

Systematic Acquisition of Reading and Writing: An Exploration of Structure in Didactic Elementary Texts for German

Kay Berkling, Rémi Lavalley

Cooperative State University
Karlsruhe, Germany

remiL@singularity.fr berkling@dhbw-karlsruhe.de

Uwe Reichel

Institute of Phonetics and Speech Processing
University of Munich, Germany

reichelu@phonetik.uni-muenchen.de

Abstract

The work presented here is part of a larger study that looks at the impact that structured teaching materials have on improving orthographic skills. The first step is an analysis of the current status of linguistic structures in common primary readers (primers) since their key purpose is to lead children along a systematic learning path towards becoming proficient readers and writers. Using text processing tools detecting morphemes, syllables and sentence structures, several popular primers are automatically examined with respect to their systematic approach to language at the phonics/syllable, word and sentence levels. It can be shown that there is little apparent structure in today's school texts at any of the examined levels when compared to some older schoolbooks (1877, 1904) that have a systematic and explicit progression at all examined levels.

1 Introduction

Over the last decade the number of children in Germany referred to language pathologists (Logopädie) has increased manifold according to WiDO AOK (health insurance) (2012; 2013) to reach an all-time high of 25% for boys aged around six. The cause for this phenomenon has not been well studied. Our hypothesis is that an improved structure in learning materials at the word and sentence level may improve orthographic skills. As a first step towards answering this question, this paper looks at the given structure in popular first grade readers (primers). To study the actual impact of structure on performance is beyond the scope of this paper but has been covered in part by other publications in this series of work (Berkling and Pflaumer, 2014;

Berkling et al., 2015). This line of research is further motivated by results stemming from a large-scale analysis of orthographic skills in a corpus from first grade until eighth grade. In order to understand why certain spelling errors persist, more analysis is clearly needed to show the impact of teaching methods on acquisition of orthography.

Much work has been done in studying various approaches for English reading and writing acquisition. The National Reading Panel (Donald N. Langenberg et al, 2000) and more up-to-date studies by Galuschka (2014) have published meta-reviews of major comparative and quantitative studies, realizing that phonics¹ is a vital component in reading instruction. It has been shown in the Anglophone research by Stahl (1989) that relying solely on the whole-word approach is to the detriment of the weaker students. It has also been shown by Steffler (2001) that making structure explicit leads to efficient learning results. To our knowledge, no comparative work has been done on the German language. Phonics in its complexity is also not known as a method for reading/writing acquisition in Germany. In that sense, phonics is not apparent in the elementary readers while other methods prevail.

Currently, in Germany, there are mainly 2-3 popular methods in use. One such method is called "Lautiermethode" and refers to the theory that words can be sounded out one letter at a time (Reichen, 2008; Brügelmann, 2014). By synthesizing the sounds, the word is supposed to be read in its entirety. This works approximately for words like "Oma" or "Banane" that is an imported word but deemed "lautgetreu", meaning one grapheme corresponds to one phoneme. However, this approach fails to generalize to typical German structures in which a single grapheme <e> can stand

¹The explicit and systematic (including clear sequence and scope) instruction of patterns in phoneme-grapheme correspondence)

for at least three semantically distinct pronunciations as a function of its position within syllable and word.

Another popular method says that reading acquisition takes place at the syllable level as this is more natural for children to bridge from the spoken language to the written representation (Röber-Siekmeyer, 2004). However, if the syllables are not taught well, then the learner pronounces the unstressed syllable in a stressed manner, thereby not recognizing the sound of the resulting word, for example by reading "Mutter" as /möt.te:r/ instead of /mʊtɐ/.

Finally, the whole word approach is widely used, even if not explicitly stated. It assumes that the child, given enough practice, will memorize the words as a unit. In general, experts say that first grade learning and teaching is based on a mix of all these methods and their respective effectiveness is child dependent.

Primers (*Fibel* in German) for first grade traditionally have the primary purpose of leading children in a systematic manner towards learning to read and write. As a first step towards understanding these methodologies and their progression, the goal of this work is to analyze and document selected structures and their progression in primary readers. The analysis looks at the lexical level through syllable and morpheme occurrence, word structures at the syllable level and complexity progression at the sentence level. Texts are analyzed for a number of books, including two older ones from 1877 and 1904 that show explicit structure unlike any of the ones used today.

The rest of the paper proceeds as follows. Section 2 motivates the proposed levels of analysis. After a brief overview of German phonics system and sentence complexity levels in Section 3, Section 4 will explain the tools used for analysis. Section 5 will introduce the data and Section 6 reports on the results, comparing the various different texts. Section 7 discusses the results.

2 Motivation

Elementary primers have traditionally had the purpose of serving as a guide for learners of a reading and writing system. In this function, they should naturally display a thought-out systematic approach on how they move from simple towards increasingly complex materials, considering the problem of inert knowledge (what children theo-

Grüner Abschnitt.

Die Laute und ihre Verbindung in Wörtern, Sätzen u. Lehrsätzen. Lautreue Schreibung.

A. Nur lange Vokale.

<p>Erste Stufe. Das Anfügen (a uersich) Schreiben (i uersich) und Lesen der Selbstlaute a, e, i, o, u, ei. (Nurstufe hat vorgearbeitet.) Der zu findende Laut ist die erste Silbe des Normalwortes. 4-6</p> <p>Normalwörter: i-da, e-mit, ei, u(u)-ren, o-ßen, a-dof.</p> <p>Zweite Stufe. Verbindung der Selbstlaute mit dehnbaren Mitlauten zu Wörtern mit ein- und zweilautigen,</p>	<p>offenen Silben. Der neue Laut steht am Anfang des Normalwortes. 6-15</p> <p>Normalw.: ka-ka, se-te, au-ka, re-ke, fe-ke, ma-ma, eu-ke, sch-ke, ha-ke.</p> <p>Dritte Stufe. Dreilautige, geschlossene Silben.</p> <p>Normalw.: wein. — w, ein.</p> <p>a) Einübung des w in Wörtern mit zweilautigen Silben.</p> <p>b) ein: kein, lein, nein, rein u. et: feil, heil, weil u.</p>
<p>Vierte Stufe. Ableitung des neuen Lautes vom Ende des Normalwortes. (Erstrechtlich auch auf die 3. Stufe.) Fortsetzung der dreilautigen Silben. . . 17—</p> <p>Nw. maus, rauch. (weich)</p> <p>Fünfte Stufe. Das tonlose e am Schluß und in leichten Endungen. 19—</p> <p>Nw. schule, eier, reifen.</p> <p>Sechste Stufe. Die Umlaute o und ü und das ie. Der neue Laut ist der zweite im Normalworte. 22—</p> <p>Nw. löwe, schäfer, wieje.</p> <p>Siebente Stufe. Die kleinen Druckbuchstaben der dehnbaren Laute. 25—</p> <p>Die Übungswörter treten zu Gruppen der Sätze und sachlichen Wortverbindungen allmählich zurück.</p> <p>Achte Stufe. Die Großbuchstaben der dehnbaren Laute. 30—</p> <p>1. Gruppe: D, N, H. } Im Anschluß an 2. Gruppe: C, G, G. } 3. Gruppe: S, Sch, N. } jede Gruppe 4. Gruppe: R, M, B. } Sätze. 5. Gruppe: F, U, S, D.</p>	<p>Neunte Stufe. Die Zischlaute. 32-42</p> <p>1. Zahnzischlaute. 32-34</p> <p>a) Nw. Daunen — D, d. 32</p> <p>I. Einübung des kleinen Buchstaben: d (die ersten drei Reihen der Wörtergruppe). II. Einübung des Großbuchstaben: D (die beiden folgenden Reihen). III. Einübung der Geschlechtsörter der, die, das.</p> <p>b) Nw. Tafel — T, t. 33</p> <p>I. Kleinbuchstabe (2 Reihen). II. Großbuchstabe (2 Reihen). III. Einübung ei. IV. Lesefrüchchen Nr. 1. Die Tafel. 33</p> <p>c) Vergleichung beider Zischlaute (Phonetische Übung. Auch zur Stärkung der Wortdeutlichkeit und für die Rechtschreibung): Dorf, toje u.</p> <p>d) Lesefrüchchen Nr. 2 und 3. Schöne Meier. Der Rhein. 34</p> <p>e) Einübung des Z und des ff. 35</p> <p>Nw. Star — S, s.</p> <p>I. Wörtergruppe. II. Einübung des ff.</p> <p>2. Lesefrüchlaute. 36-37</p> <p>a) Nw. Wäcker — W, w. 36</p> <p>I. Klein-, II. Großbuchstabe. III. Der Laut ff. IV. Lesefrü-</p>

Figure 1: Index pages for a primer around 1900

retically know but are unable to use) which can be reduced by usage in different contexts (Bereiter and Scardamalia, 1985).

When looking at old primers, this type of progression is often made explicit within the index or names of the chapters as shown in Figure 1. Today's books do not explicitly have a progression apart from the order of introduction of the graphemes. The goal of the work presented here is to visualize and render explicit any inherent progression that can be found at the syllable, morpheme or sentence structure level, keeping in mind that a systematic approach, including explicitness is important according to a study by Steffler (2001) to ensure the learners' grasp of linguistic structures. Finally, a look at progression should include a look at the complexity that can be reached.

It is well known, that learning builds on previous knowledge and that each new material should have small consecutive steps building on each other (Martin and Rose, 2005; Leong, 1998). As a consequence, complexity grows incrementally and builds on the knowledge of the preschooler, which is mostly based on syllable structure of the language and progresses from there as described by Siekmeyer (2009).

For the German language, the most common syllabic pattern is the "Trochee" type of word

(2-syllable length, first syllable stressed, second one unstressed). A comparable example for both English and German is the word "double" or "doppelt". The top 10,000 German words in newspapers include about 16% of pure 2-Syllable Trochees that follow this pattern. The 100 most used words cover around 45% in a standard text. The rest of the words are constructed words, such as compound words or those containing prefix or suffix attached to words that contain the Trochee pattern, where students have to generalize the pattern to unseen words – these make up about 45% of the top 10,000 words according to an automated analysis by Berkling (2014).

The remaining words, like "Auto", "Banane", "Tiger", "Portemonnaie" are imported vocabulary and follow different orthographic patterns that are not directly comparable to German. This is why the study of the Trochee is at the center of the word-pattern analysis.

Research questions to answer by analyzing the data are therefore the following: 1) Are the patterns of the German language (in form of Trochee) occurring in some structured form? 2) Is there repetition at syllable and morpheme level to support the cognitive process by achieving automation as described by McLaughlin (1990)? 3) Is the vocabulary embedded within sentence structures that exercise grammatical structures, including morpheme endings?

3 The Structure of German Text

In this section, the underlying theory concerning syllable types for the German language is explained. The structures underlying the study of word repetition and sentence complexity are also thereby defined and motivated.

3.1 Syllables

The German language distinguishes between three major classes of Trochee that will be described next.

- **Type 1: C-V-C-red ("b-e-t-en"):** This type of Trochee is the simplest form (phonemes: Consonant-Vowel-Consonant-reduction), comparable in that sense to the "cat", "hat", "mat" vocabulary used in the first steps of English phonics lessons. The 1-syllable form CVC derives directly from the 2-syllable Trochee ("gab" – as past tense of "geben"). In this case, the morpheme

boundary is within the Trochee. CVCred can not be reduced to a 1-syllable words when there is no morpheme boundary within the Trochee, examples are "bird" ("Vogel") or "vase" ("Vase"). Identifiable features of this word type are the tense/long vowel (V) followed by a single consonant sound and a reduction syllable that contains the letter <e> pronounced as schwa /ə/.

- **Type 2: C-v-C1-red ("B-e-tt-en"):** The second type of word is one of the most difficult features of the written language to master by learners of the orthographic system (phonemes: Consonant - short vowel - Consonant - reduction). It distinguishes itself from the first type only by the feature tense/lax of the vowel, perceived as shortness in this form. As in the first type, there is only one consonant sound in the center of the word. However, in the orthography the tenseness of the preceding vowel is denoted by duplicating the center consonant letter. Orthographically, there are regularities such as <tt>, <nn> and irregularities such as <ng>, <tz>, <ck>, or <sch> and <ch> that need to be mastered. The 1-syllable form CvC1 derives directly from the 2-syllable Trochee ("Betten" (plural) – "Bett" (singular)). In this case, the morpheme boundary is within the Trochee. Not all CvC1red will reduce to 1-syllable words, such as the word for "rattle" ("Rassel").

Identifiable features of this word type are the lax/short vowel (v) followed by a single consonant sound and a reduction syllable containing the letter <e> pronounced as schwa /ə/.

- **Type 3: C-v-C1-C2-red ("r-a-s-t-e-n"):** The third type of word is easier than the second one for the learner, as there are two distinct consonant sounds in the center of the word to help denote the tenseness of the preceding vowel. No orthographic particularities need to be mastered. The 1-syllable form CvC1C2 derives directly from the 2-syllable Trochee ("rasten" (verb) – "Rast" (noun)). In this case, the morpheme boundary is within the Trochee. Not all CvC1C2red will reduce to 1-syllable words, such as the word for *turn*: "Wende". Identifiable features of this word

type are the lax/short vowel (v) followed by two consonant sounds and a reduction syllable containing the letter <e> pronounced as schwa /ə/.

- **Other categories** of words include foreign words and high-frequency words as well as constructed words (compound words and those including prefix and suffix). These are not considered separately for the purpose of this study as they decompose either into the aforementioned formats or do not pertain to German phonics rules.

3.2 Sentences

There are numerous publications regarding the definition of sentence readability (Glöckner et al., 2006; Sitbon and Bellot, 2008; DuBay, 2008). Nelson (2012) is an example of a recent overview of such measures. Most of them have been designed for English. Regarding German language, little research has been done on that subject. We can cite (Hancke et al., 2012) where the authors used a number of different features to determine sentence readability: average number of words, characters, syllables, lexical features (noun and verb token ratios, textual lexical diversity, ...), syntactic features (number of noun or verb phrases, average length of a noun phrase, ...), language models (trained with texts for children vs. newspapers) and morphological features (ratio of finite verbs, compounds, ...).

However, the goal of the study presented here is not exactly to measure a sentence readability, but to describe the progression in the structures' complexity (word-level clues are considered separately, as described in Subsection 3.1). Hence, for the purpose of written language acquisition a progression as described by Clahsen (1982; 1988) for L1 acquisition in children is chosen as a first approach. (A more detailed analysis of sentence structures can be found in Berkling (2014)). For the purpose of this study, sentence structures are distinguished at the following levels:

- **LEVEL 1:** One-word utterances without counting articles, so it covers simple noun phrases such as "Peter" and "eine Katze" (a cat) and imperative verbs "Lauf!" ("Run!").
- **LEVEL 2:** Two-word utterances without counting articles, so simple noun phrases (as described for level 1) + Verb in present tense

("Peter isst" ("Peter is eating"). "Eine Katze läuft" ("A cat is running") or noun phrases including adjectives not followed by a verb: "die kleine Katze" ("the little cat").

- **LEVEL 3:** Common sentences, including Adjectives, Adverbs, ... So, noun phrases (including complex ones, such as article + adjective + noun) + Verb with or without Object: "Die kleine Katze." ("the little cat."). "Die kleine Katze läuft" ("The little cat is running"). "Ich gehe nach Hause" ("I am going home").

- past and future tenses, built with a conjugated auxiliary (2nd position) and an infinitive (future) or past participle (past tense) form of the meaningful verb, located at the end of the sentence: "Wir werden nächsten Sommer nach Spanien fahren". ("We *will go* to Spain next Summer"). "Du *hast* spät in der Nacht gearbeitet". ("You *have worked* late in the night").
- modal verbs: used to express obligation, ability or will. These verbs are conjugated and located in the second position in the sentence, then the verb on which the modality applies is located at the end of the sentence: "Du *sollst* morgen nicht kommen" ("You *should not come* tomorrow").
- compound verbs: conjugated verb is in 2nd position, the particle is at the end of the sentence: "Die Sonne geht immer früher auf" ("The sun *rises* earlier and earlier") – The compound verb "aufgehen" (to rise) has a different meaning than the verb "gehen" (to go).

- **LEVEL 5:** complex sentences. Depending on the kind of clause, verb positions follow different rules. Thus, we have distinguished two subcategories, to reflect different complexities: the coordinate considered as an independent sentence, where the verb is located in 2nd position (in both of the clauses). e.g., "Ich bin fertig *und* ich gehe jetzt in die Schule" ("I *am* ready *and* I *go* to the school now"), or in subordinate clauses with the verb located at the end of the subordinate clause: "Ich denke, *dass* er intelligent ist" ("I *think that* he *is* smart").

Sentence containing several levels are classified as the highest: "Ich werde ein Buch kaufen, weil ich viel lese." ("I will buy a book, because I read a lot.") – is thus considered as Level 5.

4 System

Three software modules as shown in Figure 2 are built to analyze the primary reader texts. A synthesizer, that performs automatic syllabification and morpheme boundary segmentation called Balloon, a word classification system (Berkling and Reichel, 2014) that builds on Balloon output, and a syntactic parser (Petrov et al., 2006) followed by a sentence classifier.

4.1 Balloon and Lexical Analysis

The word counter uses the output of Balloon, which works roughly as follows, for details see (Reichel, 2012).

The Grapheme-phoneme (G2P) conversion is carried out by a C4.5 decision tree (Quinlan, 2003). Syllable boundaries are placed in front of each sonority minimum, and their locations are subsequently adjusted in case German syllable phonotactics is violated. Word stress is again assigned by a C4.5 tree that predicts for each syllable whether it is stressed or not. Part of speech (POS) labels were assigned by a Markov tagger that additionally makes use of information stored in word suffix strings. Relevant suffix strings are extracted by means of an adaptation of the peak and plateau algorithm of Nascimento (1998). As POS inventory the Stuttgart Tübingen tag set (Schiller et al., 1995) is used.

The morphological analysis yields a flat segmentation of a word into morphemes and their morpheme classes. Each word is therefore decomposed into phonemes, syllables, stress markers and morpheme boundaries that are then used by the morpheme and syllable counter. The word counter simply counts new and previously seen syllables and morphemes for each page and then plots new vs. previously seen numbers for each page. The word classifier uses syllable and morpheme boundaries to filter out the 2-syllable words that match the Trochee types and their 1-syllable derivatives as described in Section 2 by their unique identifying features in a rule based system.

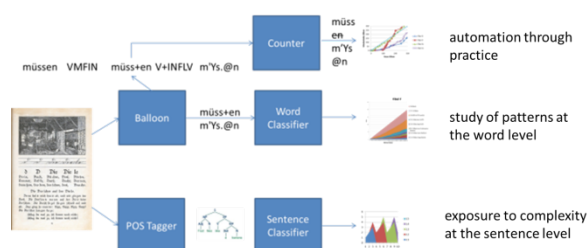


Figure 2: System Modules for calculating Features

4.2 Sentence Classifier

The sentence classifier takes its input from the Berkley's parser for German, with *-tokenize* (to use the integrated tokenizer) and *-accurate* (favours accuracy over speed) options. A small set of simple rules - including ones designed to overcome some mistakes made by the parser - is used to assign levels according to Section 3.3 in decreasing order:

- looking for subordinate or coordinate clauses; this information is generally provided by the Berkley's parser. However, to increase recall (some of these clauses were not tagged as such by the parser), it has here been re-inforced by an upper layer using a list of words introducing subordinate clauses (dass, ob, ...) or coordinate clauses (und, oder, ...) and looking for verbs positions around these specific words in order to make the distinction between "Ich habe einen Bruder und eine Schwester." ("I have a brother and a sister.") and "Ich spiele und ich **arbeite**." ("I play and I **work**.").
- looking for auxiliary in 2nd position and past participle or infinite verb at the end of the sentence for level 4. Or finite verb in second position and particle at the end of the sentence.
- number of words and their part of speech for levels 1 and 2.

5 Data

Various primary texts were transferred into electronic format up to page 50. The primers were chosen for their popularity and opposing methodologies ranging from no apparent methodology towards syllable, grapheme, and whole word approach. The two old primers from 1877 and 1904

have been analyzed only in part. Their makeup is different from today's primers. Separate sections are devoted to grapheme introduction, word-level training and sentence level training within progressively complex texts (as measured by sentence and word structure complexity). These sections have been extracted to demonstrate progression. Table 1 lists the available data.

Readers	Primary Methodology	Number of Words	Number of pages
Primer A	Whole word	1259	First
Primer B	Syllable strict	593	50
Primer C	Phonics 1877	580	pages
Primer D	Syllable	1010	(except
Primer E	Grapheme	1265	old
Primer F	Grapheme	932	primers)
Primer G	None	1328	
Primer H	Phonics 1904	320	

Table 1: Primary Readers Used in Study.

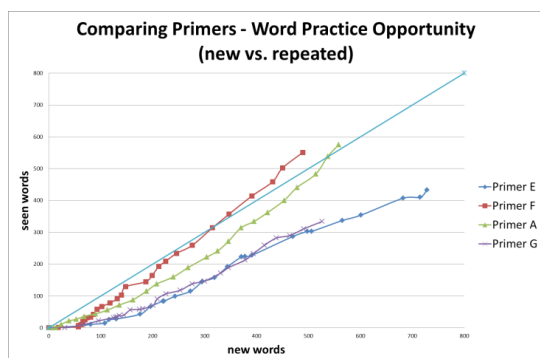


Figure 3: Repetition of words for practice, comparing four primers.

6 Evaluation

Due to limited space, only the results for selected readers are displayed where they show prominent differences. Results are reported for morpheme and syllable repetition, progression of Trochee types, and sentence complexities.

6.1 Ability to Practice on Words and Syllables

Figures 3 and 4 show how many words or syllables are repeated or new for each page. There is quite a large difference in training when comparing Primers F & A with E & G, where there is substantially less repetition.

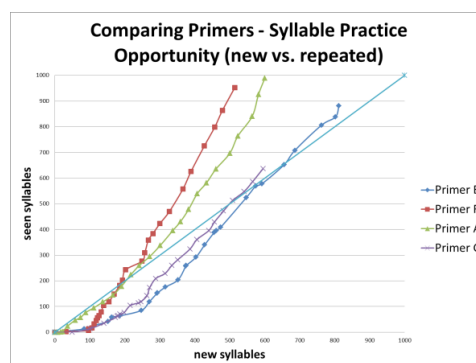


Figure 4: Repetition of syllables for practice, comparing four primers.

6.2 Progression towards Complexity at word level

The three diagrams in Figures 5 and 6 (See Appendix) show three different versions of progression for the different Trochee types and their 1-syllable derivatives. The figures are generated by looking at the distribution across the word types for every 5 consecutive words. Each slice on the x-axis therefore represents five words. The y-axis counts the number of word types for that slice from 1-5.

6.3 Progression at the Sentence Level

As depicted in the Appendix, Figures 7 to 9 show the progression in sentence complexity throughout the primers for three different books. Figure 7 depicts Primer F and some progression in the sentence structure can be seen throughout the reader. At the beginning of the book, there are lots of simple utterances (mainly Level 1, so Noun or Article+Noun), no Level 5 sentences (subordinate and coordinate clauses) and a few Level 3 sentences. As we proceed, we can observe more and more Level 3 sentences and the appearance of Level 4 sentences (verb tricks), while the number of Level 1 sentences decreases significantly. At the end of the book, there are lots of Level 4 and 5 sentences and almost no simple ones.

In contrast, Figure 8 depicts Primer E with no progression at all in the sentences structure. All the levels are merged all along the book, meaning children learn a few words and directly use them in complex sentences. Level 3 is the highest level of complexity reached, significantly diminishing the syntactic exercises that more complex sentences offer.

Figure 9 shows Primer from 1904 depicting the

progression taken from the special section on practicing sentence complexity. This progression is visible in the diagram. At the beginning of the analyzed part of the book, it can be observed that most of the sentences belong to Level 3. Then their number is slightly decreasing as we browse to the end of the book. Meanwhile there are an increasing number of Level 4 and 5 sentences. The number of high complexity sentences is significantly larger here than observed in contemporary primers.

7 Conclusion

Looking at primers through a quantitative lens revealed large differences in primary reading material for first graders who are introduced to reading with these materials. It is important to be aware of such differences that may not be apparent immediately. Today, these effects are not studied in quantitative, systematic manner and teachers are not aware of the detailed particularities of their materials. Obviously, these primers are only one component of many materials chosen by teachers to work with the children in training their reading and writing skills. This type of analysis covers only one aspect of the classroom dynamics. However, it would be desirable to see a clear progression in the examined material and an explicit goal of the skills that the students in first grade should be able to reach along with a defined progression in that direction. Based on our work, there is no evident answer as to which methodology has a clearer progression at different levels. None of the contemporary primers show the distinctive marks of the two chosen older versions that contain an apparent progression at the word and the sentence level that takes the elementary student from the simple to the complex. Looking at the older primers, much more complexity was demanded from first graders at the end of the school year. It may be that in order to reach that level more practice opportunities were supplied and explicit progression at various levels was needed.

Future work will have to study whether there is a cause and effect here. The old approach provides a stark contrast to our primers in use today and today's primers differ significantly even among each other regarding practice and progression. None of the modern readers seem to spend the time (when compared to the older primers) on the typical German word structure of the Trochee. This work has

shown that there are differences between readers, some of which are not well studied, and that there is a need to study the consequences this material has on student's learning and achievements.

References

- Carl Bereiter and Marlene Scardamalia. 1985. Cognitive coping strategies and the problem of "inert" knowledge. *Thinking and learning skills: Current research and open questions*, 2:65–80.
- Kay Berkling and Nadine Pflaumer. 2014. Phontasia - a Phonics Trainer for German Spelling in Primary Education. In *Workshop on Child Computer and Interaction WOCCI*, pages 33–38, Singapore.
- Kay Berkling and Uwe Reichel. 2014. Der phonologische Zugang zur Schrift im Deutschen. In *Symposium Deutsch Didaktik, Sektion 7, Orthographie.*, Basel, CH.
- Kay Berkling, Nadine Pflaumer, and Rémi Lavalley. 2015. German phonics game using speech synthesis: A longitudinal study about the effect on orthography skills. In *Workshop on Speech and Language Technology in Education SLaTE*, pages 167–172, Leipzig, Germany.
- Hans Brügelmann. 2014. *Kinder auf dem Weg zur Schrift: Eine Fibel für Lehrer und Laien*. Libelle : Wissenschaft. Libelle-Verl., Bottighofen, 9th edition.
- Harald Clahsen. 1982. *Spracherwerb in der Kindheit: Eine Untersuchung zur Entwicklung der Syntax bei Kleinkindern*, volume 4 of *Tübinger Beiträge zur Linguistik. Series A, Language development*. G. Narr, Tübingen.
- Harald Clahsen. 1988. Parameterized grammatical theory and language acquisition. *Linguistic theory in second language acquisition*, pages 47–75.
- Donald N. Langenberg et al. 2000. National Institute of Child Health and Human Development. 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.
- William H. DuBay. 2008. The principles of readability. 2004. *Costa Mesa: Impact Information*, 76.
- Katharina Galuschka, Elena Ise, Kathrin Krick, and Gerd Schulte-Körne. 2014. Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PloS one*, 9:2.
- Ingo Glöckner, Sven Hartrumpf, Hermann Helbig, Johannes Leveling, and Rainer Osswald. 2006. An architecture for rating and controlling text readability. In *KONVENS*, pages 32–35, Konstanz, Germany.

- Christine Göpner-Reinecke. 2013. Immer mehr Kinder brauchen vor dem Schulstart Hilfe beim Sprechenlernen. Accessed: September 14, 2015, http://www.wido.de/meldung_archiv+M50c012ae424.html.
- Julia Hancke, Sowmya Vajjala, and Detmar Meurers. 2012. Readability Classification for German using Lexical, Syntactic, and Morphological Features. In *COLING*, pages 1063–1080, Mumbai, India.
- Rémi Lavalley and Kay M. Berkling. 2014. Data exploration of sentence structures and embellishments in german texts: Comparing children’s writing vs literature. In *KONVENS*, pages 241–247, Hildesheim, Germany.
- Deborah J. Leong. 1998. Scaffolding emergent writing in the zone of proximal development. *Literacy*, 3.2:1.
- James R Martin and David Rose. 2005. Designing literacy pedagogy: scaffolding democracy in the classroom. *Continuing Discourse on Language*, pages 251–280.
- Barry McLaughlin. 1990. Restructuring. *Applied linguistics*, 11.2:113–128.
- Mario A. Nascimento and Adriano C.R. da Cunha. 1998. An experiment stemming non-traditional text. In *Proc. SPIRE’98*, pages 74–80, Santa Cruz de La Sierra, Bolivia.
- Jessica Nelson, Charles Perfetti, David Liben, and Meredith Liben. 2012. Measures of text difficulty: Testing their predictive value for grade levels and student performance. *Council of Chief State School Officers, Washington, DC*.
- Slav Petrov, Leon Barrett, Romain Thibaux, and Dan Klein. 2006. Learning accurate, compact, and interpretable tree annotation. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 433–440.
- John Ross Quinlan. 2003. *C4.5: programs for machine learning*. The Morgan Kaufmann series in machine learning. Morgan Kaufmann, Amsterdam [u.a.], 5th edition.
- Uwe Reichel. 2012. PermA and Balloon: Tools for string alignment and text processing. In *Proc. Interspeech*, pages 1874–1877, Portland and Oregon.
- Jürgen Reichen. 2008. *Hannah hat Kino im Kopf: Die Reichen-Methode Lesen durch Schreiben und ihre Hintergründe für LehrerInnen, Studierende und Eltern*. Heinevetter, Hamburg, 5th edition.
- Christa Röber-Siekmeyer. 2004. Die Berücksichtigung des kindlichen Sprachwissens für den Schrifterwerb. In *Geschriebene Sprache. Strukturen, Erwerb, didaktische Modellbildungen. Schriftenreihe der PH Heidelberg Beltz Verlag Wissenschaft*, pages 129–144.
- Christa Röber-Siekmeyer. 2009. *Die Leistungen der Kinder beim Lesen- und Schreibenlernen: Grundlagen der silbenanalytischen Methode ; ein Arbeitsbuch mit Übungsaufgaben*. Schneider-Verl. Hohengehren, Baltmannsweiler.
- A. Schiller, S. Teufel, and C. Thielen. 1995. Guidelines fuer das Tagging deutscher Textcorpora mit STTS, url = <http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf>, accessed: sept. 15, 2015, address = Stuttgart.
- Laurianne Sitbon and Patrice Belloc. 2008. A readability measure for an information retrieval process adapted to dyslexics. In *Second international workshop on Adaptive Information Retrieval (AIR 2008 in conjunction with IiX 2008)*, pages 52–57.
- Steven A. Stahl and Patricia D. Miller. 1989. Whole language and language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research*, 59.1:87–116.
- Dorothy J. Steffler. 2001. Implicit cognition and spelling development. *Developmental Review*, 21.2:168–204.
- Andrea Waltersbacher. 2012. Jeder vierte Junge zur Einschulung in Sprachtherapie. Accessed: September 14, 2015, http://www.wido.de/meldung_archiv+m50f798ala5a.html.

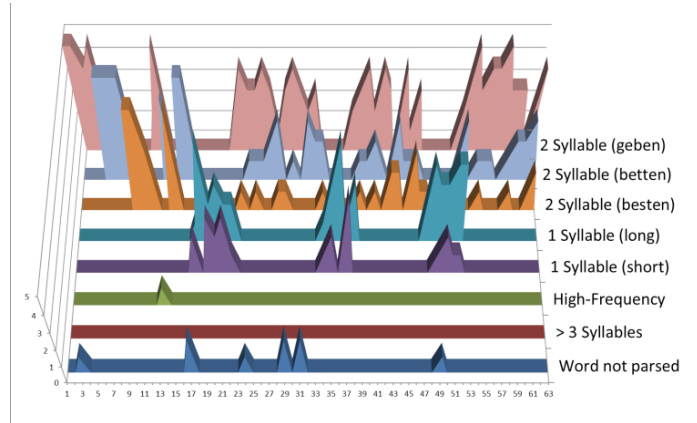


Figure 5: An excerpt from Primer H from 1904, showing progression from 2-syllable types “geben” to “betten” to “besten” to 1-syllable types and then proceeding to mix all variants in training. (x-axis = 5 word slice, y-axis = number of words of that type).

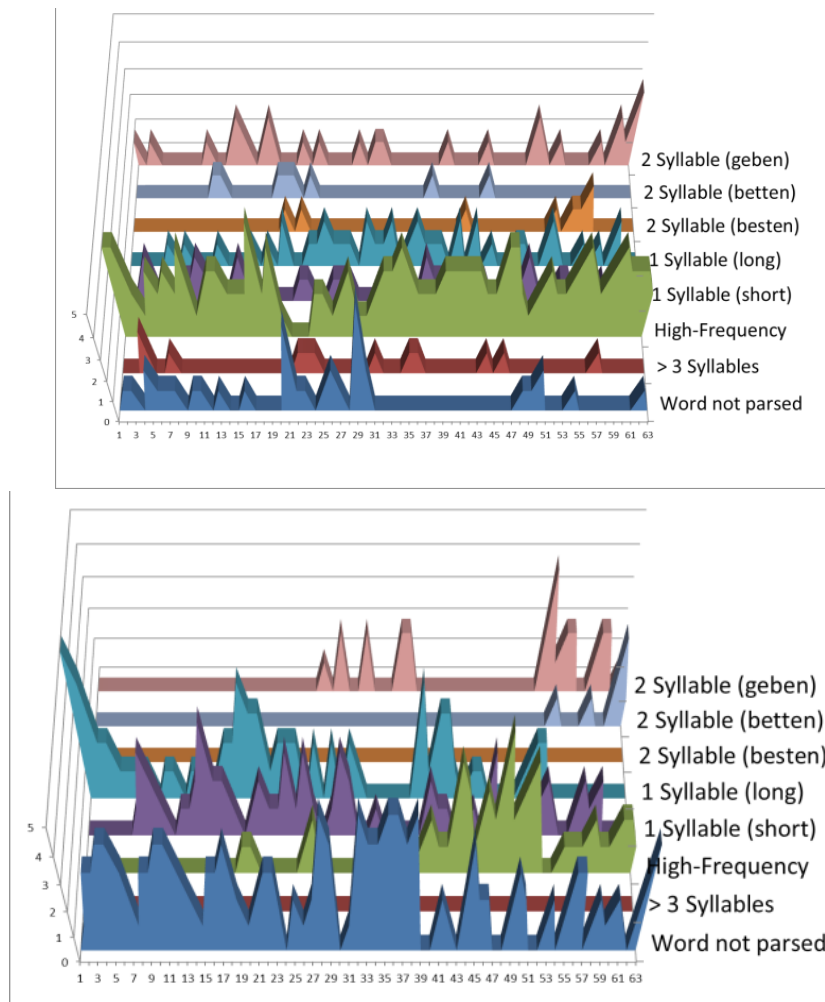


Figure 6: Primers E (at the top) and D show two different ways of using words. While E spends more time on High-Frequency words, Primer D spends time on syllables that are not real words. (x-axis = 5 word slice, y-axis = number of words of that type).

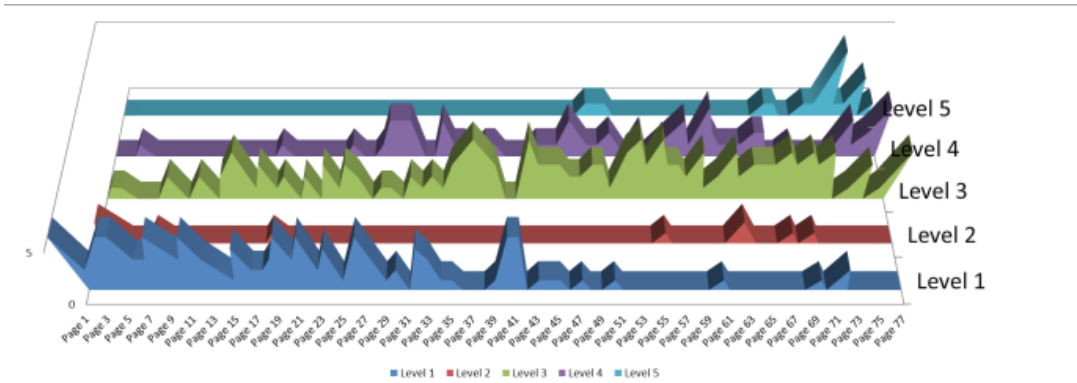


Figure 7: Progression of sentence level complexities for Primer F. Some progression to Level 5 is visible with little practice at level 5.

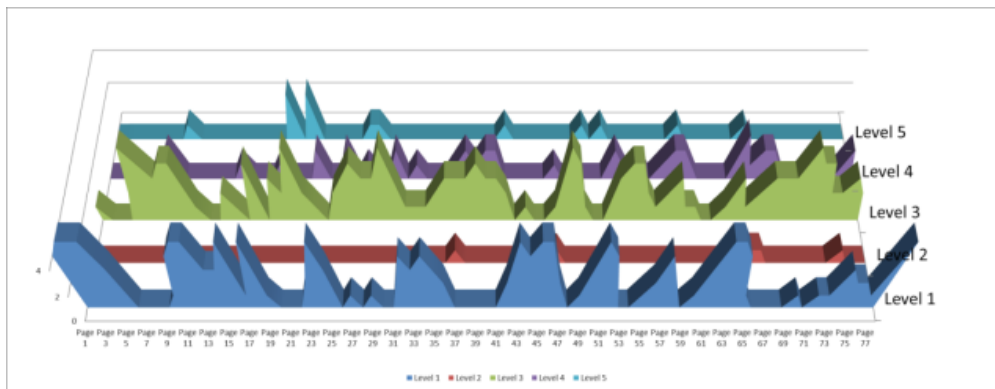


Figure 8: Progression of sentence level complexities for Primer E. Level 3 and 1 are exercised but there is no visible progression.

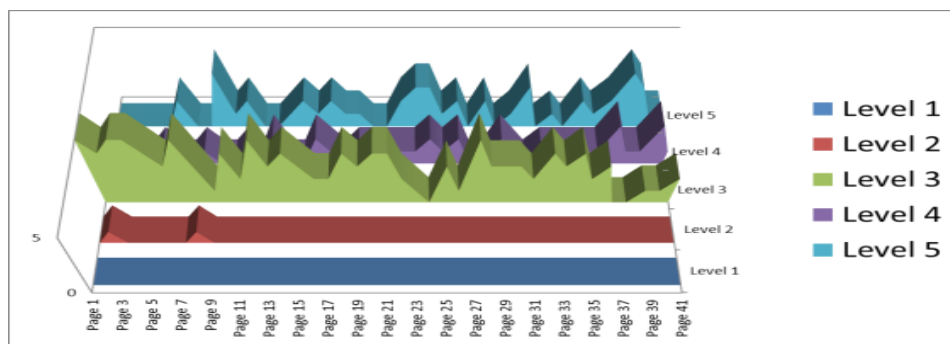


Figure 9: Progression of sentence level complexities for Primer H (1904) from Level 3 to Level 5 with increasingly significant practice at Level 5.