, and English as an inconsistent orthography was a fairly straightforward one, for both theoretical and practical reasons. While the two orthographies differ with respect to the consistency of their grapheme-phoneme-correspondences, they are highly similar in many other respects (Landerl, 2006): First, apart from the German umlaut letters, and the additional ß which do not existent in English, the two orthographies use the same set of letters. Second, both orthographies use multiletter graphemes (e.g., *sch* and *sh* as in *Schiff* and *ship*), and both orthographies use silent letters (e.g., the h in *Zahn* and the t in *ballet* are silent). Third, given the common Germanic origin of the two languages, the phonological structure of many of their words is highly comparable. Thus, both German and English are characterised by complex syllables with frequent consonant clusters in word initial and word final positions (e.g., *Optimist* and *optimist*). Fourth, there are a considerable number of words that are shared by both orthographies, so called cognates, with identical meaning, identical or highly similar spelling, and highly similar pronunciation (e.g., *Sommer* and *summer*). Finally, the two languages are not only comparable on the level of single words, but also on the level of syntactic structure (at least for main clauses) (Landerl, in press). Taking advantage of the tight connectedness of the two orthographies on both the single word, and the syntactic level, we constructed highly similar sentences (e.g., *David isst eine Kiwi und zwei Bananen – David is eating a kiwi and two bananas*; see Appendix A for the complete set of experimental sentences). The sentences' structure was kept simple so that children with only two years of reading instruction could be expected to read and understand the sentences. The English version of the sentences was checked for grammatical and linguistic correctness by native speakers of English (British and American).

2.1.2 Choice of reading material

Since the aim of the research project outlined here was to study small- and large-unit processing for both words and sentences, the experiment was a sentence-reading task with each sentence containing a target word manipulated for length and frequency.

A small-unit decoding strategy will inevitably result in longer processing times for longer as opposed to shorter words. Therefore, the length effect was chosen as a marker of small-unit sublexical processing. There were four different length categories, 3-4, 5-6, 7-8, and 9-10 letters for each level of word frequency. There were occasional cases in which targets did not have the exact same number of letters in the two orthographies. However, the occasional extra letter in one orthography was counterbalanced by an extra letter in another target of the same category in the other orthography (e.g., the German word *Maschine* has one more letter than the English word *machine*, but the German *Februar* has one letter less than the English *February*; both cognates belong to the same length category of 7-8 letter words).

Based on the notion that words occurring with high frequency in a language will result in faster lexical activation than words occurring with low frequency (Just & Carpenter, 1980), the frequency effect was chosen as an index of large-unit lexical processing. An alternative would have been to use the body neighbourhood effect, but there were two reasons speaking against this option. First, as outlined in the first chapter, a genuine match between German and English on body neighbourhood size is difficult to achieve. It appears that English words tend to have more neighbours than German words, a circumstance which is likely to contribute to greater word recognition benefits in English than in German. Second, we wanted to also include polysyllabic items because we consider the exclusive use of monosyllabic words to be representative for neither German nor English. However, for longer polysyllabic words it is nearly impossible to find cognates with a large body neighbourhood.

In order to be able to properly investigate the process of phonological recoding, nonwords were used alongside high- and low frequency targets. Since introducing a lexicality factor (words vs. nonwords) next to length and frequency would have rendered analyses overcomplicated, word frequency was used as a unifying concept and included both lexicality and frequency (for a similar procedure to ours, see Chaffin, Morris, & Seely, 2001 and Lowell & Morris, 2014). The word frequency factor thus included high-frequency words, low-frequency words, and word-like nonwords. High- and low-frequency targets were cognates with nearly identical frequency counts, number of letters, number of syllables, spelling and meaning in the two orthographies (for more details on target word matching, see Table 2 of study 2). Frequencies of targets were taken from the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995) with the frequency counts based on written and spoken sources.

Unlike the items used in the cross-linguistic studies by Ziegler and colleagues (2001, 2003), target words of the present study were not restricted to regular words (i.e., words adhering to the grapheme-phoneme-correspondence rules of the respective orthographies). Out of the German high-and low-frequency words used as targets, only one out of 48 was irregular (*Volleyball*), while in English, five were irregular (i.e., *cow*, *four*, *minute*, *February*, *blueberry*). One might argue that this was an unfair disadvantage for the English readers, but there are two reasons why we believe that this argument is a weak one. First, it may not be fair that there are so many irregular words in English, but it is a fact. The exclusion of irregular words from cross-linguistic studies is thus not a genuine reflection of English orthographic reality. Second, for the crucial eye movement measures reported in study 2, items with regularisation errors (i.e., the pronunciation of an irregular word as if it was regular; e.g., pronouncing the irregular word *pint* to rhyme with the regular word *mint*) were included.

In order to be sure that the identical nonwords to be created included typical sound sequences for both orthographies, nonword targets were derived from low-frequency cognates which were not used in the experiment. Low-frequency cognates were chosen as a basis for creating the nonwords over high-frequency cognates because the nonwords were intended to represent the far end of a dimension of visual familiarity decreasing from high-frequency words, low-frequency words, through nonwords. Concretely, the nonword targets were created by exchanging a part of the original low-frequency word with a letter sequence preserving the original syllable structure (e.g. the nonword *Jeto* was derived from *Veto* / *veto* – for a complete list of the nonwords used and the low-frequency cognates they were derived from, see Appendix B). Additionally, care was taken for mean bigram frequency for the resulting nonwords to be the comparable between both orthographies (see Table 2 in study 2). In one case, for bigram frequency to be comparable between German and English, a single letter was adapted to conform to the more typical letter sequence of the two orthographies (the nonword pair *Akromul* - *acromul* was spelled with a *k* in German, but with a *c* in English). Apart from this one case, nonwords were identical for the two orthographies and mean bigram frequency was comparable. In order for the nonwords to blend into the sentences neatly, they were mostly used as proper names (e.g., *Die beste Medizin ist Nemtratect, sagt Doktor Braun* / *The best medicine is Nemtratect, says Doctor Brown*).

Importantly, a lenient scoring regime was used for nonwords in English. With grapheme-phoneme-correspondence rules being particularly inconsistent in the case of vowels (Landerl, 2006), a number of possible pronunciation for vowels were accepted (e.g., in the nonword *Nibra*, the pronunciation of the first vowel was accepted to sound as in *pig*, *Tina*, and *night*; and the second vowel was accepted to sound as in *bra*, *aroma*; see Appendix C for acceptable pronunciations for all nonwords).

2.1.3 Participant groups and participant matching

The majority of previous studies on reading development in consistent and inconsistent orthographies have naturally been done with children. The present research project additionally included young adults since this allows considering the end-point of development, too. While cross-linguistic processing differences were expected mainly in developing readers, we were interested to establish whether such differences may persist to some degree in skilled readers.

As study 1 did not involve a cross-linguistic comparison, but investigated the transition in predominant reading strategy from principal reliance on serial sublexical processing to more lexical processing in the consistent German orthography only, participant recruiting was straightforward. Children of grades 2, 3, and 4 were invited to take part in the experiment at different primary schools in Tübingen, Southern Germany, as were undergraduates studying Psychology at the University of Tübingen. All of the German participants were included in analyses for study 1, provided that their performance in the reading fluency test of the SLRT II (Lese- und Rechtschreibtest; Moll & Landerl, 2010) was average or above average for both the word and the nonword reading subtest.

For studies 2 and 3, which involved cross-linguistic comparisons of word and sentence processing between German and English, additional English-speaking participants were recruited in York, Northern England, were two of the project partners were located at the time. Given the substantial delay in initial reading development of English-speaking children described in chapter one, it was unrealistic that English grade 2 children would be able to read the sentences with sufficient ease. Therefore, the recruitment of children from York was confined to children who had received at least three years of school education (grades 3, 4, 5, and 6). As discussed in the previous chapter, matching children from different countries is difficult when formal reading education starts at a different age, which is the case with Germany and England. The approach adopted in studies 2 and 3 was a reading level match between children of the two orthographies. Previous cross-linguistic studies have tried to achieve a reading level match based on the children's performance on standardised reading tests (e.g., Goswami et al., 2001). However, as mentioned in chapter one, this approach is questionable since the reading tests are typically quite different, and are not designed for this purpose. The reading level match applied in studies 2 and 3 was based on mean processing time for the easiest category of target words. Children and adults of the two orthographies were thus matched on their gaze duration for the identical short, high-frequency words used in the experiment itself (i.e., *Kuh/cow*, *rot/red*, *Tee/tea*, *vier/four*, *Mond/moon*, *Bier/beer*, *Musik/music*, *sieben/seven*, *Hotel/hotel*, *Mitte/middle*) (for detailed information on individual participant matching, see Appendix D). As evident in Table 1 of study 2, the reading level match meant that English children were slightly older, and had received formal education for a significantly longer amount of time than their German matches.

2.1.4 Experimental paradigm

A number of previous studies have used German-English cognates and presented them as single items in naming or lexical decision tasks (e.g., Goswami et al., 2001; Landerl et al., 1997; Ziegler et al., 2001, 2003) or as constituents of a list in a continuous reading task (e.g., Frith et al., 1998; Landerl, 2000; Wimmer & Goswami, 1994). While such experimental reading paradigms do permit maximum control over target word characteristics, they do not reflect the natural reading process for two reasons. First, the presentation of isolated targets, and even the presentation of different target words in a row or separate lines, is very different from natural text or passage reading because no context is provided. In the present experimental paradigm, target words were therefore embedded in meaningful sentences. Second, the experimental reading tasks used in most previous research often involved processes other than reading (e.g., button pressing, answering comprehension questions). Eye tracking methodology was used because it allowed for online recording of word recognition processes as they occurred without interference from additional tasks. The sentence-reading experiment was programmed in such a way that participants controlled the course of the experiment with their eye movements themselves: to start a trial, participants fixated on a small smiley on the left side of the screen. When they had completed reading a sentence, they fixated a little cross at the lower right corner of the screen, and the next smiley appeared.

While the analysis of eye movement behaviour is an established means of inferring moment-to-moment cognitive processes underlying reading (Just & Carpenter, 1980; Rayner, 1998), it has not been exploited in direct cross-linguistic research before. Apart from allowing the recording of word and sentence processing in an online fashion without having to use additional tasks, the other central benefit of eye tracking is that it provides an abundance of different measures. Taken together, such different measures provide valuable information on the time-course of word or sentence processing. The three studies attached below report different eye movement measures depending on the respective research focus.

2.1.5 Reading mode

As noted in chapter one, reading accuracy has often been reported to be much lower for English children than for children learning to read in a consistent orthography, at least in the first few years of reading acquisition (e.g., Frith et al., 1998; Seymour et al., 2003). Given that our reading material contained nonwords of considerable length and difficulty (e.g., *Nemtratect*, *Jempromisp*), it was important to be able to monitor reading accuracy. Reading mode was thus aloud. Admittedly even developing readers of grades 2 or 3 typically read silently rather than orally, and the reading process as measured by the present sentence-reading task was therefore not as natural as it could have been. Another disadvantage of reading aloud is that oral articulation may lead to more head movements than silent reading. Excessive head movements are undesirable during eye tracking because they lessen recording accuracy.

However, the present experiment was programmed in such a way that a re-calibration was required in cases where the position of the gaze had shifted. In total, we considered the advantages of oral reading to outnumber the disadvantages. It was crucial to be able to exclude misread items from analyses because including misread items may distort the data. There are two directions in which such a distortion is conceivable: either, misread items were not analysed thoroughly enough. The recorded processing time would then underestimate the actual processing needed. The alternative scenario is that misread items are read again and again. In this case, the recorded inflated processing time would indeed reflect the actual processing difficulty. However, there is no straightforward way to distinguish between the two possibilities. The methodologically sound way is therefore to tell participants to read carefully and avoid making mistakes (which we did), and exclude misread items afterwards.

As outlined in chapter one, there are differences between silent and oral reading (e.g., Radach et al., 2012; Rayner, 1998; Vorstius et al., 2014). Importantly however, it has also been noted that both reading modes represent a similar cognitive processes (Juel & Holmes, 1981).

The following sections will very briefly state the research questions and the main findings of the three studies. The findings of study 1 will be reported in a little more detail because the published article contained an error in the data. Even though reanalyses with corrected data largely confirmed all relevant findings, for the sake of completeness, corrected results are stated below. A corrigendum to the original article has been submitted and accepted by the journal.

2.2 The transition from sublexical to lexical processing in German orthography

A first study focused on investigating the development of word recognition processes in the consistent German orthography (Rau et al., 2014; publication 1 at the end of this document). Specifically, it aimed to investigate the transition in predominant reading strategy from principal reliance on serial sublexical processing to more lexical processing, which is considered characteristic of reading development in consistent orthographies. This was accomplished by comparing the length effect for words of differing visual familiarity between typically developing readers of different age and proficiency levels. There were four different length categories, but since this would have rendered analyses too complex, the 3-4 and 5-6, as well as the 7-8 and the 9-10 letter categories were collapsed to simply make up short (3-6) and long (7-10) items. The main dependent variable to reflect word recognition was gaze duration (i.e., the sum of all fixation durations during first-pass reading of the target word), because it is assumed to encompass all stages of word recognition including meaning activation (Juhasz & Rayner, 2003; Reichle, Pollatsek, Fisher, & Rayner, 1998). Furthermore, the study also reported *z*-transformed gaze durations in order to control for possible effects of overadditivity.

As early reading in consistent orthographies is mainly based on sublexical decoding, we expected a strong reliance on serial phonological recoding in the least experienced group of readers quite independent of word familiarity. For the more advanced primary school children of grades 3 and 4, we expected the degree of phonological recoding to continually decline with increasing word familiarity, because word recognition should become increasingly lexicalised in the course of reading development. For skilled adults, we expected to find only little, if any, evidence for sublexical recoding in words, which should be read by direct lexical access. In contrast, evidence for serial processing should be evident in nonword reading.

Unfortunately, there was an error in the data on which analyses in the published paper were based. The correct values are displayed in Figure 2. Importantly, results do not change substantially when reanalysing the data, and neither does the overall line of argument presented in the paper. However, since details in the data differ slightly from those reported in the published paper, the now-correct analyses are reported below.

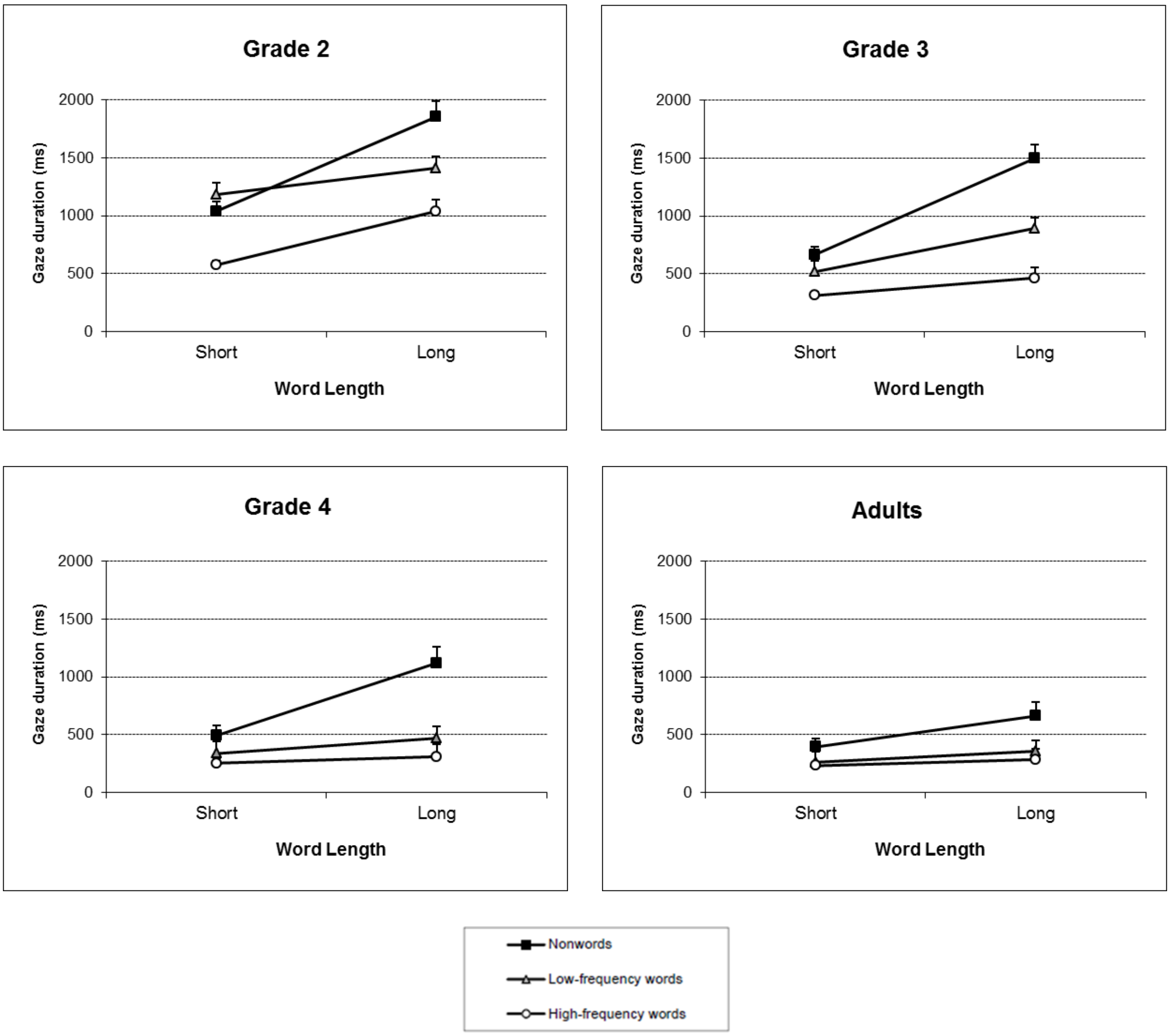


Figure 2. Word length effects as a function of word familiarity for readers of differing reading experience. *Note*. Bars indicate standard errors.

For all groups of readers, for both gaze duration and *z*-transformed gaze duration, the size of the word length effect differed as a function of word familiarity. This interaction was less pronounced in the least experienced group of readers in terms of effect size (an updated version of the values published in the article is provided in Table 1).

Table 1

Mean Length Effects for Each Level of Familiarity (SD) and Statistical Characteristics of the Familiarity x Length Interaction for the Four Groups

Group	LE HF words	LE LF words	LE nonwords	*F*	*p*	η_p^2
			Gaze duration (ms)			
Grade 2	464 (649)	229 (483)	813 (417)	5.38	.01	.29
Grade 3	153 (126)	376 (207)	833 (423)	35.85	< .001	.68
Grade 4	57 (64)	133 (132)	627 (228)	56.37	< .001	.82
Adults	54 (67)	96 (59)	268 (90)	41.20	< .001	.71
			Standardised gaze duration (*z*-values)			
Grade 2	.40 (.50)	.28 (.44)	.79 (.33)	5.58	.01	.30
Grade 3	.20 (.13)	.50 (.27)	1.03 (.34)	51.42	< .001	.75
Grade 4	.09 (.11)	.21 (.19)	1.04 (.33)	61.99	< .001	.84
Adults	.18 (.20)	.30 (.18)	.85 (.32)	43.41	< .001	.72

Note. LE = length effect; HF = high-frequency; LF = low-frequency.

There are two ways of figuring out this rather complex interplay of word length, word familiarity, and reading experience. The first is to compare effects of word length between words of differing familiarity for each group of readers separately. This is the viewpoint for which concrete hypotheses were formulated, and results met expectations very well. Consistent over analyses on both raw and standardised gaze duration, length effects for words of differing familiarity were largely comparable for the least experienced readers (except that the length effect was larger for nonwords than for low-frequency words), but increased with decreasing word familiarity for the more experienced readers.

The second is to compare the development of the length effect between the four groups separately for each word type. Since one would expect less experienced readers to rely more strongly on the process of phonological recoding than more experienced readers (at least in the case of reasonably familiar words), length effects should decrease continually from grade 2 through adults. In the actual publication, we had not put forward such concrete hypotheses, but were interested to uncover possible effects of overadditivity in this between-group comparison. There were indeed differences between analyses on raw versus standardised gaze duration (not shown in Table 1). While the length effect for high-frequency words decreased quite clearly with increasing reading experience for raw gaze duration, this decrease was much less clear when considering standardised gaze duration. Thus, while the length effect was numerically larger in grade 2 children than in all other groups, it was only significantly larger than in grade 4 children for *z*-transformed gaze duration. For low-frequency words, length effects were mostly comparable between groups for both raw and standardised gaze duration (except for a marginally larger effect in grade 3 than in grade 4 readers for gaze duration, and a larger effect in grade 3 children than in adults for *z*-transformed gaze duration). However, an interesting difference between analyses using raw and standardised values was again seen for nonwords. For raw gaze duration, the length effect was comparable between children of all grades, but less pronounced in adults than in children. In contrast to this, nonword length effects did not differ between groups for standardised gaze duration.

Findings from study 1 suggest German grade 2 children to apply serial sublexical decoding as a default reading strategy to most items, and more experienced readers to increasingly rely on direct lexical access in word recognition. Similar to a number of previous studies (e.g., Martens & de Jong, 2006; Zoccolotti et al., 2005, 2009), we used the word length effect as an indicator of serial sublexical decoding. It is important to note in this context that the developmental decrease in length effect, other than reflecting a decreasing importance of serial sublexical processing in favour of more parallel lexical processing, could be reflective of an increasing efficiency in serial sublexical processing (cf. Whitney & Cornelissen, 2005). Indeed, the developmental comparison of the length effect as a function of word type provided evidence supportive of this idea: in the analyses using standardised processing times, the length effects for high-frequency words and nonwords were largely comparable for readers of differing reading experience. The decrease in these length effects in the analyses using raw processing times is thus likely to be caused by such an increase in efficiency of serial sublexical processing. However, it should go without saying that the two processes are not mutually exclusive: the increase in efficiency of serial sublexical recoding and the transition from predominant use of this sublexical reading strategy to predominant use of lexical word recognition are likely to develop in parallel.

2.3 The time-course of cross-linguistic differences in word processing

The second study investigated the time-course of cross-linguistic differences in word recognition by comparing indicators of small-unit processing (length effects) and large-unit processing (frequency effects) in child and adult readers of the consistent German and the inconsistent English orthography (Rau, Moll, Snowling, & Landerl, 2015; publication 2 at the end of this document). In order to be able to localise possible cross-linguistic effects within the time-course of word recognition, study 2 reports four different measures of target word processing time: first fixation duration, gaze duration, re-reading time, and total reading time.

Following the predictions posited by psycholinguistic grain size theory (Ziegler & Goswami, 2005), we expected small-unit bottom-up processing to be more prevalent in the German readers, and large-unit top-down processing to be more prevalent in the English readers. Since skilled readers of both orthographies should rely on large units when reading, we expected cross-linguistic differences to be more pronounced in children than in adults. We expected cross-linguistic differences in word processing to be most pronounced for the central measure of word processing, gaze duration.

A first interesting finding concerned the relationship between phonological recoding ability and automatic lexical recognition in the four groups of participants. The standardised reading tests used for establishing average reading skill consisted of a nonword list and a word list for both orthographies, with nonword reading reflecting phonological recoding ability, and word reading reflecting lexical word recognition ability. In German participants, the association between both subtests of the reading fluency test (SLRT II, Lese- und Rechtschreibtest II; Moll & Landerl, 2010) was strong for both children ($r = .85$, $p < .001$) and adults ($r = .65$, $p < .01$). Interestingly, there was a clear correlation between the two subtests of the reading fluency test (TOWRE, Test of Word Reading Efficiency; Torgesen, Wagner, & Rashotte, 1999) in English children ($r = .79$, $p < .001$), but no such association between phonological recoding and lexical word recognition in English adults ($r = .02$, n.s.).

Consistent with previous research (e.g., Frith et al., 1998; Landerl, 2000; Seymour et al., 2003), reading accuracy was higher in German than in English children. The difference was particularly pronounced for low-frequency words and nonwords with English children typically making twice as many reading errors than German children (see Table 2). Note that nonword reading accuracy was low in English readers despite lenient scoring. An unexpected finding was that, for the longest length category, the error rate in English adults was substantially higher than in German adults. Previous studies have not reported differences in adult reading accuracy between consistent and inconsistent orthographies because accuracy is typically not considered a relevant measure in skilled reading.

Table 2

Error Percentages for the Initial Twelve Conditions and the Four Groups.

		Children		Adults	
		German	English	German	English
Familiarity	Length				
High-freq words	3 - 4 letters	1.3	0.0	0.0	0.0
	5 - 6 letters	0.0	0.0	0.0	0.0
	7 - 8 letters	1.3	3.2	1.0	0.0
	9-10 letters	3.2	7.7	0.0	0.0
Low-freq words	3 - 4 letters	7.1	13.5	0.0	1.0
	5 - 6 letters	0.6	13.5	0.0	0.0
	7 - 8 letters	7.1	21.2	1.0	0.0
	9-10 letters	8.3	18.6	0.0	0.0
Nonwords	3 - 4 letters	3.2	12.2	0.0	2.0
	5 - 6 letters	8.3	21.8	2.9	2.0
	7 - 8 letters	12.8	29.5	1.0	3.9
	9-10 letters	25.0	55.1	9.8	25.5

Note. High-freq = high-frequency; low-freq = low-frequency.

For the reasons explained in chapter one, we planned to exclude misread items from analyses, and to make analyses more easily interpretable, we planned to collapse the 3-4 and 5-6 letter, and the 7-8 and 9-10 letter length categories into two categories, short (3-6 letters) and long (7-10 letters) words. However, since this procedure would have implied substantial data loss in the case of English children, we decided to adapt the procedure in two ways: First, reading errors were classified into minor and major errors, with only major errors being excluded from eye movement analyses (for more details on the minor vs major reading error distinction and examples, see Appendix E). Second, we completely excluded the 9-10 letter length category and rearranged the remaining categories so that 3-5 letter words formed the short, and 6-8 letter words formed the long condition.

The first important finding regarding word processing time was that the time-course of word recognition differed in an interesting way between children of the two orthographies.

While the length effect just tended to be stronger in German than in English children for first fixation duration, it was clearly more pronounced in German children for gaze duration. However, and importantly, the reverse pattern was evident in re-reading times – in this late measure of processing time, English children showed a stronger length effect for low-frequency words and nonwords. In turn, in the combined measure of total reading time, there was no difference between children of the two orthographies.

These cross-linguistic differences in the time-course of word recognition appeared despite the fact that the German and English children were highly comparable with regards to reading skill: they were matched on basic word reading ability, and their total word reading times were equally comparable. However, they achieved this equal outcome in total word processing effort quite differently: While German children relied on small-unit processing early in word recognition, English children applied small-unit decoding only upon re-reading – possibly when experiencing difficulties integrating an unfamiliar word into the sentence context.

For adults, word processing times were more comparable between the two orthographies for high- and low-frequency words, but differed in the case of nonwords. Thus, for both first-pass and total reading time, processing times for nonwords were longer in English than in German adults. The greater difficulty English adults were facing in the processing of nonwords fits the higher proportion of errors they were committing in processing the later excluded longest category of 9-10 letter nonwords. Since processing times of German and English adults did not differ for neither high-frequency, nor for low-frequency words throughout the whole time-course of word recognition, and adult cross-linguistic differences were thus restricted nonword processing, it may be more appropriate to speak of a cross-linguistic lexicality rather than a frequency effect here.

We were astonished to find a cross-linguistic difference in total processing time for nonwords for adults, but not for children. This may reflect the 17-month advantage English children had over German children as far as length of formal reading instruction is concerned, or the increased efforts regarding systematic teaching of phonics in British schools in recent years, or a combination of both these factors.

2.4 Same same, but different: word and sentence reading in German and English

After establishing cross-linguistic differences in word processing in study 2, study 3 aimed to investigate whether orthographic consistency would exert an influence beyond the level of single word recognition (Rau, Moll, Snowling, & Landerl, submitted; submitted manuscript at the end of this document). To this end, study 3 compared the matched German and English children and adults of study 2 with regards to a number of both local word-based and global sentence-based eye tracking parameters. Sentences containing a nonword were excluded since we were interested in studying the processing of typical sentences.

There have been attempts to compare global text processing between consistent and inconsistent orthographies (e.g., Dürrwächter, Sokolov, Reinhard, Klosinski, & Trauzettel-Klosinski, 2010; Hutzler & Wimmer, 2004), but such previous comparisons were made across different studies which were comparable neither in terms of reading material, nor in terms of participant characteristics. Such comparisons have suggested that orthographic inconsistency causes English readers to regress more often than readers of German or Italian. While this is certainly a plausible argument, the informative value of comparisons across studies is limited.

Generally speaking, we expected a more small-unit bottom-up processing pattern in readers of German, and a more large-unit top-down processing pattern in readers of English, for both word and sentence processing. Particularly, in local word processing, we expected higher first-pass reading times, a higher number of first-pass fixations, and a higher probability to fixate a word in the first pass in the German readers. At the same time, we expected higher re-reading times, a higher number of re-reading fixations, and more skipping in the English readers. In global sentence processing, as suggested previously (Hutzler & Wimmer, 2004), we expected more regressions in readers of English than in readers of German. Since study 2 had shown cross-linguistic word processing differences in adults mostly for nonwords, and study 3 did not analyse sentences containing nonwords, we expected to find cross-linguistic differences in word or sentence processing mainly for children.

In children, results met expectations very well: For local word processing, first-pass reading time was higher in German children, re-reading time was higher in English children, and total word reading time was comparable for the two orthographies (Figure 3, left two columns). In line with this, the probability to refixate a word in first pass was higher in German than in English children, and general refixation probability did not differ between them.

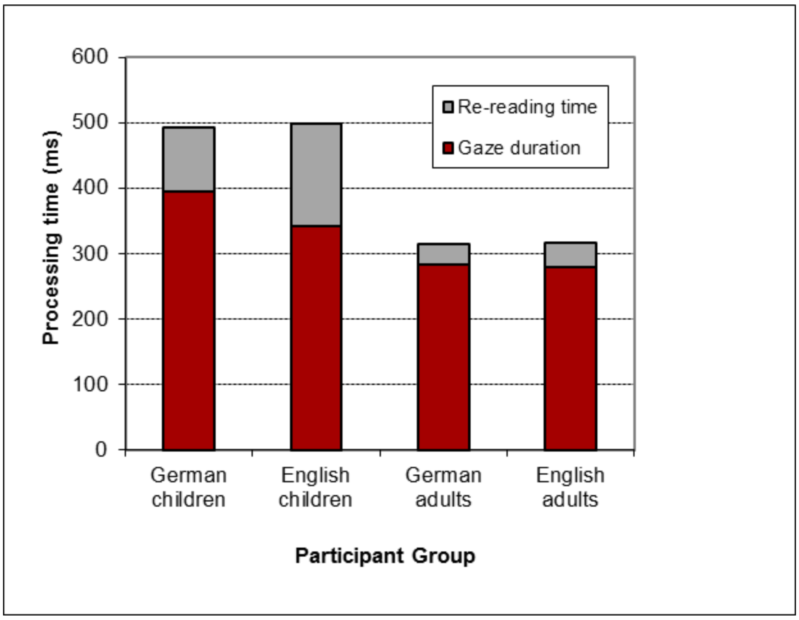


Figure 3. Decomposition of total word processing time into early gaze duration and late re-reading time.

As predicted, the probability to skip a word was clearly higher in English children. Global sentence processing did not differ between children of the two orthographies with regards to mean fixation duration, total sentence processing time, or total number of fixations per sentence. This is not surprising given that participants were intended to have comparable levels of reading skill.

English children made more regressions between words than German children as predicted. To illustrate, Figure 4 shows the differing eye movement pattern for a matched pair of children reading the same sentence: the German child (above) read the sentence in a strict left-to-right manner, placing exactly one fixation on every word. The English child (below) placed her fixations less systematically, skipping two words of the sentence, but also making two regressions to previously fixated words.

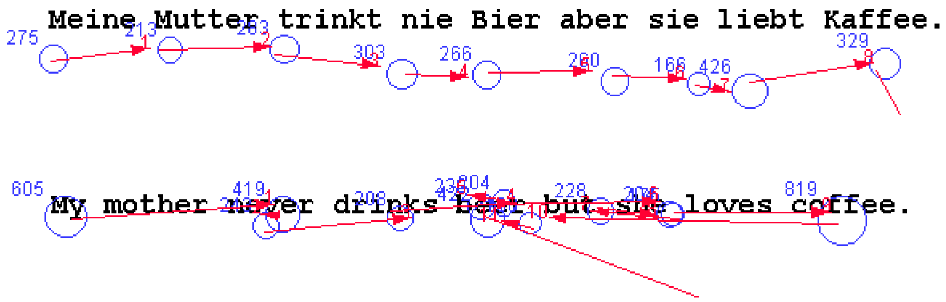


Figure 4. Eye movement pattern of a matched pair of children.

Unlike in children, first fixation duration, gaze duration, and re-reading time did not differ between German and English adults, and in line with the data in children, neither did total reading time (Figure 5, right two columns). Consistent with the children's data, the probability to refixate a word in first-pass was higher in German than in English adults, while general refixation probability did not differ between them. Finally, and also consistent with the children's data, skipping probability was nearly twice as high in English as in German adults. For global sentence processing, the adult findings were very similar to the children's findings in showing no cross-linguistic difference in mean fixation duration, total sentence reading time, or total number of fixations per sentence. However, and importantly, the relative number of interword regressions (i.e., percentage of regressions) was higher for English than for German adults.

Because of the differences in orthographic consistency, we expected more small-unit bottom-up processing in readers of German, and more large-unit top-down processing in readers of English for both word and sentence processing, with more pronounced cross-linguistic differences in children than in adults. While children of the two orthographies were comparable with regards to the *outcome* of both word and sentence processing, they *achieved* this highly similar outcome in a substantially different way. German children showed the expected small-unit reading style with more diligent first-pass reading and less re-reading time, while English children showed the expected large-unit reading style with more skippings and more regressions. Interestingly, the considerable cross-linguistic processing differences as found in children largely persisted in the adult readers.

3 General Discussion

For the most part of the past century, reading research was dominated by empirical evidence and theoretic accounts derived from studies involving native speakers of English reading in their mother tongue. The implicit assumption was that these English-based hypotheses, models, and theories on typical and atypical reading and reading development would be universally valid for other alphabetic orthographies, too. Toward the end of the past century however, there was increasing empirical evidence (e.g., Öney & Goldman, 1984; Thorstad, 1991; Wimmer & Goswami, 1994) suggesting reading and its development to depend on a number of factors which differ between the Anglo-American language area and other language areas.

In an ambitious and influential cross-linguistic study investigating the development of initial reading competence, Seymour and colleagues (2003) showed that the ability to read simple words and nonwords was delayed by more than two years for the English-speaking sample when compared to a large number of other European orthographies. While factors such as age at onset of formal reading instruction, or the method of reading instruction, certainly contribute to explaining these large differences in initial reading acquisition, the most important factor has consistently been identified to be orthographic consistency. Thus, it has very convincingly been argued that a single reading routine is sufficient for beginning readers in consistent, but not for beginning readers in inconsistent orthographies, and that the need to acquire and apply different reading procedures from the very start of reading acquisition is at the core of the delay in initial reading development in inconsistent orthographies (e.g., Seymour et al., 2003; Aro & Wimmer, 2003; Ziegler & Goswami, 2005).

To investigate and contrast typical word and sentence processing in the consistent German and the inconsistent English orthography was the aim of the research project presented here. The scope of the respective studies increased from study 1, which investigated the development of word recognition processes in German, to study 2, which compared word processing between developing and skilled readers of German and English, to study 3, which compared word and sentence processing between the two orthographies. For our sentence reading experiment, we used eye tracking because it permits for online data recording in a fairly natural reading situation. Another major motivation for using eye tracking was that we could record data for local word and global sentence processing in a single session. Most importantly, eye movement data provides a range of measures which can be used to study how a specific variable’s effect unfolds over the time-course of word recognition.

3.1 The transition from sublexical decoding to parallel lexical word recognition in German

Study 1 investigated the development of typical word recognition in the consistent German orthography. Early reading in consistent orthographies is considered to be mainly reliant on serial sublexical decoding. But since the acquisition and the application of such highly reliable grapheme-phoneme-correspondence rules are straightforward, phonology-based word recognition is soon complemented by parallel lexical word recognition for words which have been read a number of times before. To study this transition from serial sublexically-based to more parallel lexically-based word recognition, effects of word length were studied for words of differing familiarity in groups of differing age and reading experience. The naturally large differences in general reading ability between the different age groups were accounted for by analysing standardised processing times alongside to analysing raw processing times.

As expected, and consistent over both analyses, length effects for words of differing familiarity were largely comparable for the least experienced readers, but increased with decreasing word familiarity for the more experienced readers. The pattern of results was somewhat distinct for the youngest group, but differed only quantitatively between the more experienced groups of readers. Specifically, grade 2 readers appeared to apply serial sublexical decoding as a default reading strategy to most words, except for short high-frequency words which they quite obviously read by direct lexical access. Very interestingly, the transition between serial sublexical decoding and parallel lexical word recognition was not only evident in the form of decreasing length effects between groups, or evident in the form of a less pronounced length effect for more familiar words within one and the same group. It was even evident within the same word type, within the same group. Particularly, this concerns the grade 2 children's reading of low-frequency words. *Short* low-frequency words were read equally slowly as short nonwords suggesting that they were read by phonological decoding. However, for *long* low-frequency words, processing time was clearly lower than for long nonwords, so a certain amount of lexical processing must have played a role. The term partial decoding is perhaps a suitable one in this context, since grade 2 children obviously *started* reading long low-frequency words by serial decoding, but *completed* reading by lexical access.

The more experienced readers from grade 3 onwards showed the expected decrease in length effect for words of decreasing familiarity suggestive of the fact that the more familiar a word, the less the involvement of phonological processes. The fact that grade 3 children showed a much smaller length effect for high-frequency words than for low-frequency words and nonwords, together with the fact that their length effect for high-frequency words was no larger than that for the more experienced readers of grade 4 and adults, clearly shows competent lexical reading in the case of such highly familiar words. The trend for a smaller length effect in more experienced readers than grade 3 readers as found for low-frequency words reflects the increasing number of words for which detailed lexical representations are directly available in more experienced readers.

To sum up, study 1 has shown the developmental transition from sublexical decoding to parallel lexical word recognition to be modulated by word familiarity. Grade 2 readers read highly familiar words by direct lexical access. For less familiar words, they showed a clear reliance on phonological word recognition which was complemented by lexical processes in the case of longer words. For the more experienced readers from grade 3, the involvement of sublexical phonological word recognition decreased with increasing word familiarity. Importantly, this pattern was confirmed when the generally higher processing times in younger readers were controlled for.

Even though controlling for the generically different levels of reading skill between the groups did not change the main results, it is nevertheless important. When comparing the involvement of phonological processing in reading words of differing familiarity between groups, results for analyses using both raw and standardised processing times did show differences between groups when it came to *word* reading. Thus, when reading words of both high and low frequency, less experienced readers rely on phonological processes to a larger degree than more experienced readers even after their generally lower word reading ability has been controlled for. In contrast, differences between groups for *nonword* reading were only found for raw processing times. The smaller nonword length effect in adults than in children of all ages is indicative of an increase in efficiency of sublexical decoding in the course of reading development. The comparable length effect over all groups found after controlling for differences in general word processing, however, is indicative of the fact that in nonword reading, all groups relied on phonological recoding to the same extent.

3.2 The time-course of word and sentence processing in German and English

Studies 2 and 3 compared eye movement behaviour in word and sentence processing cross-linguistically. German and English participants were carefully matched on basic word reading ability. Since the standardised reading tests as used to establish average or above-average reading ability were not comparable between the two languages, we chose to use the easiest category of target words of the experiment itself as a basis for participant matching.

In line with the predictions of grain size theory (Ziegler & Goswami, 2005), we expected readers of the consistent German orthography to rely more on small-unit processing, and readers of the inconsistent English orthography to rely more on large-unit processing. As suggested by de Jong (2006), we expected cross-linguistic differences to be more pronounced in developing than in skilled readers. Findings of studies 2 and 3 will be discussed together in the following, starting with the children's findings.

3.2.1 Developing readers: different approach, but identical outcome

In children, orthographic consistency did not impact on total word processing effort, whereas it clearly impacted on the way in which this total effort was achieved. German children applied more small-unit processing in first-pass reading, and consequently needed less time for re-analysis. Quite clearly, there is limited need for re-reading when during first-pass a word has been carefully decoded in a serial fashion, and grapheme-phoneme-correspondences are highly consistent. On the contrary, English children attempted to use larger units in first-pass reading, applying more careful small-unit processing only in second-pass reading. The increased need for re-analysis was necessary in cases where the subsequent sentence context did not confirm their first reading attempt. The advantage of eye tracking and a sentence-reading paradigm over the use of vocal reaction times and single word reading becomes very obvious when it comes to uncovering such subtle cross-linguistic differences in the time-course of word recognition processes. When vocal reaction times for single target items are used, the response is either correct or incorrect, and there is no re-analysis of single targets, because no context is provided.

While study 2 established cross-linguistic differences in the time-course of word recognition for especially chosen target words, study 3 sought to enlarge the perspective by investigating the influence of orthographic consistency on local word and global sentence processing parameters for *all* words of the presented sentences. Importantly, findings of study

3 fit and complement those of study 2 very well. As intended by the reading level match procedure between readers of the two orthographies, children did not differ with respect to the *outcome* of either local total word processing time or global total sentence processing time. Importantly, they did differ in their *approach* to reading words and sentences. German children devoted more resources to careful first-pass processing, a trend which was evident in their higher initial processing times, higher number of first-pass fixations, higher first-pass refixation probability, and lower skipping probability. This approach of rather slow and meticulous first-pass reading makes frequent regressions and long re-reading times unnecessary. English children showed a quite different approach to word and sentence processing. They read in a less systematic, more fragmented way with lower first-pass reading times, more word skippings, more regressions, and higher re-reading times.

The crucial point is that great care was taken to match children of the two orthographies on basic word reading ability. Indeed, their highly comparable level of reading ability is reflected in their identical total sentence processing times. The reported cross-linguistic differences are therefore truly reflecting different *approaches* to text processing, and not some unintended difference in reading *skill*.

Quite interestingly, the different pattern of eye movements as shown by German and English children in word and sentence processing fits the distinction of plodders and explorers established in dyslexia research (Olson et al., 1985). German children displayed more of a plodder, and English children more of an explorer style of reading. Besides the eye movement data, which show a more diligent step-by-step way of processing for the consistent orthography, and a more context-seeking, back-and-forth way of processing for the inconsistent orthography, this plodder/explorer pattern is also evident in the reading accuracy data. Given that surface dyslexia is more often reported in dyslexic readers of consistent orthographies (Bergmann & Wimmer, 2008; de Luca et al., 1999), while phonological dyslexia is more often reported in dyslexic readers of the inconsistent English orthography (Castles & Coltheart, 1993), the more frequent reading style shown by dyslexics of consistent and inconsistent orthographies appears to reflect the typical reading style shown by typically developing readers of their respective orthography.

3.2.2 Skilled readers: different approaches to word and sentence processing persist

It has been proposed that reading development converges between consistent and inconsistent orthographies as word recognition in consistent orthographies becomes increasingly reliant on lexical processes during the course of reading development (de Jong, 2006). We therefore expected cross-linguistic differences in word and sentence processing to be more pronounced in children than in adults.

On the level of specified target words, this was indeed the case. Other than for children, orthographic consistency exerted hardly an influence on the processing of matched high- and low-frequency words. However, there was a clear cross-linguistic difference when it came to nonword processing: not only was reading accuracy for the category of 9-10 letter nonwords much lower in English than in German adults; their reading times were also much higher. A conceivable explanation for the specific nonword reading difficulty in English adults is that their approach to word recognition generally involves more reliance on contextual cues than that of German adults, a top-down strategy which works well in the case of familiar words, but is less adequate in the case of nonwords.

While the more large-picture cross-linguistic comparison of word and sentence processing in study 3 did not show German and English adults to differ to quite the same extent as the children, the plodder-explorer distinction based on skipping and regressions rates fits the adults' eye movement pattern just as well as the children's. Thus, both the skipping rate and the between-word regression rate were lower in German than in English adults.

To sum up, the pattern of a more diligent small-unit plodder-like reading style in readers of German, and a more context-seeking large-unit explorer-like reading style in readers of English was evident in both developing and highly experienced readers of the two orthographies.

3.2.3 Skilled readers of English: no nonsense, please

As noted above, English adults took longer to process nonwords than German adults and were less accurate in reading the category of 9-10 letter nonwords which was excluded from eye movement analyses. What is particularly interesting about this finding is that total processing time for nonwords differed between adults, but not between children of the two orthographies. One possibility is that the non-existing cross-linguistic nonword processing difference in children is the result of a more stringent phonics regime in British schools in recent years, of which the English children of our sample took advantage, but not the adults.

Additionally, it is telling that phonological recoding ability and automatic lexical word recognition were closely linked in English children, but not at all in English adults of the present sample (recall that the performance on the nonword reading and the word reading subtests of the TOWRE correlated in English children, but not in English adults). While the inclusion of a large number of exception words in the TOWRE's sight word efficiency list makes it likely that both children and adults employed different reading procedures for reading the nonword list and the word list (namely, phonological recoding and lexical retrieval, respectively), it is not unlikely that the increased efforts at phonics teaching have brought about this tight association of decoding skill and automatic word recognition in English children. What speaks in favour of this is that the association between word and nonword reading in the German reading fluency test (Moll & Landerl, 2010) was strong in German children and adults, both of whom learned reading through phonics.

3.3 Eye movements as a reflection of reading skill in cross-linguistic research?

Many previous studies have used eye movement parameters as useful indices characterising the development of reading skill from beginning to skilled readers (e.g., English: McConkie et al., 1991; Finnish: Häikiö et al., 2009; German: Huestegge et al., 2009). Thus, Huestegge and colleagues (2009; readers of German) have identified high refixation rates as a general characteristic for less developed reading skills; Hawelka, Gagl, and Wimmer (2010; readers of German) took high skipping rates as indicative of orthographic whole-word recognition and thus as characteristic of skilled reading; and Olson and colleagues (1985; readers of English) argued that less developed readers make more fixations and regress more often than more developed readers. With respect to the cross-linguistic findings reported here, a somewhat ambiguous picture emerges: The fact that German children show a higher refixation probability in first-pass and show less word skipping than English children would appear to suggest less developed reading skill on their part, while the higher regression rate in English children would suggest the opposite.

Importantly, the data presented here has shown such fundamental differences in word and sentence processing between readers of the consistent German and the inconsistent English orthography that it is not appropriate to relate these cross-linguistic processing differences to cross-linguistic differences in reading skill. Given that similar cross-linguistic processing differences were found in adults, too, there appear to be persistent and reliable differences in reading pattern between consistent and inconsistent orthographies. It is therefore appropriate to conclude that comparisons between less developed and more developed readers must be done within, and not across orthographies.

3.4 Limitations

To every experimental study, there are of course a number of methodological drawbacks. In the following, the limitations applying to the present studies will be discussed.

First, the CELEX frequency counts which were used to identify high- and low-frequency words for the two orthographies were not specifically devised for the use in children (Baayen et al., 1995). CELEX lexical database was chosen because it existed for both German and English, and target frequencies could thus be matched between the two orthographies. The frequency counts were obtained from sources such as adult literature, newspapers, and magazines, and from transcripts of adult spontaneous speech. Naturally, frequency counts collected for adults are not necessarily equally valid for primary school children. However, we used the combined frequency counts for written and spoken sources expecting that these would better correspond to the actual frequencies for children than the frequency counts for written sources only. When choosing the high- and low-frequency targets, we did however pay attention to only include words that could reasonably be expected to be high- or low-frequency for children, too (e.g., the cognate *Hamster* / *hamster* is a low-frequency target according to CELEX frequency counts and was not included in our experiment because we thought it may not be low-frequency for children).

Second, the words immediately preceding and following target words were not matched between the two orthographies. Given the existence of parafoveal-on-foveal and spillover effects, this is clearly suboptimal for the analyses involving specific target words. However, in many cases the sentence frame was so similar for the two languages that words n-1 and n+1 were in fact comparable (e.g., the matched sentences for the target word *Gazelle* / *gazelle* were <u>*Die*</u> *Gazelle* <u>*lebt*</u> *in Afrika* / <u>*The*</u> *gazelle* <u>*lives*</u> *in Africa*). In a similar vein, it was not possible to perfectly match the sentences with regards to number of words (which was slightly higher for English), or average word length (which was higher for German). There is thus the possibility that the findings of study 3 were in part caused by such confounds. For instance, skipping was much more frequent in readers of English than in readers of German, and this could partly be due to the fact that English words were slightly shorter than German words. However, the average number of letters per sentence was no different between the two orthographies, and while the possibility cannot be excluded that the subtle differences in sentence construction have influenced the reading pattern, cross-linguistic processing differences were solid enough to suggest that they are genuine.

Third, we were not able to collect data concerning the method of reading instruction participants had been taught with. This is not detrimental in the case of the German participants because phonics-based teaching methods are unchallenged as the standard method of reading instruction. It would have been more interesting to know more about the method of reading instruction in the case of the English participants. Unfortunately, for the child participants, there was no opportunity to ask at their schools, because the children were recruited during a summer sports camp and we were thus not in contact with the schools themselves. However, there is no reason to expect that these children were learning to read in a different way from the increasingly phonics-based method practised in England since the late 1990s. Aside from the fact that the method of reading instruction was probably not fundamentally different, at least between children of the two orthographies, it has been suggested in the literature that orthographic consistency and instructional methods are linked (e.g., Frith et al., 1998; Landerl, 2000). If this is the case, the actual method of reading instruction is not the central factor anyhow.

Fourth, the present experiment asked participants to read the sentences out loud because it was important to control for reading accuracy of specified target words. Undeniably, reading is typically done silently, and one could therefore question whether reading aloud was appropriate. However, the inclusion of long and difficult nonwords necessitated the oral reading procedure to be able to exclude those words which were not read correctly. While there is no doubt that eye movement behaviour differs to a certain extent depending on reading mode (e.g., Rayner, 1998; Vorstius et al., 2014), Radach and colleagues (2012) report an interesting cross-linguistic finding in this context: The difference between silent and oral reading was more pronounced in American than in German children. It should be noted, though, that the sentences they used had not originally been created for the cross-linguistic comparison, and were thus not identical. Also, in their review article, the authors provide no details on the age or grade level of German and American children in question. However, assuming that the difference between reading modes is indeed larger for English than for more consistent orthographies, it may be considered problematic that participants of the present studies had to read aloud, because this may have meant more of a change of eye movement behaviour in English as opposed to German readers. However, it has previously been shown that readers make slightly more and longer fixations and shorter saccades in oral than in silent reading (Rayner, 1998). Since this pattern corresponds more to the plodder style of reading more strongly associated with the German readers of our studies, the possible stronger distortion for English than for German as brought about by forcing them to read orally would have meant that the shift towards a more plodder-like small-unit reading approach as induced by reading mode was larger in our English participants than in our

German participants. Importantly, however, since it were the *German* readers that were characterised by the plodder style of reading, the possibly larger difference between silent and oral reading modes in English readers cannot have influenced results in the wrong direction.

Fifth, the reading level matching procedure used for studies 2 and 3 were based on participants' mean gaze duration (i.e., first-pass reading time) for the easiest category of target words. Gaze duration was chosen because it is considered to comprise all stages of word recognition and is thus the most widely reported measure. The category of short, high-frequency words as a basis for participant matching was chosen because even the youngest readers of grades 2 and 3 of both countries can be expected to read such highly familiar words by direct lexical access. However, when considering the time-course of word recognition over all words of the sentences, it emerged that first-pass reading times were higher in German children, and re-reading times were higher in English children. It is therefore possible that our reading level match worked well for first-pass, but not for total processing time. Importantly however, when comparing mean total reading time for short high-frequency words between children of the two orthographies, the numerically higher total reading times of the English children were statistically not significantly higher (mean total reading time high-frequency words: German children = 378 (114) ms vs English children = 431 (139) ms; $t(48) = 1.47$, $p = .15$). In adults, total reading time was also comparable (German adults = 267 (60) ms vs English adults = 255 (59) ms, $t(30) = .60$, n.s.).

Finally, the research presented here has used a cross-sectional and no longitudinal design, which would be the preferred way to studying developmental processes. The reasons for choosing a cross-sectional design were mostly practical: the project partner in England was not available over an extended period of time, and neither was the eye tracking laboratory in Germany.

3.5 Outlook

There of course remain a number of open questions which would be interesting for future studies to address.

The present research project is the first to have compared word and sentence processing between a consistent and an inconsistent alphabetic orthography using carefully matched reading material on the sentence level. We believe that this methodological approach may provide good grounds for further research questions. For instance, with regards to the distinct plodder / explorer pattern of eye movement behaviour shown by German and English readers, it would be interesting to establish whether the two reading styles would emerge in larger groups of developing readers of the same age within both orthographies. Since the original plodder / explorer distinction was found within a group of English dyslexic children, it is likely that both types of reading style would be present in typically developing and dyslexic readers of both consistent and inconsistent orthographies. If the relative frequencies of occurrence of the two types of readers were studied longitudinally, it would furthermore be possible to determine the stability of such an individual reading style profile. Of course, the reading material of the present research project would also permit to compare sentence processing in dyslexic readers of the two orthographies, which would take the plodder / explorer distinction in reading style back to where it originated from (namely, dyslexia research; cf., Olson et al., 1985).

Another interesting research endeavour would be to try and establish whether German-English bilinguals process nearly identical sentences in a similar way in both languages, or whether they adapt their eye movement behaviour to the demands of the orthography they are reading.

On a final note, the cross-linguistic comparison of the reading process between alphabetic orthographies of differing consistency can certainly make important contributions to a universal science of reading as called for by Share (2008, 2014). However, in the future, an increasing number of comparisons between alphabetic and non-alphabetic scripts would be desirable. There are already some pioneering studies in this field such as a very interesting study investigating effects of universal and script-dependent factors by comparing the development of eye movements of English and Chinese children and adults (Feng, Miller, Shu, & Zhang, 2009). Ultimately, studies such as this one are needed to create an empirical basis for the development of truly universal theories of reading and reading development. As a first step in this direction, the development of a universally applicable model for cross-linguistic and cross-script research would be of great practical use.

3.6 Conclusion

The process of reading has been suggested to differ as a function of orthographic consistency (orthographic depth hypothesis: Frost, Katz, & Bentin, 1987; psycholinguistic grain size theory: Ziegler & Goswami, 2005). More specifically, small-unit bottom-up processing should be more prevalent in readers of consistent orthographies, and large-unit top-down processing should be more prevalent in readers of inconsistent orthographies.

In a series of studies, the present work has provided support of this claim: Eye movement data from readers of the consistent German orthography have shown an early reliance on sublexical recoding and a continuous lexicalisation of word recognition in the course of reading development. In two direct cross-linguistic comparisons, the time-course of word and sentence processing was found to differ as a function of orthographic consistency. While readers of German showed a plodder-like reading style with careful and accurate first-pass processing and relatively little time devoted to re-reading, readers of English showed an explorer-like reading style with many more word skipping and regressive eye movements.

In conclusion, orthographic consistency influences the processing of both words and sentences, in both developing and skilled readers.

References

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review, 105*, 678-723.

Aro, M. (2006). Learning to read: The effect of orthography. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 531-550). Hove: Erlbaum.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*, 621-635.

Australian Government. Department of Education, Science, and Training (2005). *Teaching reading. Report and recommendations. National enquiry into the teaching of literacy*. Barton, Australia: Department of Education, Science, and Training.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Linguistic Data Consortium. Philadelphia, PA: University of Pennsylvania.

Bergmann, J., & Wimmer, H. (2008). A dual-route perspective on poor reading in a regular orthography: Evidence from phonological and orthographic lexical decisions. *Cognitive Neuropsychology, 25*, 653-676.

Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research, 51*, 84-92.

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's eye movements during reading. In D. Gilchrist & S. Everling (Eds.), *Oxford handbook on eye movements* (pp. 643-662). Oxford: University Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults (2009). *Vision Research, 49*, 1583-1591.

Borgwaldt, S. R., Hellwig, F. M., & de Groot, A. M. B. (2005). Onset entropy matters – Letter-to-phoneme mappings in seven languages. *Reading and Writing, 18*, 211-229.

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictor of reading difficulty: An evaluation using the Potsdam sentence corpus. *Journal of Eye Movement Research, 2*, 1-12.

Brown, G. D. A. & Deavers, R. P. (1999). Units of analysis in nonword reading: Evidence from children and adults. *Journal of Experimental Child Psychology, 73*, 208-242.

Brush up on your English. (n.d.). In *A. Word. A. Day*. Retrieved from https://www.wordsmith.org/awad/english2.html

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

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago, IL: University of Chicago Press.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*, 1398-1407.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., … & Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science, 23*, 678-686.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 149-180.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 225-235.

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

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In M.I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 3-37). Hillsdale, NJ: Erlbaum.

Coltheart, M. (2006). Dual route and connectionist models of reading: an overview. *London Review of Education, 4*, 5-17.

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

Cunnings, I., & Clahsen, H. (2007). The time-course of morphological constraints: Evidence from eye-movements during reading. *Cognition, 104*, 476-494.

Davies, R., Cuetos, F., & Glez-Seijas, R. M. (2007). Reading development and dyslexia in a transparent orthography: A survey of Spanish children. *Annals of Dyslexia, 57*, 179-198.

De Jong, P. F. (2006). Units and routes of reading in Dutch. *Developmental Science, 9,* 441-442.

De Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading, 11*, 55-71.

De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: an eye movement study of developmental dyslexia. *Brain and Language, 80*, 617-626.

De Luca, M., di Pace, E., Judica, A., Spinelli, D., & Zoccolotti, P. (1999). Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia, 37*, 1407-1420.

Department for Education and Employment (1998). *The National Literacy Strategy: Framework for reaching*. Suffolk, England: Author.

Duncan, L. G., Seymour, P. H. K., & Hill, S. (1997). How important are rhyme and analogy in beginning readers? *Cognition, 63*, 171-208.

Dürrwächter, U., Sokolov, A. N., Reinhard, J., Klosinski, G., & Trauzettel-Klosinski, S. (2010). Word length and word frequency affect eye movements in dyslexic children reading in a regular (German) orthography. *Annals of Dyslexia, 60*, 86-101.

Ellis, N. C. & Hooper, A. M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics, 22,* 571-599.

English orthography. (n.d.). In *Wikipedia – the free encyclopedia*. Retrieved from http://en.wikipedia.org/wiki/English_orthography#History

Feng, G., Miller, K., Shu, H., & Zhang, H. (2009). Orthography and the development of reading process: An eye-movement study of Chinese and English. *Child Development, 80*, 720-735.

Frith, U. (1985). Beneath the surface of a developmental dyslexia. In K. Patterson, J. C. Marshall & M. Coltheart (Eds.), *Surface dyslexia: Cognitive and neuropsychological* *studies of phonological reading* (pp. 301-330). London: Lawrence Erlbaum Associates.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading, 2,* 31-54.

Frost, R. (1994). Prelexical and postlexical strategies in reading: evidence from a deep and a shallow orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 116-129.

Frost, R. (2012a). Towards a universal model of reading. *Behavioral and Brain Sciences, 35,* 263-279.

Frost, R. (2012b). A universal approach to modelling visual word recognition and reading: Not only possible, but also inevitable. *Behavioral and Brain Sciences, 35*, 310-329.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 104-115.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42,* 73-83.

Goswami, U., Gombert, J. E., & Fraca de Barrera, L. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics*, *19*, 19-52.

Goswami, U., Porpodas, C., & Wheelwright, S. (1997). Children's orthographic representations in English and Greek. *European Journal of Psychology of Education, 3*, 273-292.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *Journal of Memory and Language, 45*, 648-664.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics, 24*, 235-247.

Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition*, *123*, 280-292.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 201*, 167-181.

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.

Hatzidaki, A., Gianneli, M., Petrakis, E., Makaronas, N., & Aslanides, I. M. (2011). Reading and visual processing in Greek dyslexic children: An eye-movement study. *Dyslexia, 17*, 85-104.

Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition, 115*, 367-379.

Holopainen, L., Ahonen, T., & Lyytinen, H. (2002). The role of reading by analogy in first grade Finnish readers. *Scandinavian Journal of Educational Research, 46*, 83-98.

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research, 49*, 2948-2959.

Hutzler, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language, 89*, 235-242.

Hutzler, F., Ziegler, J. C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition, 91*, 273-296.

Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures to study global text processing. In R. Radach, J. Hyönä, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313-334). Amsterdam: Elsevier.

Hyönä, J., & Olson, R. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1430-1440.

Juel, C., & Holmes, B. (1981). Oral and silent reading of sentences. *Reading Resarch Quarterly, 16*, 545-568.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation duration in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1312-1318.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension, *Psychological Review 87*, 329-354.

Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 157-166.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz, L. (Eds.), *Orthography, Phonology, Morphology, and Meaning* (pp. 67-84). Amsterdam: Elsevier North Holland Press.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153-168.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General, 135*, 12-35.

Kuperman, V., & van Dyke, J. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65*, 42-73.

Landerl, K. (2000). Influences of orthographic consistency and reading instruction on the development of nonword reading skills. *European Journal of Psychology of Education, 15*, 239-257.

Landerl, K. (2006). Reading acquisition in different orthographies: Evidence from direct comparisons (pp. 513-530). In M. Joshi & P.G. Aaron (Eds.), *Handbook of orthography and literacy*. Hove: Erlbaum.

Landerl, K. (in press). Reading acquisition in German language. In C. Perfetti & L. Verhoeven (Eds.), *Reading acquisition across languages and writing systems*. Cambridge: Cambridge University Press.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and orthographic spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology, 100*, 150-161.

Landerl, K., & Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics, 21*, 243-262.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition, 63*, 315-334.

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

Laxon, V., Masterson, J., & Moran, R. (1994). Are children's representations of words distributed? Effects of orthographic neighbourhood size, consistency and regularity of naming. *Langauge and Cognitive Processes, 9*, 1-27.

Learning to read. (n.d.). In *Wikipedia – the free encyclopedia*. Retrieved from http://en.wikipedia.org/wiki/Learning_to_read

Lowell, R., & Morris, R. (2014). Word length effects on novel words: Evidence from eye movements. *Attention, Perception, and Psychophysics, 76*, 179-189.

Mann, V. & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and English children. *Reading and Writing: An interdisciplinary journal, 15*, 653-682.

Martens, V. E. G. & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language, 98*, 140-149.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. In J. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). London: The Macmillan Press.

Moll, K., Fussenegger, B., Willburger, E., & Landerl, K. (2009). RAN is not a measure of orthographic processing: Evidence from the asymmetric German orthography. *Scientific Studies of Reading, 13*, 1-25.

Moll, K., & Landerl, K. (2010). *Lese- und Rechtschreibtest (SLRT II)*. Bern: Huber.

National Institute of Child Health and Human Development (NICHD) (2000). *Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*, NIH publication no. 00-4754. Washington, DC: US Government Printing Office.

Norris, D. (2013). Models of visual word recognition. *Trends in Cognitive Sciences, 17*, 517-524.

Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In G. E. McKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* 4th ed. (pp. 1-6). New York: Academic Press.

Olson, R. K., Wise, B., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the posttraining development of word recognition. *Scientific Studies of Reading, 1*, 235-253.

Öney, B. & Goldman, S. R. (1984). Decoding and comprehension skills in Turksish and English: Effects of the regularity of grapheme-phoneme correspondences. *Journal of Educational Psychology, 76*, 557-568.

Paap, K. R., & Noel, R. W. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53*, 13-24.

Pagliuca, G., Arduino, L. S., Barca, L., & Burani, C. (2008). Fully transparent orthography, yet lexical reading aloud: The lexicality effect in Italian. *Language and Cognitive Processes, 23*, 422-433.

Patel, T. K., Snowling, M. J., & de Jong, P. F. (2004). A cross-linguistic comparison of children learning to read in English and Dutch. *Journal of Educational Psychology, 96*, 785-797.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., Cotelli, M., Cossu, G., Corte, F., Lorusso, M., Pesenti, S., Gallagher, A., Perani, D., Price, C., Frith, C. D., & Frith, U. (2000). A cultural effect on brain function. *Nature Neuroscience, 3*, 91-96.

Perry, C., Ziegler, J. C., & Zorzi, M. (2014). When silent letters say more than a thousand words: An implementation and evaluation of CDP++ in French. *Journal of Memory and Language, 72*, 98-115.

Phonics. (n.d.). *In Wikipedia – the free encyclopedia*. Retrieved from http://en.wikipedia.org/wiki/Phonics

Plaut, D. C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. *Cognitive Science, 23*, 543-568.

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

Porpodas, C. D. (2006). Literacy acquisition in Greek: Research review of the role of phonological and cognitive factors. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 189-200). Hove: Erlbaum.

Radach, R., Günther, T., & Huestegge, L. (2012). Blickbewegungen beim Lesen, Leseentwicklung und Legasthenie, *Lernen und Lernstörungen, 1*, 185-204.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology, 16*, 3-26.

Rau, A. K., Moll, K., Snowling, M., & Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology, 130*, 92-105.

Rau, A. K., Moll, K., Snowling, M., & Landerl, K. (submitted). Same same, but different: word and sentence reading in German and English.

Rau, A. K., Moeller, K., & Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: An eye tracking study. *Scientific Studies of Reading*, 18, *224-233.*

Rau, A. K. (2015). Corrigenda. *Scientific Studies of Reading, 19*, 86-87.

doi*:* 10.1080/10888438.2015.970875

Rayner, K. (1985). The role of eye movements in learning to read and reading disability. *Remedial and Special Education, 6*, 53-60.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K., & Duffy, S. D. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity and lexical ambiguity. *Memory and Cognition, 14*, 191-201.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton Jr., C. (1989). Eye movements and on-line language comprehension processes, *Language and Cognitive Processes, 4*, 21-49.

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading, 7*, 219-237.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125-157.

Rose, J. (2006). *Independent review of the teaching of early reading*. Nottingham: DfES Publications.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-174.

Seymour, P. H. K., & Duncan, L. G. (1997). Small unit versus large unit theories of reading acquisition. *Dyslexia, 3*, 125-134.

Seymour, P. H. K. & Elder, L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology, 3*, 1-36.

Share, D. L (1995). Phonological recoding and self-teaching: *sine qua non* of reading acquisition. *Cognition, 55*, 151-218.

Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin, 134*, 584-615.

Share, D. L. (2014). Alphabetism in reading science. *Frontiers in Psychology, 5*: 752. doi: 10.3389/fpsyg.2014.00752

Signorini, A. (1997). Word reading in Spanish: A comparison between skilled and less skilled beginning readers. *Applied Psycholinguistics, 18*, 319-344.

Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition, 30*, 139-181.

Taylor, S. E. (1965). Eye movements while reading: Facts and fallacies. *American Educational Research Journal, 2*, 187-202.

Thorstad, G. (1991). The effect of orthography on the acquisition of literacy skills. *British Journal of Psychology, 82*, 527-537.

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

Trauzettel-Klosinski, S., Koitzsch, A. M., Dürrwächter, U., Sokolov, A. N., Reinhard, J., & Klosinski, G. (2010). Eye movements in German-speaking children with and without dyslexia when reading aloud. *Acta Ophthalmologica, 88*, 681-691.

Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: Implications for reading development and theory. *Memory & Cognition, 18*, 559-567.

Treiman, R., Mullenix, J., Bjeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124*, 107-136.

Van den Boer, M., de Jong, P. F., Haentjens-van Meeteren, M. M. (2013). Modelling the length effect: Specifying the relation with visual and phonological correlates of reading. *Scientific Studies of Reading, 17*, 243-256.

Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition, 22*, 458-485.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latencies. *The Quarterly Journal of Experimental Psychology, 50A*, 439-456.

Wesseling, R., & Reitsma, P. (2000). The transient role of explicit phonological recoding for reading acquisition. *Reading and Writing, 13*, 313-336.

Whitney, C. & Cornelissen, P. (2005). Letter-position encoding and dyslexia. *Journal of Research in Reading, 28*, 274-301.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: word recognition in English and German children. *Cognition, 51*, 91-103.

Wimmer H., & Hummer, P. (1990). How German-speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics, 11*, 349-368.

Wyse, D., & Goswami, U. (2008). Synthetic phonics and the teaching of reading. *British Educational Research Journal, 34*, 691-710.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*, 3-29.

Ziegler, J. C., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. *European Journal of Cognitive Psychology, 12*, 413-430.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12*, 379-384.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology*, *86*, 169-193.

Zoccolotti, P., de Luca, M., di Filippo, G., Judica, A., & Martelli, M. (2009). Reading development in an orthographically regular language: Effects of length, frequency, lexicality and global processing ability. *Reading and Writing*, *22*, 1053-1079.

Zoccolotti, P., de Luca, M., di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language, 93*, 369-373.

Appendix A

Experimental sentences in German and English

Target Frequency	German Sentences	English Sentences
High-freq word	Onkel Thomas hat eine Kuh und drei Hennen.	Uncle Thomas has a cow and three hens.
High-freq word	Die Äpfel in der Box sind rot und grün.	The apples in the box are red and green.
High-freq word	Eric trinkt Tee und Susi trinkt Milch.	Eric drinks tea and Suzie drinks milk.
High-freq word	Der Priester aß vier Tomaten.	The priest ate four tomatoes.
High-freq word	Nachts sehen wir den Mond und die Sterne.	At night, we see the moon and the stars.
High-freq word	Meine Mutter trinkt nie Bier aber sie liebt Kaffee.	My mother never drinks beer but she loves coffee.
High-freq word	Die Musik im Radio ist laut aber gut.	The music on the radio is loud but good.
High-freq word	Simon ist sieben Jahre alt.	Simon is seven years old.
High-freq word	Das neue Hotel öffnet im Juli.	The new hotel will open in July.
High-freq word	Kann ich in der Mitte sitzen, fragt Nora.	Can I sit in the middle, asked Nora.
High-freq word	Ich hasse das Wetter im Winter es sei denn es schneit.	I hate the weather in winter except when it is snowing.
High-freq word	Sechzig Sekunden sind eine Minute, sagt Tina.	There are sixty seconds in a minute, says Tina.
High-freq word	Produkte guter Qualität haben oft einen hohen Preis.	Products of good quality often have a higher price.
High-freq word	Großvater hat ein Problem mit seinem Knie.	Grandfather has a problem with his knee.
High-freq word	Vater reparierte die Maschine letztes Wochenende.	Father repaired the machine last weekend.
High-freq word	Das Wetter im Februar ist nicht so gut wie das im Juni.	The weather in February is not as good as it is in June.
High-freq word	Der neue Minister spricht sehr laut im Parlament.	The new minister speaks very loudly in parliament.
High-freq word	Mein Vater liebt Fußball und meine Mutter liebt Golf.	My father loves football and my mother loves golf.

High-freq word	Kann ich eine Zigarette rauchen, fragte der Elektriker.	Can I smoke a cigarette, asks the electrician.
High-freq word	Der alte Professor vergaß seine Notizen.	The old professor forgot his notes.
High-freq word	Dies ist eine komplett neue Situation, sagte der Präsident.	This is a completely new situation, said the president.
High-freq word	Die Trompete ist das lauteste Instrument im Orchester.	The trumpet is the loudest instrument in an orchestra.
High-freq word	Der Monat September folgt auf den Monat August.	The month of September follows the month of August.
High-freq word	Dies ist eine interessante Diskussion, sagt der Student.	This is an interesting discussion, says the student.
Low-freq word	Der Dieb stahl Vaters Axt und seinen Hammer.	The thief stole father's axe and his hammer.
Low-freq word	Der Angler fing einen Aal und drei Krabben.	The angler caught an eel and three crabs.
Low-freq word	Hast du den Piep gehört, fragte Julia.	Have you heard the beep, asked Julia.
Low-freq word	David isst eine Kiwi und zwei Bananen.	David is eating a kiwi and two bananas.
Low-freq word	Lisa spielt Oboe und ihr Bruder spielt Trompete.	Lisa plays the oboe and her brother plays the trumpet.
Low-freq word	Dennis ist aus dem Koma erwacht, sagt Doktor Müller.	Dennis has awoken from his coma, says Doctor Miller.
Low-freq word	Dieser Kaffee hat ein gutes Aroma, sagt Tanja.	This coffee has a good aroma, says Tanya.
Low-freq word	Onkel Thomas sitzt im Kajak mit dem Paddel in der Hand.	Uncle Thomas sits in the kayak with the paddle in his hands.
Low-freq word	Kevin aß eine Mango im Park.	Kevin ate a mango in the park.
Low-freq word	Sarahs neuer Bikini ist blau und pink.	Sarah's new bikini is blue and pink.
Low-freq word	Bienen machen aus Nektar Honig, sagt Peter.	Bees make honey from nectar, says Peter.
Low-freq word	Patrick geht auf Safari nach Afrika.	Patrick is going on a safari in Africa.
Low-freq word	Der große Pianist gab ein Konzert in Rom.	The great pianist gave a concert in Rome.
Low-freq word	Die Gazelle lebt in Afrika.	The gazelle lives in Africa.
Low-freq word	Die Nadel von Toms Kompass ist zerbrochen.	The needle of Tom's compass is broken.
Low-freq word	Der Bischof hat ein Aquarium daheim.	The bishop has an aquarium at home.

Low-freq word	Dein Bruder ist ein Optimist, sagt Martin.	Your brother is an optimist, says Martin.
Low-freq word	Der junge Flamingo hat graue Federn.	The young flamingo has grey feathers.
Low-freq word	Robert ist ein typischer Pessimist, warnt mich seine Mutter.	Robert is a typical pessimist, his mother warns me.
Low-freq word	Harrys Erklärung war nicht gerade plausibel aber dafür kreativ.	Harry's explanation was not very plausible, but it was creative.
Low-freq word	Paula aß eine Blaubeere und Kim einen Apfel.	Paula ate a blueberry and Kim ate an apple.
Low-freq word	Der junge Schimpanse lebt in einem Käfig im Zoo.	The young chimpanzee lives in a cage at the zoo.
Low-freq word	Der Salamander isst Würmer und Insekten.	The salamander eats worms and insects.
Low-freq word	Laura ist gut im Volleyball und Marc ist gut im Tennis.	Laura is good at volleyball and Marc is good at tennis.
Nonword	Für seine Party dekoriert Paul Bix den Garten mit roten Laternen.	For his party, Paul Bix is decorating the garden with red lanterns.
Nonword	Mein Freund Daniel Koo geht in die Oper.	My friend Daniel Koo goes to the opera.
Nonword	Oskar und Ogo sind beste Freunde.	Oscar and Ogo are best friends.
Nonword	Linda Etom hat lange blonde Haare.	Linda Etom has long blonde hair.
Nonword	Der junge Eskimo Malk schläft in seinem Iglu.	The young Eskimo, Malk, is sleeping in his igloo.
Nonword	Der Millionär nannte seine Söhne Jeto und Jakob.	The millionaire named his sons Jeto and Jacob.
Nonword	Michaels jüngere Schwester Nibra ist ein Jahr alt.	Michael's younger sister Nibra is one year old.
Nonword	Der Prinz findet ein Schwert im Bomet Palast.	The prince finds a sword in the Bomet Palace.
Nonword	Der türkische Minister Fadim grüßt seinen Kollegen aus Island.	The Turkish minister Fadim greets his colleague from Iceland.
Nonword	Lillys Hamster Birsul ist hungrig.	Lilly's hamster Birsul is hungry.
Nonword	Sofie Paturn trinkt ihren Kaffee mit Milch und Zucker.	Sophie Paturn drinks her coffee with milk and sugar.
Nonword	Tausende Touristen flogen nach Magril letzten Sommer.	Thousands of tourists flew to Magril last summer.
Nonword	Das Land Harilla ist eine kleine Insel inmitten des Ozeans.	The land of Harilla is a small island in the midst of the ocean.
Nonword	Mein Kollege ist ein Akromul, sagt Vater.	My colleague is an acromul, says father.

Nonword	Vor hundert Jahren war Fiapard eine spanische Kolonie.	A hundred years ago, Fiapard was a Spanish colony.
Nonword	Nick geht ins Nogobond Theater.	Nick goes to the Nogobond Theatre.
Nonword	Elisabeth Karbinol ist relativ gut in der Schule.	Elisabeth Karbinol is relatively good at school.
Nonword	Hanna ist allergisch gegen Crumafix und Pollen.	Hannah is allergic to crumafix and pollen.
Nonword	Robin studiert Bagratorn und Henry studiert Medizin.	Robin studies bagratorn and Henry studies medicine.
Nonword	Anna findet ein Paar Ohrringe im Kontinast Park.	Anna finds a pair of earrings in the Kontinast Park.
Nonword	Emma geht in den Trapedest Zoo.	Emma goes to the Trapedest Zoo.
Nonword	Die beste Medizin ist Nemtratect, sagt Doktor Braun.	The best medicine is Nemtratect, says Doctor Brown.
Nonword	Tim ist Blopassant und Max ist Katholik.	Tim is a Blopassant and Max is a Catholic.
Nonword	Doktor Schmidt lebt im Jempromisp Gebäude.	Doctor Smith lives in the Jempromisp Building.

Note. High-freq word = high-frequency word. Low-freq word = low-frequency word.

Note. Target words are underlined.

Appendix B

Construction of word-like nonwords

Low-frequency cognates in German / English	Resulting nonwords
Zoo / zoo	Koo
Box / box	Bix
Ego / ego	Ogo
Atom / atom	Etom
Malz / malt	Malk
Veto / veto	Jeto
Kobra / cobra	Nibra
Fabel / fable	Fadim
Komet / comet	Bomet
Konsul / consul	Birsul
Saturn / saturn	Paturn
Magnet / magnet	Magril
Gorilla / gorilla	Harilla
Leopard / leopard	Fiapard
Akrobat / acrobat	Akromul / acromul
Karneval /carnival	Karbinol
Vagabund / vagabond	Nogobond
Kruzifix / crucifix	Crumafix / crumafix
Vibration / vibration	Bagratorn / bagratorn
Propeller / propeller	Trapedest
Kontinent / continent	Kontinast
Protestant / protestant	Blopassant
Kompromiss / compromise	Jempromisp
Architekt / architect	Nemtratect

Appendix C

Acceptable pronunciations for the nonwords in English

Nonword	Accepted pronunciations in English
Bix	(it)
Koo	(loo)
Ogo	O (oboe, off) go (so)
Etom	E (Eton, elephant) tom (t'm - silent o, from)
Jeto	Je (jelly, Jesus) to (toe)
Malk	Malk (map, chalk, malt)
Bomet	Bom (bomb, boa) met (metronome, may)
Fadim	Fad (add, father), im (dim, deem)
Nibra	Ni (pig, Tina, night) bra (bra, aroma)
Birsul	Bir (bird) sul (sulking, s'l - silent u)
Paturn	Pat (rat, p'turn - silent a) urn (turn)
Magril	Mag (magazine, m'gril - silent a) gril (ill)
Acromul	A (at, ape) cro (crow, actor) mul (mulberry)
Fiapard	Fia (fiat, fire) pard (hard)
Harilla	Ha (happy, harp) rilla (it)
Crumafix	Crum (gum, cruel) a (trauma) fix (tick)
Nogobond	No (no, not) go (so) bond (pond)
Karbinol	Karb (carbohydrate) bi (bicycle, bin) nol (n'l - silent o, pole)
Bagratorn	Bag (lag, bargain) ra (rabbit, Basra) torn (ornament)
Kontinast	Kon (conservative) ti (tin, tie) nast (nasty, nag)
Trapedest	Two syll.: Trape (grape) dest (zest)
	Three syll.: Tra (traffic, tray, tr' – silent a) pe (pest, peas, p') dest (zest)
Blopassant	Blo (blow, plop) pass (passive, p'ss - silent a) ant (and, 'nt - silent a)
Jempromisp	Jem (gem) pro (proactive, promise) misp (it)
Nemtratect	Nem (net) tra (tractor, tr'tect – silent a) tect (sect)

Appendix D

Individual participant matching

English participants				German participants			
Participant number	GD short HF words	Age (months)	School (months)	Participant number	GD short HF words	Age (months)	School (months)
Children							
101	148	130	59	31	178	122	47
94	230	120	59	26	247	113	41
103	237	119	47	20	254	135	45
106	243	111	46	24	255	125	47
98	255	124	59	42	255	107	26
89	261	132	71	25	262	125	47
118	262	106	35	48	269	107	25
93	268	127	59	32	282	105	26
114	272	94	28	65	288	98	17
115	276	99	35	23	290	118	45
113	282	96	35	44	291	103	26
109	287	112	41	34	291	101	26
110	290	119	47	47	295	106	26
105	298	115	47	27	300	119	47
100	298	129	59	22	312	115	45
95	299	131	59	59	322	96	17
102	307	108	47	19	323	125	46
97	342	124	53	33	345	103	26
96	351	130	53	67	369	94	17
116	356	96	28	52	372	103	27
107	396	111	47	37	393	103	26
91	407	135	71	61	411	96	18
117	438	106	35	56	419	95	17
111	440	115	47	39	421	113	26
112	503	100	34	60	481	93	17
Mean	310	116	48		317	109	31

English participants				German participants			
Participant number	GD short HF words	Age (months)	School (months)	Participant number	GD short HF words	Age (months)	School (months)
Adults							
76	182	264		16	177	240	
77	185	252		1	185	312	
73	186	240		12	187	312	
81	193	312		17	205	312	
70	203	276		5	219	300	
75	203	240		6	230	312	
78	212	288		4	230	456	
74	213	312		7	232	312	
79	231	240		13	245	252	
84	255	300		14	245	240	
82	286	252		3	250	300	
80	290	228		2	252	252	
72	295	276		11	253	252	
87	308	432		15	259	252	
85	329	264		9	303	348	
86	335	288		10	357	300	
Mean	244	279			239	297	

Appendix E

Reading error classification

Error category	Error type	Sample reading error
Minor error	Regularisation	Flamingo for [Flaming-go] in G kay-yuk for kayak in E
	Wrong stress assignment	Kompass for Kompass in G pianist for pianist in E
	Wrong decoding of one grapheme	Herilla for Harilla in G Herilla for Harilla in E
	One phoneme missing	Patur for Paturn in G Crumfix for Crumafix in E
	Reverse order	Patrun for Paturn in G Karbilno for Karbinol in E
	Intrusion of one phoneme	Bombet for Bomet in G Entom for Etom in E
Major error	Wrong decoding of more than one grapheme	Marilt for Magril in G Karabol for Karbinol in E
	Another word read instead of target word	oben for Oboe in G patient for pianist in E
	Word read instead of nonword target	Biss for Bix in G Jet for Jeto in E

Part II

Publications and submitted manuscript

Scientific Studies of Reading, 00:1–10, 2013

ISSN: 1088-8438 print/1532-799X online
DOI: 10.1080/10888438.2013.857673

Routledge
Taylor & Francis Group

RESEARCH REPORT

The Transition From Sublexical to Lexical Processing in a Consistent Orthography: An Eye-Tracking Study

Anne K. Rau
University of Tuebingen

Korbinian Moeller
Knowledge Media Research Center Tuebingen, Germany

Karin Landerl
University of Graz

We studied the transition in predominant reading strategy from serial sublexical processing to more parallel lexical processing as a function of word familiarity in German children of Grades 2, 3, 4, and adults. High-frequency words, low-frequency words, and nonwords of differing length were embedded in sentences and presented in an eye-tracking paradigm. The size of the word length effect was used as an indicator of serial sublexical decoding. When controlling for the generally higher processing times in younger readers, the effect of length over reading development was not direct but modulated by familiarity: Length effects were comparable between items of differing familiarity for Grade 2, whereas from Grade 3, length effects increased with decreasing familiarity. These findings suggest that Grade 2 children apply serial sublexical decoding as a default reading strategy to most items, whereas reading by direct lexical access is increasingly dominant in more experienced readers.

Reading is an indispensable yet complex skill that becomes smooth and automatized only with experience. Most theories of reading development have typically emphasized the distinction between a rule-based decoding procedure and a lexical retrieval procedure. According to Share's (1995) self-teaching hypothesis, the process of phonological recoding serves as an item-based mechanism for the buildup of word-specific representations. Particularly, reading of unfamiliar

Correspondence should be sent to Anne K. Rau, Department of Psychology, University of Tuebingen, Schleichstrasse 4, 72076 Tuebingen, Germany. E-mail: anne.rau@uni-tuebingen.de

words is claimed to depend on phonological recoding to a large extent, whereas recognition of familiar words is claimed to require relatively little phonological assistance (Share, 2008). Word-specific orthographic representations appear to be acquired after a relatively small number of successful decoding encounters (e.g., de Jong & Share, 2007; Reichle & Perfetti, 2003).

Studies involving consistent orthographies have often used the word length effect (i.e., the systematic increment in processing time associated with increasing number of letters) to characterize the course of this transition from phonological recoding to lexical reading (e.g., Zoccolotti, de Luca, di Filippo, Judica, & Martelli, 2009; Zoccolotti et al., 2005). The word length effect has been found to generally decrease with increasing reading experience (typical readers: e.g., Huestegge, Radach, Corbic, & Huestegge, 2009; Spinelli et al., 2005; Zoccolotti et al., 2009; Zoccolotti et al., 2005; dyslexic readers: di Filippo, de Luca, Judica, Spinelli, & Zoccolotti, 2006; Hawelka, Gagl, & Wimmer, 2010; Hutzler & Wimmer, 2004; Martens & de Jong, 2006; Spinelli et al., 2005). Apart from reading experience, the word length effect also interacts with lexicality and frequency. Thus, the length effect has been shown to be larger in nonwords than in words (e.g., de Luca, Borrelli, Judica, Spinelli, & Zoccolotti, 2002; Martens & de Jong, 2006; Weekes, 1997; Zoccolotti et al., 2009) and larger in low-frequency words than in high-frequency words (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Dürrwächter, Sokolov, Reinhard, Klosinski, & Trauzettel-Klosinski, 2010; Zoccolotti et al., 2009). Following the line of argument that the size of the word length effect is indicative of the degree of serial sublexical decoding, the finding of stronger word length effects in less experienced readers has been taken to reflect the fact that they rely on sequential processing to a larger degree than more experienced readers (e.g., van den Boer, de Jong, & Haentjens-van Meeteren, 2013; Zoccolotti et al., 2009). Similarly, the fact that the word length effect is typically larger in nonwords than in words and larger in low-frequency words than in high-frequency words has been interpreted to result from the fact that the recognition of less familiar items requires a larger degree of phonological processing than the recognition of more familiar items (e.g., Share, 1995, 2008).

In the present study the transition in predominant reading strategy from serial sublexical processing to more lexical processing was examined in German, a consistent orthography, by investigating the interaction between word length and word familiarity in different age and proficiency levels (Grades 2, 3, and 4, and young adults). The length effect was measured as the difference in processing time between long and short target items. Word familiarity was used as a unifying concept to include two factors that are commonly varied independently, namely, lexicality and frequency. Specifically, high-frequency words, low-frequency words, and nonwords of different length were presented as targets.

Two important extensions were made to the existing literature: First, using eye-tracking allowed for online recording of word processing parameters on a single word basis. Target words were presented within meaningful sentences to mimic natural reading as closely as possible. We report gaze duration[1] (i.e., the sum of all fixation durations on a word during first-pass reading) as a measure of processing time, which is assumed to encompass all stages of word recognition including meaning activation (Juhasz & Rayner, 2003; Reichle, Pollatsek, Fisher, & Rayner, 1998). Second, by performing a *z*-transformation on gaze durations, we controlled for

[1]Gaze duration was highly correlated with the number of fixations in first-pass reading for participants of all groups; correlation coefficients ranged from .86 for long nonwords through .96 for short low-frequency words (all *p*s < .01). We therefore refrained from reporting number of fixations/saccades.

potential effects of overadditivity. As pointed out by Faust, Balota, Spieler, and Ferrano (1999), the effect due to any experimental manipulation will naturally be larger for subjects with relatively slow responses than for subjects with relatively fast responses. Therefore, comparing specific effects (e.g., length effects) between groups of fundamentally different performance levels may produce spurious interactions. This approach has been applied in a few earlier studies on reading development (di Filippo et al., 2006; van den Boer et al., 2012; Zoccolotti et al., 2009), but to our knowledge, the present study is the first to compare the development of length effects as a function of familiarity, which analyzes z-transformed eye-tracking data.

We expected strong length effects for all word types in Grade 2 readers. For the more advanced readers of Grades 3 and 4, we expected a continuous decline in length effect from nonwords, to low-frequency words, through high-frequency words. For skilled adult readers, we expected only small length effects for high- and low-frequency words but clear length effects for nonwords. No fundamental differences between analyses using raw and standardized processing times were expected for such comparisons within groups. However, we were particularly interested in uncovering possible effects of overadditivity in the context of between-group comparisons of length effects.

METHOD

Participants

All 63 participants were native speakers of German with adequate reading ability and normal or corrected-to-normal vision. Four different groups of readers were assessed: children of Grades 2 ($n = 14$), 3 ($n = 18$), and 4 ($n = 13$), and young adults ($n = 18$; mean ages were 7;10, 8;8, 10;2, and 24;6 years, respectively). Data from six additional children were excluded because of reading difficulties.

Apparatus

Eye movements were recorded using an EyeLink 1000 eye-tracker (SR Research, Toronto, Canada) with a sampling rate of 1000 Hz. Single-line sentences were presented on the centre line of a 20-in. monitor (120 Hz refresh rate, 1024 × 768 resolution). Standard 9-point calibration at the beginning of the experiment ensured a spatial resolution of less than 0.5 degrees of visual angle.

Materials and Design

There were 72 experimental sentences, each containing a target stimulus of varying familiarity and length.[2] Twenty-four high-frequency words (mean occurrence of 140/million according

[2]We did not experimentally control for words immediately preceding and following the targets. However, both length and frequency of words n–1 and n+1 were fairly comparable between different experimental conditions. Between seven and 10 of 12 words n–1 and n+1 per condition were very high frequency in terms of CELEX frequency counts (occurrences > 1000/million; these were words such as direct and indirect articles, conjunctions such as und [and], conjugated

to CELEX; Baayen, Piepenbrock, & Gulikers, 1995) served as items of high familiarity, 24 low-frequency words (mean occurrence of 2.6/million) as items of medium familiarity, and 24 word-like nonwords as items of low familiarity. The nonwords were obtained by exchanging one or two syllables of low-frequency words not used in the present experiment with a letter sequence preserving the original phonotactic structure (e.g., *Birsul* was derived from *Konsul*). Half of the items of each category were short (3 to 6 letters), half were long items (7 to 10 letters).[3] Frequencies within item categories were kept constant across word length as far as possible. Length and frequency of targets were thus uncorrelated ($r = .07$, $p = .65$). In line with the procedure of other studies, target stimuli never appeared as sentence-initial or sentence-final words (e.g., Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner, Kambe, & Duffy, 2000). Nonwords were mostly used as proper names (e.g., "Oskar und **Ogo** sind beste Freunde"; "Oscar and **Ogo** are best friends"). Sentences were displayed using the unproportional font Courier New (bold, font size 14 point) in black against a white background. Participants were seated approximately 70 cm away from the monitor in a dimly lit room. The order of the sentences was pseudo-randomized within subjects to make sure no more than two sentences including a nonword appeared after another. There were four different orders between subjects.

Procedure

Participants were seated in front of the screen with their forehead leaned against a forehead rest. They were instructed to read the sentences aloud (allowing to check for reading accuracy) at their normal pace without making mistakes and to look at a cross at the lower right corner of the screen after having completed reading. They were informed that some sentences contained "made-up words." During six practice sentences, errors were corrected by the experimenter.

At the beginning of each sentence, there was a small smiley on the left side of the centre line that participants had to fixate. When fixation of the smiley was identified by the eye-tracking device within 5,000 ms from trial onset, the following sentence was presented with its first word appearing at the location of the smiley. When no fixation on the smiley was detected, the 9-point calibration cycle was repeated. Sentences disappeared upon fixation of the small cross at the bottom right of the screen. Voice recording allowed for checking reading accuracy afterward. The experimental session lasted between half an hour for adult subjects and an hour for the youngest children.

verb forms of high-frequency verbs such as geht [goes], sagt [says], hat [has]). The only condition with a lesser amount of very high-frequency words was the word immediately preceding short nonwords. Because many of the nonwords were used as proper names in the sentences, the words preceding short nonwords were often familiar first names (e. g., Paul, Linda), which are not listed in CELEX. Similarly, word length of words n–1 and n+1 was mostly comparable for most conditions (namely, approximately four letters/word) apart from the words preceding long high-frequency words and short nonwords (5.1 and 5.3 letters/word, respectively). On the whole, words n-1 and n+1 were fairly comparable in terms of frequency and length between all conditions with the exception of n–1 for short nonwords. However, because there was nothing peculiar about processing times for short nonword targets, we do not think that the longer and less frequent words in position n–1 necessarily exerted a detrimental influence on processing times of these words.

[3] Stimuli are available on request.

RESULTS

Target Word Reading Accuracy

In line with previous research on reading development in consistent orthographies (e.g., Frith, Wimmer, & Landerl, 1998; Seymour, Aro, & Erskine, 2003), reading accuracy was generally high for all reading groups (mean percentages of correct readings ranged mostly between 90 and 100%, except for long nonwords for which accuracy ranged from 71% for Grade 2 children to 94% for adults). By excluding incorrect readings from our analysis, we were able to ensure that the processing times reflected in children's eye movements were not negatively influenced by reading failure.

Target Word Processing Times

Fixations shorter than 80 ms were excluded from analysis (see also Blythe, Liversedge, Joseph, White, & Rayner, 2009; Hawelka et al., 2010, for a similar procedure). Figure 1 presents the mean gaze durations for the three item types separately for each age group. Two repeated-measures analyses of variance (ANOVAs) were computed, one on raw gaze duration and one on *z*-transformed gaze duration. Individual means and standard deviations of each participant

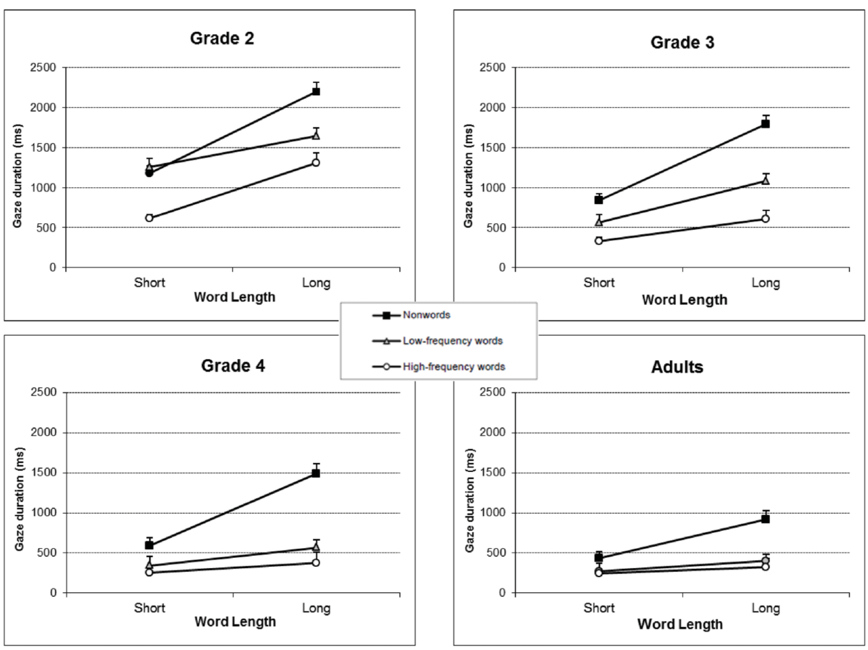


FIGURE 1 Word length effects as a function of word familiarity for the four groups of readers. *Note.* Bars indicate standard errors.

TABLE 1
4 × 3 × 2 Analyses of Variance on Raw and Standardized Gaze Durations

	Gaze Duration			z-Transformed Gaze Duration		
Factor	*df*	*F*	η_p^2	*df*	*F*	η_p^2
Reading Experience	3	25.09***	.56	3	1.38 *ns*	—
Familiarity	2	292.24***	.83	2	670.00***	.92
Length	1	468.69***	.89	1	838.38***	.93
Reading Experience × Familiarity	3	11.26***	.36	3	17.75***	.47
Reading Experience × Length	6	22.45***	.53	6	.53 *ns*	—
Familiarity × Length	2	45.91***	.44	2	122.13***	.67
Reading Experience × Familiarity × Length	6	2.42*	.11	6	4.97***	.20

$^{*}p < .05$. $^{***}p < .001$.

were used to perform *z*-transformation. For raw gaze duration, all main effects as well as all interactions were highly significant (Table 1, upper part).

For *z*-transformed gaze duration, results were mostly replicated (Table 1, lower part) with two exceptions: First, by definition, the effect of reading experience was no longer significant. Second, the two-way interaction of reading experience and length was no longer significant either. All significant main effects and two-way interactions have been described previously and are therefore not be discussed further. Of importance, the two-way interactions were qualified by a significant Reading Experience × Familiarity × Length interaction in both analyses.

To disentangle this interaction we first computed the interaction of familiarity and length separately for each group (Table 2). For raw gaze duration, the Familiarity × Length interaction was highly significant for all groups but the youngest for which it was marginally significant. When

TABLE 2
Mean Length Effects for Each Level of Familiarity (*SD*) and Statistical Characteristics of the Familiarity × Length Interaction for the Four Groups

Group	*Length Effect High-Freq Words*	*Length Effect Low-Freq Words*	*Length Effect Nonwords*	*F*	*p*	η_p^2
			Gaze duration (ms)			
Grade 2	693 (702)	390 (573)	1018 (255)	3.91	.06	.23
Grade 3	277 (176)	521 (242)	950 (484)	20.04	< .001	.54
Grade 4	119 (67)	223 (100)	899 (480)	27.62	< .001	.70
Adults	74 (67)	130 (45)	484 (159)	82.38	< .001	.83
			Standardized gaze duration (*z* values)			
Grade 2	.61 (.50)	.44 (.52)	1.05 (.36)	5.64	< .01	.30
Grade 3	.35 (.20)	.70 (.36)	1.24 (.55)	23.96	< .001	.59
Grade 4	.20 (.11)	.38 (.16)	1.46 (.56)	46.73	< .001	.80
Adults	.24 (.20)	.42 (.17)	1.47 (.21)	224.62	< .001	.93

Note. freq = frequency.

considering z-transformed gaze duration, the Familiarity × Length interaction was significant for all groups including the youngest.

For each group of readers, we subsequently computed the length effects for each level of familiarity (see Table 2 for means; note that length effects for all word types differed significantly from zero for all groups, all ps $< .05$) and submitted these to repeated-measures ANOVAs comparing the three length effects separately for each participant group. In Grade 2 children, pairwise comparisons indicated the length effect to be significantly larger for nonwords than for low-frequency words (gaze duration: $p < .01$, z-transformed gaze duration: $p = .001$), whereas the length effect for high-frequency words did not differ significantly from the length effect for low-frequency words and nonwords (all ps $> .10$). In the case of Grade 3 children, Grade 4 children, and adults, length effects increased significantly from high- to low-frequency words, and from low-frequency words to nonwords (gaze duration: all ps $< .05$, z-transformed gaze duration: all ps $\leq .01$).

To characterize how the length effect for the three word types develops with increasing reading expertise, we computed the Reading Experience × Length interaction separately for each word type.

For gaze duration, the interaction of reading experience and length was highly significant for all types of words. For high-frequency words, post hoc Scheffé tests showed that Grade 2 children had a larger length effect for high-frequency words than all other groups (all ps $< .05$), whereas the more experienced readers did not differ significantly from another. For low-frequency words, Grade 3 children's length effect was significantly larger than that in adults ($p < .01$) and tended to be larger than that in Grade 4 children ($p = .07$). For nonwords, the length effect was comparable between children of all grades, whereas it was significantly less pronounced in adults than in children (all ps $< .05$).

For standardized gaze duration, the Reading Experience × Length interaction was also significant for all word types (F values and effect sizes were lower than for gaze duration, though). For high-frequency words, post hoc Scheffé tests largely confirmed the findings from the analysis on gaze duration; Grade 2 children's length effect for high-frequency words was larger than that in adults and Grade 4 children (both ps $< .01$) and tended to be larger than that in Grade 3 children ($p = .09$). Also largely consistent with what was found for gaze duration, length effects for low-frequency words were mostly comparable between groups, apart from a trend for a larger length effect in Grade 3 children as compared to Grade 4 children ($p = .09$). For nonwords, results differed from those for gaze duration: Post hoc Scheffé tests revealed length effects to be largely comparable between groups. The only trend pointed in a somewhat surprising direction: The length effect for nonwords tended to be smaller in Grade 2 children than in adults ($p = .08$).

In summary, consistent over analyses on both raw and standardized gaze duration, length effects for words of differing familiarity were largely comparable for the least experienced readers but increased with decreasing word familiarity for the more experienced readers. The two analyses also gave similar results as far as the between-group comparison of length effects for high- and low-frequency words were concerned. An important difference between the two analyses was found for the within-group comparison of nonword length effects: Whereas the length effect for nonwords was significantly smaller in adults than all children's groups when considering raw gaze duration, length effects did not differ significantly between groups when considering standardized gaze duration.

DISCUSSION

The current study aimed to investigate the transition in predominant reading strategy from serial sublexical processing to more parallel lexical processing as a function of word familiarity in German children of Grades 2, 3, 4, and adults. High-frequency words, low-frequency words, and nonwords of differing length were embedded in sentences and presented in an eye-tracking paradigm. The size of the word length effect was used as an indicator of serial sublexical decoding. We accounted for the large differences in general reading ability between the different age groups by analyzing standardized processing times alongside to analyzing raw processing times.

The most important result, which was found across both analyses, was that the effect of length over reading experience was not direct but modulated by word familiarity. Specifically, length effects for words of differing familiarity were largely comparable for the least experienced readers but increased with decreasing word familiarity for the more experienced readers.

Grade 2 children showed length effects which were fairly comparable between words of differing familiarity (i.e., length effects differed neither between high-frequency words and nonwords nor between high- and low-frequency words). This least experienced group of readers appeared to apply serial sublexical decoding as a sort of default reading strategy to most words. An interesting case was their reading of low-frequency words: For short low-frequency words, their processing times were equally slow as for short nonwords, a fact that is indicative of serial sublexical decoding. For long low-frequency words, however, their processing times were clearly lower than for long nonwords, a fact that suggests a certain amount of lexical processing. We argue that Grade 2 readers *started* reading low-frequency words by serial decoding, *followed* by lexical access. In the case of short items, lexical access was no faster than the decoding procedure (reading times were therefore identical for low-frequency words and nonwords). Yet long low-frequency words were long enough for lexical access to be faster than full serial decoding. Alternatively, after having decoded the initial part of long low-frequency words, children may have simply guessed the remaining part. It is thus likely that Grade 2 readers read long low-frequency words (and probably also long high-frequency words) by a strategy best described as partial decoding. The fact that short high-frequency words were processed much faster than short low-frequency words and short nonwords suggests that these items were read by direct lexical access. This interpretation is consistent with previous findings showing early evidence for lexical access for highly familiar words in beginning readers of consistent orthographies (e.g., Zoccolotti et al., 2005).

The more experienced readers from Grade 3 showed the predicted continuous decline in length effect from nonwords, low-frequency words, through high-frequency words. The extent to which these sufficiently skilled readers relied on serial sublexical decoding thus clearly increased with decreasing word familiarity. Grade 4 readers showed a remarkably similar pattern to that of adults with significant but less marked length effects for high- and low-frequency words and strong length effects for nonwords. In fact, even Grade 3 readers showed a virtually competent word reading profile which differed from that of skilled adults only quantitatively. The fact that Grade 3 children showed a clearly smaller length effect for high-frequency than for low-frequency words and nonwords, combined with the fact that their length effect for high-frequency words was no larger than that found for Grade 4 children and adults clearly indicates competent word recognition in the case of such highly familiar words. The trend for a decreasing length effect from

Grade 3, Grade 4 through adults as found for low-frequency words reflects the increasing number of words for which detailed lexical representations are directly available in more experienced readers.

We employed the length effect as an indicator of serial sublexical decoding. It should be noted that besides reflecting a decreasing importance of serial sublexical word processing in favour of more parallel lexical processing, the developmental decrease in length effect could merely reflect an increase in efficiency of serial sublexical processing (Whitney & Cornelissen, 2005). There was indeed evidence in the data supportive of this idea: The length effect for nonwords was comparable for readers of all groups in the analysis using standardized processing times—the clear decrease in length effect for nonwords in the analysis using raw processing times is thus likely to be caused by such an increase in efficiency of serial sublexical processing. However, it is important to note that the two processes are not mutually exclusive: The increase in efficiency of serial sublexical recoding and the transition from predominant use of this sublexical reading strategy to predominant use of lexical word recognition are likely to develop in parallel.

ACKNOWLEDGMENTS

Many thanks to Jessica Geier and Sandra Hillian-Tress for helping with data acquisition and preparation.

REFERENCES

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database, Release 2* [CD-ROM]. Philadelphia, PA: Lingusitic Data Consortium, University of Pennsylvania.

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.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, *49*, 1583–1591.

de Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading*, *11*, 55–71.

de Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye-movement study of developmental dyslexia. *Brain and Language*, *80*, 617–626.

di Filippo, G., de Luca, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2006). Lexicality and stimulus length effects in Italian dyslexics: the role of the overadditivity effect. *Child Neuropsychology*, *12*, 141–149.

Dürrwächter, U., Sokolov, A. N., Reinhard, J., Klosinski, G., & Trauzettel-Klosinski, S. (2010). Word length and word frequency affect eye movements in dyslexic children in a regular (German) orthography. *Annals of Dyslexia*, *60*, 86–101.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferrano, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*, 777–799.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading*, *2*, 31–54.

Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition*, *115*, 367–379.

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, *49*, 2948–2959.

Hutzler, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language*, *89*, 235–242.

Juhasz, B., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1312–1318.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*, 262–284.

Martens, V. E. G., & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language*, *98*, 140–149.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology*, *53A*, 1061–1080.

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading*, *7*, 219–237.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*, 151–218.

Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin*, *134*, 584–615.

Spinelli, D., de Luca, M., di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length effect in word naming in reading: Role of reading experience and reading deficit in Italian readers. *Developmental Neuropsychology*, *27*, 217–235.

van den Boer, M., de Jong, P. F., Haentjens-van Meeteren, M. M. (2012). Lexical decision in children: Sublexical processing or lexical search? *The Quarterly Journal of Experimental Psychology*, *65*, 1214–1228.

van den Boer, M., de Jong, P. F., Haentjens-van Meeteren, M. M. (2013). Modelling the length effect: Specifying the relation with visual and phonological correlates of reading. *Scientific Studies of Reading*, *17*, 243–256.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latencies. *The Quarterly Journal of Experimental Psychology*, *50A*, 439–456.

Whitney, C., & Cornelissen, P. (2005). Letter-position encoding and dyslexia. *Journal of Research in Reading*, *28*, 274–301.

Zoccolotti, P., de Luca, M., di Filippo, G., Judica, A., & Martelli, M. (2009). Reading development in an orthographically regular language: Effects of length, frequency, lexicality and global processing ability. *Reading and Writing*, *22*, 1053–1079.

Zoccolotti, P., de Luca, M., di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language*, *93*, 369–373.

Journal of Experimental Child Psychology 130 (2015) 92–105

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently

Anne K. Rau [a,*], Kristina Moll [b,c], Margaret J. Snowling [b,d], Karin Landerl [a,e]

[a] Department of Psychology, University of Tübingen, 72076 Tübingen, Germany
[b] Department of Psychology, University of York, Heslington, York YO10 5DD, UK
[c] Department of Child and Adolescent Psychiatry, Psychosomatics, and Psychotherapy, Ludwig Maximilians University Munich, 80337 München, Germany
[d] Department of Experimental Psychology and St. John's College, University of Oxford, Oxford OX1 3JP, UK
[e] Department of Psychology, University of Graz, 8010 Graz, Austria

ARTICLE INFO

Article history:
Received 5 June 2014
Revised 15 September 2014

Keywords:
Word recognition
Eye movements
Length effect
Frequency effect
Cross-linguistic
Orthographic consistency

ABSTRACT

The current study investigated the time course of cross-linguistic differences in word recognition. We recorded eye movements of German and English children and adults while reading closely matched sentences, each including a target word manipulated for length and frequency. Results showed differential word recognition processes for both developing and skilled readers. Children of the two orthographies did not differ in terms of total word processing time, but this equal outcome was achieved quite differently. Whereas German children relied on small-unit processing early in word recognition, English children applied small-unit decoding only upon rereading—possibly when experiencing difficulties in integrating an unfamiliar word into the sentence context. Rather unexpectedly, cross-linguistic differences were also found in adults in that English adults showed longer processing times than German adults for nonwords. Thus, although orthographic consistency does play a major role in reading development, cross-linguistic differences are detectable even in skilled adult readers.

* Corresponding author.
E-mail address: anne.rau@uni-tuebingen.de (A.K. Rau).

http://dx.doi.org/10.1016/j.jecp.2014.09.012
0022-0965/

Introduction

The current study aimed to investigate cross-linguistic differences in word recognition in an eye-tracking paradigm. In particular, we examined indicators of small-unit processing (length effects) and large-unit processing (frequency effects) in child and adult readers of the consistent German and the inconsistent English orthographies.

Alphabetic orthographies differ with respect to the consistency with which letters map onto sounds, and it has long been established that reading development progresses more slowly in inconsistent orthographies compared with consistent orthographies (e.g., Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013; Frith, Wimmer, & Landerl, 1998; Seymour, Aro, & Erskine, 2003; Öney & Goldman, 1984). This difference is not only quantitative but also qualitative; beginning readers of inconsistent orthographies such as English have particular difficulties in phonological decoding. Thus, nonword reading is relatively poor in beginning readers of inconsistent orthographies (e.g., Frith et al., 1998; Wimmer & Goswami, 1994). Even though factors such as a language's syllabic complexity, method of reading instruction, and age at onset of formal instruction have all been shown to exert an influence on the development of reading, the crucial factor in explaining the differences in reading acquisition appears to be orthographic consistency (e.g., Aro & Wimmer, 2003; Landerl, 2005; Seymour et al., 2003).

In the context of psycholinguistic grain size theory (Ziegler & Goswami, 2005), the lower rate at which reading development progresses in inconsistent orthographies such as English is explained by the need to develop reading strategies targeting psycholinguistic units at a variety of grain sizes. For children learning to read a consistent orthography, decoding at the smallest linguistic grain size of the phoneme is sufficient during the early phase of reading development. However, this strategy is notoriously unreliable in English, forcing young readers to develop a variety of reading strategies using different linguistic grain sizes such as onsets, rimes, syllables, and whole words in order to cope with the complexities of the English orthography. Therefore, it appears that reading acquisition in an inconsistent orthography is delayed for two reasons: (a) because of the need to develop a number of reading strategies targeting different grain sizes and (b) because both the acquisition and the successful application of grapheme–phoneme correspondence rules are harmed by their inherent inconsistent nature.

It is crucial to note that even early word recognition in consistent orthographies is not *entirely* reliant on small grain sizes (e.g., Burani, Marcolini, & Stella, 2002; Davies, Cuetos, & Glez-Seijas, 2007); conversely, readers of less consistent orthographies are far from being *entirely* reliant on larger grain sizes (e.g., Duncan, Seymour, & Hill, 1997; Goswami, Ziegler, Dalton, & Schneider, 2001). Indeed, differences are relative rather than absolute. To properly test the relative importance of grain size across languages, cross-linguistic studies are of special importance. Following this strategy, a growing number of studies have directly compared reading in different orthographies.

First, nonword reading has consistently been reported to be better in consistent orthographies than in inconsistent orthographies (e.g., Aro & Wimmer, 2003; Frith et al., 1998; Landerl, 2000; Mann & Wimmer, 2002; Seymour et al., 2003; Thorstad, 1991), reflecting the fact that serial decoding of small-unit grapheme–phoneme correspondences is more readily available to beginning readers of consistent orthographies.

Second, a high percentage of reading errors in beginning readers of less consistent orthographies constitute refusals and word substitutions for both words and nonwords (e.g., Ellis & Hooper, 2001; Frith et al., 1998; Seymour et al., 2003), suggesting an inability to apply small-unit grapheme–phoneme correspondences. In contrast, beginning readers of more consistent orthographies have been reported to produce mainly nonword errors, often reflecting minor deviations from the correct pronunciation resulting from a small-unit decoding strategy (e.g., Ellis & Hooper, 2001; Seymour et al., 2003).

Third, direct comparisons have shown stronger word length effects in consistent orthographies than in inconsistent orthographies, also indicating stronger reliance on systematic decoding procedures (Ellis & Hooper, 2001; Goswami, Gombert, & Fraca de Barrera, 1998).

Finally, there is stronger evidence for the use of lexical analogies at the rime level in English than in more consistent orthographies. Thus, the facilitatory effect of orthographic neighborhood size was found to be larger in English children than in German children (Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Körne, 2003) and adults (Ziegler, Perry, Jacobs, & Braun, 2001), indicating more large-unit processing on the part of the English readers. In addition, length effects were found to be reduced in high-neighborhood words as compared with low-neighborhood words in English but not in German, suggesting that the German readers continued to process smaller units even though large-unit information was available. Likewise, English (and, in one study, French) children showed a clear benefit for nonwords that could be read by analogy to existing words (e.g., bicket; real-word analogue: ticket) as compared with phonologically identical nonwords for which no such analogy was evident (e.g., bikket), a benefit that was either less pronounced or absent in beginning readers of more consistent orthographies (Goswami, Porpodas, & Wheelwright, 1997; Goswami, Ziegler, Dalton, & Schneider, 2003; Goswami et al., 1998).

Readers of English have also been shown to exhibit stronger effects of frequency and lexicality than readers of more consistent orthographies (Frith et al., 1998; Landerl, Wimmer, & Frith, 1997; Paulesu et al., 2000; Wimmer & Goswami, 1994). Stronger effects of frequency/lexicality are to be expected for readers who predominantly process larger units than for readers who predominantly process smaller units.

Although differences between orthographies are probably most prominent during development, some studies suggest that they are still evident among skilled adult readers (Paulesu et al., 2000; Ziegler et al., 2001). Interestingly, in a cross-linguistic comparison of English and Italian readers, Paulesu et al. (2000) could even show significant differences in brain activation; while the Italian readers showed greater activation in brain regions associated with sublexical phonological processing, the English readers showed greater activation in brain regions associated with lexical and semantic processing.

Notwithstanding the emergent pattern of results, cross-linguistic studies face particular methodological problems associated with matching reading materials as well as participants. Equating reading materials is easier between languages that are closely related linguistically, such as English and German, than between more distant languages, such as English and Finnish. The Germanic languages English and German share numerous linguistic features with respect to phonology, morphology, and syntax, but they are different with respect to orthographic consistency, which is higher in German (Borgwaldt, Hellwig, & De Groot, 2005). This difference has been exploited by a number of studies presenting words that are similar or identical in the two languages with respect to spelling, pronunciation, meaning, and frequency of occurrence as well as nonwords derived from such cognates (see Landerl, 2005, for a review).

In the current study, we have also used this approach, but with the innovation of eye tracking to allow for online recording of word recognition processes as they occur. The analysis of eye movement measures is an established means of inferring moment-to-moment cognitive processes underlying reading (Just & Carpenter, 1980; Rayner, 1998), but it has not yet been exploited in direct cross-orthography research. To localize cross-orthographic effects within the time course of word recognition, three different measures of processing time are reported: (a) the duration of the first fixation on a target item, reflecting an early component of word recognition; (b) gaze duration (i.e., the summed duration of all fixations on a word prior to a saccade to another word), a central measure in that it encompasses all stages of word recognition, including meaning activation (Juhasz & Rayner, 2003; Reichle, Pollatsek, Fisher, & Rayner, 1998); and (c) rereading time (i.e., the sum of all fixation durations on a word occurring after first-pass reading), a late measure that incorporates only fixations after completion of initial processing of the word and that reflects the integration of a word into the sentence context (Cunnings & Clahsen, 2007; Kuperman & Van Dyke, 2011).

Target words were manipulated in length and frequency, with the length effect serving as an index of small-unit processing and the frequency effect serving as an index of large-unit processing. Alongside high- and low-frequency words, the word frequency factor also included word-like nonwords (for a similar procedure, see Chaffin, Morris, & Seely, 2001, and Lowell & Morris, 2014).

The current study also used an innovative method of matching the German and English participants. Because children in Germany begin formal reading instruction at an older age than those in the United Kingdom, it is only possible to match children on either age or grade level, thereby accepting differences in one of the two. In the current study, we matched German and English participants in two age groups (children and adults) on basic word reading ability by matching their mean gaze duration for the short high-frequency words of our target word set. Although the reading level match design as often used in dyslexia research (i.e., comparing typically developing children with older dyslexic children with limitations in cognitive processing) has rightly been criticized (e.g., Van den Broeck & Geudens, 2012), our approach applied a reading level match design to typical reading development in cross-linguistic research.

For gaze duration, the central eye-tracking measure of word processing, we expected to find stronger evidence for small-unit processing on the part of the German readers and stronger evidence for large-unit processing on the part of the English readers. Thus, we expected a stronger length effect in German and a stronger frequency effect in English, with more pronounced differences in children than in adults. We were also interested to see whether such cross-orthographic differences would already appear in the early measure of first fixation duration and whether they extend to the late processing measure of rereading time.

Method

Participants

From a relatively large pool of 118 participants, we excluded those with below average word reading ability as established by standardized reading fluency tests (German: SLRT II [Lese- und Rechtschreibtest] by Moll & Landerl, 2010; English: TOWRE [Test of Word Reading Efficiency] by Torgesen, Wagner, & Rashotte, 1999). Of the remaining 109 participants, we identified 25 German–English pairs of children matched on gaze duration for short high-frequency words as well as 16 German–English matched pairs of adults. All participants were native speakers of their respective languages and had normal or corrected-to-normal vision. German children were recruited from local elementary schools and attended Grades 2, 3, and 4. English children were recruited during a summer sports camp and had completed Grade 3, 4, 5, or 6 at the time of testing. Informed consent was provided by the parents, and children received a small gift in appreciation of their participation. The participating adults were students of the German and English universities at which the experiment took place. Participant characteristics are presented in Table 1.

Apparatus

Eye movements were recorded using an EyeLink 1000 tower mount eye tracker in Germany and an EyeLink 1000 desktop mount eye tracker in England (SR Research, Toronto, Canada). Single line sentences were presented in the center line of a 20-inch monitor (120-Hz refresh rate, 1024 × 768 resolution) in both countries. Although reading was binocular, only movements of the dominant

Table 1
Participant matching between the two orthographies.

	German	English	*t*	*p*
Children				
Age (months)	109 (12)	116 (13)	0.97[a]	.06[c]
School attendance (months)	31 (12)	48 (12)	5.05[a]	<.001[c]
Gaze duration for short HF words (ms)	317 (69)	310 (79)	0.34[a]	.73
Adults				
Age (months)	297 (54)	279 (48)	0.99[b]	.33
Gaze duration for short HF words (ms)	239 (45)	244 (55)	0.27[b]	.79

Note. Standard deviations are in parentheses. HF, high-frequency.
[a] $df = 48$.
[b] $df = 30$.
[c] Two-tailed.

eye were recorded. Standard 9-point calibration at the beginning of the experiment ensured a spatial resolution of less than 0.5° of visual angle.

Materials and design

The 72 target items were embedded in sentences closely matched for the two languages (e.g., "Der große Pianist gab ein Konzert in Rom"/"The great pianist gave a concert in Rome"). Syntactic structure was simple, and the language was appropriate for primary school children. Most of the words used in the sentences were German–English cognates (i.e., words with identical meaning and highly similar orthography and phonology in both languages).

Each sentence contained a target word carefully matched for the two orthographies in frequency and length (see Table 2 for target word characteristics; note that the differences between German and English [bigram] frequency counts were all nonsignificant, all *t*s < 1). High- and low-frequency targets were cognates with nearly identical CELEX frequency counts (Baayen, Piepenbrock, & Gulikers, 1995), number of letters, number of syllables, spelling, and meaning in the two orthographies. Nonword targets were derived from low-frequency cognates that were not used in the current experiment by replacing one or two syllables with a letter sequence preserving the original phonotactic structure (e.g., the nonword "Jeto" was created from the low-frequency cognate "Veto"/"veto"). Nonwords were matched on bigram frequency and were used as proper names within the sentences (e.g., "Der Millionär nannte seine Söhne Jeto und Jakob"/"The millionaire named his sons Jeto and Jacob"). Length categories were short items (three to five letters) and long items (six to eight letters). Frequencies within item categories were kept constant across word length as far as possible. Thus, length and frequency of our targets were uncorrelated for both orthographies (German: $r = .09$; English: $r = .05$).

The target word was never the first or last word of a sentence. Sentence order was pseudo-randomized to ensure that no more than two sentences containing a nonword appeared after another. There were four different orders between participants.

Procedure

After completion of the standardized test of reading fluency measuring the number of words and nonwords read correctly within a specified time, participants were seated in front of the screen with their forehead up against a forehead rest to restrict head movements. Participants were instructed to read the sentences aloud at their normal pace without making mistakes and to immediately look at a small cross in the lower right corner of the screen after having completed reading. They were told that some sentences would contain "made-up words" and that they should pay attention to read these words correctly. We chose to have participants read the sentences aloud in order to be able to check reading accuracy. Sentences were displayed using the unproportional font Courier New (bold, 14-point font size) in black against a white background. Participants were seated 70 cm away from the monitor in a dimly lit room.

Reading errors were corrected by the experimenter only during the six practice sentences. At the beginning of each sentence, participants needed to fixate a small smiley on the left side of the center

Table 2
Mean CELEX frequency counts (high-frequency/low-frequency) and mean bigram frequencies (nonwords).

Condition	Sample items	German	English
Short high-frequency words	Rot/red; Musik/music	115 (108)	126 (88)
Long high-frequency words	Winter/winter; Problem/problem	162 (97)	165 (147)
Short low-frequency words	Axt/axe; Kajak/kayak	2.1 (4.3)	3.1 (3.3)
Long low-frequency words	Bikini/bikini; Pianist/pianist	2.8 (1.3)	82.4 (1.6)
Short nonwords	Bix, Etom, Nibra	1159 (545)	1259 (573)
Long nonwords	Paturn, Fiapard, Karbinol	1613 (609)	1622 (577)

Note. Standard deviations are in parentheses.

line. When fixation of the smiley was identified by the eye-tracking device within 5000 ms of trial onset, the following sentence was presented with its first word appearing at the location of the smiley. When no fixation on the smiley was detected, the 9-point calibration cycle was repeated. Sentences disappeared as soon as the small cross at the bottom right of the screen was fixated. The experimenter noted down reading errors. In addition, a voice recorder was used to double-check critical passages. Child participants could take a break whenever needed. The experimental session lasted between half an hour for adult participants and an hour for the youngest children.

Results

Reading accuracy

Table 3 displays the percentage of reading errors for the four groups. Because excluding all incorrectly read target words from eye-tracking analyses would have resulted in quite substantial data loss, we categorized reading errors into minor and major. Minor errors were fairly close to the correct pronunciation and, thus, were included in eye-tracking analyses. Reading errors were considered as minor in the case of wrong stress assignment, reverse order of otherwise correctly identified phonemes, or misidentification (or omission/intrusion) of a maximum of one phoneme. The percentage of major errors is reported in parentheses in Table 3.

Consistent with predictions, overall reading accuracy was clearly higher in German children than in English children in the case of low-frequency words, $t(48) = 3.89$, $p < .001$, Cohen's $d = 1.10$, and nonwords, $t(48) = 3.55$, $p = .001$, $d = 1.01$. Moreover, the types of errors differed between the two orthographies; the dominant error type among the German children were minor errors (84.1% of total error rate), whereas the English children produced more comparable amounts of minor and major deviations (55.7% and 44.3% of total error rate) from the target item. Reading accuracy was generally high in both German and English adults.

Eye-tracking data screening

Fixations shorter than 80 ms were not analyzed. Target words were excluded when there was a major reading error, when there were problems with calibration accuracy, or when a participant failed to read the complete sentence. Targets that received no fixation in first-pass reading were excluded from analyses for all measures of processing time. Total data loss was 11.0% for German children, 13.6% for English children, 11.6% for German adults, and 11.1% for English adults.

Processing time

For children and adults separately, we computed $2 \times 3 \times 2$ analyses of variance (ANOVAs) with length (short or long) and frequency (high-frequency words, low-frequency words, or nonwords) as within-participants factors and orthography (German or English) as a between-participants factor. Dependent measures of processing time were (a) first fixation duration, (b) gaze duration, and (c) rereading time.[1] When the results of independent-samples t tests are reported, Bonferroni-corrected p values are two-sided throughout.

Children

Descriptive statistics are provided in Fig. 1, and ANOVAs for the three variables are reported in Table 4. The upper part of the table shows main effects of word length and word frequency as well as the interaction between the two. These effects have been reported previously for different

[1] Measures of processing time are typically closely related with the respective measures of number of fixations. The analyses for gaze duration, rereading time, and total reading time as reported below were also carried out for number of first-pass fixations, number of rereading fixations, and total number of fixations. Results largely replicated those reported for the measures of processing time with one exception, namely that the three-way interaction of orthography, length, and frequency, which was significant for adults' rereading time, was not significant for adults' number of rereading fixations, $F(2, 60) = 2.15$, $p = .13$.

Table 3
Error percentage per group and condition.

Condition	Children		Adults	
	German	English	German	English
Short high-frequency words	0.8 (0.8)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Long high-frequency words	0.4 (0.0)	1.6 (0.0)	0.6 (0.6)	0.0 (0.0)
Short low-frequency words	5.2 (1.6)	14.4 (7.2)	0.0 (0.0)	0.6 (0.6)
Long low-frequency words	2.8 (0.0)	14.8 (5.2)	0.6 (0.0)	0.0 (0.0)
Short nonwords	2.0 (0.8)	12.4 (6.4)	0.0 (0.0)	1.9 (0.6)
Long nonwords	9.6 (0.1)	23.6 (10.8)	2.5 (0.6)	3.1 (1.3)

Note. Percentages of major errors are in parentheses.

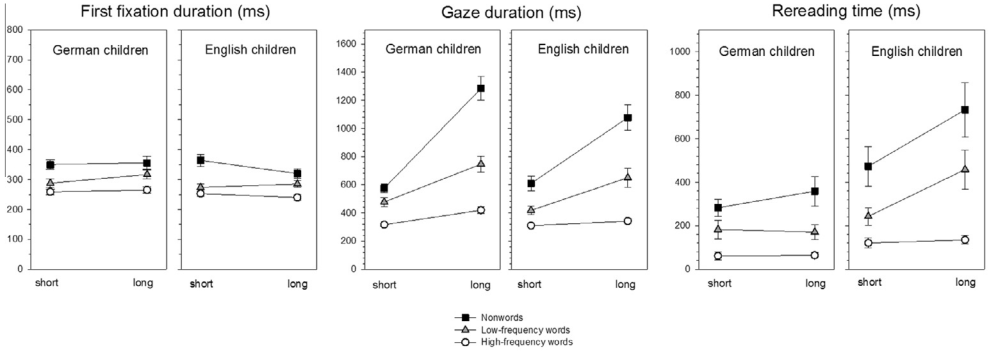


Fig. 1. Three measures of word processing time as a function of length and frequency for German and English children. Bars indicate standard errors.

orthographies (e.g., De Luca, Borrelli, Judica, Spinelli, & Zoccolotti, 2002; Huestegge, Radach, Corbic, & Huestegge, 2009; Martens & de Jong, 2006; Rau, Moeller, & Landerl, 2014; Weekes, 1997; Zoccolotti, De Luca, Di Filippo, Judica, & Martelli, 2009) and mainly provide a validation for our selection of reading materials. The highly significant length effects meant that gaze duration and rereading times were higher for long words than for short words (note that there was no length effect for first fixation duration), and the highly significant frequency effects meant that for all three variables processing times continuously increased from high-frequency words, to low-frequency words, to nonwords. The significant interactions between the two factors meant that length effects increased with decreasing frequency for gaze duration and rereading times, whereas for first fixation duration the length effect was significant for low-frequency words only.

In the current context, we were mainly interested in effects that included the crucial orthography factor. These effects are displayed in the lower part of Table 4.

For the early measure of processing time, *first fixation duration,* the predicted Orthography × Length interaction was marginally significant and indicated a minimally larger length effect in German children (14 ms) than in English children (−15 ms) (see Fig. 1, left panel). There was no Orthography × Frequency interaction and no three-way interaction for this early measure.

The predicted Orthography × Length interaction was more evident for the central measure of word recognition, *gaze duration* (Fig. 1, middle panel). Thus, the overall length effect was significantly more pronounced for German children than for English children (360 vs. 244 ms), $t(48) = 2.61$, $p = .01$, $d = 0.74$. Similar to the results for first fixation duration, we found no Orthography × Frequency interaction and no Orthography × Length × Frequency interaction.

For the late measure of word processing, *rereading time* (Fig. 1, right panel), there was a highly significant main effect of orthography, with longer rereading times for English children than for German children (361 vs. 187 ms). As for gaze duration, the interaction of orthography and length was highly

Table 4
2 × 3 × 2 ANOVAs on three different measures of word processing time: Children.

Effect	First fixation duration		Gaze duration		Rereading time	
	F	η^2_p	F	η^2_p	F	η^2_p
Length	.01	.00	184.52***	.79	17.57***	.27
Frequency	70.98***	.60	149.89***	.76	38.78***	.45
Length × Frequency	4.08*	.08	46.57***	.49	4.94**	.09
Orthography	1.30	.03	2.22	.04	8.29**	.15
Orthography × Length	3.73[a]	.07	6.79**	.12	10.12**	.17
Orthography × Frequency	0.34	.01	0.27	.01	3.38*	.07
Orthography × Length × Frequency	0.85	.02	1.99	.04	2.55[b]	.05

[a] $p = .06$.
[b] $p = .08$.
* $p \leqslant .05$.
** $p \leqslant .01$.
*** $p \leqslant .001$.

significant, but the direction was reversed; for rereading time, the length effect was clearly more pronounced for English children than for German children (163 vs. 22 ms), $t(48) = 3.18$, $p < .01$, $d = 0.90$. Table 4 also indicates a reliable Orthography × Frequency interaction. Although Fig. 1 suggests a stronger frequency effect among English children than among German children, the frequency effect was highly significant for both German children, $F(2, 48) = 25.26$, $p < .001$, $\eta^2_p = .51$, and English children, $F(2, 48) = 20.09$, $p < .001$, $\eta^2_p = .46$. At the same time, rereading time was significantly lower in German children than in English children for high-frequency words (62 vs. 128), $t(48) = 3.12$, $p < .01$, $d = 0.89$, low-frequency words (177 vs. 351 ms), $t(48) = 2.53$, $p < .05$, $d = 0.72$, and nonwords (321 vs. 603 ms), $t(48) = 2.50$, $p < .05$, $d = 0.71$. Importantly, the predicted two-way interactions were qualified by a marginally significant three-way interaction of orthography, length, and frequency. Separate analyses for the two orthographies showed that the Length × Frequency interaction was significant in English children, $F(2, 48) = 5.20$, $p < .01$, $\eta^2_p = .18$, but not in German children, $F(2, 48) = 1.17$, $p = .32$. As evident in Fig. 1, English children's length effect for high-frequency words was not significant ($t < 1$, *ns*), but it was significant for both low-frequency words, $t(24) = 3.36$, $p < .01$, $d = 0.67$, and nonwords, $t(24) = 3.36$, $p < .01$, $d = 0.67$.

To exclude the possibility that the generally higher rereading times of English children as compared with German children simply reflected their lower reading accuracy, we reran the children's analysis on rereading times on the correctly read items only (i.e., this time also excluding targets with minor errors). Importantly, results exactly mirrored those reported above.

In a final analysis, we wanted to examine whether the cross-orthographic effects observed for gaze duration and rereading time would cancel each other out. An additional ANOVA was run for total reading time, which constitutes the sum of these two measures and is thought of as a general comprehension measure (e.g., Boston, Hale, Kliegl, Patil, & Vasishth, 2008). Indeed, no effects involving orthography were found for this combined measure.

In summary, the length effect tended to be stronger in German children than in English children for first fixation duration. As predicted, the length effect was more pronounced in German children than in English children for gaze duration. Interestingly, the reverse pattern emerged for rereading time; in this late measure of processing time, for low-frequency words and nonwords, the length effect was more pronounced in English children. The combined measure of total reading time did not differ between child readers of the two orthographies, suggesting that the effects on first-pass and rereading times cancelled each other out.

Adults

Descriptive statistics are provided in Fig. 2, and ANOVAs for the three variables are reported in Table 5.

The upper part of Table 5 shows significant main effects of word length and word frequency for all three processing measures. The Length × Frequency interaction was significant for gaze duration only.

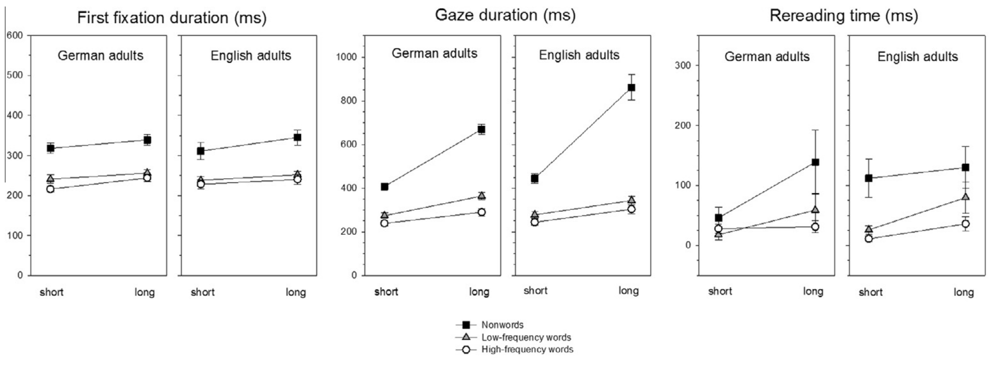


Fig. 2. Three measures of word processing time as a function of length and frequency for German and English adults. Bars indicate standard errors.

Table 5
2 × 3 × 2 ANOVAs on three different measures of word processing time: Adults.

Effect	First fixation duration		Gaze duration		Rereading time	
	F	η_p^2	F	η_p^2	F	η_p^2
Length	15.17[***]	.34	178.80[***]	.86	8.92[**]	.23
Frequency	79.63[***]	.73	287.74[***]	.91	17.10[***]	.36
Length × Frequency	0.55	.02	73.76[***]	.71	2.82[a]	.09
Orthography	0.00	.00	3.25[a]	.10	0.24	.01
Orthography × Length	0.02	.001	3.86[a]	.11	0.25	.008
Orthography × Frequency	0.11	.004	10.16[***]	.25	0.75	.02
Orthography × Length × Frequency	0.60	.02	6.72[**]	.18	4.18[*]	.12

[a] $p \leqslant .08$.
[*] $p \leqslant .05$.
[**] $p \leqslant .01$.
[***] $p \leqslant .001$.

Because the focus of the current study was on processing differences between readers of different orthographies, in the following we concentrate on effects involving the crucial orthography factor that are displayed in the lower part of Table 5.

For *first fixation duration*, there was no effect of orthography whatsoever. Fig. 2 (left panel) shows that early processing time was highly comparable between German and English adults.

For *gaze duration*, there was a marginally significant main effect of orthography, indicating a trend for longer first-pass processing times in English adults than in German adults (413 vs. 374 ms). Unlike what was found for children, the Length × Orthography interaction for gaze duration was only marginally significant with a trend for a stronger overall length effect for English adults than for German adults (181 vs. 135 ms). The significant Frequency × Orthography interaction indicated comparable gaze durations for German and English adults for both high- and low-frequency words (both ts < 1, *ns*) but longer gaze durations for English adults for nonwords (653 vs. 539 ms), $t(30) = 2.99$, $p < .01$, $d = 1.05$. Importantly, the marginally significant main effect of orthography and the (marginally) significant two-way interactions were qualified by a significant three-way interaction of length, frequency, and orthography (Fig. 2, middle panel). Independent-samples t tests revealed length effects to be comparable between German and English adults for high- and low-frequency words, ts(30) = 0.73 and 1.01, both *ns*, but the length effect for nonwords was significantly more pronounced in English adults than in German adults (417 vs. 262 ms), $t(30) = 2.74$, $p = .01$, $d = 0.97$.

Unlike in children, there was no main effect of orthography on *rereading time* (Fig. 2, right panel). Although neither the Orthography × Length interaction nor the Orthography × Frequency interaction was significant, there was a significant three-way interaction. When considering the two groups separately, the Length × Frequency interaction turned out to be significant for German adults, $F(2, 30) = 4.86$, $p < .05$, $\eta_p^2 = .25$, but not for English adults, $F(2, 30) = 1.32$, $p = .28$. The length effect for German adults was not significant for high-frequency words ($t < 1$) but was marginally significant for low-frequency words, $t(15) = 2.01$, $p = .06$, $d = 0.50$, and significant for nonwords, $t(15) = 2.19$, $p < .05$, $d = 0.55$. As evident in Fig. 2, standard errors for the nonword rereading times were remarkably high. A closer inspection of individual participants' means by condition indicated extremely high values for one English participant and two German participants (i.e., their rereading times for long nonwords were ∼600–700 ms, whereas those for all others ranged from 0 to just over 200 ms). When rerunning the ANOVA on rereading times excluding those three participants, the main effect of orthography was significant, $F(1, 27) = 6.71$, $p < .05$, $\eta_p^2 = .20$, indicating generally longer rereading times in English adults than in German adults (50 vs. 26 ms). This main effect was qualified by a significant Orthography × Frequency interaction, $F(2, 54) = 4.08$, $p < .05$, $\eta_p^2 = .13$. Post hoc *t* tests for independent samples showed that rereading times for German and English adults were comparable for high-frequency words (20 vs. 18 ms, $t < 1$) but were higher in English adults for both low-frequency words, $t(27) = 3.79$, $p = .001$, $d = 1.39$, and nonwords, $t(27) = 2.21$, $p < .05$, $d = 0.82$. Unlike for the analysis including participants with extreme nonword rereading times, there was no three-way interaction.

Again, we ran an additional ANOVA to uncover whether any cross-orthographic effects remained for *total reading time.* The only effect involving orthography that was significant was the Orthography × Frequency interaction, $F(2, 60) = 5.47$, $p < .01$, $\eta_p^2 = .15$. Independent-samples *t* tests showed that total reading times between German and English adults did not differ for either high-frequency words (294 vs. 297 ms, $t < 1$, *ns*) or low-frequency words (357 vs. 364 ms, $t < 1$, *ns*). For nonwords, however, total reading times were higher among English adults than among German adults (775 vs. 631), $t(30) = 2.12$, $p < .05$, $d = 0.75$.

In summary, first fixation duration did not differ between German and English adults. For gaze duration, German and English adults' length effects for high- and low-frequency words did not differ either. For nonwords, however, the length effect was more pronounced in English adults than in German adults. For rereading time, German adults showed a marginally significant length effect for low-frequency words and a significant length effect for nonwords, whereas English adults showed no differential effect of length on words of differing frequency. When excluding three participants with extremely high rereading times for nonwords, rereading times were higher in English adults than in German adults for low-frequency words and nonwords. Importantly, total reading times were comparable between adults of the two orthographies for both high- and low-frequency words but were higher in English adults than in German adults for nonwords.

Discussion

The current study set out to extend the empirical evidence on differences in word recognition processes as a function of orthographic consistency. Clearly, direct cross-linguistic studies are the preferred way to compare reading across different orthographies. At the same time, such cross-linguistic comparisons face important methodological problems relating to the difficulty of equating reading material across languages, matching children across countries with differing educational systems, and choosing an experimental task reflecting the natural process of reading.

First, comparing the linguistically closely related orthographies of German and English ensured that the presented reading material could be matched well (but note that even careful matching of the reading material does not imply that the general level of difficulty between conditions is equally comparable between orthographies). Second, we matched participants in terms of reading experience on the basis of their ability to read simple short high-frequency words. It can be assumed that this basic word category is processed via direct lexical access even by relatively young readers in the more consistent German orthography (Burani et al., 2002; Rau et al., 2014). By using this matching procedure, we accepted that the English children had a 17-month advantage in formal reading instruction. However, the current study focused on qualitative word processing differences rather than quantitative word processing differences as a function of orthographic consistency. Thus, we report the first study to apply eye tracking in a cross-linguistic reading paradigm. The chief advantage of eye movement recording was that it allowed for a detailed investigation of the time course of possible cross-orthographic differences in word processing. The current findings, hence, are more informative than earlier studies in single orthographies that have already suggested that there may be fine-grained differences in eye movement patterns depending on orthographic depth (e.g., Hutzler & Wimmer, 2004; Trauzettel-Klosinski et al., 2010).

Overall, we expected stronger evidence for small-unit processing on the part of the German readers and stronger evidence for large-unit processing on the part of the English readers—particularly for the central measure of word processing, gaze duration. We also expected differences between orthographies to be more pronounced in children than in adults.

A first indication that German and English children may process identical words using units of different sizes was provided by their distribution of minor versus major reading errors. The higher proportion of minor reading errors in German children than in English children is indicative of their stronger reliance on small-unit decoding.

As for processing times, different measures showed different effects of orthographic consistency over the time course of word processing. Because the predicted stronger length effects for German than for English just began to show for first fixation duration among children, orthographic consistency does not seem to have much influence on early word recognition processes. This only very slight cross-linguistic effect is actually in line with the fact that standard effects of length and frequency for the early first fixation duration measure were less pronounced in terms of effect size than for the later gaze duration measure. According to Perry, Ziegler, and Zorzi (2010), the components of early processing include basic letter identification, letter coding, and graphemic parsing. Because neither letter identification nor letter coding can be expected to differ between German and English, the difference in length effect reported here, albeit subtle, is most likely due to differences in graphemic parsing. Speculatively, this could be because the attentional window is larger in English than in German as a consequence of the necessity to process larger segments.

In line with predictions, children's gaze durations indicated stronger effects of word length in German than in English, suggesting stronger reliance on systematic phonological decoding in the consistent orthography compared with the inconsistent orthography during first-pass reading. Interestingly, the cross-linguistic effect for first-pass reading was reversed for the late rereading measure; for rereading time, the overall length effect was stronger in English children than in German children, particularly in the case of nonwords. Importantly, the generally higher rereading times and the stronger length effects on the part of the English children did not merely reflect their lower reading accuracy; excluding words with minor reading errors from analyses did not change these effects.

With regard to frequency, results were less conclusive than those for the length effect. Contrary to prediction, there was no cross-linguistic effect of frequency on children's gaze duration, and although the predicted interaction of orthography and frequency was significant for children's rereading times, it is hard to interpret this because rereading times increased with decreasing frequency in a similar fashion for children of both orthographies and rereading times were higher for English children than for German children for all types of words. Thus, although frequency plays a major role during the whole time course of children's word processing, the interplay of length and frequency late in children's word processing suggests that German children apply small- and large-unit processing in a comparable way for words of differing frequency, whereas English children apply more large-unit processing for high-frequency words and apply more small-unit processing for low-frequency words and nonwords. The influence of frequency on children's word recognition, therefore, is best understood as depending on both orthographic consistency and length.

On a merely speculative note, the surprisingly similar frequency effect in German and English children could result from frequency-weighted phonological representations; if representations in the phonological output lexicon are (similar to representations in the orthographic lexicon) frequency-weighted, one would expect a frequency effect even when word recognition is achieved by assembled phonology rather than addressed phonology.

In sum, although the outcome in terms of total word processing time was no different for German and English children, the means by which this outcome was achieved differed substantially. In first-pass reading, German children relied on small-unit phonological decoding to a larger extent than

English children. This explains not only their larger length effect in first-pass reading but also their much lower rereading times; when during the first pass a word is read by small-unit decoding (which is highly reliable in the consistent German orthography), there is not much need for rereading. English children, on the other hand, seemed to use units of larger grain size and more lexically based strategies during first-pass reading. Only when the subsequent sentence context did not confirm their first reading attempt (i.e., mainly when confronted with an unfamiliar low-frequency word or a nonword) did the eyes move back to the target item and a more systematic decoding procedure was applied. Previous cross-linguistic studies (e.g., Frith et al., 1998; Goswami et al., 2001, 2003; Landerl et al., 1997) were ill-equipped to detect such fine-grained processing differences because they mostly presented single words instead of sentences and used vocal reactions times instead of eye tracking (thereby not allowing for between-word regressions).

In line with our predictions, the reading performance of German and English adults was comparable with respect to reading accuracy and first fixation duration. For gaze duration, the main measure of word recognition, adult length effects were again comparable for high- and low-frequency words, but unexpectedly English adults showed a more pronounced length effect for nonwords than German adults. The most likely explanation for this is that adult readers of both orthographies were relying on small-unit processing when reading nonwords in the first pass but that the German readers are more experienced in this process due to greater practice in everyday reading. Readers of German are likely to be more used to reading longer words with complex morphology (e.g., compounds are typically spelled in one word in German, whereas they are typically spelled in two words in English; e.g., Schulbus vs. school bus). With the greater experience that skilled readers of German have in using small units and in reading long words, they were apparently more successful in dividing long nonwords into manageable units than their English counterparts.

With respect to late word recognition processes, there was evidence for more small-unit processing in nonwords (and, to a lesser degree, in low-frequency words) in German adults but not in English adults. Because neither the exclusion of three participants with extremely high rereading times for nonwords nor the analysis on the related measure on number of rereading fixations indicated a differential effect of length on adult readers of the two orthographies, we consider this strong rereading length effect for nonwords as shown by German adults, but not by English adults, to be less informative than another unexpected cross-linguistic finding in adult readers, namely that English adults showed higher total processing times than their German counterparts for nonwords (i.e., for the measure combining early and late word recognition processes). Importantly, because processing times of German and English adults throughout the whole time course of word recognition were comparable for both high- and low-frequency words, and cross-linguistic differences were confined to a first-pass reading difficulty for long nonwords and an overall processing difficulty for nonwords on the part of the English adults, it may be more appropriate to be speaking of a lexicality effect rather than a frequency effect in this case.

Thus, it appears that there are cross-orthographic word processing differences even between skilled adult readers when it comes to decoding nonwords. It is notable that no such difference in total processing time was found for children, a finding that may reflect the English samples' advantage with regard to length of formal reading instruction, the increased efforts regarding systematic teaching of phonics in British schools during recent years, or a combination of both these factors.

Conclusion

We found the predicted evidence for stronger length effects in German children's first-pass reading than in English children's first-pass reading. While frequency effects were more comparable between German and English children, English adults took longer to process nonwords than German adults. Thus, the current study showed cross-linguistic differences in the time course of word recognition during sentence reading between German and English children and adults. Children of the two orthographies did not differ in terms of total word processing time, but this equal outcome was achieved differently; German children relied on small-unit processing early in the course of word recognition, whereas English children did so particularly during rereading. Cross-linguistic differences were also found in adult readers. For both first-pass reading and total reading time, English adults showed longer processing times than German adults for nonwords. Thus, although orthographic consistency does play a major role in early reading and reading development, cross-linguistic differences are detectable even in skilled adult readers.

Acknowledgments

The second author has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013 under grant agreement ELDEL PITN-GA-2008-215961-2). We thank Rachel Kirmond and Calvin Wong for their support in eye movement recordings.

References

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*, 621–635.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database, Release 2 [CD–ROM]*. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Borgwaldt, S. R., Hellwig, F. M., & De Groot, A. M. B. (2005). Onset entropy matters: Letter-to-phoneme mappings in seven languages. *Reading and Writing, 18*, 211–229.

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam sentence corpus. *Journal of Eye Movement Research, 2*(1), 1–12.

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

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*, 1398–1407.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 225–235.

Cunnings, I., & Clahsen, H. (2007). The time-course of morphological constraints: Evidence from eye-movements during reading. *Cognition, 104*, 476–494.

Davies, R., Cuetos, F., & Glez-Seijas, R. M. (2007). Reading development and dyslexia in a transparent orthography: A survey of Spanish children. *Annals of Dyslexia, 57*, 179–198.

De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye-movement study of developmental dyslexia. *Brain and Language, 80*, 617–626.

Duncan, L. G., Seymour, P. H. K., & Hill, S. (1997). How important are rhyme and analogy in beginning readers? *Cognition, 63*, 171–208.

Ellis, N. C., & Hooper, A. M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics, 22*, 571–599.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading, 2*, 31–54.

Goswami, U., Gombert, J. E., & Fraca de Barrera, L. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics, 19*, 19–52.

Goswami, U., Porpodas, C., & Wheelwright, S. (1997). Children's orthographic representations in English and Greek. *European Journal of Psychology of Education, 3*, 273–292.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *Journal of Memory and Language, 45*, 648–664.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics, 24*, 235–247.

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research, 49*, 2948–2959.

Hutzler, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language, 89*, 235–242.

Juhasz, B., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 1312–1318.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language, 65*, 42–73.

Landerl, K. (2000). Influences of orthographic consistency and reading instruction on the development of nonword reading skills. *European Journal of Psychology of Education, 15*, 239–257.

Landerl, K. (2005). Reading acquisition in different orthographies: Evidence from direct comparisons. In M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 513–530). Hove, UK: Lawrence Erlbaum.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German–English comparison. *Cognition, 63*, 315–334.

Lowell, R., & Morris, R. (2014). Word length effects on novel words: Evidence from eye movements. *Attention, Perception, and Psychophysics, 76*, 179–189.

Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and English children. *Reading and Writing, 15*, 653–682.

Martens, V. E. G., & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language, 98*, 140–149.
Moll, K., & Landerl, K. (2010). *Lese- und Rechtschreibtest (SLRT II)*. Bern, Switzerland: Huber.
Öney, B., & Goldman, S. R. (1984). Decoding and comprehension skills in Turkish and English: Effects of the regularity of grapheme–phoneme correspondences. *Journal of Educational Psychology, 76*, 557–568.
Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., et al (2000). A cultural effect on brain function. *Nature Neuroscience, 3*, 91–96.
Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology, 61*, 106–151.
Rau, A. K., Moeller, K., & Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: An eye-tracking study. *Scientific Studies of Reading, 18*, 224–233.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.
Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*, 125–157.
Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174.
Thorstad, G. (1991). The effect of orthography on the acquisition of literacy skills. *British Journal of Psychology, 82*, 527–537.
Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency (TOWRE)*. Austin, TX: Pro-Ed.
Trauzettel-Klosinski, S., Koitzsch, A. M., Dürrwächter, U., Sokolov, A. N., Reinhard, J., & Klosinski, G. (2010). Eye movements in German-speaking children with and without dyslexia when reading aloud. *Acta Ophthalmologica, 88*, 681–691.
Van den Broeck, W., & Geudens, A. (2012). Old and new ways to study characteristics of reading disability: The case of the nonword-reading deficit. *Cognitive Psychology, 65*, 414–456.
Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latencies. *Quarterly Journal of Experimental Psychology A, 50*, 439–456.
Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition, 51*, 91–103.
Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3–29.
Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12*, 379–384.
Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*, 169–193.
Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., & Martelli, M. (2009). Reading development in an orthographically regular language: Effects of length, frequency, lexicality, and global processing ability. *Reading and Writing, 22*, 1053–1079.

Abstract

The current study compared eye movement patterns during word and sentence processing in a consistent and an inconsistent alphabetic orthography. German and English children and adults matched on word reading ability read matched sentences while eye movements were recorded. German children showed a more small-unit plodder-like reading style with more diligent first-pass reading and less re-reading time. English children showed a more large-unit explorer-like reading style with more skippings and more regressions. Similar cross-linguistic processing differences persisted in the adult readers. Orthographic consistency therefore impacts upon both local word recognition and global sentence processing, in both developing and skilled readers.

Same same, but different: Word and sentence reading in German and English

The present study investigated cross-linguistic differences in eye movement patterns in word and sentence processing in two alphabetic orthographies with different degrees of letter-sound consistency. We expected to find evidence for more small-unit bottom-up processing in readers of the consistent German orthography, and more large-unit top-down processing in readers of the inconsistent English orthography for both the local word and the global sentence processing level. In other words, in German readers, we expected to find an eye movement pattern which in dyslexia research has been described as a plodder reading style, while in English readers, we expected to find more of an explorer reading style (cf. Olson, Kliegl, Davidson, & Foltz, 1985).

Previous research on typically developing English-speaking readers of different ages has shown consistently that eye movement behaviour changes as a function of increasing reading experience: as reading skill increases, fixation durations and sentence reading times decrease, as does the number of fixations. While the probability for making more than one fixation on the same word decreases, the probability that a word is not fixated but skipped in first pass increases with reading skill.

Another indicator of reading efficiency is how often the reader needs to move the eyes against the normal reading direction and make regressive right-to-left saccades. For number or proportion of regressions, or regression probability, results have been slightly less consistent: while some studies showed a decrease in number of regressions in the first few years of reading instruction, constant numbers in higher grades, and again a decrease in adults (Buswell, 1922; Rayner, 1985), other studies showed a constant number of regressions among children of different ages, and only a decrease in adults (Blythe, Liversedge, Joseph, White, & Rayner, 2009; McConkie et al., 1991; Taylor, 1965). Yet another pattern was reported recently by Vorstius, Radach, and Lonigan (2014) who found an *increase* in the proportion of interword regressions from grade 1 to 5. A likely explanation for these seemingly inconsistent results is that former studies have typically combined both intra- and interword regressions in their measures, even though these two types of regressions probably have very different reading functions. Intraword regressions indicate careful first-pass word processing, while interword regressions, besides indicating processing difficulties due to grammatical and semantic complexities, may further indicate a more advanced speeded reading style, which needs to be corrected if it was too hasty. A plausible assumption is that the number of intraword regressions decreases in the course of reading development, and is replaced by an increased number of interword regressions so that the cumulated regression rate appears to remain constant (cf. Radach, Günther, & Huestegge, 2012).

Crucially, the development of basic eye movement characteristics in sentence or passage reading has been described in a remarkably similar way not only for English, but also for orthographies with more consistent grapheme-phoneme-correspondences (e.g., Finnish: Blythe, Häikiö, Bertram, Liversedge, & Hyönä, 2011; German: Huestegge, Radach, Corbic, & Huestegge, 2009). In a sense this is surprising as reading acquisition has repeatedly been shown to be slower in inconsistent orthographies (e.g., Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013; Frith, Wimmer, & Landerl, 1998; Seymour, Aro, & Erskine, 2003). It thus appears that basic eye movement characteristics develop in a similar way for both consistent and inconsistent orthographies when comparing the development *within* orthographies. However, given the differing rate of initial reading acquisition, it is plausible to expect the development of eye movement patterns to differ *between* consistent and inconsistent orthographies.

In the context of cross-linguistic processing differences, it has been proposed that readers of consistent orthographies preferentially use smaller units when reading, while readers of inconsistent orthographies preferentially use larger units which often help to disambiguate the inconsistencies on the individual letter-sound level (Psycholinguistic Grain

Size Theory; Ziegler & Goswami, 2005). There is indeed evidence supportive of this claim from a number of cross-linguistic studies on single word recognition: While a stronger word length effect in readers of consistent orthographies has been taken to indicate their stronger reliance on systematic decoding procedures (e.g., Ellis & Hooper, 2001; Goswami, Gombert, & Fraca de Barrera, 1998; Rau, Moll, Snowling, & Landerl, 2014), stronger orthographic neighbourhood (rime or body-N) effects in readers of less consistent orthographies have been taken to indicate more large-unit processing on their part (e.g., Goswami et al., 1998; Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Körne, 2003). Importantly, although such cross-linguistic word processing differences are likely to be most pronounced in developing readers, some studies suggest that they are still evident in skilled adult readers (Paulesu et al., 2000; Rau et al., 2014; Ziegler, Perry, Jacobs, & Braun, 2001).

Previous studies on cross-linguistic processing differences have all concentrated on local word recognition for single target items (e.g., Frith et al., 1998; Rau, Moeller, & Landerl, 2014) and those studies which have examined global sentence or text processing have only done so in a single orthography (English: McConkie et al., 1991; Valle, Binder, Walsh, Nemier, & Bands, 2013; Vorstius et al., 2014; Finnish: Blythe et al., 2011; German: Hawelka, Gagl, & Wimmer, 2010; Huestegge et al., 2009; Trauzettel-Klosinski et al., 2010; Greek: Hatzidaki, Gianneli, Petrakis, Makaronas, & Aslanides, 2011; Italian: De Luca, di Pace, Judica, Spinelli, & Zoccolotti, 1999). In an attempt to compare global text processing across orthographies and across studies, Hutzler and Wimmer (2004) have compared their own data on 13-year old German-speaking dyslexics and age-matched typical readers with data sets from two older studies, one with 12-year old Italian dyslexics and age-matched controls (de Luca et al., 1999), and one with 11-year old English-speaking dyslexics and age-matched controls (Lefton, Nagle, Johnson, and Fisher, 1979). Mean fixation duration did not differ a great deal between the Italian dyslexics and controls in the study by de Luca and colleagues (290 vs. 230 ms, respectively), but was clearly more pronounced in Hutzler and

Wimmer's Austrian dyslexics than in their controls (360 vs. 190 ms, respectively); a finding the authors attributed to the greater complexity of syllables in German as compared to Italian. An interesting case was the comparison of proportion of regressions between orthographies (unfortunately, none of the studies specifies whether their regression measures pertain to intra- or interword regressions, or a composite of the two).

The percentage of regressions of English-speaking readers in the study by Lefton and colleagues was higher for both dyslexic and control children (35 and 15 %, respectively) than that of the two German-speaking groups in the Hutzler and Wimmer study (16 and 9 %, respectively), while the Italian dyslexic and control children in the study by de Luca and colleagues both had a percentage of regressions of 19 %. Hutzler and Wimmer suggested that the lower orthographic consistency of English may cause English-speaking children to regress more often than German or Italian children; a conclusion which is surely plausible. Crucially, however, comparisons made across different studies which are comparable neither in terms of reading material, nor in terms of participant age or reading ability are obviously suboptimal (a fact which the authors of the study in question do acknowledge).

Thus, it is crucial that eye movement characteristics be compared between English and more consistent orthographies in direct cross-linguistic studies using an appropriate procedure for matching reading materials as well as participants, and this was the main aim of the current study. To our knowledge, this is the first study to directly compare eye movement behaviour between typically developing readers of a consistent and an inconsistent alphabetic orthography on both the local word processing, and the global sentence processing level. We were interested to see whether differences in orthographic consistency exert an influence beyond the level of pure word recognition. Several previous cross-linguistic studies have compared German and English. This comparison is ideal because the two languages differ with respect to orthographic consistency, but have common Germanic roots – a fact which permits the use of highly comparable target stimuli (i.e., cognates with nearly identical phonology, orthography, and semantics). Importantly, previous studies have only ever compared single word processing (e.g., Frith et al., 1998; Rau et al., 2014; Wimmer & Goswami, 1994; Ziegler et al., 2001). Since the two languages are not only comparable on the level of single words, but also on the level of syntactic structure (at least for main clauses) (Landerl, in press), we constructed highly similar sentences which we presented to German and English children and adults. The sentences mostly consisted of cognates and were highly comparable with regards to their syntactic structure (see Appendix).

Participants were carefully matched on basic word reading ability as indexed by their mean gaze duration for a set of short, high-frequency cognates included in the experimental reading material. Eye tracking allowed for online recording of sentence-reading processes as they occurred. We report a number of local word-based variables which were computed over all words of the matched sentences, as well as a number of global sentence-based variables which were computed over all sentences.

Our intention was to establish whether cross-linguistic processing differences caused by differences in orthographic consistency go beyond the level of single word recognition. We predicted that differences in orthographic consistency would impact upon local word processing and global sentence processing alike. Generally speaking, we expected more small-unit bottom-up processing in readers of German, and more large-unit top-down processing in readers of English. Note that these assumptions of orthography-dependent differences in reading style are reminiscent of an influential classification of dyslexic readers' eye movements based on the percentages of interword regressions and word-skipping saccades (Olson et al., 1985). Accordingly, so called "plodders" display relatively few regressions between words or word-skipping forward movements, but move steadily forward with relatively frequent forward saccades. At the other end of the dimension, "explorers" display more regressions to previous words, more forward word-skipping movements, and fewer intraword and word-to-word progressive movements. This classification has later been suggested to fit the distinction between developmental surface vs. developmental phonological dyslexia as proposed by Castles and Coltheart (1993) with surface dyslexia corresponding to the plodders', and phonological dyslexia to the explorer's eye movement pattern (de Luca et al., 1999). It has further been suggested that skilled readers can be divided into similar styles (Rayner, 1998), and we were interested to establish the applicability of the plodder/explorer distinction in the context of cross-linguistic processing particularities.

More specifically, in local word processing, we expected a more plodder-like processing style of German readers to become manifest in higher first-pass reading times, a higher number of first-pass fixations, and a higher first-pass refixation probability (note that these measures include intraword regressions). At the same time, we expected a more explorer-like processing style of English readers to become manifest in higher re-reading times and a higher number of re-reading fixations in English readers (these measures include interword regressions). Further, we expected their more explorer-like reading style to induce a higher skipping probability in English than in German readers.

In global sentence processing, we did not expect mean fixation duration to differ between orthographies, because German and English both have complex syllables (Seymour et al., 2003), and previously reported cross-linguistic differences in mean fixation duration have been attributed to differences in syllabic complexity (Hutzler & Wimmer, 2004). As participants were carefully matched in terms of basic word reading ability, we did not expect differences in total sentence reading time or total number of fixations between German and English readers. Importantly however, as suggested by Hutzler and Wimmer, we expected that the lower orthographic consistency of English would cause English readers to regress more often than German readers. Since we expected possible cross-linguistic differences to diminish with increasing reading expertise, we expected to find cross-linguistic differences mainly in children.

Method

Participants

From a relatively large pool of 118 participants, we excluded those with below average word reading ability as established by standardized reading fluency tests (German: SLRT II by Moll & Landerl, 2010; English: TOWRE by Torgesen, Wagner, & Rashotte, 1999). Out of the remaining 109 participants, we identified 25 German-English pairs of children matched on gaze duration for a number of short, high-frequency words featuring in the sentences[1] as well as 16 German-English matched pairs of adults. Basic word reading ability as defined by mean gaze duration for the short, high-frequency cognates was thus comparable for German and English children (317 vs. 310 ms, $t(48) = 0.34$, *n.s.*) and German and English adults (239 vs. 244 ms, $t(30) = 0.27$, *n.s.*). English children were slightly older than German children (116 vs. 109 months of age, $t(48) = 0.97$, $p = .06$), and had a significant advantage in terms of length of school attendance (48 vs. 31 months, $t(48) = 5.05$ $p < .001$). Age was comparable for adults of the two orthographies (297 vs. 279 months of age for German and English adults, respectively, $t(30) = 0.99$, *n.s.*). All participants were native speakers of their respective languages and had normal or corrected-to-normal vision. German children were recruited from local elementary schools and attended grades 2, 3, and 4. English children were recruited during a summer sports camp and had completed grades 3, 4, 5, or 6 at time of testing. Informed consent was provided by the parents, and children received a small gift in appreciation of their participation. The participating adults were students of the German and English universities at which the experiment took place.

Apparatus

Eye movements were recorded using an EyeLink 1000 tower mount eye tracker in Germany, and an EyeLink 1000 desktop mount eye tracker in England (SR Research, Toronto). Single line sentences were presented one by one using the unproportional font Courier New (bold, font size 14 pt) in black against a white background on the centre line of a

20-inch monitor in both countries. Standard 9-point calibration at the beginning of the experiment ensured a spatial resolution of less than 0.5 degrees of visual angle.

Materials and Design

There were 72 sentences in total. In the context of a different research question, 24 of the sentences contained a nonword and were excluded from current analyses since we were interested in the processing of normal sentences. The sentences were closely matched between the two languages (e.g., Die Äpfel in der Box sind rot und grün / The apples in the box are red and green – see Appendix for the complete set of sentences). The average number of words per sentence was slightly higher for English, (8.2 vs. 7.7, $t(94) = 1.25$, $p = .22$), while the average number of letters per word was higher for German (4.8 vs. 4.4, $t(94) = 2.92$, $p < .01$). Importantly, however, the average number of letters/sentence was no different between the two orthographies (36.6 vs. 35.2, $t(94) = .88$, $p = .38$).

Results

Data screening

Fixations shorter than 80 ms were not analysed. Words which received no fixation in first-pass reading were excluded from analyses for all measures of processing time. Words/sentences were excluded when there were problems with calibration accuracy, when a participant failed to immediately fixate the small cross at the bottom of the screen after having finished reading, or when a participant did not finish the whole sentence. For local word-based parameters, we furthermore excluded the first and the last word of each sentence (for a similar procedure, see Vorstius et al., 2014), and words which were not read correctly. For local word-based parameters, data loss was 10.3% and 11.0% for German and English children, and 5.5% and 4.9% for German and English adults; for global sentence-based parameters, data loss was 7.7% and 8.3% for German and English children, and 4.7% and 4.6% for German and English adults.

Analyses

Because of very obvious large overall differences in processing speed between children and adults, possible cross-linguistic differences in local word processing and global sentence processing were examined separately for the two age groups. We computed univariate ANOVAs with orthography as a fixed factor and a range of word-based and sentence-based parameters as dependent variables.

Children

Local word-based parameters.

The results on local word-based parameters presented in the upper part of Table 1 are fully in line with predictions. While first-pass reading times were higher in German children (i.e., first fixation duration; gaze duration), the mean re-reading time was higher in English children, resulting in comparable total word reading times for the two orthographies. The same held for number of fixations: while the number of first-pass fixations was higher in German, and the number of re-reading fixations higher in English children, the total number of fixations was no different between the two groups. In line with this, the probability to refixate a word in first pass was clearly higher in German than in English children, while the general refixation probability did not differ between them. Finally, and also consistent with predictions, skipping probability was clearly more pronounced in English than in German children.

Please insert Table 1 here

Global sentence-based parameters.

The lower part of Table 1 shows the cross-linguistic comparison of global sentence-processing parameters for children. As expected, mean fixation duration, total sentence reading time and total number of fixations did not differ between children of the two orthographies. Importantly, and in line with predictions, English children made more interword regressions than German children, both in absolute and in relative terms (i.e., when taking into consideration the respective number of saccades).

Adults

Local word-based parameters.

The cross-linguistic comparison of local word-processing parameters for adults is shown in the upper part of Table 2. Unlike in children, first fixation duration, gaze duration, and re-reading time did not differ between German and English adults, and in line with the data in children, neither did total reading time. When considering number of fixations, as in children, German adults made more fixations than English adults in first-pass reading. However, unlike in children, the number of re-reading fixations was equal for German and English adults, meaning that German adults had a higher total number of fixations than English adults. Consistent with what was found for children, the probability to refixate a word in first-pass was higher in German than in English adults, while general refixation probability did not differ between adult readers of the two orthographies. Finally, consistent with the children's data, skipping probability was nearly twice as high in English as in German adults.

Please insert Table 2 here

Global sentence-based parameters.

The lower part of Table 2 shows the cross-linguistic comparison of global sentence-processing parameters for adults. In analogy to what was found in children, German and English adult readers showed similar mean fixation duration, total sentence reading time, and total number of fixations. As expected, English adults made more interword regressions in reading the sentences than German adults, but this difference in absolute number of regressions failed statistical significance. Importantly however, the relative number of regressions (i.e., percentage of regressions) was significantly higher in English than in German adults.

Discussion

To date, sentence processing has only ever been studied in single orthographies, meaning that previous cross-linguistic comparisons had to be made across studies using different reading material, different eye tracking devices, and participants of different age and reading ability. The current study investigated cross-linguistic differences in eye movement behaviour for local word and global sentence processing in a direct comparison of German and English. Participating children and adults were matched on basic word reading ability. The reading material itself was also thoroughly matched between orthographies and comprised 48 highly similar simple sentences. Because of the differences in orthographic consistency, we expected more small-unit bottom-up processing in readers of German, and more large-unit top-down processing in readers of English for both words and sentences, with more pronounced cross-linguistic differences in children than in adults.

In children, results met expectations perfectly well: While the overall processing outcome for both the local word, and the global sentence level was no different for children of the two orthographies, there were distinct cross-linguistic differences with respect to how this equal outcome was achieved. German children took more time to process both words and sentences in first-pass reading and consequently needed less time for re-analysis. In contrast to this, English children were faster in initial processing, but made more regressions between words, and also took more re-reading time. In line with this, first-pass refixation probability was higher for German than for English children, whereas general refixation probability was comparable between them. Also in line with their generally more diligent first-pass processing, the probability to skip a word was markedly less pronounced in German than in English children.

Thus, children of both orthographies did not differ with regards to the *outcome* of sentence processing: they differed neither in mean fixation duration, total sentence reading time, nor in the total number of fixations they made. However, the *way* in which they achieved this highly similar outcome differed in an important way: German children appeared to put more effort in careful first-pass processing, a trend which was reflected in their higher initial processing times, higher number of first-pass fixations, higher first-pass refixation probability, and lower skipping probability. This approach of rather slow and meticulous first-pass reading makes frequent regressions and long re-reading times unnecessary. English children, on the contrary, appeared to read in a more fragmented way. They were much more likely to skip words and took less time for thorough first-pass processing. They did, however, regress much more often and took more time for second-pass reading than the German children.

Importantly, the children in our study were carefully matched on basic word reading ability, and their identical total sentence processing time corroborates their highly comparable level of reading skill. The cross-linguistic differences we found are therefore reflecting genuinely different approaches to word and sentence processing rather than different levels of reading skill.

The plodder-explorer distinction as described for dyslexic readers by Olson and colleagues (1985) fits remarkably well with the orthography-dependent differences in sentence processing reported here, with German children exhibiting a plodder, and English children exhibiting an explorer style of reading. If we concede that surface dyslexia is more often reported in struggling readers of consistent orthographies (Bergmann & Wimmer, 2008; de Luca et al., 1999), while phonological dyslexia is more often reported in struggling readers of the inconsistent English orthography (Castles & Coltheart, 1993), the more frequent reading style shown by dyslexics of consistent and inconsistent orthographies appears to reflect the typical reading style shown by typically developing readers of their respective orthography.

Importantly, the considerable cross-linguistic processing differences we found in children were still present in a similar way in adult readers. Although German and English adults did not differ to quite the same extent as did the children, German adults made more fixations in first-pass and in total, and were more likely to make a refixation in first-pass, while English adults showed a higher skipping probability and a higher percentage of regressions. The pattern of a more small-unit plodder-like reading style in readers of German, and a more large-unit explorer-like reading style in readers of English is thus also evident in experienced readers.

The present study has shown that German and English children and adults with comparable reading ability process highly similar reading material with identical overall processing effort, but fundamentally different reading styles. While German readers exhibit more of a plodder, English readers exhibit more of an explorer approach of reading. Orthographic consistency therefore impacts upon reading development and skilled reading in a similar way.

References

Bergmann, J., & Wimmer, H. (2008). A dual-route perspective on poor reading in a regular orthography: Evidence from phonological and orthographic lexical decisions. *Cognitive Neuropsychology, 25*, 653-676.

Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research, 51*, 84-92.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research, 49*, 1583-1591.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago, IL: University of Chicago Press.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*, 1398-1407.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 149-180.

De Luca, M., di Pace, E., Judica, A., Spinelli, D., & Zoccolotti, P. (1999). Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia, 37*, 1407-1420.

Ellis, N. C., & Hooper, A. M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics, 22*, 571-599.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading, 2*, 31-54.

Goswami, U., Gombert, J. E., & Fraca de Barrera, L. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics, 19*, 19-52.

Hatzidaki, A., Gianneli, M., Petrakis, E., Makaronas, N., & Aslanides, I. M. (2011). Reading and visual processing in Greek dyslexic children: An eye-movement study. *Dyslexia, 17*, 85-104.

Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition, 115*, 367-379.

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research, 49*, 2948-2959.

Hutzler, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language, 89*, 235-242.

Landerl, K. (in press). Reading acquisition in German language. In C. Perfetti & L. Verhoeven (Eds.) *Reading acquisition across languages and writing systems*. Cambridge: Cambridge University Press.

Lefton, L. A., Nagle, R. J., Johnson, G., & Fisher, D. F. (1979). Eye movement dynamics of good and poor readers: Then and now. *Journal of Reading Behavior, 11*, 319-328.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. In J. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). London: The Macmillan Press.

Moll, K., & Landerl, K. (2010). *Lese- und Rechtschreibtest (SLRT II)*. Bern: Huber.

Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In G. E. McKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (pp. 1-6). New York: Academic Press.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., ... & Frith, U. (2000). A cultural effect on brain function. *Nature Neuroscience, 3*, 91-96.

Radach, R., Günther, T., & Huestegge, L. (2012). Blickbewegungen beim Lesen, Leseentwicklung und Legasthenie. *Lernen und Lernstörungen, 1*, 185-204.

Rau, A. K., Moeller, K., & Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: An eye-tracking study. *Scientific Studies of Reading*, *18*, 224-233.

Rau, A. K., Moll, K., Snowling, M. J., & Landerl, K. (2014). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology, 130,* 92-105.

Rayner, K. (1985). The role of eye movements in learning to read and reading disability. *Remedial and Special Education, 6*, 53-60.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-174.

Taylor, S. E. (1965). Eye movements while reading: Facts and fallacies. *American Educational Research Journal, 2*, 187-202.

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

Trauzettel-Klosinski, S., Koitzsch, A. M., Dürrwächter, U., Sokolov, A. N., Reinhard, J., & Klosinski, G. (2010). Eye movements in German-speaking children with and without dyslexia when reading aloud. *Acta Ophthalmologica, 88*, 681-691.

Valle, A., Binder, K. S., Walsh, C. B., Nemier, C., & Bands, K. E. (2013). Eye movements, prosody, and word frequency among average- and high-skilled second-grade readers. *School Psychology Review, 42*, 171-190.

Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition, 22*, 458-485.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition, 51*, 91-103.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin, 131*, 3-29.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12*, 379-384.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*, 169-193.

Footnote

[1] The German/English cognates used for participant matching were Kuh/cow, rot/red, Tee/tea, vier/four, Mond/moon, Bier/beer, Musik/music, sieben/seven, Hotel/hotel, Mitte/middle.

WORD AND SENTENCE READING IN GERMAN AND ENGLISH

Table 1

Comparison of Local Word Processing and Global Sentence Processing Parameters between German and English Children.

Local word-based variables	German (*n* = 25)	English (*n* = 25)	*F*(1, 48)	η_p^2
Measures of first-pass processing				
Mean first fixation duration (ms)	268 (32)	251 (28)	4.09*	.08
Mean gaze duration (ms)	394 (70)	342 (55)	8.67**	.15
Mean number of first-pass fixations	1.43 (.26)	1.18 (.18)	14.81***	.24
First-pass refixation probability	.33 (.12)	.23 (.07)	12.00***	.20
Skipping probability	.06 (.04)	.14 (.06)	28.19***	.37
Measures of second-pass processing				
Mean re-reading time (ms)	98 (43)	157 (78)	10.93**	.19
Mean number of re-reading fixations	.37 (.16)	.53 (.27)	7.04**	.13
Measures of total processing				

Mean total word reading time (ms)	492 (93)	498 (115)	.05	-
Mean total number of fixations	1.80 (.37)	1.72 (.41)	.52	-
General refixation probability	.49 (.14)	.47 (.13)	.33	-
<u>Global sentence-based variables</u>				
Mean fixation duration/sentence (ms)	271 (27)	263 (28)	1.20	-
Mean fixation duration/sentence (ms)	271 (27)	263 (28)	1.20	-
Total sentence reading time (ms)	3991 (871)	3956 (1200)	.01	-
Total number of fixations/sentence	14.88 (3.13)	15.05 (3.92)	.03	-
Mean number of interword regressions/sentence	1.26 (.50)	2.46 (1.03)	27.03***	.36
Mean percentage of interword regressions/sentence	.09 (.03)	.17 (.06)	36.99***	.44

Note. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$.

Table 2

Comparison of Local Word Processing and Global Sentence Processing Parameters between German and English Adults.

Local word-based variables	German (n = 16)	English (n = 16)	$F(1, 30)$	η_p^2
Measures of first-pass processing				
Mean first fixation duration (ms)	241 (23)	241 (24)	.003	-
Mean gaze duration (ms)	283 (25)	280 (33)	.08	-
Mean number of first-pass fixations	1.06 (.11)	.89 (.10)	20.28***	.40
First-pass refixation probability	.17 (.05)	.12 (.03)	10.99**	.27
Skipping probability	.13 (.07)	.25 (.08)	20.99***	.41
Measures of second-pass processing				
Mean re-reading time (ms)	32 (37)	36 (26)	.19	-
Mean number of re-reading fixations	.13 (.15)	.13 (.12)	.000	-
Measures of total processing				

Mean total word reading time (ms)	315 (46)	317 (36)	.02	-
Mean total number of fixations	1.20 (.22)	1.02 (.20)	5.67*	.16
General refixation probability	.27 (.11)	.22 (.08)	2.02	-
Global sentence-based variables				
Mean fixation duration/sentence (ms)	243 (26)	239 (27)	.13	-
Total sentence reading time (ms)	2399 (450)	2207 (430)	1.52	-
Total number of fixations/sentence	10.05 (1.87)	9.37 (1.67)	1.20	-
Mean number of interword regressions/sentence	.68 (.61)	1.01 (.63)	2.29	-
Mean percentage of interword regressions/sentence	.07 (.05)	.12 (.06)	6.65*	.18

Note. * $p \le .05$; ** $p \le .01$; *** $p \le .001$.

WORD AND SENTENCE READING IN GERMAN AND ENGLISH

Appendix

Matched Sentences

German	English
Ich hasse das Wetter im Winter es sei denn es schneit.	I hate the weather in winter except when it is snowing.
Harrys Erklärung war nicht gerade plausibel aber dafür kreativ.	Harry's explanation was not very plausible, but it was creative.
Robert ist ein typischer Pessimist, warnt mich seine Mutter.	Robert is a typical pessimist, his mother warns me.
Die Musik im Radio ist laut aber gut.	The music on the radio is loud but good.
Großvater hat ein Problem mit seinem Knie.	Grandfather has a problem with his knee.
Dies ist eine interessante Diskussion, sagt der Student.	This is an interesting discussion, says the student.
Meine Mutter trinkt nie Bier aber sie liebt Kaffee.	My mother never drinks beer but she loves coffee.
Laura ist gut im Volleyball und Marc ist gut im Tennis.	Laura is good at volleyball and Marc is good at tennis.
Das neue Hotel öffnet im Juli.	The new hotel will open in July.
Lisa spielt Oboe und ihr Bruder spielt Trompete.	Lisa plays the oboe and her brother plays the trumpet.

WORD AND SENTENCE READING IN GERMAN AND ENGLISH

Der junge Schimpanse lebt in einem Käfig im Zoo.	The young chimpanzee lives in a cage at the zoo.
Dennis ist aus dem Koma erwacht, sagt Doktor Müller.	Dennis has awoken from his coma, says Doctor Miller.
Kann ich eine Zigarette rauchen, fragte der Elektriker.	Can I smoke a cigarette, asks the electrician.
Der große Pianist gab ein Konzert in Rom.	The great pianist gave a concert in Rome.
Die Nadel von Toms Kompass ist zerbrochen.	The needle of Tom's compass is broken.
Sechzig Sekunden sind eine Minute, sagt Tina.	There are sixty seconds in a minute, says Tina.
Produkte guter Qualität haben oft einen hohen Preis.	Products of good quality often have a higher price.
Paula aß eine Blaubeere und Kim einen Apfel.	Paula ate a blueberry and Kim ate an apple.
Die Äpfel in der Box sind rot und grün.	The apples in the box are red and green.
Die Trompete ist das lauteste Instrument im Orchester.	The trumpet is the loudest instrument in an orchestra.
Sarahs neuer Bikini ist blau und pink.	Sarah's new bikini is blue and pink.
Nachts sehen wir den Mond und die Sterne.	At night, we see the moon and the stars.
Der Salamander isst Würmer und Insekten.	The salamander eats worms and insects.

WORD AND SENTENCE READING IN GERMAN AND ENGLISH

Onkel Thomas hat eine Kuh und drei Hennen.	Uncle Thomas has a cow and three hens.
Der neue Minister spricht sehr laut im Parlament.	The new minister speaks very loudly in parliament.
Bienen machen aus Nektar Honig, sagt Peter.	Bees make honey from nectar, says Peter.
Der Angler fing einen Aal und drei Krabben.	The angler caught an eel and three crabs.
Hast du den Piep gehört, fragte Julia.	Have you heard the beep, asked Julia.
Vater reparierte die Maschine letztes Wochenende.	Father repaired the machine last weekend.
Kevin aß eine Mango im Park.	Kevin ate a mango in the park.
Dies ist eine komplett neue Situation, sagte der Präsident.	This is a completely new situation, said the president.
Patrick geht auf Safari nach Afrika.	Patrick is going on a safari in Africa.
Dein Bruder ist ein Optimist, sagt Martin.	Your brother is an optimist, says Martin.
Der Priester aß vier Tomaten.	The priest ate four tomatoes.
David isst eine Kiwi und zwei Bananen.	David is eating a kiwi and two bananas.
Die Gazelle lebt in Afrika.	The gazelle lives in Africa.

WORD AND SENTENCE READING IN GERMAN AND ENGLISH

Das Wetter im Februar ist nicht so gut wie das im Juni.	The weather in February is not as good as it is in June.
Eric trinkt Tee und Susi trinkt Milch.	Eric drinks tea and Suzie drinks milk.
Der Dieb stahl Vaters Axt und seinen Hammer.	The thief stole father's axe and his hammer.
Der Bischof hat ein Aquarium daheim.	The bishop has an aquarium at home.
Onkel Thomas sitzt im Kajak mit dem Paddel in der Hand.	Uncle Thomas sits in the kayak with the paddle in his hands.
Der Monat September folgt auf den Monat August.	The month of September follows the month of August.
Simon ist sieben Jahre alt.	Simon is seven years old.
Mein Vater liebt Fußball und meine Mutter liebt Golf.	My father loves football and my mother loves golf.
Dieser Kaffee hat ein gutes Aroma, sagt Tanja.	This coffee has a good aroma, says Tanya.
Der junge Flamingo hat graue Federn.	The young flamingo has grey feathers.
Kann ich in der Mitte sitzen, fragt Nora.	Can I sit in the middle, asked Nora.
Der alte Professor vergaß seine Notizen.	The old professor forgot his notes.