

Progression in Materials for Learning to Read and Write - a Cross-Language and Cross-Century Comparison of Readers

Kay Berkling¹, Uwe Reichel²

¹Computer Science, Cooperative State University, Karlsruhe, Germany

²Research Institute of Linguistics, Hungarian Academy of Sciences, Budapest, Hungary

berkling@dhw-karlsruhe.de, uwe.reichel@nytud.mta.hu

Abstract

This work is part of a larger effort about understanding the effects of didactic materials on acquisition of reading and writing. In this paper, the focus is on progression in primers (beginner readers). Texts are analyzed in terms of complexity, as measured by entropy between letters and phonemes, from the point of view of reading or writing that text. The assumption is that a good teaching text would start low and increase gradually in complexity. At the same time, languages have different requirements on progression depending on their orthographic depth. The goal of this work is to compare readers in various languages for the beginning stages of reading skill acquisition. We show that there is a difference in difficulty across languages and a large span of approaches in primers of approaching this final difficulty.

Index Terms: orthography acquisition, text analysis, entropy, orthographic depth, text difficulty, text progression

1. Introduction

Acquisition of reading and writing are interrelated. Over the years, a number of studies have looked at the basal abilities needed to predict these skills in children to obtain early diagnostics. In this area, publications for English language have dominated the research, exemplified by [1, 2, 3]. This focus may be problematic, since English orthography presents an outlier for readers and writers [4] due to its orthographic depth [5, 6] and the results may not be generalizable to other languages. Depending on the language, reading and writing can be more or less difficult to acquire [7, 8]. Dyslexia, and some reading or writing difficulties manifest themselves differently depending on orthographic depth [9]. Deeper orthographies take more context at phoneme, syllable and morpheme or even sentence grammatical knowledge into account to understand the relationship between phoneme and grapheme sequences. Some studies regarding reading and writing acquisition exist that compare the process across languages [10, 11, 12, 13, 14, 5]. Few authors look at spelling acquisition [15] or focus on interaction between the two skills during acquisition [16, 17].

For German, Fricke [18] looks at basal skills that can predict both reading and writing. Landerl [19] looks at specific orthographic pattern acquisition for spelling in German and Dutch. Roth [20] looks at long-term effects on orthographic skills for German after training basal skills. Suggate [21] studies long-term effects of intervention on reading. Röber-Siekmeyer [22] focuses on orthographic acquisition process in multilingual children. Similar studies are reflected in the literature for the French language [23] and Italian in comparison to

English [24].

Part of the problem with having a detailed understanding of the reading and writing skill acquisition is the inability to prevent the so-called Matthew effect that separates students at a very early stage (Kindergarten) into those that excel versus those that stagnate. Early diagnostics, improved learning materials and early training in basal skills are repeatedly shown to be of importance in the above mentioned studies. The effects persist however [25, 26, 27]. Looking at spelling skills by children in Germany and France, there is reason to believe that there is some room for improvement in the presentation of didactic materials in the beginning years [28, 29, 30].

Two important lines of thought in the literature lead us to examine reading materials for first graders. Firstly, the self-teaching theory that builds on a strong interaction between reading and writing process based on the textual interactions. The theory states that through reading, orthographic skills and reading skills are self-taught with practice. Some of the key literature in this area over the years has been [31, 17, 32, 33, 34, 35]. It is therefore important to study the reading materials that are supplied to children in the first years of their schooling. In fact, the study by Ziegler [32] seems to suggest that bad reading materials (just like learning impediments) will have a long-term negative impact on skill acquisition in the same way that inherent learning disabilities have. Bad learning material may simulate a disability and could lead to negative acquisition effects in the same way as inherent learning disabilities. Apart from [36] we are not aware of any other related work that analyzes progression in reading materials at the early stages of acquisition in a quantitative manner.

The second very important strand of thought in the literature relates to the regularity in learning material that are required to support optimal learning [37, 38, 2, 39]. This includes training or drill as an important component [40, 41]. A look at the reading materials for regularity and incremental learning is therefore an important step towards evaluating the quality of learning materials.

The rest of the paper is structured as follows. Section 2 introduces the texts that will be studied. Section 3 will explain the theory and implementation of the automatic tools to generate the features that we are studying for each of the texts. Section 4 will discuss the results followed by a conclusion and future work discussion in 5.

2. Corpora

In order to analyze progression in primary readers in various languages, we picked a selection for English, French, German

and Italian to have a spectrum from rather flat orthography to deep orthography, with German being in the middle.

2.1. English

In the US, current teaching of the orthographic system has moved from first grade into the preschools, even into pre-kindergarten. It is therefore not feasible to look at first grade readers. The current state-of-the-art practice is the reading of easy literature that is chosen by the children themselves from a larger set of stories that is often supplied as leveled readers by the publishers, or books that form part of the reading workshop [42].

In order to study a text that follows the phonics approach [43], a book from the turn of the century was chosen for the analysis. It gives us the opportunity to look at the progression within a single text.

For the purpose of this paper, the 'McGuffey's First Eclectic Reader' [44] was chosen because it was the first widely used textbook in the US, follows a leveled phonics approach and is still in use today. In stark contrast to today's books, this one is able to describe the methodology in one page.

This First Reader may be used in teaching reading by any of the methods in common use; but it is especially adapted to the Phonic Method, the Word Method, or a combination of the two.

I Phonic Method.—First teach the elementary sounds and their representative, the letters marked with diacriticals, as they occur in the lessons; then, the formation of words by the combination of these sounds. For instance, teach the pupil to identify the characters a, o, n, d, g, r, and th, in Lesson I, as the representatives of certain elementary sounds; then teach him to form the words at the head of the lesson, then other words, as nag, on, and, etc. Pursue a similar course in teaching the succeeding lessons. Having read a few lessons in this manner, begin to teach the names of the letters and the spelling of words, and require the groups, "a man," "the man," "a pen," to be read as a good reader would pronounce single words.

II When one of the letters in the combinations ou or ow, is marked in the words at the head of the reading exercises, the other is silent. If neither is marked, the two letters represent a diphthong. All other unmarked vowels in the vocabularies, when in combination, are silent letters. In slate or blackboard work, the silent letters may be canceled.

III Word Method.—Teach the pupil to identify at sight the words placed at the head of the reading exercises, and to read these exercises without hesitation. Having read a few lessons, begin to teach the names of the letters and the spelling of words.

IV Word Method and Phonic Method Combined.—Teach the pupil to identify words and read sentences, as above. Having read a few lessons in this manner, begin to use the Phonic Method, combining

it with the Word Method, by first teaching the words in each lesson as words; then the elementary sounds, the names of the letters, and spelling.

V Teach the pupil to use script letters in writing, when teaching the names of the letters and the spelling of words.

The preface of the current version gives the following teaching instructions that are relevant to the text analysis:

1. Words of only two or three letters are used in the first lessons. Longer and more difficult ones are gradually introduced as the pupil gains aptness in the mastery of words.
2. A proper gradation has been carefully preserved. All new words are placed at the head of each lesson, to be learned before the lesson is read. Their number in the early lessons is very small, thus making the first steps easy. All words in these vocabularies are used in the text immediately following.

The text is obtained from the online Gutenberg Project as a txt file and is therefore fully digitized for the automated analysis.

2.2. French

Similar to English, the study of learning materials is of great importance in a deep orthography such as French [30]. In a study published only as a technical report [45] several readers in the first grade have been analyzed for their methodology. We chose a book that follows a syllable based approach most closely related to the Phonics approach in English in that it prescribes a leveled approach to reading (and therefore writing). 'Léo et Léo' [46] and 'Kimamila (1&2)' [47, 48] were chosen for this study and digitized.

2.3. German

For German, a selection of readers was taken that represent different methods of teaching. Each of these have been digitized.

Syllabic Method: This method assumes that reading is best taught starting with the syllable. Therefore, the book starts with teaching syllables instead of letters in isolation. Usually, this method also distinguishes stressed from unstressed syllables. An example of such a primer is 'ABC der Tiere' [49].

Analytic-Synthetic Method: This method assumes that there is more or less a 1-1 correspondence between phoneme and grapheme with some more nuances that are postponed to a later stage, without following a phonics approach. Examples for this approach are 'Jo-Jo' [50] and 'Kunterbunt' [51]. The books are usually complemented with additional materials that also work on the syllable method and most of them have since been updated accordingly.

Phonics Method: The phonics approach emphasizes regularity over simplification. There are old primers from the turn of the century that exemplify this method [52]. The description of the method spans the first pages of the primer. It is similar to the 'McGuffey' reader explanation, therefore we choose not to translate it for this paper. A small excerpt is given in German below.

"Der Aufbau des Lesestoffes ruht auf der Lautlehre (Phonetik). Nur an einigen Stellen

war aus praktischen Gründen eine Abweichung notwendig Die meisten Fibeln bringen schon auf den allerersten Seiten dreilautige Silben wie 'nun', 'nein', 'mein', 'neun' und erschweren damit ohne Grund besonders den weniger Begabten das Erlernen des Lesens ganz bedeutend. Viel interessanter, übersichtlicher und leichter zu lesen und aufzuschreiben sind Wörter wie 'Leo', 'Lisa' 'Leine', 'Lina'. Erst müssen wir eine gewisse Gewandtheit im Zusammenziehen zweilauter Silben erzielt haben, ehe wir zu den dreilautigen übergehen dürfen. Dann machen diese nur geringe Mühe, zumal sie mit reichlichem Lautmaterial und planmäßig eingeführt werden können.

2.4. Italian

In Italian, it was difficult to find essays evaluating various readers. The problem of reading and spelling acquisition is much less prominent in a shallow orthography. The chosen text for analysis is a first-grade reader 'A come ...' by a well-known publisher [53]. The first 50 pages were digitized.

2.5. Comparison of Language Corpora

Looking at the Corpora in various languages by inspection, different approaches are clearly apparent in the text. There are four main approaches; the first two follow the phonics approach of teaching regularities and word structures along with the introduction of new graphemes:

1. Slow progression with some practice (including grapheme, syllable and word structure)
2. Slow progression with a lot of practice (including grapheme, syllable and word structure)
3. Progression based on introduction of new letters
4. No apparent progression at all (Italian text is simply a reader)

From easiest orthography to most difficult one, we see that phonics comes into play more so for English and French than for Italian. Italian book shows no progression other than practice through texts. There may be books that have a different approach on the market that we are not aware of. The English text shows a slow progression with some practice of new concepts within texts before moving on to the next level. There is a clear difference between the two French texts. While both have a clear and structured progression, with every page introducing new items that are then practiced extensively on that page and used on subsequent pages. 'Léo et Léa' places much more emphasis on practice within the book than 'Kimamila'. The former has a similar concept as the 'McGuffey' reader. Looking at the lexical complexity in Table 1 we can see that the two readers behave similarly in this respect, a lot of practice for the introduced words (low $LD = \text{types/tokens}$). The latter is more closely comparable to 'ABC der Tiere', a simple syllable approach to German that emphasizes practice and simplicity at the beginning. The other chosen German primers relate more closely to the Italian approach of treating the orthography as having a simple phoneme-grapheme correspondence. They focus mainly on introducing new letters without taking context into account. (Newer versions of these primers incorporate syllable based structure into their progression.) Table 1 does not show a large difference in lexical diversity at the word level. In

Title	Lang.	Tokens	Types	LD	SD
ABC der Tiere	DE	3164	1059	.33	.07
JoJo	DE	1411	454	.32	.20
Tinto	DE	1580	719	.45	.30
Kunterbunt	DE	1347	508	.37	.26
Tobi	DE	1378	558	.40	.27
Alte Fibel	DE	3641	1466	.40	.09
Kimamila	FR	3245	1403	.43	.06
Leo et Lea	FR	6486	1325	.20	.11
Acome	IT	2298	798	.34	.06
McGuffey	EN	7009	737	.10	.11

Table 1: Statistics for each of the analyzed primers. LD=Language Diversity (Types/Tokens). SD=Syllable Diversity, using counts for types and tokens at syllable level.

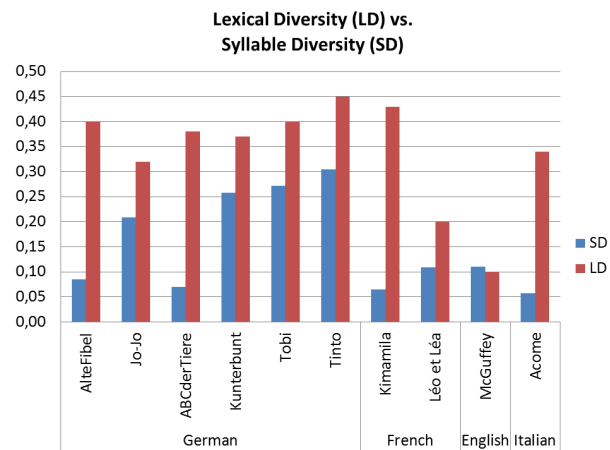


Figure 1: Statistics for each of the analyzed primers. LD=Language Diversity (Types/Tokens). SD=Syllable Diversity, using counts for types and tokens at syllable level.

German, practice may be accomplished at the syllable instead of the word level - more easily than in English. Therefore, the LD measure is insufficient in measuring this aspect of practice. Looking at the SD measure (types/tokens of syllable) it can be seen that there is a striking difference between the old primer and the new ones in that fewer syllables are being reused to form more new words when compared to any of the new readers.

3. Automated Data Analysis

3.1. Alignment

The alignment of a grapheme sequence v and a canonic transcription w is derived from their Levenshtein distance, i.e. the minimum edit costs to transform v into w . The edit operations are substitution, deletion and insertion. The sequences v and w don't stem from a shared vocabulary, as an example, substituting grapheme 'X' by phoneme /X/ is not a zero-substitution. Therefore, a naive cost function assigning cost 0 to zero-substitutions and 1 to the other operations cannot be applied. Instead we make use of the more flexible PermA approach [54] defining edit costs c in terms of conditional probabilities. The cost of substituting grapheme v_i by phoneme w_j (i and j refer to their position in the respective sequence)

is thus given as follows: $c(v_i, w_j) = 1 - P(w_j|v_i)$. Insertion and deletion costs are derived by introducing empty graphemes and phonemes ' '. The grapheme deletion cost is thus defined as $c(v_i, -) = 1 - P(-|v_i)$, and the phoneme insertion cost as $c(-, w_j) = 1 - P(w_j|-)$. Details on smoothing and permutation techniques to derive the underlying grapheme-phoneme co-occurrence counts can be found in [54].

In order to gain a better view of the alignment output, there are examples given for each of the languages processed by the system. Some limitations to the computation of the entropy are evident. For example, *They;D_- + eI* will result in the sonogram pairs (T,D)(h,_) (e,_) (y,+eI). A more didactic alignment would have been (Th,D)(ey,+eI). However, the same method is used consistently throughout the texts in all languages and the result is a valid indicator of the entropy. (Future work will include a post-processing to align phoneme with graphemes, such as <th> or <ey> rather than letters.)

English:

T h e y;D _ _ ' +eI
a r e;' +A: _ _
a l l;' +O: _ l
i n;' +I n
t h e;D _ ' +@
s h a d e;S _ ' +eI d _
n o w;n ' +aU _

French.

l e;l @
c h a t;_ S a _
I l;i l
s ' a f f a l e;s _ a _+. f a l _
s u r;s y R
l e;l @
l i t;l i _
P u i s;p H i _
i l;i l
s e;s @
l a v e;l a v _

German:

H a f e r;h ' +a:+. f 6 _
H e i d e;h _ ' +aI+. d @
H e r d e;h ' +e:6 _+. d @
H a l l e;h ' +a _ l+. @
D e r;d ' +e:6 _
D o m;d ' +o: m
i s t;?' +I s t
h o c h;h ' +o: _ x

Italian:

g l i;_ L ' +i
a l l u c i;' +a _ ll+. u+. tS i
s i;s ' +i
t o c c a;t ' +o _ kk+. a
s o l l e t i c a;s o _ ll+. ' e+. t i+. k a
p i z z i c a;p ' +i _ tts+. i+. k a

3.2. Measure of Difficulty

According to Seymour [5, p. 146], European languages have different orthographic depth that can be sorted into levels of difficulty. Using the entropy measure suggested by [55] this ranking can be quantified.

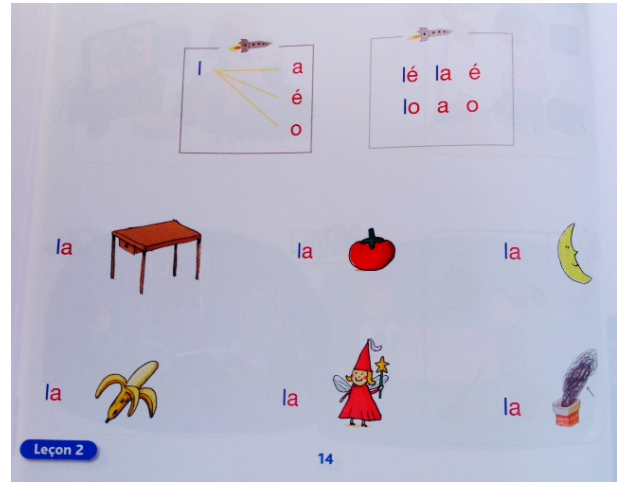


Figure 2: First page in French primer with low entropy.

We estimate the orthographic depth with the following measures at the phonemic levels. Conditional Entropy, H , can be seen as a measure of complexity for the process of reading (the reader transcribes graphemes into phonemes) $H(PHO|GRA)$ and writing (the reader transcribes phonemes into graphemes) $H(GRA|PHO)$. PHO stands for the set of phonemes, while GRA stands for the set of graphemes. For beginner readers, this set contains the number of graphemes seen so far in their studies. Therefore, starting with a lower number of graphemes, that have a clear phoneme-grapheme correspondence would result in a low entropy measure. Figure 2 shows a set $GRA = \{<l>, <a>, <e>, <o>\}$ and $PHO = \{/l/, /a/, /e/, /o/\}$, the combination of these exhibiting an entropy of 0, given the reading task and the information that the child has learned up to now. With each additional page in the book, more graphemes and phonemes and more complex interrelationships between the two sets are introduced. The entropy measure reflects how much additional information is needed in order to choose the correct correspondence or the uncertainty given previous knowledge. In languages where phoneme, syllable, and morpheme information is needed for the transcription process, this measure will be higher.

Entropy complexity for reading is given by Equation 1:

$$H(PHO|GRA) = - \sum_{x \in GRA} p(x) \cdot \sum_{y \in PHO} p(y|x) \cdot \log p(y|x) \quad (1)$$

Entropy complexity for spelling is given by Equation 2:

$$H(GRA|PHO) = - \sum_{y \in PHO} p(y) \cdot \sum_{x \in GRA} p(x|y) \cdot \log p(x|y) \quad (2)$$

$$H_{read}(n) = \sum_{i=0}^n H(PHO_n|GRA_n) \text{ and} \quad (3)$$

$$H_{spell}(n) = \sum_{i=0}^n H(GRA_n|PHO_n), \quad (4)$$

where PHO_n and GRA_n contain all sonographs seen in the first n words. Figure 7 compares the Entropy measure between

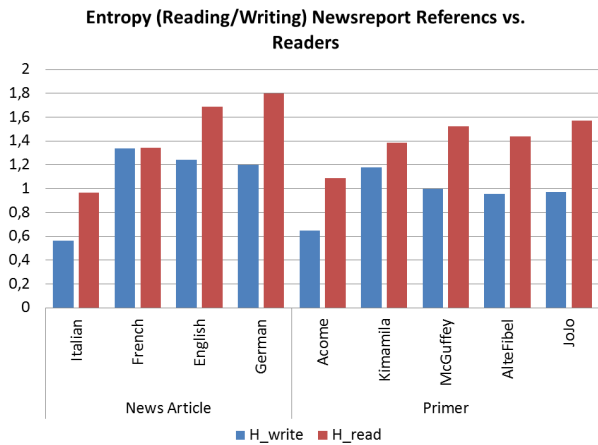


Figure 3: Comparative final H_{read} and H_{write} for selected readers compared to a random newspaper article of around the same length as the primers

selected primers and randomly chosen newspaper articles representing normal language entropy. It can be seen that the languages have different complexity and that the chosen primers reach similar levels. The results section will depict H_{read} and H_{spell} as a function of n , the number of words seen so far, in increments of 5.

3.3. Measure of Progression

After quantifying complexity of languages according to their orthographic depth with the estimate of entropy measure, the question of progression is to see how the beginning books ramp up to reach this complexity. We look at three measurements that can be examined visually for the curves shown in the results section.

Starting Point: A low start in complexity shows that the material is simplified for both reading and writing complexity.

Steepness/Slope: Slower rise reflects amount of practice that a primer provides for a given level of complexity.

Figure 4 and 5 plots Entropy for a non didactic random newspaper text for each of the languages from both reading and writing perspective.

4. Results

Reading:

Figure 6 depicts German readers in their progression from the point of view of reading. 'Tinto' starts high and remains flat. 'Kunterbunt' starts low and progresses to a high level. 'Jo-Jo' starts high and remains flat. The old reader is similar without going quite as high. 'Tobi' and 'ABC der Tiere' start low and move up slowly. They will be used to compare their slope of learning to readers in other languages.

Figure 7 shows that 'Tobi' has a much faster progression than any of the other readers. The Italian reader is an outlier in that it has no progression but starts at the level of Italian orthographic complexity and remains level. 'Kimamila' and 'McGuffey' choose a rapid ascent, showing little practice at each level, yet more than the German reader. There are some steps visible in the progression. In comparison, 'Léo et Léa' spend a lot more time on the lower level of complexity before moving up slowly.

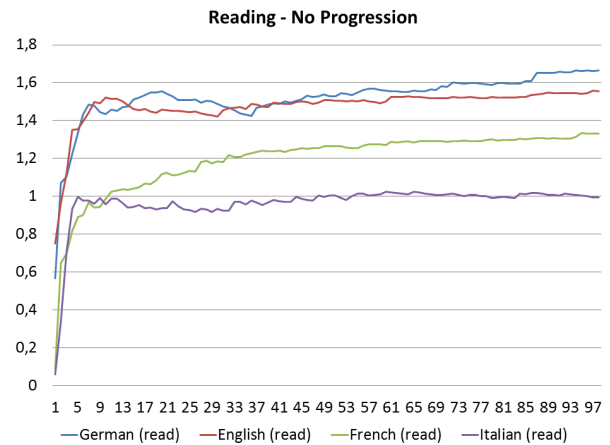


Figure 4: Entropy without progression from reading point of view.

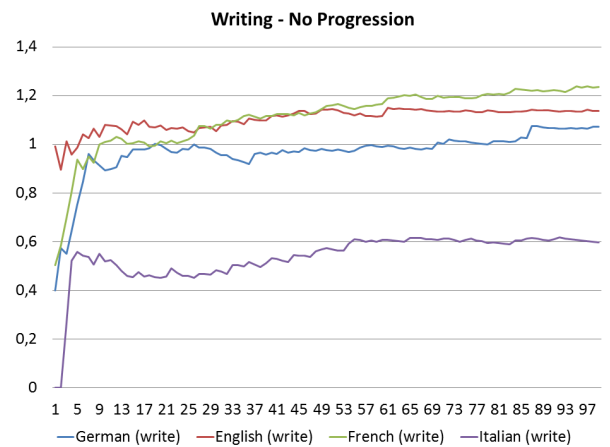


Figure 5: Entropy without progression from writing point of view.

Writing:

German readers differ in their speed and steepness of progression as shown in Figure 8. The old primer has a clearly visible staircase that attains the full height of German complexity. 'Tinto' starts high and moves in steps to the full complexity. 'Kunterbunt' starts low and has the steepest ascent. 'Tobi' and 'ABC der Tiere' are the reader that have a steady and slow climb in complexity. They are compared to the readers in other languages. Figure 9 shows that even though writing has a comparative complexity for the elementary texts as French and English, the progression in the text is much more intense even for the reader with the slowest progression. It can also be seen that there are two very different approaches for French. 'Léo et Léa' spends a lot more time practicing at a simpler level than 'Kimamila' that moves forward in complexity steadily. 'McGuffey' reader for English chooses a middle path between the two. 'Acome' is an outlier in that it has no progression but reflects the very low level of orthographic complexity of Italian.

Figure 10 depicts the scatter plots for both reading and writing view of slope vs. intercept, showing the readers that fit into each of the quadrants as marked. Léo et Léa is the only primer that starts slow and adds practice at each stage for both direc-

		Acome	Jo-Jo	Kunterbunt	Alte Fibel	Léo et Léa	Tinto	Tobi	Kimamila	McGuffey	ABCd.T.
Read	slope	0.05	0.08	0.10	0.10	0.12	0.14	0.21	0.22	0.23	0.32
	int.	0.80	0.88	0.92	0.75	-0.23	0.75	0.33	-0.26	-0.01	-0.44
Write	slope	0.03	0.07	0.04	0.08	0.16	0.10	0.22	0.29	0.18	0.23
	int.	0.44	0.29	0.54	0.42	-0.38	0.38	-0.37	-0.78	-0.36	-0.68

Table 2: Intercept and steepness of learning curve as defined by the best fit equation: $y = slope * \ln(x) + int.$ ($= intercept$). Low numbers in both relate to a slow start and a lot of practice at each level of difficulty before moving on. A high intercept with low slope can mean a steep entry at the full level of complexity.

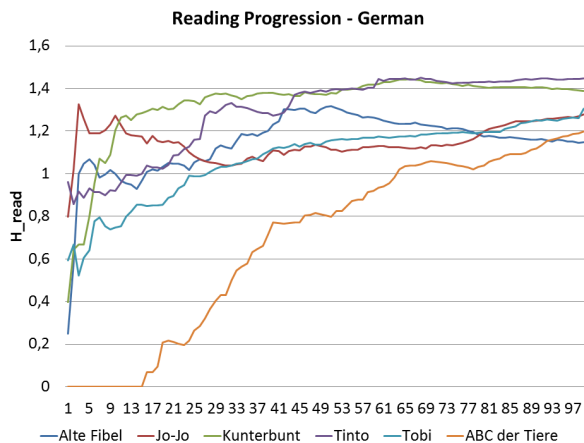


Figure 6: Reading progression for all German readers.

tions (reading and writing). All other readers can be interpreted with respect to that. Most primers have a steep entry for both reading and writing. 'Kimamila' and 'ABC der Tiere' are similar in that they have a leveled start but lack the practice that the green quadrant books offer. 'Tobi' has a slightly more difficult entry and could also benefit from more practice.

5. Discussion

In this paper we have analyzed the beginnings of several readers in their progression with respect to letter and phoneme transcription. These didactic texts were compared across languages and across first-grade readers that are designed as learning materials. We have shown that there are different levels of difficulty in texts for various languages. The assumption is that didactic primers that are designed to teach reading and writing should start with a low level of complexity and rise slowly, with lots of practice. The analysis of the German primers in comparison to primers in other languages has shown that the German primers progress at a much more rapid pace to more complex levels than their counterparts, even taking into account that German orthography may not have the same final complexity as some of the other languages with deeper orthography. In addition, German primers offer a lot less reading practice than any other of their counterparts in other languages. For the German readers we also found a measure of distinguishing an older reader from newer ones by looking at the syllable level repetition. It was shown that the reader, despite having a steep learning curve at the Entropy level, has an intensive repetition rate of syllables with which it forms new words. There are lessons to be learned from this work, presenting new ideas on how Ger-

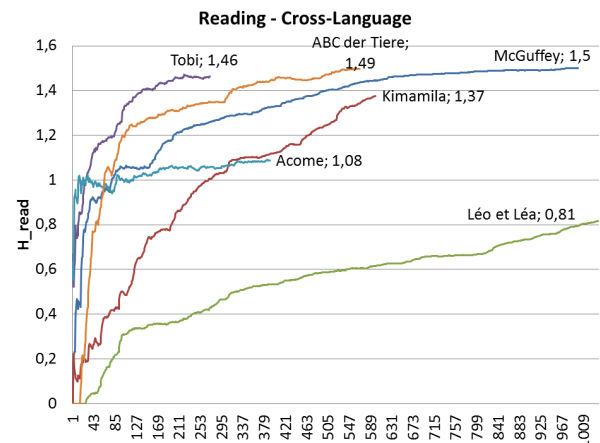


Figure 7: Reading progression for readers across languages, including the German reader with the lowest slope.

man primers can be adapted to offer more practice at phoneme, grapheme, syllable and word level to obtain a less steep learning curve for children in first grade. There is also a lot more work to be done to analyze primers in more detail and correlate their effects to the spelling abilities of children over time.

6. Acknowledgements

We wish to thank the unfunded weekend warriors for pursuing scholarly research no matter what!

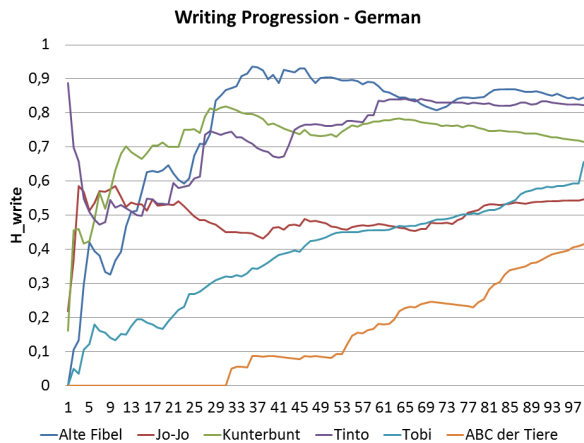


Figure 8: Writing progression for all German readers.

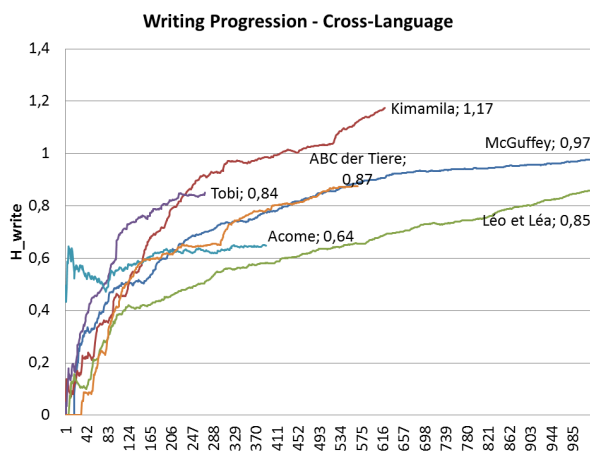


Figure 9: Writing progression for readers across languages, including the German reader with the lowest slope.

7. References

- [1] A. Castles and M. Coltheart, "Is there a causal link from phonological awareness to success in learning to read?" *Cognition*, vol. 91, no. 1, pp. 77–111, 2004.
- [2] M. Melby-Lervag, S. A. Lyster, and C. Hulme, "Phonological skills and their role in learning to read: a meta-analytic review," *Psychological bulletin*, vol. 138, no. 2, pp. 322–352, 2012.
- [3] C. Hulme and M. J. Snowling, "Learning to Read: What We Know and What We Need to Understand Better," *Child development perspectives*, vol. 7, no. 1, pp. 1–5, 2015.
- [4] D. L. Share, "On the Anglocentricities of current reading research and practice: the perils of overreliance on an outlier orthography," *Psychological bulletin*, vol. 134, no. 4, p. 584, 2008.
- [5] P. H. K. Seymour, M. Aro, and J. M. Erskine, "Foundation literacy acquisition in European orthographies," *British Journal of Psychology*, vol. 94, no. 2, pp. 143–174, 2003.
- [6] F. Richlan, "Functional neuroanatomy of developmental dyslexia: the role of orthographic depth," *Frontiers in human neuroscience*, vol. 8, p. 347, 2014.
- [7] H. Wimmer, H. Mayringer, and K. Landerl, "Poor reading: A deficit in skill-automatization or a phonological deficit?" *Scientific Studies of Reading*, vol. 2, no. 4, pp. 321–340, 1998.
- [8] M. Aro and H. Wimmer, "Learning to read: English in comparison to six more regular orthographies," *Applied Psycholinguistics*, vol. 24, no. 04, pp. 621–635, 2003.
- [9] K. Landerl, H. Wimmer, and U. Frith, "The impact of orthographic consistency on dyslexia: A German-English comparison," *Cognition*, vol. 63, no. 3, pp. 315–334, 1997.
- [10] J. C. Ziegler and U. Goswami, "Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory," *Psychological bulletin*, vol. 131, no. 1, p. 3, 2005.
- [11] J. C. Ziegler, D. Bertrand, D. Tóth, V. Csépe, A. Reis, L. Fáisca, N. Saine, H. Lyytinen, A. Vaessen, and L. Blomert, "Orthographic depth and its impact on universal predictors of reading a cross-language investigation," *Psychological Science*, vol. 21, no. 4, pp. 551–559, 2010.
- [12] M. Caravolas, A. Lervag, P. Mousikou, C. Efrim, M. Litavsky, E. Onochie-Quintanilla, N. Salas, M. Schoffelova, S. Defior, M. Mikulajova, G. Seidlova-Malkova, and C. Hulme, "Common patterns of prediction of literacy development in different alphabetic orthographies," *Psychological Science*, vol. 23, no. 6, pp. 678–686, 2012.
- [13] K. Landerl, F. Ramus, K. Moll, H. Lyytinen, P. H. T. Leppänen, K. Lohvansuu, M. O'Donovan, J. Williams, J. Bartling, J. Bruder *et al.*, "Predictors of developmental dyslexia in European orthographies with varying complexity," *Journal of Child Psychology and Psychiatry*, vol. 54, no. 6, pp. 686–694, 2013.
- [14] K. Moll, F. Ramus, J. Bartling, J. Bruder, S. Kunze, N. Neuhoff, S. Streiftau, H. Lyytinen, P. H. Leppänen, K. Lohvansuu, D. Tóth, F. Honbolygó, V. Csépe, C. Bogliotti, S. Iannuzzi, J.-F. Démonet, E. Longeras, S. Valdois, F. George, I. Soares-Boucaud, M.-F. Le Heuzey, C. Billard, M. O'Donovan, G. Hill, J. Williams, D. Brandeis, U. Maurer, E. Schulz, S. van der Mark, B. Müller-Myhsok, G. Schulte-Körne, and K. Landerl, "Cognitive mechanisms underlying reading and spelling development in five European orthographies," *Learning and Instruction*, vol. 29, pp. 65–77, 2014.
- [15] K. Landerl and H. Wimmer, "Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up," *Journal of Educational Psychology*, vol. 100, no. 1, pp. 150–161, 2008.
- [16] L. C. Ehri, "The development of spelling knowledge and its role in reading acquisition and reading disability," *Journal of learning disabilities*, vol. 22, no. 6, pp. 356–365, 1989.
- [17] A. E. Cunningham, "Accounting for children's orthographic learning while reading text: do children self-teach?" *Journal of experimental child psychology*, vol. 95, no. 1, pp. 56–77, 2006.
- [18] S. Fricke, M. Szczerbinski, A. Fox-Boyer, and J. Stackhouse, "Preschool Predictors of Early Literacy Acquisition in German-Speaking Children," *Reading research quarterly*, vol. 51, no. 1, pp. 29–53, 2016.
- [19] K. Landerl and P. Reitsma, "Phonological and morphological consistency in the acquisition of vowel duration spelling in Dutch and German," *Journal of Experimental Child Psychology*, vol. 92, no. 4, pp. 322–344, 2005.
- [20] E. Roth and W. Schneider, "Langzeiteffekte einer Förderung der phonologischen Bewusstheit und der Buchstabenkenntnis auf den Schriftspracherwerb," *Zeitschrift für Pädagogische Psychologie*, vol. 16, no. 2, pp. 99–107, 2002.
- [21] S. P. Suggate, "A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions," *Journal of learning disabilities*, vol. 49, no. 1, pp. 77–96, 2016.
- [22] C. Röber-Siekmeier, "Die Entwicklung orthographischer Fähigkeiten im mehrsprachigen Kontext," in *Didaktik der deutschen Sprache*, ser. UTB, U. Bredel, Ed. F. Schöningh, 2003, vol. 1, pp. 392–403.
- [23] L. Sprenger-Charolles, L. S. Siegel, and P. Bonnet, "Reading and spelling acquisition in French: The role of phonological mediation and orthographic factors," *Journal of experimental child psychology*, vol. 68, no. 2, pp. 134–165, 1998.

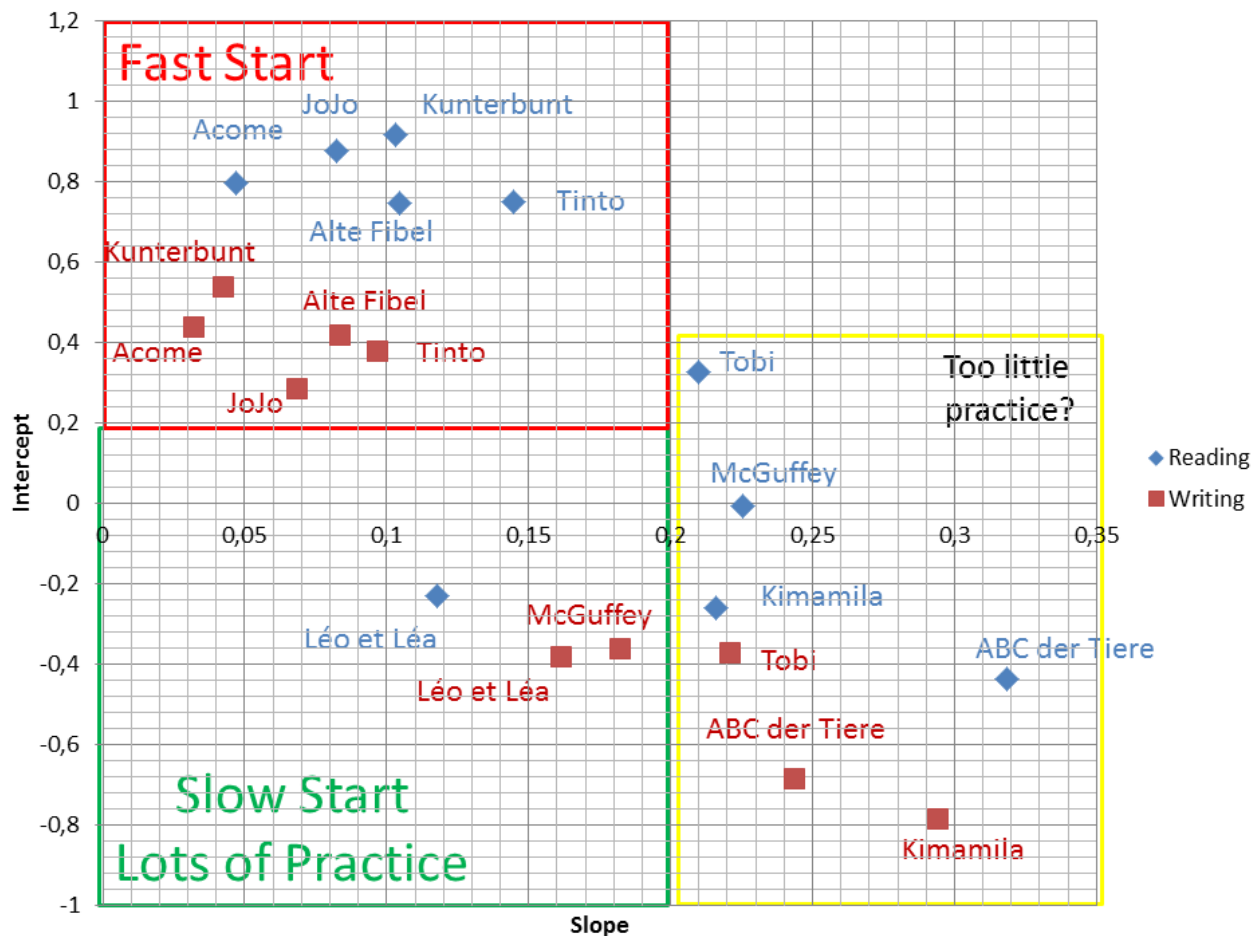


Figure 10: Slope vs. Intercept from Table 2 comparing progression in quantitative manner for primers.

- [24] C. V. Marinelli, C. Romani, C. Burani, and P. Zoccolotti, "Spelling Acquisition in English and Italian: A Cross-Linguistic Study," *Frontiers in psychology*, vol. 6, p. 1843, 2015.
- [25] K. E. Stanovich, "Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy," *Reading research quarterly*, vol. 21, no. 4, pp. 360–407, 1986.
- [26] M. Pfost, J. Hattie, T. Dorfler, and C. Artelt, "Individual Differences in Reading Development: A Review of 25 Years of Empirical Research on Matthew Effects in Reading," *Review of educational research*, vol. 84, no. 2, pp. 203–244, 2014.
- [27] A. Protopapas, R. Parrila, and P. G. Simos, "In Search of Matthew Effects in Reading," *Journal of learning disabilities*, 2014.
- [28] K. Berkling and R. Lavalley, "WISE: A Web-Interface for Spelling Error Recognition Description and Evaluation of the Algorithm for German," in *International Conference of the German Society for Computational Linguistics and Language Technology*, ser. GSCL. Gesellschaft für Sprachtechnologie und Computerlinguists, 2015, pp. 87–96.
- [29] R. Lavalley, K. Berkling, and S. Stüker, "Preparing children's writing database for automated processing," in *Language Teaching, Learning and Technology*, ser. LTLT, K. Berkling and W. Lenhard, Eds. (ISCA) International Speech Communication Association, 2015, pp. 9–15. [Online]. Available: \url{@http://isca-speech.org/archive/ltl2015/lt15_009.html}
- [30] J. Deauvieu, J. Reichstadt, and J.-P. Terrail, *Enseigner efficacement la lecture: Une enquête et ses implications*. Paris: O. Jacob, impr. 2015, cop. 2015.
- [31] A. E. Cunningham, K. E. Perry, K. E. Stanovich, and D. L. Share, "Orthographic learning during reading: Examining the role of self-teaching," *Journal of experimental child psychology*, vol. 82, no. 3, pp. 185–199, 2002.
- [32] J. C. Ziegler, C. Perry, and M. Zorzi, "Modelling reading development through phonological decoding and self-teaching: implications for dyslexia," *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 369, no. 1634, p. 20120397, 2014.
- [33] M.-I. Bosse, N. Chaves, P. Largy, and S. Valdois, "Orthographic learning during reading: The role of whole-word visual processing," *Journal of Research in Reading*, vol. 38, no. 2, pp. 141–158, 2015.
- [34] R. Tucker, A. Castles, A. Laroche, and S. H. Deacon, "The nature of orthographic learning in self-teaching: Testing the extent of transfer," *Journal of experimental child psychology*, vol. 145, pp. 79–94, 2016.
- [35] S. C. Pritchard, M. Coltheart, E. Marinus, and A. Castles, "Modelling the implicit learning of phonological decoding from training on whole-word spellings and pronunciations," *Scientific studies of reading*, vol. 20, no. 1, pp. 49–63, 2016.
- [36] K. Berkling, U. Reichel, and R. Lavalley, "Untersuchung der Eignung von Fibeln für einen systematischen Schriftspracherwerb – Analyse von Fibeltexten durch automatische Sprach-verarbeitungsmethoden," in *Gesellschaft für Empirische Bildungsforschung*, ser. GEBF, 2014.

- [37] H. C. Wang, A. Castles, L. Nickels, and K. Nation, "Context effects on orthographic learning of regular and irregular words," *Journal of experimental child psychology*, vol. 109, no. 1, pp. 39–57, 2011.
- [38] H. C. Wang, A. Castles, and L. Nickels, "Word regularity affects orthographic learning," *Quarterly journal of experimental psychology (2006)*, vol. 65, no. 5, pp. 856–864, 2012.
- [39] L. G. Duncan, S. L. Castro, S. Defior, P. H. Seymour, S. Bailie, J. Leybaert, P. Mousty, N. Genard, M. Sarris, C. D. Porpodas, R. Lund, B. Sigurethsson, A. S. Thornrainsdottir, A. Sucena, and F. Serrano, "Phonological development in relation to native language and literacy: variations on a theme in six alphabetic orthographies," *Cognition*, vol. 127, no. 3, pp. 398–419, 2013.
- [40] G. W. Lauth and M. Grünke, "Interventionen bei Lernstörungen," *Monatsschrift Kinderheilkunde*, vol. 153, no. 7, pp. 640–648, 2005.
- [41] W. L. Heward, "Ten Faulty Notions About Teaching and Learning That Hinder the Effectiveness of Special Education," *The Journal of Special Education*, vol. 36, no. 4, pp. 186–205, 2003.
- [42] F. Serafini, *The reading workshop: Creating space for readers*. Portsmouth, NH: Heinemann, 2001.
- [43] M. Adams, "Phonics and beginning reading instruction," *Reading, language, and literacy: Instruction for the twenty-first century*, pp. 3–23, 1994.
- [44] W. H. McGuffey, *McGuffey's eclectic primer: For young children, designed to precede Wm. H. McGuffey's eclectic readers*. [S.l.]: Nabu Press, 2010. [Online]. Available: [\url{http://www.gutenberg.org/cache/epub/14640/pg14640.txt}](http://www.gutenberg.org/cache/epub/14640/pg14640.txt)
- [45] O. Espinoza, A.-M. Bruno, and J. Deauvieu. (2013) Lecture au CP: un effet-manuel considérable. Saint Quentin en Yvelines. [Online]. Available: [\url{http://www.uvsq.fr/medias/fichier/rapport-enquete-lecture_1384503420148-pdf}](http://www.uvsq.fr/medias/fichier/rapport-enquete-lecture_1384503420148-pdf)
- [46] M. Sommer, L. Gaudin, T. Cuche, and G. Barril, *Léo et Léa: Méthode de lecture CP [6 vol.]*. Paris: Belin, DL 2009, cop. 2009.
- [47] I. Le Guay, *Je lis avec Kimamila 1*, ser. Un monde à lire. Paris: Nathan, DL 2012, cop. 2012.
- [48] I. Le Guay, M. Billet, I. Nicolle, and J.-N. Rochut, *Je lis avec Kimamila 2*, ser. Un monde à lire. Paris: Nathan, 2012.
- [49] R. Handt and K. Kuhn, *ABC der Tiere 1*. Offenburg: Mildenerger, 2010-2011.
- [50] N. Namour, G. J. Renk, W. Metze, K. Rahm, K. Stäpeler, S. Kierzek, M. Schramm, J. Arnold, F. Zauleck, and L. Lemke, *Jo-Jo Fibel: Ein Leselehrgang*, 2008th ed. Berlin: Cornelsen, 2011-.
- [51] H. Bartnitzky, *Kunterbunt / Fibel*, 1st ed. Stuttgart and Leipzig: Klett, 2009.
- [52] F. Stöwesand, *Lesebuch der Kleinen: Schreiblese- und Normalwortmethode, den Grundsätzen der Phonetik und mit Berücksichtigung der Schwachbegabten: Ausgabe A in zwei Teilen. Für Volksschulen: Erstes und zweites Schuljahr.*, 3rd ed. Magdeburg: Verlag von C.E.Klotz, 1903, vol. 1.
- [53] P. Seregini, M. Caimi, and C. Sirtori, *A come... più 1. Con e-book. Con espansione online. Per la Scuola elementare 1: Letture*. Pearson, 2015.
- [54] U. Reichel, "PermA and Balloon: Tools for string alignment and text processing," in *Proc. Interspeech*, 2012, pp. 1874–1877.
- [55] K. A. Spencer, "Feedforward, -backward, and neutral transparency measures for British English," *Behavior Research Methods*, vol. 41, no. 1, pp. 220–227, 2009.