Visual word processing in the context of learning to read

Eberhard-Moscicka, Aleksandra K

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-152371
Dissertation
Published Version

Originally published at:
Eberhard-Moscicka, Aleksandra K. Visual word processing in the context of learning to read. 2015, University of Zurich, Faculty of Arts.
Neurocognitive mechanisms of learning to read: print tuning in beginning readers related to word-reading fluency and semantics but not phonology

Aleksandra K. Eberhard-Moscicka, Lea B. Jost, Margit Raith, Urs Maurer

Department of Psychology, Division of Cognitive Neuroscience, University of Zurich, Switzerland
Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland

Abstract
During reading acquisition children learn to recognize orthographic stimuli and link them to phonology and semantics. The present study investigated neurocognitive processes of learning to read after one year of schooling. We aimed to elucidate the cognitive processes underlying neural tuning for print that has been shown to play an important role for reading and dyslexia. A 128-channel EEG was recorded while 68 (Swiss-) German monolingual first grade children (mean age: 7.6) performed a one-back task with different types of letters and false-font strings. Print tuning was indexed by the N1 difference in the ERPs between German words and false-font strings, while the N1 lexicality effect was indexed by the difference between German words and pseudowords. In addition, we measured reading fluency, rapid automatized naming, phonological awareness, auditory memory span, and vocabulary. After one year of formal reading instruction N1 print tuning was clearly present at the group level, and could be detected at the individual level in almost 90% of the children. The N1 lexicality effect, however, could not be reliably found. On the cognitive level, next to word-reading fluency, also vocabulary was associated with N1 print tuning, but not measures reflecting phonological processing. These results demonstrate the presence of print tuning in the first year of

reading acquisition and its development at the individual level. Moreover, individual differences in print tuning are not only related to word-reading fluency, but also to semantic knowledge, indicating that top-down modulation of print tuning are rather of semantic than phonological nature.

1. Introduction
Reading acquisition is a major step in child development. Some children learn reading without significant problems; some, however, face substantial difficulties (National Center for Education Statistics, 2011). Together with reading training a new functional network emerges at the neural level. This functional network comprises a set of bilaterally distributed areas in occipito-temporal, temporo-parietal, as well as precentral and inferior frontal regions (see e.g., Dehaene et al., 2010; Price & Mechelli, 2005; Schlaggar & McCandliss, 2007; Turkeltaub et al., 2002) Yet, some of these neural processes have an individual trajectory (Parviainen et al., 2006), and may be related to developmental differences in cognitive reading components.

One example of a functional correlate for reading is the N1 (or N170) coarse-tuning for print. This neural tuning for print occurs early during visual processing (at about 150 to 200 ms after stimulus onset) and is indexed by an increased N1 component of the ERP for visual words compared to control stimuli, such as symbol strings (Bentin et al., 1999; Maurer et al., 2006). Showing posterior negativity and fronto-central positivity it follows the P1 component, and originates from bilateral occipito-temporal regions, typically with left hemisphere dominance (e.g., Bentin et al., 1999; Brem et al., 2005; Brem et al., 2009; Parviainen et al., 2006; Tarkiainen et al., 1999). Development of the N1 print tuning has been shown to follow an inverted U curve with initial increase and its subsequent decrease with age (Brem et al., 2009; Maurer et al., 2006). To date, it has been extensively studied at the group level (e.g., Araújo et al., 2012; Brem et al., 2009; Brem et al., 2010; Maurer et al., 2006; Maurer et al., 2007). However, to our knowledge, only one study (Parviainen et al., 2006) investigated this neural correlate at the individual response level. At the
group level Maurer et al. (2006) found that the N1 print tuning, absent in kindergarten children (despite their considerable letter knowledge), emerged in the second grade, after children mastered initial reading abilities. Most impressively, Brem and colleagues (2010) recently demonstrated that the N1 print tuning, absent in illiterate kindergarteners, emerged after a short grapheme-phoneme training. Importantly, this print sensitivity declined after discontinuation of the training. Hence, the authors suggested that without continuing practice, the sensitivity for print gained with the grapheme-phoneme training could not be consolidated. Contrary to these robust group effects, using an individual level approach in an MEG study, Parviainen et al. (2006) found word-specific sources only in slightly more than half of the first graders assessed. Considering the broad evidence on rapid emergence of N1 print tuning at the group level, the question arises whether the word-specific sources were really absent in the first grade children assessed by Parviainen et al. (2006), or whether they were undetectable due to insensitivity of MEG to deep or radial sources (Michel et al., 2009), which may contribute to print tuning in children. Moreover, regarding the fact that N1 specialization for print occurs early in the course of learning to read (Brem et al., 2009; Brem et al., 2010; Maurer et al., 2005a; Maurer et al., 2007a), it seems critical to better understand this process in the first year of reading training, and further, to investigate it not only at the group but also at the individual response level. Contrary to the group analysis, the individual subject analysis offers a unique opportunity to investigate this phenomenon more deeply, hence allowing to draw more precise conclusions which are not restricted to group averages (Salmelin et al., 1996).

Whereas N1 print tuning shows consistent differences in amplitudes between letter and symbol strings, and therefore, is a robust neurophysiological marker for print specialization (e.g., Araújo et al., 2012; Brem et al., 2005; Brem et al., 2009; Maurer et al., 2005a; Maurer et al., 2006; Maurer et al., 2007a; Maurer et al., 2011), studies do not agree in finding differences between different kinds of letter strings. Possible reasons for these inconsistencies are the typically small effect sizes, the influence of several factors on stimulus processing (such as
fine-tuning for orthographic patterns, i.e., bigram frequencies or frequencies of higher order letter combinations), as well as task demands that might influence phonological and lexical-semantic processing (Maurer et al., 2006).

Concerning lexical-semantic properties, it has been shown that low-frequency words evoked stronger brain responses than high-frequency words in adults (Assadollahi & Pulvermuller, 2003; Hauk & Pulvermuller, 2004), but not in pre-adolescent children (Araújo et al., 2012). Furthermore, results of ERP studies that investigated visual processing of words and pronounceable pseudowords are equivocal, with some studies indicating a significant difference in N1 between these two types of stimuli (Hauk et al., 2006b; McCandliss et al., 1997) and other studies finding no such differences (Araújo et al., 2012; Bentin et al., 1999; Kast et al., 2010). In addition to the factors discussed above, also developmental aspects and properties of the writing system may contribute to inconsistent results across studies (Maurer et al., 2005b).

Most interestingly, Maurer et al. (2006) found that, while learning to read, second-grade children developed an N1 specialization for words over pseudowords, with larger amplitudes for words. This lexicality effect was, however, absent in kindergarten children (Maurer et al., 2005a) and in adult readers of shallow orthographies when the same task was used (Maurer et al., 2005a; Maurer et al., 2006). According to these results, it is possible that this N1 difference in processing words compared to pseudowords is pronounced at the initial stages of learning to read, and diminishes with growing literacy. Investigating first grade children would allow testing whether both coarse-tuning and the lexicality effect have already developed, or whether the lexicality effect develops only later in life. Moreover, a large group of children with a wide range of reading skills allows further examination of correlations between reading skills and the size of the lexicality effect.

There is a growing number of studies that provide support for the assumption of impaired N1 print tuning in dyslexic children (e.g., Araújo et al., 2012; Maurer et al., 2007a) and adults (e.g., Helenius et al., 1999; Mahé et al., 2012). Whereas electrophysiological tuning in the left occipito-temporal cortex for words compared to symbol strings had
emerged in the time course between kindergarten and the second grade in good readers, it was reduced in children who developed dyslexia (Maurer et al., 2007a). Also functional neuroimaging meta-analytic results support these findings and emphasize the importance of an early recruitment of the left occipito-temporal region in the development of reading, and an early failure of this engagement in dyslexia (Richlan et al., 2011). Given that in shallow orthographies dyslexia is defined by reading fluency (Wimmer & Schurz, 2010), and given the attenuated N1 print tuning in dyslexic subjects, Maurer et al. (2007a) proposed print tuning to index reading fluency. However, the question arises whether, next to reading fluency, there are other cognitive reading components that are related to print tuning.

As yet, the best predictors of early reading achievement are commonly recognized to be phonological awareness, basic letter knowledge and rapid automatized naming (e.g., Pennington & Lefly, 2001; Puolakanaho et al., 2008). These precursors of reading, predicting subsequent learning to read, can therefore be called cognitive reading components (Vellutino et al., 2007). Moreover, while there is not much debate about the central role of phonological skills in the course of learning to read, a growing body of empirical work indicates the importance of additional cognitive reading components, such as vocabulary and comprehension (Nation & Snowling, 2004). Ricketts and colleagues (2007) examined the relationship between reading skills and oral vocabulary in a large population of children between 8 and 9 years. They found that oral vocabulary predicted reading comprehension and exception word-reading but did not predict text reading accuracy, nonword-reading and regular word-reading. Hence, there are a number of cognitive reading components that are linked to reading, which in turn is reflected in N1 print tuning. Given that some studies indicated a modulation of visual word processing by phonology (Shaywitz et al., 2004; Parviainen et al., 2006; Xue et al., 2006), such a modulation could also be expected for print tuning. Despite considerable number of studies on N1 print tuning, to our knowledge, no study discussed the direct effects of cognitive reading components on this neurophysiological marker for print processing.
In the present study, we investigated a sample of 68 (Swiss-) German first graders (mean age: 7.6), with a wide variety of reading skills, in order to answer the following questions: In a first set of questions about the development of visual word processing at the neural level we tested (I) whether N1 coarse-tuning for print had developed by the end of the first grade; (II) whether the word-specific occipito-temporal N1 specialization was present in all first graders on an individual basis; (III) and whether the N1 lexicality effect had emerged after one year of schooling. In a second set of questions we investigated cognitive correlates of print tuning by testing (IV) whether word-reading fluency predicts the size of print tuning across children with a wide range of reading skills; (V) whether there are other cognitive reading components, next to word-reading fluency, that explain additional variance in print tuning; and finally (VI) which combination of cognitive reading components explains print tuning best?

2. Methods
2.1. Participants
We report data of 68 monolingual (Swiss-) German-speaking children at the end of first grade between the age of 6.7 and 8.5 years (27 females and 41 males; 63 right and 5 left handed). From an original group of 70 subjects, two subjects were excluded either due to low signal-to-noise ratio or due to poor behavioral performance (omission of all targets in one condition). All subjects had normal or corrected-to-normal vision and a scaled score that was equal or above 6 (corresponding to estimated non-verbal IQ equal or above 80) in a block design task of an intelligence scale for children (Petermann & Petermann, 2010, see below). The assessed children took part in a larger longitudinal study about learning English as a foreign language, and the data presented here are from the session before learning English. The study protocol was in agreement with the local ethics committee. Participants’ parents filled out a background questionnaire screening for a history of neurological diseases and psychiatric disorders. Consent was obtained orally from children and in written form from their parents.
2.2. Procedure
All children participated in a behavioral and an EEG session administrated on separate days with the mean difference between the two sessions of 11.06 days (±11.13). The behavioral session took about 1.5 hours and took place either at schools (in a separate room provided by schools), at the Department of Psychology at the University of Zurich or at participants’ homes. The EEG session was administrated using one of two identical portable EEG systems (Electrical Geodesics, Inc, EGI). The recording was approximately 3.5 hours long and was administrated either in a separate room provided by schools or in the EEG laboratory at the Department of Psychology at the University of Zurich. Before using a room at schools, a standard quality check was applied in order to ensure the absence of 50 Hz noise. As compensation, children received a book voucher of 40 CHF and a written report about their reading skills.

2.3. Behavioral assessment
During the cognitive session the child was seated opposite the experimenter and performed a set of behavioral language tests (for a detailed list of subtests see Table 1). Behavioral measures assessed different aspects of language processing, such as word-reading fluency (Landerl, Wimmer, & Moser, 2006; Moll & Landerl, 2010), pseudoword-reading fluency (Landerl et al., 2006; Moll & Landerl, 2010), rapid automatized naming (RAN; one syllable animals, three syllable animals, digits, Landerl, 2001; lower case letters, Landerl et al., 2013), phonological awareness (phoneme deletion and pseudoword segmentation, Stock et al., 2003), sentence-reading fluency (Mayringer & Wimmer, 2005), vocabulary (vocabulary, Petermann & Petermann, 2010), auditory memory span (digit span forwards and backwards, Petermann & Petermann, 2010; pseudoword span, Landerl et al., 2013), spelling (Moll & Landerl, 2010) and non-verbal IQ (block design, Petermann & Petermann, 2010). The spelling task occurred to be too difficult for the first graders, and therefore was not used for further analysis. Non-verbal abilities, measured by the block design subtest, served as approximation of participants’ non-verbal IQ (Landerl et al.,
2013) that was used as an exclusion criterion (scaled score < 6). The other measures were used as potential predictors in the regression approach (see below).

2.4. EEG session
Participants were seated 80 cm away from the computer screen and pressed a mouse button whenever a stimulus occurred twice in a row (Fig. 1). This one-back task, assessing rapid specialization for print, was part of a larger session that included several experiments presented in a pseudo-randomized order. Children were allowed to take breaks between experiments and were monitored for compliance with a digital camera. The stimuli presented were familiar German words (high frequency of occurrence in the text books of children aged 6-8, $M = 161.86$/Mio, ChildLex Lexical Database, Schroeder et al., 2015), false-font strings (matched with German words, see below), English words and pseudowords (pronounceable in German and English, created from letters that appeared in German and English stimuli). Due to the limited number of English words that we expected children to know at the follow-up session after one year of learning English (not part of this study, which focuses on children before learning English), we limited the number of items per condition to 14. The 14 stimuli per condition were repeated six times (84 stimuli per condition) and presented in six blocks (the order of conditions was counterbalanced), and in addition 12 immediate repetitions serving as targets were presented in each condition. We used the blocked design to be consistent with previous studies (e.g., Maurer et al., 2005a, 2005b, 2006, 2007a). The stimuli were presented in black (Arial, bold, font size 28) and occurred in the center of a white rectangular box (85 mm x 47 mm) in the middle of a grey background. The stimulus duration was 500 ms followed by a mean interstimulus interval of 1500 ms (jittered between 1250-1750 ms). The stimuli were matched for string length and contained 3.9 letters/false-font characters on average (range: 3-5; average length and height: 31.9 mm x 7mm). In addition, German words, pseudowords and English words were matched for number of letters, frequency of letters and number of syllables. Moreover, bigram frequency was not
significantly different ($p = .304$) between German words ($M = 31965.36$) and pseudowords ($M = 25980.89$) according to a text corpus encompassing 6-8 year old children’s print language in German (ChildLex Lexical Database, Schroeder et al., 2015). English words were part of the longitudinal design, and will not be discussed in this paper as children did not have knowledge of English yet. Pseudowords were meaningless, pronounceable stimuli. False-font characters were designed for the purpose of this study, and special care was taken to ensure that each alphabetical letter had its unique false-font correspondent.

Fig.1. One-back task with German words, false-font strings, English words and pseudowords.

2.5. Electrophysiological recording and analysis
A 128-channel EEG (HydroCel GSN, EGI) was recorded against the Cz reference, at a sampling rate of 250 Hz, with high-pass (0.1 Hz) and low-pass (100 Hz) filter settings. Impedance was kept below 50 kΩ. Raw data was preprocessed using BESA software. The continuous EEG was corrected for eye blinks after channels with extensive artifacts were spline interpolated. Corrected files were digitally lowpass (0.3 Hz) and highpass filtered (30 Hz) and segmented (-150 ms prior and 850 ms following the stimulus). Furthermore, trials with artifacts exceeding 180 μV (max-min difference) in any channel were automatically excluded before averaging. The mean accepted trials (with standard deviations in parentheses) averaged in the final analysis (per child and condition) were as follows: German words 62.88 (14.44), false-font strings 64.10 (12.06) and pseudowords 62.04 (13.07). The data was later transformed to the average reference (Lehmann & Skrandies, 1980) and the recording reference was used as an electrode. In the final preprocessing step the ERPs were corrected for the constant 20 ms
delay (as revealed by a timing test using a photo sensor) and averaged separately according to non-target stimuli for the four conditions.

Using Brain Vision Analyzer Software, the individual ERPs were baseline corrected and Global Field Power (GFP) together with grand means for all four condition stimuli were computed. Coarse-tuning was indexed by the N1 difference in the ERPs between German words and false-font strings. The lexicality effect was indexed by the N1 difference in the ERPs between German words and pseudowords. In order to acquire the most robust N1 measure, we analyzed the average values of the N1 segment (180-296 ms) at left (LOT) and right (ROT) occipito-temporal clusters. This N1 segment was defined by the two GFP minima from the average of word and false-font grand means (Maurer et al, 2005a). Clusters were built from 13 occipito-temporal electrodes on each lateral site (LOT: E50, E57, E58, E59, E63, E64, E65, E66, E68, E69, E70, E73, E74; ROT: E82, E83, E84, E88, E89, E90, E91, E94, E95, E96, E99, E100, E101; for illustration see Fig. 2 & 3). Corresponding electrodes in 10-10 system (Luu & Ferree, 2000) are P7, P9, TP9, PO7, PO9, O1 (LOT) and P8, P10, TP10, PO8, PO10, O2 (ROT).

2.6. Statistical analysis

We investigated print tuning in first grade children (mean age: 7.6) at the group level (I) as well as at the individual response level (II). For the group analysis, an ANOVA on the N1 amplitude (mean segment) was computed with within subject factors wordlike (German words vs false-font strings) and laterality (left vs right).

For the individual subject analysis, we performed a one-tailed unpaired t-test at each of the two electrode clusters, separately for each individual subject. This analysis was computed by comparing the average mean segment intervals of the N1 print tuning (180-296 ms) for German words against false-font strings by taking the trial-to-trial variance into account. The degrees of freedom were adjusted for each subject depending on the number of trials averaged in the ERP.
To test for the presence of lexicality effect (III), we performed an ANOVA on the N1 amplitude (mean segment) with within subject factors *lexicality* (German words vs pseudowords) and *laterality* (left vs right). The behavioral measures used for the regression analyses (IV, V, VI) consisted of several z-transformed and averaged subtests (Table 1). As indicated in the introduction section, behavioral language measures predicting subsequent learning to read, can also be called cognitive reading components. As such, the predictors used in the regression model will be further called cognitive reading components. Before analyzing the effects of cognitive reading components on print tuning, the measures showing high inter-correlations with word-reading fluency were excluded (*r* > .80, Table 1). Thus, pseudoword-reading fluency and sentence-reading fluency were not further investigated. In order to test for the influence of cognitive reading components (word-reading fluency, rapid automatized naming, phonological awareness, auditory memory span and vocabulary) on print tuning (left occipito-temporal electrode cluster), a set of linear regression analyses was computed (described in the result section). The choice of the left occipito-temporal cluster relies upon previous investigations on N1 print tuning (e.g., Bentin et al., 1999; Brem et al., 2005; Brem et al., 2009; Parviainen et al., 2006; Tarkiainen et al., 1999). For the behavioral analysis, separate repeated measure ANOVAs were computed for accuracy and reaction time with within subject factors *wordlike* (German words vs. false-font strings) and *lexicality* (German words vs. pseudowords), respectively.
Table 1. Behavioral language measures used in the regression approach

<table>
<thead>
<tr>
<th>MEASURES</th>
<th>M (SD)</th>
<th>corr. with word-reading fluency</th>
<th>corr. with Word.-false-fon diff. at LOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word-reading fluency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLRT I, words (correct/1min)</td>
<td>35.83 (21.00)</td>
<td>1.00</td>
<td>-.28</td>
</tr>
<tr>
<td>SLRT-II, words (correct/1min)</td>
<td>30.38 (15.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLRT I, short text (correct/1min)</td>
<td>48.92 (33.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudoword-reading fluency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLRT I, pseudowords (correct/1min)</td>
<td>18.33 (9.30)</td>
<td>.91†</td>
<td>-.22</td>
</tr>
<tr>
<td>SLRT-II, pseudowords (correct/1min)</td>
<td>26.24 (9.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid automatized naming</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAN, one syllable animals (time in sec)</td>
<td>71.12 (18.06)</td>
<td>-.43</td>
<td>.06</td>
</tr>
<tr>
<td>RAN, three syllable animals (time in sec)</td>
<td>88.66 (24.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAN, lower case letters (time in sec)</td>
<td>41.22 (11.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAN, digits (time in sec)</td>
<td>42.21 (12.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonological awareness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAKO, phoneme deletion (correct items/max: 7)</td>
<td>4.75 (1.77)</td>
<td>.24</td>
<td>.03</td>
</tr>
<tr>
<td>BAKO, pseudoword segm. (correct items/max: 8)</td>
<td>4.85 (1.74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence-reading fluency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLS (correct sentences/3mins)</td>
<td>18.88 (9.49)</td>
<td>.94†</td>
<td>-.26</td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAWK-IV, vocabulary (raw score)</td>
<td>27.31 (5.93)²</td>
<td>.26</td>
<td>-.34</td>
</tr>
<tr>
<td>Auditory memory span</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAWIK-IV, digit span backwards (raw score)</td>
<td>5.79 (1.19)²</td>
<td>.42</td>
<td>-.10</td>
</tr>
<tr>
<td>HAWIK-IV, digit span forwards (raw score)</td>
<td>6.59 (0.97)²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudoword span (raw score)</td>
<td>4.03 (1.02)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 excluded from regression analyses, as r > .80.
2 standard score: vocabulary 11.38 (2.34); digit span (backwards and forwards) 10.51 (2.00).

3. Results
3.1. Behavioral data
The average accuracy and reaction time to target stimuli for German words, false-font strings and pseudowords are reported in Table 2. According to repeated measure ANOVAs subjects responded faster to false-font strings than to German words (wordlike, \( F(1, 67) = 4.80, p = .032 \)), but were more accurate in responding to German words (wordlike, \( F(1, 67) = 7.00, p = .010 \)). In addition, participants responded more accurately to German words than to pseudowords (lexicality,
$F(1, 67) = 8.91, p = .004)$, and tended to respond faster to German words than to pseudowords ($\text{lexicality, } F(1, 67) = 3.31, p = .073$).

**Table 2. Behavioral results for target stimuli**

<table>
<thead>
<tr>
<th></th>
<th>Accuracy (%) ±SD</th>
<th>Reaction time (ms ±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>German words</td>
<td>71.91 (23.48)</td>
<td></td>
</tr>
<tr>
<td>False-font strings</td>
<td>64.57 (23.22)</td>
<td></td>
</tr>
<tr>
<td>Pseudowords</td>
<td>66.16 (22.56)</td>
<td></td>
</tr>
</tbody>
</table>

3.2. ERP data

3.2.1. Development of coarse-tuning for print (I, II)

**Development of print tuning at the group level (I)**

The two-way ANOVA on the N1 amplitude revealed that the amplitudes for German words compared to false-font strings were significantly more negative ($\text{wordlike, } F(1, 67) = 318.45, p < .001$), indicating that at the group level the N1 print tuning developed after only one year of reading instruction. This effect was highly significant at occipito-temporal electrodes ($t(\text{max}) = -14.14, p < .001$, Fig. 2) and showed the typical N1 distribution with posterior negativity and fronto-central positivity. Neither the main laterality effect ($\text{laterality, } F(1, 67) = 2.34, p = .130$) nor the interaction between condition and laterality ($\text{wordlike x laterality interaction, } F(1, 67) = 1.07, p = .305$) were significant. The results for the difference between pseudowords and false-font strings are presented in the supporting material (see S1).

**Development of print tuning at the individual level (II)**

According to unpaired one-tailed t-tests 60 (88.2%) out of 68 first grade children (mean age: 7.6) exhibited presence of print tuning irrespective of hemisphere (left/right/bilateral). Respectively, 9 participants showed print tuning only on the left, and 7 only on the right hemisphere.
According to Fisher’s exact test children who lacked print tuning tended to be more frequently found amongst poor (4 out of 12) than normal (4 out of 52) readers (\(p = .081\), for graphical illustration of this result, and the description of poor and normal reading criteria see Fig. 3). Thus, the odds ratio of the absence of print tuning was 6.25 times higher in poor compared to normal readers.

![ERP waveforms and N1 segment maps for German words and false-font strings (180-296 ms).](image)

Fig.2. ERP waveforms and N1 segment maps for German words and false-font strings (180-296 ms).

![Print tuning across poor and normal readers.](image)

Fig.3. Print tuning across poor and normal readers.

Poor and normal poor readers were categorized according to standardized scores of word-reading fluency and short text reading (Landerl, et al., 2006). Subjects were classified as normal readers if their score exceeded the 30\(^{th}\) percentile in either of the two subtests and the 20\(^{th}\) percentile in the other. Consequently, subjects scoring below the 20\(^{th}\) percentile in either of the two subtests and below the 30\(^{th}\) percentile in the other were classified as poor readers.
3.2.2. Development of the lexicality effect (III)

ANOVA on the lexicality effect missed the trend level (lexicality, $F(1, 67) = 2.59, p = .113$), hence indicating no significant difference in processing German words and pseudowords in 7.6 year old first graders. However, when testing for German word – pseudoword difference separately at each hemisphere site, the difference between words and pseudowords was significant over the right hemisphere cluster ($F(1, 67) = 4.42, p = .039$, Fig. 4). Moreover, the lexicality effect measured over the right occipito-temporal cluster showed marginally significant correlation with word-reading fluency ($r(68) = .23, p = .058$), which indicated that the larger the lexical effect the slower the child tended to read.

![Fig.4. ERP waveforms and N1 segment maps for German words and pseudowords (180-296 ms).](image)

3.2.3. Cognitive underpinnings of print tuning (IV, V, VI)

In order to test the influence of word-reading fluency on print tuning, we performed a linear regression analysis. The result indicated that more fluent word reading was associated with larger print tuning ($R^2 = .076, p = .023$, Fig. 5).
Further, we wanted to test whether there are additional relevant variables that explain variance in print tuning. As such, word-reading fluency was forced entered and the remaining cognitive reading components (rapid automatized naming, phonological awareness, auditory memory span and vocabulary) were added in an additional block in a forward regression model. The result showed that, beyond word-reading fluency, vocabulary explained 7.8% additional variance in print tuning ($\Delta R^2 = .077$, $p = .005$).

When entering all language measures including word-reading fluency in one block (backward regression method), word-reading fluency and vocabulary significantly explained variation in print tuning ($R^2 = .153$, $p = .005$, Table 3). As such, a combination of word-reading fluency and vocabulary explained the largest part of the entire variance in print tuning.

Fig.5. Scatterplot of word-reading fluency and N1 coarse-tuning.

![Scatterplot of word-reading fluency and N1 coarse-tuning](image)
Table 3. Multiple regression analysis using stepwise procedure (backwards)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>SE $B$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>final step</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>-1.25</td>
<td>1.70</td>
<td></td>
</tr>
<tr>
<td>word-reading fluency</td>
<td>-0.62</td>
<td>0.37</td>
<td>-0.20</td>
</tr>
<tr>
<td>vocabulary</td>
<td>-0.15</td>
<td>0.06</td>
<td>-0.29</td>
</tr>
</tbody>
</table>

$^1p < .10, ^2p < .05$

Variables initially excluded due to multicollinearity ($r > .80$): pseudoword-reading fluency and sentence-reading fluency.

Variables added at the initial step: word-reading fluency, rapid automatized naming, phonological awareness, auditory memory span and vocabulary.

4. Discussion

Given the importance of reading in nowadays society the question how children learn to read has gained large interest (e.g., Nation & Snowling, 2004; Pennington & Lefly, 2001; Puolakanaho et al., 2008; Ricketts et al., 2007; Vellutino et al., 2007). Recent developmental studies using cognitive neuroscience methods pointed to the important role of neural print tuning in reading acquisition (e.g., Brem et al., 2010; Maurer et al., 2005a; Maurer et al., 2006; Maurer et al., 2007a; Parviainen et al., 2006). Taking advantage of a large sample of first grade children, we investigated individual differences in print tuning and related it to cognitive reading components.

We found a very robust print tuning that developed after only one year of formal reading instruction, which was further present at the individual level in almost 90% of the children. The lexicality effect, however, could not be reliably detected. On the cognitive level, next to word-reading fluency, also vocabulary was associated with print tuning.

4.1. N1 coarse-tuning – rapid emergence

At the group level, words elicited larger N1 component than false-font strings, indicating the presence of print tuning in 7.6 year old first grade children. Print tuning was not only indicated by larger amplitudes for German words than false-font strings, but also by shorter peak latencies (see S2). Considering that print tuning was found to be absent in kindergarten children in a previous study (Maurer et al., 2006) our result
suggests that it has developed after one year of reading training in school (Swiss children start formal reading training in first grade). Given the results of other EEG studies showing print tuning in second or higher grades (Araújo et al., 2012; Brem et al., 2009; Maurer et al., 2006), the present result extends those findings, by providing EEG based evidence on neural tuning for print being present already in first grade children. Brem et al. (2010) reported rapid print tuning emergence even in illiterate kindergarteners after short grapheme-phoneme training, which was followed by its prompt decline after discontinuation of the training. Hence, we propose that sensitivity for print develops as early as the very initial stages of reading acquisition, but that continuing practice appears to be the critical factor for its stable establishment.

Furthermore, most previous developmental studies on print tuning used symbol strings (as a control condition) that were clearly distinct from letter strings (e.g., Araújo et al., 2012; Brem et al., 2005; Brem et al., 2009; Maurer et al., 2005a; Maurer et al., 2005b; Maurer et al., 2006; Maurer et al., 2007a; Parviainen et al., 2006). Despite better control for low-level visual differences with the use of false-font strings in the present study, print tuning effects were still robust, suggesting that the effects shown in previous studies were not due to low-level differences (see also Brem et al., 2010).

An important result of the present study was that print tuning could be detected at the individual level in almost 90% of the children, not just at the group level as in previous studies (e.g., Araújo et al., 2012; Brem et al., 2005; Brem et al., 2009; Brem et al., 2010; Maurer et al., 2005a; Maurer et al., 2006; Maurer et al., 2007a). The only previous study that investigated print tuning at the individual level (Parviainen et al., 2006) found print tuning only in slightly more than half of the first grade children. Given that the children were of the same age as in the present study the discrepancy in results may be explained by a combination of developmental effects and the different techniques used. Parviainen et al. (2006) used MEG which is considered to be less sensitive to deep or radial sources (Michel, et al., 2009). Possibly, print tuning in children is more strongly dominated by radial sources for which only EEG, but not MEG is sensitive. Print tuning in adults, however, may also result from
tangential sources, which can be detected by both EEG and MEG (Tarkiainen et al., 1999). An additional difference between the two studies arises from the presentation mode, which was blocked in our study and randomized in the Pariviainen et al. (2006) study. It is therefore possible that the blocked presentation enhanced print processing through top-down modulation. However, further research is needed in order to clarify which of these explanations hold.

The eight children who exhibited absence of print tuning tended to be disproportionately more often found amongst poor than normal readers, suggesting that the absence of print tuning may rely upon poor reading performance. This is in accordance with previous studies that emphasized the importance of this neural correlate for print in the development of unimpaired reading performance (e.g., Araújo et al., 2012; Maurer et al., 2007a).

4.2. N1 lexicality effect
The N1 lexicality effect could not be reliably found in 7.6 year old first graders, as it missed the trend level in the main analysis, and was detectable only over the right hemisphere cluster. As yet, studies comparing word and pseudoword processing have displayed inconsistent findings. Whereas some researchers found significant differences in N1 between these two types of stimuli (Hauk et al., 2006b; Maurer et al., 2006; McCandliss et al., 1997; Sereno, Rayner, & Posner, 1998), other studies found no difference in processing words and pseudowords (Araújo et al., 2012; Bentin et al., 1999; Kast et al., 2010). Among differences in tasks (Xue et al., 2008), experimental parameters (Xue et al., 2008), or differences in orthographic depth of the languages used (Maurer et al., 2005b), the developmental stage of reading acquisition may also influence the degree of the lexicality effect (Maurer et al., 2006). The latter was indicated by an N1 difference between words and pseudowords in a repetition detection task in second graders (Maurer et al., 2006), but not in adults of the same language (Maurer et al., 2005a; Maurer et al., 2006), and not in non-reading kindergartners (Maurer et al., 2005a). However, contrary to our findings the second grade children (Maurer et al., 2006) showed the
lexicality effect over the left hemisphere similar to adult readers of a deep orthography (Maurer et al., 2005b). This atypical right lateralization that further tended to diminish with better reading performance may reflect a temporary phase in the development of the N1 lexicality effect at the initial stages of learning to read.

4.3. N1 coarse-tuning and its underlying cognitive mechanisms
Taking advantage of the large number of participants assessed, we tested the association between coarse-tuning and various cognitive reading components using the regression approach. A first result indicated that word-reading fluency predicted print tuning across children with a wide range of reading skills. The faster the child read, the larger the amplitude difference between German words and false-font strings was. This association was driven by the correlation between word-reading fluency and the N1 response to German words ($r(68) = -.232, p = .056$), while the correlation between word-reading fluency and the N1 response to false-font strings was not significant ($r(68) = .003, p = .980$). The association between word-reading fluency and N1 print tuning is in accordance with Maurer et al. (2007a) who proposed print tuning to index word-reading fluency, as it was reduced in dyslexic readers, which is in agreement with other dyslexia studies (Araújo et al., 2012; Helenius et al., 1999; Mahé et al., 2012). The present result extends these findings by indicating that this effect is not driven by a clinical group, but that it is present in a sample of children with a wide range of reading skills. Therefore, word-reading fluency appears to be a stable measure to explain print tuning in the first school year, even across children ranging from very poor to most fluent readers.

From the range of four additional cognitive reading components only vocabulary contributed to the variance in print tuning, but not rapid automatized naming, nor phonological awareness, nor auditory memory span. Whereas vocabulary can be seen as a measure of semantic knowledge (Vellutino et al., 2007), the remaining measures can be considered to reflect performance on phonological tasks (Landerl et al., 2013; Moll et al., 2009). The absence of significant correlations between
print tuning and any phonological measures is interesting, given the role of phonology in theories of reading (Ehri et al., 2001) and dyslexia (Bradley & Bryant, 1978). Moreover, it has been suggested that the development of print tuning is modulated by phonological top-down influence (phonological mapping hypothesis, Maurer & McCandliss, 2007b), which could explain reduced print tuning in dyslexia (Maurer et al., 2007a). Yet, the present results suggest that the top-down influence at the initial stages of learning to read might be of semantic rather than phonological nature.

While one model of visual word processing suggests that regions in the ventral occipito-temporal (vOT) cortex become tuned for letters and letter groups in the course of learning (Dehaene et al., 2005), a recent model implementing a predictive coding framework (Price & Devlin, 2011) proposes that top-down phonological and semantic predictions interact with bottom-up visual input in the vOT cortex in reading. As N1 tuning has been associated with word-specific processes in the Visual Word Form Area in the vOT, the present results indicate that such predictions are derived from semantic rather than from phonological information. Such semantic top-down influence may not only increase activation of word processing, but may also inhibit activation of false-font processing, as vocabulary showed a negative correlation with the N1 response to German words ($r(68) = -.134, p = .274$), and a positive correlation with the N1 response to false-font strings ($r(68) = .177, p = .148$), even though both these correlations were not significant.

In agreement with a previous study (Maurer et al., 2005b), we suggest that the N1 difference between words and false-fonts in the present study consists of two processes: one dominating process reflecting general perceptual expertise for orthographic strings, and the other one reflecting lexical-semantic processing, as indicated by the correlation of print tuning with vocabulary and the marginally significant lexical effect. Possibly, the influence of phonological and semantic top-down modulation, as postulated in the Price & Devlin model (2011), may further depend on the developmental stage of reading or on task demands, with potentially larger phonological modulation in more experienced readers or in more explicit phonological tasks.
Indeed, previous studies indicated a relation between phonological skills and visual word activation (Shaywitz et al., 2004; Xue et al., 2006). Phonological training has been shown to increase activation in the vOT cortex in adult normal readers (Xue et al., 2006) and in children with a reading disability (Shaywitz et al., 2004). Moreover, phonological skills were shown to modulate the size of letter-string activation in first grade children, even though the correlation was of negative direction (Parviainen et al., 2006). It is thus possible that the phonological measures included in this study did not capture the critical aspects of phonological processing that are relevant for print tuning. Nevertheless, the two measures included were not correlated with print tuning, neither separately (phoneme deletion $r(68) = .004, p = .974$; pseudoword segmentation $r(68) = .044, p = .722$), nor as a combined measure ($r(68) = .029, p = .812$).

Although RAN can be considered a measure of phonology, its relation to reading fluency is worth noting. RAN measures are commonly recognized to be amongst precursors of early reading achievement (e.g., Pennington & Lefly, 2001; Puolakanaho, et al., 2008). Given the relation between word-reading fluency and print tuning, it is interesting that RAN was not correlated with print tuning. While RAN has been called a ‘fluency in phonology measure’ (Moll et al., 2009), word-reading fluency could thus be called ‘fluency in reading’. This would mean that not fluency per se is relevant for print tuning, but rather fluency that involves converting letters, letter groups, or even entire words into spoken language. Although we did not assess grapheme-phoneme conversion directly, word-reading fluency may be the closest measure in this study that indicates connectivity (Wimmer & Schurz, 2010) or binding of phonology and orthography (Blomert, 2011).

While the relation between vocabulary and print tuning can be interpreted as an influence of semantic knowledge, there might also be an alternative explanation. Children with high exposure to written material might increase their vocabulary and at the same time develop enhanced print tuning. Data from a background questionnaire, however, do not support such an interpretation, as the number of books at home was not significantly correlated with print tuning ($r(68) = -.151, p = .219$).
Moreover, after the variable ‘books at home’ was forced entered in a forward regression analysis, vocabulary still significantly predicted the residual variance N1 print tuning ($\Delta R^2 = .100$, $p = .008$). We thus suggest that the association between vocabulary and print tuning in the present study reflects influence of semantic knowledge.

Moreover, although word-reading fluency has been shown to be associated with print tuning, a large part of the variance remained unexplained. Some of this variance could be attributed to semantic knowledge. Nevertheless, further studies are needed to identify additional factors that influence print tuning. The two factors identified in this study suggest that conversion of print to speech (reflected by word-reading fluency) and semantic knowledge (indicated by vocabulary) are associated with the development of print tuning, and that higher and lower levels of language processing interact early on during reading.

Taken together, this study demonstrates the presence of print tuning in the first year of reading acquisition and its development at the individual level. Moreover, individual differences in print tuning are not only related to word-reading fluency, but also to semantic knowledge, thereby contributing to our knowledge of the cognitive characteristics of print tuning. Such knowledge is critical to better understand how children learn to read, and why some children face difficulties at the beginning of reading training.

5. Supporting information
S1. N1 pseudoword – false-font difference

For the analysis on the difference between pseudowords and false-font strings, an ANOVA on the N1 amplitude (mean segment of interest 180-296 ms) was computed with within subject factors condition (pseudowords words vs false-font strings) and laterality (left vs right).

The two-way ANOVA on the N1 amplitude revealed that the amplitudes for pseudowords compared to false-font strings were significantly more negative (condition, $F(1, 67) = 278.59$, $p < .001$). Neither the main laterality effect ($laterality F(1, 67) = 0.70$, $p = .404$) nor the interaction between condition and laterality ($condition \times laterality$ interaction,
$F(1, 67) = 0.00, p = .994$) were significant. This result resembled the N1 print tuning pattern. The figure below illustrates the ERP waveforms and N1 segment maps for pseudowords and false-font strings (180-296 ms).

![ERP waveforms and N1 segment maps](image)

S2. Peak latency analysis
To investigate differences in peak latencies between German words and false-font strings, and German words and pseudowords we performed automatic peak detection (with the use of Brain Vision Analyzer Software) for the time segment of interest (180-296 ms). Further, we computed separate ANOVAs for the latency in print tuning (wordlike: German words vs false-font strings) and the latency in the lexicality effect (lexicality German words and vs. pseudowords) on the left occipito-temporal cluster. The peak latency analysis revealed a significant latency difference between German words and false-font strings (wordlike, $F(1, 67) = 11.30, p < .001$). However, this latency difference was not
significantly correlated with neither word-reading fluency ($r(68) = -.085, p = .493$) nor vocabulary ($r(68) = .025, p = .841$).

The peak latency analysis for the lexicality effect showed a similar pattern with a significant latency difference between German words and pseudowords (lexicality, $F(1, 67) = 4.11, p = .047$) and no significant correlations with neither word-reading fluency ($r(68) = .020, p = .871$) nor vocabulary ($r(68) = -.019, p = .880$).

Acknowledgement
This research was supported by the Swiss National Science Foundation (SNF-128610).

References


