

Emergence of representations through repeated training on pronouncing novel letter combinations leads to efficient reading



Atsuko Takashima^{a,*}, Iris Hulzink^a, Barbara Wagensveld^b, Ludo Verhoeven^a

^a Radboud University, Behavioural Science Institute, PO Box 9104, 6500 HE Nijmegen, The Netherlands

^b Studio Lakmoes, Statenlaan 8, De Kleine Campus, BG lokaal 0.2, 6828 WE Arnhem, The Netherlands

ARTICLE INFO

Article history:

Received 17 December 2015

Received in revised form

9 May 2016

Accepted 13 May 2016

Available online 15 May 2016

Keywords:

Reading

Learning

Serial decoding

Holistic decoding

fMRI

Angular gyrus

Supramarginal gyrus

Left middle temporal gyrus

ABSTRACT

Printed text can be decoded by utilizing different processing routes depending on the familiarity of the script. A predominant use of word-level decoding strategies can be expected in the case of a familiar script, and an almost exclusive use of letter-level decoding strategies for unfamiliar scripts. Behavioural studies have revealed that frequently occurring words are read more efficiently, suggesting that these words are read in a more holistic way at the word-level, than infrequent and unfamiliar words. To test whether repeated exposure to specific letter combinations leads to holistic reading, we monitored both behavioural and neural responses during novel script decoding and examined changes related to repeated exposure. We trained a group of Dutch university students to decode pseudowords written in an unfamiliar script, i.e., Korean Hangul characters. We compared behavioural and neural responses to pronouncing trained versus untrained two-character pseudowords (equivalent to two-syllable pseudowords). We tested once shortly after the initial training and again after a four days' delay that included another training session. We found that trained pseudowords were pronounced faster and more accurately than novel combinations of radicals (equivalent to letters). Imaging data revealed that pronunciation of trained pseudowords engaged the posterior temporo-parietal region, and engagement of this network was predictive of reading efficiency a month later. The results imply that repeated exposure to specific combinations of graphemes can lead to emergence of holistic representations that result in efficient reading. Furthermore, inter-individual differences revealed that good learners retained efficiency more than bad learners one month later.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

How do we become skilled readers? In order to read a word, we need to be able to map the prints to the corresponding sounds, a process that is referred to as decoding. Multiple models have been proposed to understand how we translate orthography to phonology (e.g. [Ans et al. \(1998\)](#), [Coltheart \(2005\)](#), [Coltheart et al. \(1993\)](#), [Plaut et al. \(1996\)](#), [Seidenberg \(2007\)](#) and [Share \(2008\)](#)). The Dual Route Cascaded model (DRC; e.g. [Coltheart \(2005\)](#)) assumes that there are two routes to word reading. One route, referred to as the lexical route, is based on holistic word decoding, where the orthography on the word-level is linked to the phonology on the word-level (addressed phonology). The other route, referred to as the non-lexical route, relies on letter-by-letter conversion, where each grapheme-level representation is converted to a phoneme-level representation, and assembled as a word

(assembled phonology). Lexical decoding can only be applied in the presence of word-unit representations.

If we assume that word-level representations emerge through repeated exposure, every combination of letter strings encountered for the first time is always unfamiliar. With repeated experience, the status of the word (or the experienced letter strings) gradually changes from unfamiliar to familiar. Frequent words would be more familiar than infrequent words since the probability of the exposure to frequent words is higher. Empirical findings have revealed that naming speed is greatly influenced by word frequency (e.g. [Forster and Chambers \(1973\)](#)). Correct pronunciation is predicted by the regularity and frequency of words ([Waters and Seidenberg, 1985](#)). There is even an interaction between frequency and regularity, such that irregular words, if frequently encountered, are read with ease ([Seidenberg and McClelland, 1989](#)). One could assume that with repeated experience of reading a word, a clearly specified holistic orthographic representation and its associated phonological representation becomes available ([Share, 1995](#); [Ziegler et al., 2014](#)). According to

* Corresponding author.

E-mail address: Atsuko.Takashima@donders.ru.nl (A. Takashima).

Share (2008), the unfamiliar-to-familiar transition with experience can be interpreted as the transition from non-lexical (assembled) to lexical (addressed) route of the DRC model of reading. In a similar vein, it can be assumed that a lexical representation becomes available for a frequently occurring word to enable holistic processing once the word has become familiar after repeated occurrence.

If experience leads to emergence of holistic orthographic representations, beginning readers with very few exposures to written words would be expected to rely on grapheme-level decoding at the initial stage of reading. Evidence to support this idea comes from developmental studies examining decoding skills. During the early stages of word reading, the length of a word is closely related to its reading speed - the longer the word length, the longer the time it takes to read (Spinelli et al., 2005). This length effect, however, diminishes with proficiency (Zoccolotti et al., 2005) suggesting a shift from serial to parallel/holistic processing (van den Boer and de Jong, 2015). This is because when grapheme-level conversion is applied, the number of letters/graphemes constitutes the number of units to be decoded. In contrast, in a lexical approach, there is only one unit, the whole word, to be converted. These behavioural findings indicate that children initially apply grapheme-level conversion, but with experience, frequently occurring words are decoded holistically, leading to a diminished length effect. There is also evidence from studies with adults reading nonwords in familiar scripts that the length effect can be eliminated by only a small number of repetitions as few as four times within a session (Maloney et al., 2009). Interestingly, repeated training on decoding disyllabic pseudowords in Greek as a less familiar script by Dutch university students led to more efficient pronunciation not only of trained words but also of recombined words using the same syllables in different order. This suggests that syllable representations are also strengthened through repeated exposure to multi-syllable words (Takashima et al., 2014b).

If repeated occurrence of a specific letter combination leads to the emergence of a holistic representation, then we would expect neural responses also to be different once holistic decoding becomes possible through experience. From an energy-saving perspective of brain processing (Friston, 2010), it would be more efficient if the brain stored a pattern or a code for frequently occurring combinations of letters as one unit, such that every time this pattern is encountered again, there is no need to inefficiently decode serially. But does the brain actually code for holistic representations for frequently occurring words? And if so, can we observe differential neural decoding processes between holistic-level units and grapheme-level units? Tracking the developmental changes in children's brain activity is challenging. To circumvent this obstacle, one can simulate the developmental process over time and experience by training adults on decoding a novel script, and tracking changes both in behaviour and brain responses as a function of training. Although it should be acknowledged that adults have a well established reading system and there may be differences in the acquisition of developmental (first language) and second language lexical representations, a common underlying process change through experience can also be expected. The current study aimed to establish whether repeated exposure to a specific combination of letters is the key to the emergence of holistic representations (both on the word and syllable levels) in a healthy adult sample, and to track brain areas involved in holistic and grapheme-level decoding processes. To this end, we trained participants to read aloud a set of two-syllable pseudowords written in an unfamiliar script, namely the Korean Hangul characters.

Unlike Chinese, Japanese, or Korean Hanja characters, Hangul characters are comprised of radicals that have a transparent

orthography-phonology relationship, such that once the radical-phoneme association is learned, one can derive the sound of the characters through converting each of the radicals to their associated phonemes. Even though Hangul characters look like logograms to those unfamiliar with Korean scripts, the radicals function like letters and characters function like syllables. By recruiting participants without any knowledge of Korean, and teaching them the radical-phoneme correspondence of Hangul characters, we can measure the initial stage of decoding, when grapheme-to-phoneme conversion and compilation processes are assumed to take place. After repeated exposure to a small set of two character- (equivalent to two syllable-) pseudowords, we would expect holistic representations of these trained pseudowords to emerge. If indeed the frequently occurring combination is extracted and coded as a sub-word level unit through experience, representations on the character-level unit (equivalent to the syllable unit) will also be expected to emerge as a result of training on two-character pseudowords.

In which areas of the brain do we expect to see changes? In a meta-analysis, Jobard and colleagues mapped two different brain networks for the two reading routes suggested in the DRC model (Jobard et al., 2003). More specifically, for the lexical route, the co-activation of the visual word form area (VWFA) in the left fusiform gyrus and semantic areas, including the basal inferior temporal area, the posterior part of the middle temporal gyrus, and the triangular part of the inferior frontal gyrus seems to take place. For the non-lexical grapheme-phoneme conversion reading process, predominantly left lateralized areas in the superior temporal, supramarginal, and the inferior frontal gyrus (pars opercularis) were shown to be more involved. The distinction seemed to be based on the availability of the semantic information in these two routes.

A more recent meta-analysis by Cattinelli et al. (2013) reports three different networks involved in reading. First, the word (or semantic)-related network was observed in the left temporoparietal areas including the posterior angular, precuneus, middle temporal, and anterior part of the fusiform gyrus. Second, the pseudoword-related network was observed in the left inferior parietal gyrus at the border between the supramarginal and angular gyrus, and the right parietal cortex. Third, the difficulty-modulated network evidenced that conditions requiring longer reaction times recruited the left inferior frontal gyrus, mid cingulum, left superior parietal gyrus, and pons. The authors attributed these areas to reflect increased attentional processes.

Another meta-analysis by Taylor et al. (2013) also identified regions that correspond to different processes in word reading. They found posterior fusiform and occipito-temporal regions to be more involved in non-lexical orthographic processing, with the anterior fusiform area to code for orthographic lexicon and possibly associated semantic representation, the angular and posterior middle temporal gyri to code phonological and semantic lexicons, the inferior parietal cortex for spelling-sound conversion, and the posterior aspects of the inferior frontal gyrus to compute phonological output.

To observe whether there are neural patterns that correspond to the dual-route cascade model (Coltheart, 2005; Coltheart et al., 1993) without the influence of top-down modulation of the semantics, a study by Mei et al. (2014) trained two groups of young adults extensively for a week to read a set of Hangul characters. One group was trained on association between a character and its corresponding sound (holistic decoding/addressed phonology), and the other group was trained on each radical of the characters (grapheme-to-phoneme decoding/assembled phonology) before being trained to read characters. They showed that the addressed phonology training activated the areas such as the anterior to posterior cingulate cortex, right orbital frontal cortex, angular gyrus and middle temporal gyrus, whereas the assembled

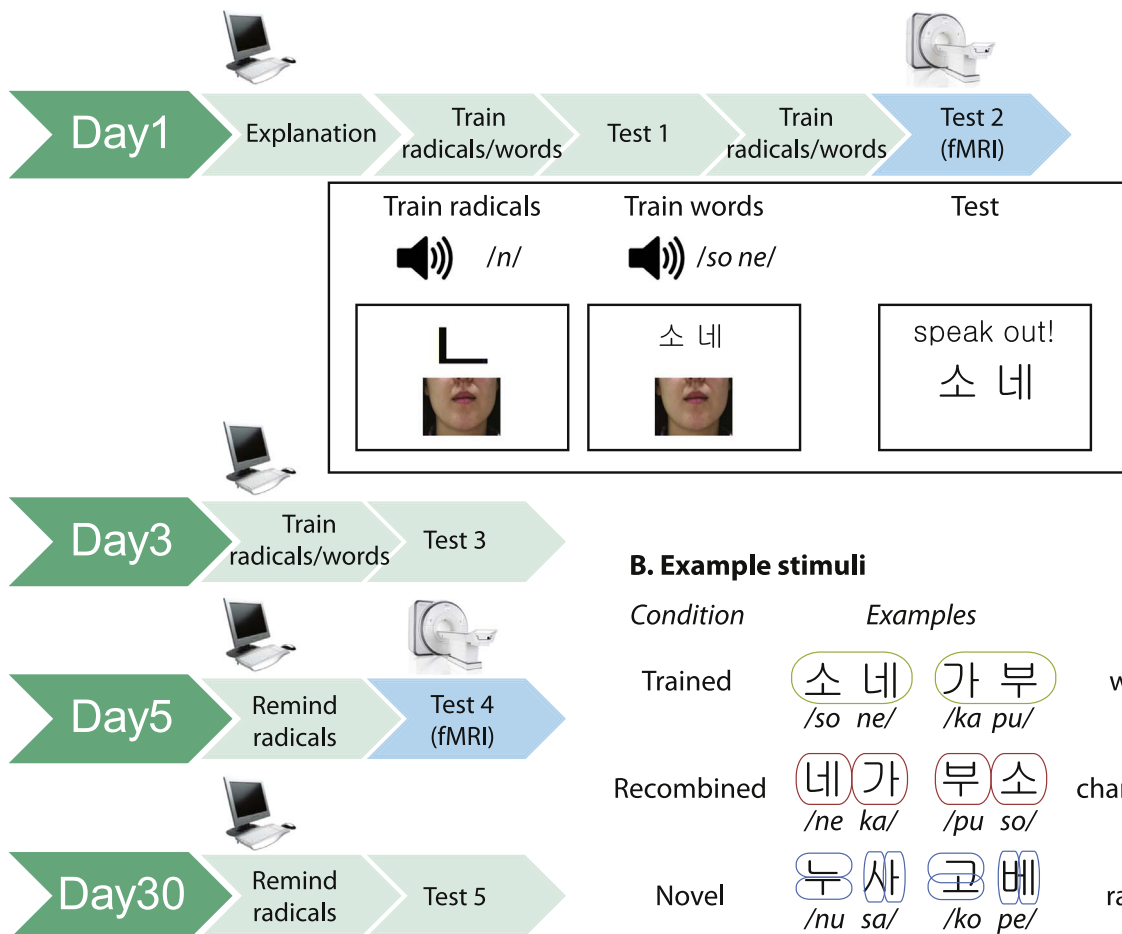
phonology training recruited areas such as the left precentral gyrus/inferior frontal gyrus and supramarginal gyrus.

The above mentioned unfamiliar script training study by Takashima et al., (2014b) using Greek letters observed a shift in the processing network of reading pseudowords as they became more familiar through repeated training. They showed that in the initial stages of decoding unfamiliar scripts, increased demands on visual attention and phonological working memory were observed. They speculated that each grapheme-phoneme converted output had to be kept in the working memory for compilation to successfully produce the whole word when decoding took place at the grapheme-phoneme level. In line with the meta-analysis by Cattinelli et al. (2013), higher activation in the occipital to superior parietal cortex, reflecting increased visual attention (Kastner and Ungerleider, 2000; Milner and Goodale, 2008), and inferior frontal cortex activity reflecting the verbal working memory load (Paulsen et al., 1993; Smith and Jonides, 1999) was observed. After extensive training on five recurring pseudowords, the study revealed activity level increases in areas that are known to code for orthography-phonology correspondence in the posterior middle temporal cortex (Gow, 2012; Hickok and Poeppel, 2007; Indefrey and Levelt,

2004; Lau et al., 2008), lexical information in the angular gyrus (Binder et al., 2003; Price, 2012; Sandak et al., 2004), and a part of the memory retrieval network in the precuneus (Daselaar et al., 2009) suggesting that this change in the brain response is reflecting emergence of lexical and sub-lexical representations. Since the pattern of activity observed for the trained condition was also evident in the recombined condition, in which only the syllables were experienced but not the actual pseudowords prior to test, the study concluded that holistic orthographic and phonological representations and sub-lexical, syllabic representations emerged through a month of repeated training. However, differences in letter-level frequency between the trained and untrained conditions may have introduced confound in this study. Furthermore, repetition of stimuli used within a session was not equal between the conditions, which could have accounted for some of the differences between the conditions observed.

In the present study, we aimed to replicate the previous study by Takashima et al., (2014b) using a controlled stimulus set to show that it is the familiarity to the *specific combination* of scripts that leads to holistic decoding and not just a mere exposure difference between certain graphemes, and furthermore to find

A. Experimental schedule



B. Example stimuli

Condition	Examples	Level
Trained	/so ne/ /ka pu/	word / lexical
Recombined	/ne ka/ /pu so/	character / syllable
Novel	/nu sa/ /ko pe/	radical / letter

Fig. 1. Experimental schedule and material. (A) Experimental schedule: participants came to the lab four times through the whole experiment (Day 1, Day 3, Day 5 and Day 30). On Day 1, they were informed about the nature of Hangeul (Korean) characters, exposed to the radicals to be used in training and test, and trained on a set of six two-character pseudowords. Each word in the Trained condition appeared on the screen accompanied by a video of its pronunciation by a young native Korean female speaker, who clearly showed the corresponding mouth movements for each sound. The test at the end of the training took place in the fMRI scanner. On Day 3, the participants were trained once again on the same set of six pseudowords, and again tested on words from the three conditions; Trained, Recombined and Novel. On Day 5, participants took the test in the scanner shortly after being reminded of the pronunciation of the radicals. On Day 30, participants were briefly reminded of the pronunciation of the radicals and subsequently performed the test on a computer. Tests 2 and 4 took place in the MRI scanner. (B) Example stimuli: Example of words in the different conditions and their pronunciations. Recombined words are recombination of characters used in the Trained condition, and Novel words are recombination of radicals used in the Trained condition.

whether there is a longer lasting effect of this holistic representation even after the training was terminated. More specifically, the first aim was to test the training effect on decoding pseudowords written with a novel script using Hangul characters by controlling for the exposure frequency of the radicals across conditions. We assumed that the pseudowords can be decoded on three levels once the basic grapheme-to-phoneme association rule had been acquired: (1) conversion and compilation on the radical level, equivalent to letter/grapheme decoding, (2) conversion and compilation on the character level, equivalent to syllable unit decoding, and (3) conversion on the word level, equivalent to holistic decoding. We hypothesized that training would cause a shift in the dominant decoding process, from grapheme-level conversion-and-compilation to word-level conversion. We compared these three levels of decoding while keeping each radical occurrence constant across three conditions: *Trained* (six two-character pseudowords that were repeatedly exposed during training), *Recombined* (untrained two-character pseudowords comprised of recombined characters from the trained condition), and *Novel* (untrained pseudowords comprising of two novel characters sharing the same radicals used in the Trained condition) (see Fig. 1B for an example and Table 1 for the whole list). In order to track changes in behavioural and neural processing with training, pronunciation tests took place in the functional magnetic resonance imaging (fMRI) scanner on two occasions, once after the initial training on Day 1, and again on Day 5 after the participants had gone through repeated exposure on Day 3 (see Fig. 1A for the overall experimental procedure). We hypothesized that Trained pseudowords will be decoded with high accuracy and speed. Due to the emergence of sub-lexical units on the character (syllable) level, we hypothesized that Recombined pseudowords will also have an advantage over Novel pseudowords. We expected a small training effect to be visible already on Day 1, but to be enhanced on Day 5 after another session of training on Day 3. On the neural level, we expected a qualitative change in the engagement of the brain areas. We expected to replicate the areas found in the prior study (Takashima et al., 2014b), namely the bilateral angular gyrus, the posterior middle temporal gyrus, and the precuneus to increase in activation pattern with training, reflecting a conversion

process on the holistic/combinatorial level (syllable/character, and word/two-character combination). If, on the other hand, training only promoted efficiency in serial-decoding, we expected performance to equally improve across all conditions, as radicals used in the Recombined and Novel conditions were trained through exposure to the Trained condition. At the neural level, only a quantitative change will be observed with training, namely a reduction in the attentional and effortful demands reflected as less engagement of the fronto-parietal attentional areas and preparation of motor programs reflected in the motor and premotor areas extending to supplementary motor area and the insula. We included a second scanning session to see if extra training boosted the difference between conditions, but also because memory consolidation literature postulates that lexical representation emerges after a time of consolidation, especially after a period of sleep (Davis et al., 2009; Takashima et al., 2014a, 2009; Thielen et al., 2015). Thus we may not observe differences at the neural level on Day 1 but this may become visible after a delay that included sleep.

The second aim of this study was to examine how stable the holistic representations are once they are established through repeated training on Days 1 and 3. For this reason, we measured the behavioural performance one month after training to investigate the long-term retention of the ability to decode the Hangul characters (Day 30). If the emerged holistic representations are stable, better performance for the Trained condition over the Novel condition one month later is expected. Furthermore, we hypothesized that greater level of activation in the areas that code for holistic representations on Day 5 would predict better performance on Day 30.

2. Method

2.1. Participants

Twenty-eight right-handed participants (6 males, mean age 22.7, range 18–27 years) with no prior knowledge of Korean or any other Asian languages were recruited from the experiment

Table 1
Stimulus material.

		Radicals (=letters)					
Consonants		ㄱ	ㅁ	ㄴ	ㅂ	ㅅ	ㄷ
Vowels		ㅏ	ㅓ	ㅣ	ㅜ	ㅛ	ㅡ
		Pseudowords: Trained, Novel1, Novel2					
List A	나 비	세 구	드 모	기 데	보 사	무 느	
	na pi	se ku	tuh mo	ki te	po sa	mu nuh	
List B	니 다	소 네	두 스	가 부	베 미	므 고	
	ni ta	so ne	tu suh	ka pu	pe mi	muh ko	
List C	마 도	게 시	브 누	노 메	수 바	디 그	
	ma to	ke si	puh nu	no me	su pa	ti kuh	
		Pseudowords: Recombined1, Recombined2					
List 1	List A'	비 무	구 드	모 세	데 나	사 기	느 보
		pi mu	ku tuh	mo se	te na	sa ki	nuh po
	List B'	다 므	네 가	스 니	부 소	미 두	고 베
		ta muh	ne ka	suh ni	pu so	mi tu	ko pe
	List C'	도 수	시 마	누 게	메 브	바 디	그 노
		to su	si ma	nu ke	me puh	pa ti	kuh no
List2	List A'	비 드	구 나	모 기	데 보	사 무	느 세
		pi tuh	ku na	moki	te po	sa mu	nuh se
	List B'	다 베	네 두	스 가	부 므	미 소	고 니
		ta pe	ne tu	suh ka	pu muh	mi so	ko ni
	List C'	도 게	시 브	누 디	메 수	바 노	그 마
		to ke	si puh	nu ti	me su	pa no	kuh ma

participation pool of Radboud University (22 Dutch, 5 German and 1 Polish). Participants had no history of neurological disorders, language-related disorders, or any contraindication for MRI scanning, and reported having normal or corrected-to-normal vision and hearing. They provided written informed consent and received course credit or were paid for their participation. Three participants were omitted from the analyses, one due to not completing the full protocol, one due to an asymptomatic mass in the anatomical scan which caused distortion in the functional images, and one due to having too few trials (< 10 on all conditions on Day 1), leaving data of 25 participants in the analyses.

2.2. Materials

Six consonant- and six vowel-radicals were chosen from the existing Korean Hangeul orthography (see Table 1). These consonant-vowel combinations resulted in 36 unique characters either in a left-right or top-bottom construction. The characters were divided into three sets, with each radical occurring twice in each of the sets (Table 1). For all three sets, 12 unique characters were combined to create six two-character pseudowords. Since the radicals have a one-to-one correspondence to phonemes, once the participants are aware of the character construction (left-right or top-bottom), the characters can be pronounced in the same way that words can be decoded in a transparent language if the letter-sound associations are known. Radicals thus served as letters (graphemes) and characters served as syllables in our experimental manipulation. All radicals were distinct and the associated phonemes existed in the participants' first language, except for one vowel for which we explicitly explained how it would be pronounced. Of the three lists (Table 1 Lists A, B and C), one list served as Trained pseudowords, and the other two lists were used as Novel pseudowords for Day 1 and Day 3 (Novel1), and Day 5 and Day 30 (Novel2). Two lists of recombination of characters were created for each of the three sets (i.e., all characters appeared in the Trained condition, but character combinations were rearranged; lists 1: A'1, B'1, C'1; lists 2: A'2, B'2, C'2), and these served as the Recombined condition on Day 1 and Day 3 (Recombined1), and Day 5 and Day 30 (Recombined2). Assignment of lists A, B and C to Trained, Novel1 and Novel2, as well as assignment of lists 1 and 2 of the Recombined condition to Recombined1 and Recombined2, were counter-balanced across participants. We limited our number of radicals used in this paradigm to six consonants and six vowels because we needed the participants to read Novel pseudowords correctly already on Day 1 after a short training session.

2.3. Procedure

On Day 1, participants first received an explanation of the nature of Korean orthography. This was followed by the training session. Upon completion of the training session, the test in the functional magnetic resonance imaging (fMRI) scan took place. The second training session (Day 3) took place two days later (one day later for one participant and three days later for another participant). Participants came back on Day 5 (four days after the initial training for all participants, except for one who came back after three days) and again on Day 30 (average 31.1 days after the first training session, range 28–37 days). In the sessions on Day 5 and Day 30, participants were reminded of the pronunciation of each radical once, just prior to the test session (Fig. 1).

2.3.1. Explanation

First, the basic nature of Hangeul characters was explained using radicals that were not used in the actual experiment. Pronunciations of two vowels and two consonants were first introduced

with a video depicting the mouth movement and the actual pronunciation by a native Korean woman. Different consonant-vowel (CV) combinations of these radicals resulting in two characters in a left-right arrangement and two characters in a top-down arrangement were then introduced together with the video clips showing the mouth movements and pronunciations.

2.3.2. Training

Training on Day 1 consisted of training radicals and training pseudowords (Fig. 1A). First, participants were exposed to six consonant radicals and six vowel radicals (Table 1) on the screen with an example pronunciation video, one at a time, where they were instructed to repeat the pronunciation after the example. Three rounds of pronunciation for each radical were administered. For each radical, participants could repeat the video as many times as they wished before proceeding to the next radical trial.

Training of pseudowords followed. Participants were exposed to a list of two-character pseudowords comprised of the 12 radicals that they had been exposed to. In the very first instance of the pseudoword training, participants were exposed to all six pseudowords from the list assigned to the Trained condition, which were presented on the screen one word at a time with an example pronunciation video. After the initial exposure, the participants viewed a recapitulation of the radical pronunciations. Subsequently, the participants received five blocks of pseudoword training, with each block containing one trial for each pseudoword from the Trained condition. This time they were encouraged to pronounce each item out loud when they were cued with the pseudoword on the screen. They could press a button to see the correct pronunciation video as many times as they wished for each trial before proceeding to the next.

A short test followed to ensure that the participants were focusing on both the word and the grapheme level of reading (Test1, see next section for details). This was followed by another five blocks of pseudoword training, two blocks of radical training and again five blocks of pseudoword training, after which they were tested inside the scanner (Test2, see next section for details).

Training on Day 3 comprised of pseudoword training (5 blocks), radical training (2 blocks) and again pseudoword training (10 blocks). After this training, participants took part in Test3. There was no explicit training on Day 5 and Day 30, but a short recapitulation of all trained radicals prior to testing. We briefly exposed the participants to the radical-phoneme association prior to the test sessions to eliminate possible repetition effect on retrieving radical-level print-to-sound association since the same radicals appeared in all conditions.

2.3.3. Test

During the test sessions, participants were cued with two-character pseudowords on the screen one at a time, and they were instructed to say the pseudoword out loud. Test1 consisted of six trained pseudowords (Trained condition), six novel pseudowords (Novel1), and six recombined pseudowords (Recombined1), resulting in total of 18 pseudowords. The pseudowords were presented one at a time, and participants were instructed to pronounce the word clearly into the microphone. The word stayed on screen until the participant pressed a button to proceed to the next word.

Test2 took place in the scanner after the training session on Day 1. The same 18 pseudowords were used as in Test1, but this time they were repeated in five blocks with each pseudoword occurring once in every block, giving a total of 90 trials. Participants were told to pronounce the pseudoword as soon as they knew the pronunciation, and that they had five seconds to respond for each trial. For each trial, one pseudoword (two Hangeul characters) in black font on a white background was presented with a cue "Speak

out!” written in English above the pseudoword. After five seconds, “Press the button to go to the next trial” appeared at the bottom of the screen. Once the participants had pronounced the pseudoword, they could choose to press the button to go to the next trial, or wait until the trial ended. The trial terminated when a button response was given, or 10 s after the onset of the pseudoword presentation if no button response was given. The next trial appeared after an inter-trial interval (ITI) that jittered between three to six seconds, during which a fixation cross was presented in the centre of the screen. The average duration of the test in the scanner was 15 min (range 13–19 min).

Test3 took place on Day 3 at the computer, again with 18 pseudowords tested on Day 1, presented once each. Test4 took place on Day 5 in the scanner and was identical to Test2 on Day 1, except that the lists used for Recombined and Novel conditions were new (Recombined2 and Novel2) and the trial order within every block was altered. Test5 took place on Day 30 at the computer, and was identical to Test4, apart from the trial order within every block being rearranged, and ITI set to 1 s.

For all tests, participants were instructed to start pronouncing the pseudoword as soon as they knew the pronunciation, with a time limit of five seconds (in reality they had ten seconds maximum for production), except for Test1 in which participants were instructed to press the space-key before initiating their production of the pseudoword. Presentation of the stimulus material was done using Presentation software (Neurobehavioral Systems; www.neurobs.com).

2.4. Data acquisition

2.4.1. Behavioural data

Voice recordings during the test were auditorily inspected and classified as correct or incorrect, and visually inspected for the reaction time (RT) of the initiation of word production, relative to the onset of the cue on the screen (voice onset time; VOT) as well as the duration of the utterance. This was scored offline using software Audacity (<http://audacity.sourceforge.net/>). Time taken to complete the pronunciation (voice offset time; VOFT) was calculated as the sum of VOT and duration. Behavioural data were analyzed using SPSS (IBM SPSS Statistics for Windows, Version 21.0).

2.4.2. MRI acquisition

fMRI data were recorded in a 3 T MR scanner (Skyra, Siemens Healthcare, Erlangen, Germany) using a 32-channel head coil. For functional images, we used a T2*-weighted gradient multi-echo planar imaging sequence with the following parameters: repetition time (TR): 2.07 s, echo times (TE): TE1 9.0 ms, TE2 19.3 ms, TE3 30 ms, TE4 40 ms, 34 slices, ascending slice order, 3.0 mm slice thickness, 0.5 mm slice gap, in plane matrix size: 64 × 64, field of view (FOV): 224 × 224 mm, leading to a voxel size of 3.5 × 3.5 × 3.5 mm, flip angle: 90°. Slices were angulated in an oblique axial manner to reach whole-brain coverage (except for a part of the parietal cortex and the cerebellum). Additionally, T1-weighted anatomical scan at 1 mm isotropic resolution were acquired with TR 2300 ms, TE 3.03 ms, flip angle 8°, and FOV 256 × 256 × 192 mm on either Day1 or Day5.

2.5. fMRI analyses

The multi-echo sequence acquired four echoes per TR for the functional scans. To allow for the use of standard fMRI pre-processing tools in SPM8 (www.fil.ion.ucl.ac.uk), these four echoes were combined using the procedure described by Poser et al. (2006). Thirty initial volumes of the functional run were used to calculate the optimal weighting of the four echoes to be combined

to one value per time point for every voxel, and this weighting matrix was applied to all the functional scans. Image pre-processing and statistical analysis were performed using SPM8 (www.fil.ion.ucl.ac.uk). The subject-mean image of the functional run after realignment was co-registered with the corresponding structural MRI using mutual information optimization. The functional images were subsequently slice-time corrected to the first slice, using SPM8's Fourier phases shift interpolation. Structural scans were segmented and both functional and structural scans were spatially normalized and transformed into the common Montreal Neurological Institute space (resampled at voxel size 2 × 2 × 2 mm), as defined by the SPM8 T1.nii template, as well as spatially filtered by convolving the functional images with an isotropic 3D Gaussian kernel (8 mm full width at half maximum).

The fMRI data were analyzed statistically using a general linear model (GLM) and statistical parametric mapping. To compare the brain activation patterns related to correct pronunciation of pseudowords for the three conditions of interest, four explanatory variables were included in the first model for each session: Trained-correct, Recombined-correct, Novel-correct, and trials-of-no-interest comprising all incorrect trials and no response trials. Although the number of trials for the trials-of-no-interest were very few (Day 1 $M=9.6$ trials, Day 5 $M=6.6$ trials), we treated mispronounced trials and non-responded trials separately from the correctly pronounced trials to circumvent possible brain activation related to retrieval of incorrect grapheme-phoneme associations, or doubts. These explanatory variables were temporally convolved with the canonical Hemodynamic Response Function (HRF) provided by SPM8. Each event was time-locked to the onset of the pseudoword presentation on the screen. Furthermore, all three correct conditions included regressors pertaining to two parametric modulators, one corresponding to repetition of each pseudoword, and another accounting for the memory strength, which was computed as the difference in response latency from the mean (mean VOFT for all correct pseudowords - VOFT for the trial). The design matrix included the six head motion regressors (three translations, three rotations). A high pass filter was implemented using a cut-off period of 128 s to remove low-frequency effects from the time series. For statistical analysis, relevant contrast parameter images were generated for each participant and then subjected to a second-level analysis (Penny et al., 2003), treating subjects as a random variable.

To test for longer lasting effects, a second model was run. In this model, Day 5 session data included two explanatory variables (correct, incorrect conditions) convolved with the canonical HRF. For correct trials, similar to the first model above, two parametric modulators were included: one corresponding to repetition of each pseudoword, and another accounting for the memory strength on Day 30, which was computed as the difference in response latency between the mean VOFT of all correct trials and the mean of all correct trials for each pseudoword across the five blocks on Day 30. If a specific pseudoword was incorrect for all five presentations on Day 30, this pseudoword was also categorized as incorrect.

Finally, we ran another model to account for the different reaction times between conditions that could potentially be a confounding factor. Following the method suggested by Taylor et al. (2014), both Day 1 and Day 5 session data were modeled with two explanatory variables (correct, incorrect conditions), with three additional parametric modulator regressors for the correct condition, the first accounting for the repetition, the second for the reaction time (VOFT), and the third for the different conditions (weighted as $-1 = \text{Trained}$, $0 = \text{Recombined}$ and $1 = \text{Novel}$). In this way, the third parametric modulator will explain the variance in the brain response after the effects of repetition and difference due to reaction times have been explained.

For all contrasts, we applied a cluster defining threshold at voxel-level $p < .001$ uncorrected, and used the family-wise error corrected cluster extent threshold of $p_{FWE} < .05$.

3. Results

3.1. Behavioural results

We assessed accuracy and reaction times (RT) as dependent variables of interest. Repeated Measures analysis of variance (rmANOVAs) were conducted to assess differences between the test days (Day 1, Day 5, Day 30) and the conditions (Trained, Recombined, Novel) for all dependent variables. Greenhouse-Geisser values are reported whenever sphericity assumption was violated. To elaborate on the main and interaction effects, we investigated their contrasts as a planned comparison test. Since all dependent variables on the RT data did not follow, or even approximate, a normal distribution as indicated by positive kurtosis and skewness values, we used a Log Transformation on the raw RT data (Keene, 1995; Ratcliff, 1993). This entails that the logarithm with base 10 was taken for these variables, which approximated a normal distribution. Since accuracy was left skewed, we utilized Square Transformation to normalize its scores (Mukherjee et al., 1998).

In addition to rmANOVAs, we conducted Pearson's correlations to address our second question on the retention of training effects on performance between Day 1 and Day 30, and Day 5 and Day 30. All correlational analyses were performed on the transformed data.

3.1.1. Accuracy

We expected performance to improve with training, thus Trained pseudowords were expected to yield the best performance and Novel pseudowords the worst, and Day 5 performance was expected to exceed performance on Day 1. If indeed a sufficiently stable holistic representation emerged, performance should furthermore remain stable at Day 30. This was investigated by a 3×3 rmANOVA with factors Test day (Day 1, Day 5 and Day 30) and Condition (Trained, Recombined and Novel). Although accuracy was very high in general (all $> 80\%$), the test yielded both a main effect of Test day and of Condition (both $p < .05$; see Table 3 for detailed statistics). No significant interaction effect was found ($p = .361$). Pairwise contrasts for Test day revealed that performance on Day 5 showed a trend towards improved performance in comparison with Day 1 ($p = .053$). In addition, performance on Day 5 was significantly better than on Day 30 ($p = .013$). This suggests overall forgetting over time, such that without further training,

Table 2.
Behavioural performance: mean (standard error of the mean).

	Condition	Day 1	Day 5	Day 30
Accuracy [% correct responses]				
	Trained	92.3 (2.0)	96.8 (1.4)	93.2 (1.4)
	Recombined	89.9 (2.9)	92.9 (1.6)	85.7 (3.6)
	Novel	86.0 (2.4)	88.1 (2.6)	81.1 (3.8)
Reaction Times [ms]				
Voice Onset Time (VOT)	Trained	2260 (130)	1870 (120)	1987 (175)
	Recombined	2404 (151)	2085 (138)	2109 (186)
	Novel	2574 (152)	2239 (142)	2313 (191)
Duration	Trained	1685 (109)	1408 (83)	1406 (90)
	Recombined	1829 (119)	1631 (105)	1675 (122)
	Novel	1929 (134)	1773 (116)	1826 (141)
Voice Offset Time (VOFT)	Trained	4008 (168)	3297 (133)	3438 (195)
	Recombined	4308 (187)	3756 (168)	3836 (214)
	Novel	4554 (176)	4059 (165)	4232 (215)

Table 3.

Statistics of the repeated measures ANOVAs on accuracy and reaction time (Test day \times condition).

Accuracy	d.f.	F-value	p-value	η_p^2
<i>Main effects</i>				
Test day	2	5.065	.010	.174
Day1 < Day5	1	4.134	.053	.147
Day1 = Day30	1	2.150	.156	.082
Day5 > Day30	1	7.265	.025	.194
Error	24			
Error	48			
Condition	2	22.591	.001	.485
Trained > Recombined	1	14.050	.001	.657
Trained > Novel	1	47.372	< .001	.664
Recombined > Novel	1	9.219	.006	.278
Error	24			
Error	48			
VOFT	d.f.	F-value	p-value	η_p^2
<i>Main effects</i>				
Test day		29.074	< .001	.548
Day1 > Day5	1	71.493	< .001	.749
Day1 > Day30	1	29.892	< .001	.555
Day5 = Day30	1	1.499	.233	.059
Error	24			
Error	48			
Condition		124.887	< .001	.839
Trained < Recombined	1			
Trained < Novel	1	200.835	< .001	.893
Recombined < Novel	1	73.371	< .001	.754
Error	24			
Error	48			
<i>Interaction effect</i>				
Test day \times condition	3.09	9.54	< .001	.284
Trained – Novel				
Day5 > Day1	1	27.617	< .001	.535
Day30 > Day1	1	27.594	< .001	.535
Trained – Recombined				
Day5 > Day1	1	22.074	< .001	.479
Day30 > Day1	1	6.169	.02	.204
Recombined – Novel				
Day5 > Day1	1	3.257	.084	.119
Day30 > Day1	1	11.275	.003	.320
Error	24			
Error	62.211			
VOFT first vs repeated blocks	d.f.	F-value	p-value	η_p^2
<i>Main effects</i>				
Test day	2	28.354	< .001	.563
Error	48			
Condition	1.483	88.255	< .001	.800
Error	32.623			
Block	1	132.876	< .001	.858
Error	22			
<i>Interaction effect</i>				
Test day \times Condition \times Block	2.99	3.719	.016	.145
Error	65.84			
Test day \times Condition	4	5.609	< .001	.203
Error	88			
Test day \times Block	2	2.593	.086	.105
Error	44			

Note: Only relevant interaction effects are displayed.

performance dropped back to the level of Day 1 (Day 1 vs. Day 30 $p = .156$). Pairwise contrasts for Condition revealed that accuracy for Trained pseudowords was significantly higher than for Recombined and Novel pseudowords, and accuracy for Recombined pseudowords was significantly better than for Novel pseudowords (all $p < .01$). Overall, training was effective, with the Trained

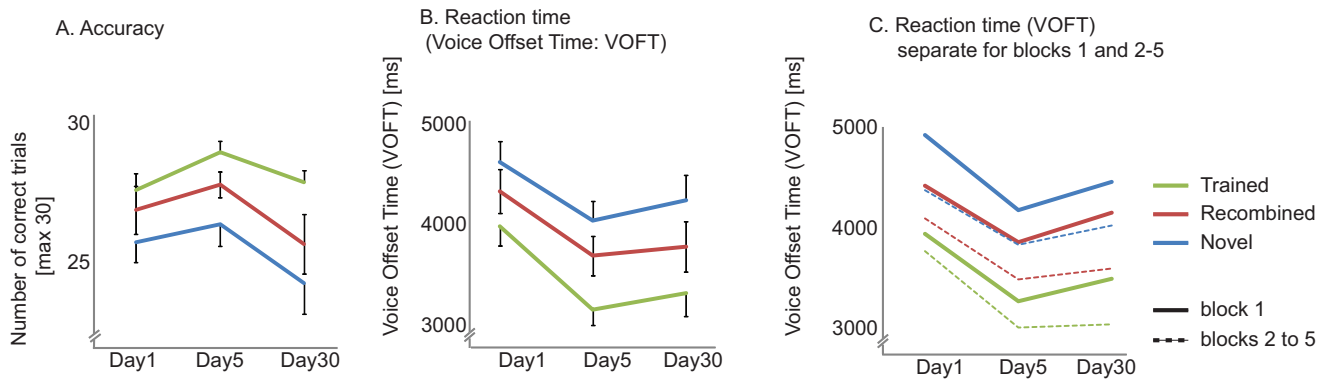


Fig. 2. Behavioural results. (A) Accuracy: mean number of words correctly pronounced during each test. Six words in each condition were presented five times each, summing to 30 trials per condition. (B) Reaction times: mean Voice offset times (VOFT=voice onset time (VOT) + duration of utterance). (C) Mean VOFT separately presented for block 1 and mean of blocks 2–5. Reaction times are in ms. Error-bars denote standard error of the mean. Green: Trained, red: Recombined, blue: Novel. For panel C, solid line: block 1, dashed line: mean of blocks 2–5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

condition showing the best performance on Day 5, however termination of training led to forgetting in general (Tables 2 and 3, Fig. 2A).

3.1.2. Reaction time

We noticed that participants had different strategies in production. Some initiated the pronunciation once they knew how to read the whole pseudoword, whereas others initiated the pronunciation as soon as the first radical was decoded, and figured out how to pronounce the rest while they were producing the pseudoword. The former strategy will result in shorter overall utterance time (duration) at the cost of longer VOTs. A Pearson's correlation between VOT and duration across conditions revealed a negative correlation between VOT and duration for all conditions when RTs were summed across the Test days (Trained $r = -.419$, $p = .037$, Recombined $r = -.479$, $p = .015$, Novel $r = -.634$, $p = .001$). Correlations between VOT and Duration across the conditions were also all very high (all $p < .001$). Thus a composite measure of VOT+duration, i.e., VOFT, would reflect overall RT better.

Using the VOFT of the correct trials as the RT measures, we tested our hypothesis. We expected RT to speed up with training from Day 1 to Day 5, and the Trained condition was expected to be faster than the Novel condition. Additionally, if a stable holistic representation emerged through training, we expected RT to remain stable over time when tested on Day 30. If, on the other hand, we observe longer RT on Day 30, this would suggest that the trained representations were not stable enough, and forgetting had occurred after the training had stopped. To test this, a 3×3 Repeated Measures ANOVA with factors Test day (Day 1, Day 5 and Day 30) and Condition (Trained, Recombined and Novel) was conducted. We found a significant main effect of Test Day and Condition, as well as the interaction of Test Day \times Condition (all $ps < .001$, Table 3). Planned pair-wise comparisons by means of Simple and Repeated contrasts were conducted and confirmed our hypotheses. RT sped up over time. VOFT was longer on Day 1 than on Day 5 ($p < .001$) and remained stable over time from Day 5 to Day 30 ($p = .233$). In addition, training shortened RT as expected. The Trained words were pronounced faster than the Novel words, with the Recombined words in between (all $ps < .001$). The difference between the Trained and Novel condition is greater on Day 5 and Day 30 than on Day 1, the same effect is observed for the difference between the Trained and Recombined condition (all $ps < .05$). The difference between the Recombined and Novel condition is marginally significant on Day 5 ($p = .084$) and significantly larger on Day 30 than on Day 1 ($p = .003$). Overall, we observed a speeding up in reading time with training as expected,

reflecting more efficient pronunciation of repeated radical combinations across testing days and conditions as a result of training. We found the same pattern of results when VOT and duration were tested separately (see Supplementary Table S1 for the details of the effects).

3.1.3. Comparison between first and the repeated trials

Since every pseudoword was repeated five times within each test, there may be a benefit of repetition. If the initial exposure has an influence on the subsequent presentations within a session, participants may become faster with repeated presentation of the same stimulus material. To test this, and to see if the effect is different between conditions across Test days, we split the mean VOFT of all correct trials for the first block and the mean of blocks 2–5. A $3 \times 3 \times 2$ rmANOVA with factors Test day (Day 1, Day 5 and Day 30), Condition (Trained, Recombined and Novel) and Block (block 1, blocks 2–5) was conducted for VOFT (Fig. 2C). Due to missing values, two participants were excluded from the analysis.

There was a main effect of Condition ($p < .001$) and Test day ($p < .001$) and a two-way interaction for Test day \times Condition ($p = .001$) as mentioned above (see Table 3 for statistics). There was also a main effect of Block showing slower RT for the first block relative to the mean of the rest of the blocks (mean block 1 = 4129 ms, mean blocks 2–5 = 3787 ms) as well as a three-way interaction ($p = .016$), and a trend for a two-way interaction Test day \times Block ($p = .086$).

To investigate this further, we conducted rmANOVAs separately for block 1 and blocks 2–5 (see Supplementary Table S2 for detailed statistics). For block 1, there was a significant main effect of both Test day ($p < .001$) and Condition ($p < .001$). Shortest response times for Day 5 and longest response times for Day 1 were observed (all $ps < .05$; mean Day 1 = 4441 ms, Day 5 = 3855 ms, Day 30 = 4091 ms) with the Trained condition being the shortest and Novel the longest (all $ps < .001$; mean Trained = 3677 ms, Recombined = 4189 ms, Novel = 4522 ms). There was no significant interaction effect ($p = .161$).

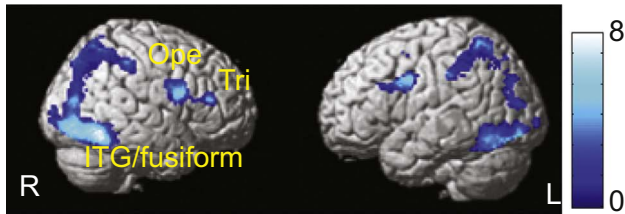
For the repeated blocks (mean RT of blocks 2–5), there was a significant main effect of both Test day ($p < .001$) and Condition ($p < .001$) as well as a significant interaction effect ($p < .001$). Day 1 was significantly slower than Day 5 or Day 30 (both $ps < .001$), but no difference between Day 5 and Day 30 ($p = .462$; mean Day 1 = 4132 ms, Day 5 = 3566 ms, Day 30 = 3665 ms), with the Trained condition being the shortest and Novel the longest (all $ps < .001$; mean Trained = 3414 ms, Recombined = 3817 ms, Novel = 4131 ms). The interaction effect was driven by significantly greater speeding up of the RT for the Trained condition from Day 1 to Day

5 compared to the Recombined or Novel conditions (both $p < .001$) and this difference sustained on Day 30 (all $p < .01$). When we compared the first block to the rest of the blocks separately for each condition, we found significant speeding up

across Test days across all conditions (all $p < .001$), with the first Block being slower than the mean of blocks 2–5 (all $p < .007$), but the interaction effect of Test day \times Condition was significant only for the Trained and Recombined conditions (Trained $p = .010$,

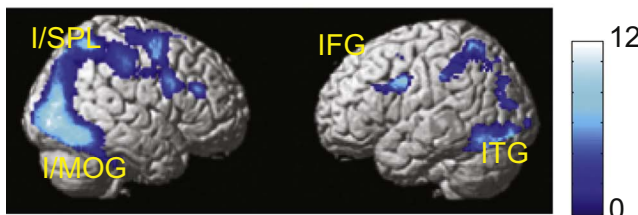
A. Day1

Novel > Trained

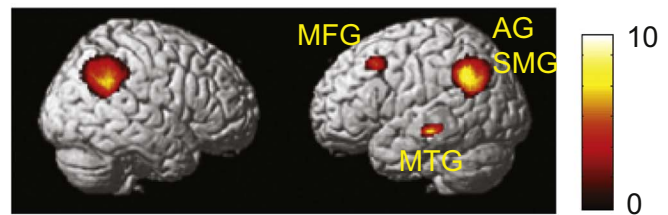


B. Day5

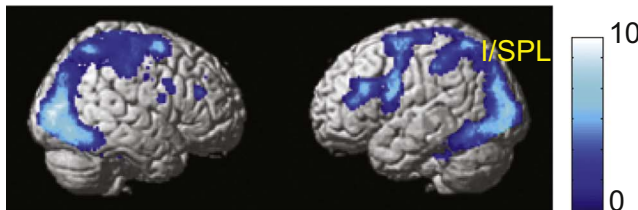
Novel > Trained



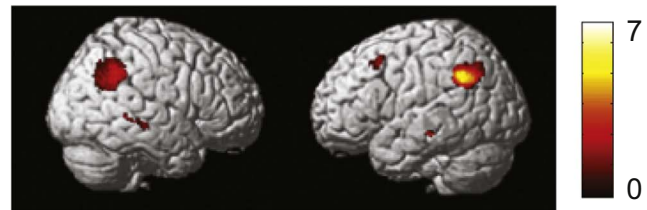
Trained > Novel



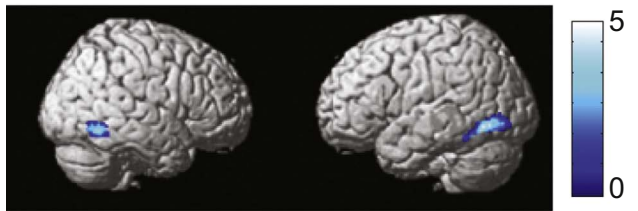
Recombined > Trained



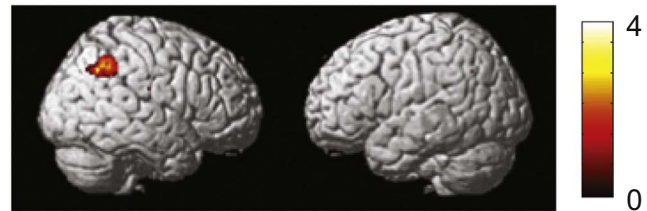
Trained > Recombined



Novel > Recombined

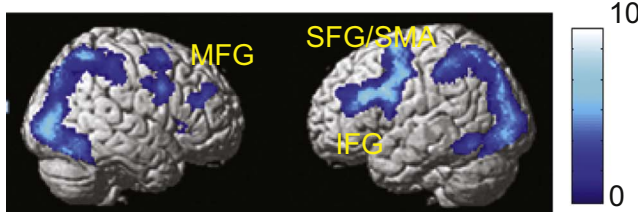


Recombined > Novel



C. Day5 activation pattern predictive of performance on Day30

Negative correlation



Positive correlation

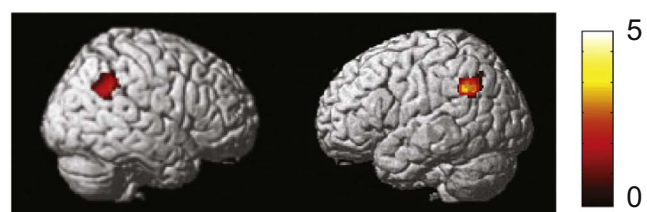


Fig. 3. Difference in brain activation patterns across the conditions and Test days. Areas that showed high activation pattern for the Novel relative to Trained condition on Day 1 (A). Differences between conditions on Day 5 (B). Activation pattern on Day 5 that was predictive of pronunciation speed on Day 30 (C). All significant clusters are superimposed on a template T1 brain (multiple comparison corrected $p_{FWE-cluster} < .05$ on the whole brain search after initial voxel-level thresholding at $p < .001$), uncorrected. Colour bars represent the height of the voxel level t-values. R: right, L: left, ITG: inferior temporal gyrus, I/SPL: inferior/superior parietal lobule, IFG: inferior frontal gyrus, Ope: Pars opercularis of the IFG, Tri: Pars triangularis of the IFG, AG: angular gyrus, SMG: supramarginal gyrus, MFG: middle frontal gyrus, I/MOG: inferior/middle occipital gyrus, MTG: middle temporal gyrus, SFG: superior frontal gyrus, SMA: supplementary motor area.

Recombined $p = .016$, Novel $p = .413$), with greater difference being observed on Day 30 relative to Day 1 or 5 for both conditions (all $ps < .03$).

3.1.4. Correlation between days on performance

Were the good learners good performers even after a delay? To test this, we calculated the Pearson's correlation between Day 5 and Day 30 performance on Accuracy and VOFT. Both performance measures showed a high correlation between the two test days (Accuracy: $r = .543$, $p = .005$, VOFT: $r = .810$ $p < .001$). Within-condition analyses also showed a high correlation between Day 5 and Day 30 (all $r > .5$, $p < .01$), except for Accuracy on the Trained condition ($r = .349$, $p = .087$), which is probably due to a ceiling effect for accuracy on Day 5. Similar results were obtained when Day 1 and Day 30 correlations were analysed.

3.1.5. Summary behavioural results

A clear effect of Training was already visible on Day 1. With additional training, performance for the Trained condition increased even more, as performance on Day 5 showed higher accuracy and faster responses. Even though pseudowords in the Recombined condition were never trained, they did benefit from having the same characters as the pseudowords in the Trained condition, in that behavioural performance in the Recombined condition was better than in the Novel condition. Our follow-up data from Day 30 revealed that there was forgetting over time, suggested by a decrease in accuracy on Day 30. For pseudowords that were correctly pronounced, however, the RTs remained stable even without further training between Day 5 and Day 30. Furthermore, the good learners on Day 1 performed better on Days 5 and 30 as well.

3.2. Imaging results

3.2.1. Effect of training: Day 1 (Trained > Novel, Novel > Trained)

First we were interested in whether we can see a difference between the Trained and Novel condition after a short initial training (Day 1 Test2). To see differential brain activity patterns, we contrasted correctly pronounced Trained pseudowords to Novel pseudowords. This revealed a set of brain areas that showed greater engagement for Novel compared to Trained pseudowords in the dorsal visual attention network (Fig. 3A, Table 4). This indicates that participants required less visual attention and motor preparation in order to read aloud Trained pseudowords than Novel pseudowords. The reverse contrast (Trained > Novel) did not show any significant clusters to suggest more engagement or different processes for the Trained relative to the Novel condition. Thus we did not observe any indication for emergence of holistic representations for the Trained pseudowords on Day 1.

To be sure that the above difference is not merely due to differences in reaction time between conditions, we ran another model that included two main regressors (correct, incorrect) and three parametric modulators for the correct responses (repetition, reaction time, and conditions) and observed the brain activation pattern that is explained by the condition modulator after the effect of repetition and reaction times were accounted for. This revealed a cluster in the right inferior temporal lobe extending to the right fusiform gyrus (peak voxel [50 -58 -12], $p = .001$, cluster size 445 voxels) to be more engaged for the Novel > Recombined > Trained condition. The reverse correlation (i.e., more activation for Trained > Recombined > Novel) did not show any significant clusters. The reaction time modulator on the other hand explained most of the variance for longer RT trials in the areas reported for the Novel > Trained contrast reported above (see Supplementary Material Fig. S1a). No

Table 4.
Significant clusters for the contrast Novel > Trained on Day 1

Cluster		Peak voxel		MNI coordinates			Anatomical areas
<i>p</i> (FWE-corr)	Size	T	Z	x	y	z	
< .001	5002	8.34	5.66	50	-58	-12	R inferior temporal gyrus
		7.24	5.22	40	-46	-20	R fusiform
		5.86	4.57	44	-34	42	R supramarginal
		5.86	4.57	42	-80	-10	R inferior occipital gyrus
		5.57	4.42	20	-64	44	R superior occipital gyrus
		5.56	4.41	18	-56	56	R superior parietal lobule
		5.29	4.26	28	-60	46	R angular gyrus
< .001	593	4.75	3.95	-22	-4	52	L superior frontal gyrus
		4.47	3.78	-8	18	40	L middle cingulate cortex
		4.39	3.72	-8	18	46	L supplementary motor area
		4.38	3.71	-8	22	42	L medial superior frontal gyrus
		4.26	3.64	0	12	52	L supplementary motor area
< .001	1316	6.07	4.68	-24	-78	-14	L lingual gyrus
		5.53	4.4	-46	-64	-14	L inferior occipital gyrus
		5.42	4.34	-40	-82	-12	L fusiform gyrus
		4.91	4.05	-44	-58	-8	L inferior temporal gyrus
		4.58	3.84	-10	-98	-4	L calcarine gyrus
		4.22	3.61	-10	-90	-8	L calcarine gyrus
		4.06	3.51	-14	-98	2	L middle occipital gyrus
		5.77	4.52	50	10	22	R inferior frontal gyrus; pars opercularis
		4.32	3.68	48	34	16	R inferior frontal gyrus; pars triangularis
< .001	767	4.27	3.65	60	8	20	R precentral gyrus
		3.93	3.42	48	6	12	R Rolandic operculum
		5.68	4.48	-26	-56	48	L superior parietal lobule
		4.92	4.05	-40	-40	40	L inferior parietal lobule
< .001	1817	4.55	3.83	-22	-66	38	L superior occipital gyrus
		4.88	4.03	-42	0	30	L precentral gyrus
		4.8	3.98	-40	4	30	L inferior frontal gyrus; pars opercularis
< .001	505	3.86	3.37	-40	16	28	L inferior frontal gyrus; pars triangularis

R: right, L: left, MNI: Montreal Neurological Institute, FWE: Family-wise error corrected.

Table 5.
Significant clusters for the contrast for the Trained, Recombined and Novel conditions on Day5.

<i>p</i> (FWE-correr)	Size	Peak voxel		MNI coordinates			Anatomical areas		
		T	Z	x	y	z			
Trained > Novel < .001	1215	10.65	6.41	−58	−56	40	L inferior parietal lobule		
		8.62	5.76	−44	−72	40	L angular gyrus		
		4.83	4	−64	−46	32	L supramarginal gyrus		
		3.91	3.41	−46	−56	18	L middle temporal gyrus		
< .001	1274	8.68	5.79	56	−62	36	R angular gyrus		
		7.57	5.36	58	−50	44	R inferior parietal lobule		
		7.33	5.26	64	−50	30	R supramarginal gyrus		
0.015	231	6.77	5.01	−58	−26	−10	L middle temporal gyrus		
0.01	251	6.47	4.88	−36	16	42	L middle frontal gyrus		
Novel > Trained 0	22,728	12.44	6.88	−40	−78	−8	L inferior occipital gyrus		
		11.02	6.52	−24	−62	50	L superior parietal lobule		
		10.75	6.44	−36	−86	0	L middle occipital gyrus		
		9.86	6.18	−26	−64	44	L inferior parietal lobule		
		9.65	6.11	−24	−74	32	L superior occipital gyrus		
		11	6.51	16	−96	0	R calcarine gyrus		
		10.32	6.32	24	−88	−4	R lingual gyrus		
		9.99	6.22	30	−68	32	R middle occipital gyrus		
		9.95	6.21	38	−84	−4	R inferior occipital gyrus		
		9.4	6.03	46	−64	−8	R inferior temporal gyrus		
		9.35	6.02	22	−64	50	R superior parietal lobule		
		7.68	5.4	48	12	24	R inferior frontal gyrus pars Opercularis		
		6.84	5.05	28	0	50	R precentral gyrus		
		6.03	4.66	48	0	54	R middle frontal gyrus		
		< .001	5563	8.68	5.79	2	12	52	R supplementary motor area
				7.32	5.26	10	18	36	R middle cingulum
				6.13	4.71	−2	6	68	L supplementary motor area
				7.89	5.49	−50	6	30	L precentral gyrus
				7.69	5.41	−28	−2	68	L superior frontal gyrus
				6.72	4.99	−40	28	24	L inferior frontal gyrus pars Triangularis
5.35	4.3			−36	34	30	L middle frontal gyrus		
5.67	4.47			10	−14	6	R thalamus		
< .001	793	4.34	3.69	20	0	20	R caudate		
		4.74	3.94	−14	−10	4	L thalamus		
		5.29	4.26	44	32	26	R inferior frontal gyrus pars Triangularis		
0.044	175	5.29	4.26	44	32	26	R inferior frontal gyrus pars Triangularis		
		5.29	4.26	44	32	26	R inferior frontal gyrus pars Triangularis		
Trained > Recombined < .001	530	7.02	5.13	−56	−50	36	L inferior parietal lobule		
		6.79	5.03	−58	−58	36	L angular gyrus		
< .001	401	6.27	4.78	62	−50	38	R angular gyrus		
		5.16	4.19	56	−52	44	R inferior parietal lobule		
Recombined > Trained < .001	25,785	10.76	6.44	28	−88	6	R middle occipital gyrus		
		8.94	5.88	22	−88	−2	R lingual gyrus		
		8.93	5.87	24	−62	48	R angular gyrus		
		8.26	5.63	26	−68	36	R superior occipital gyrus		
		7.49	5.33	40	−62	−12	R inferior occipital gyrus		
		7.31	5.25	18	−96	0	R calcarine gyrus		
		10.33	6.32	−28	−78	30	L middle occipital gyrus		
		9.82	6.17	−24	−74	38	L superior occipital gyrus		
		8.05	5.55	−18	−64	58	L inferior parietal lobule		
		7.94	5.51	−14	−70	44	L superior parietal lobule		
		7.71	5.42	−40	−78	−8	L inferior occipital gyrus		
		7.6	5.37	−12	−66	48	L precuneus		
		7.63	5.39	−6	6	56	L supplementary motor area		
		6.24	4.76	−26	−2	60	L superior frontal gyrus		
		5.93	4.61	−26	−14	70	L precentral gyrus		
		5.32	4.28	4	12	46	R middle cingulum		
		6.38	4.83	−40	26	20	L inferior frontal gyrus pars Triangularis		
		0.017	194	4.57	3.84	−48	40	16	L superior temporal gyrus
				5.11	4.16	12	−16	8	R thalamus
		Recombined > Novel 0.001	297	4.72	3.93	58	−48	44	R inferior parietal lobule
Novel > Recombined 0	327	5.02	4.11	−44	−76	−8	L inferior occipital gyrus		

Table 5. (continued)

p (FWE-correr)	Size	Peak voxel		MNI coordinates			Anatomical areas
		T	Z	x	y	z	
0.012	182	4.68	3.9	−52	−66	−8	L inferior temporal gyrus
		4.73	3.94	48	−56	−10	R inferior temporal gyrus
		3.8	3.33	42	−60	−14	R inferior occipital gyrus

R: right, L: left, MNI: Montreal Neurological Institute, FWE-corr: Family-wise error corrected

clusters that reflected shorter RTs were found. Thus most of the areas reported in the Novel > Trained condition above can be explained by higher demands in visuo-spatial attentional network, phonological working memory, and motor preparation.

3.2.2. Effect of training: Day 5

Although we did not observe a neural pattern that suggests emergence of holistic representations on Day1, we can still expect a qualitative change with more repetition and/or time (Davis et al., 2009; Takashima et al., 2014a, 2009; Thielen et al., 2015). To see if this was the case, we contrasted the Trained and Novel condition tested on Day 5. The contrast Trained greater than Novel condition revealed an activation pattern in the bilateral angular gyrus, left middle temporal gyrus, and left middle frontal gyrus suggesting emergence of holistic memory representations for the Trained condition (Fig. 3B, Table 5). The reverse contrast (Novel > Trained) showed a similar pattern to that observed on Day 1, suggesting an effortful process for the Novel condition relative to the Trained condition, engaging the visuo-spatial attentional network, phonological working memory, and motor preparation.

As with the Day 1 result to account for the reaction time difference between trials, we also tested the model with three parametric modulators (repetition, RT, and condition), especially since the areas that showed an effect for the Trained compared to the Novel condition on Day 5 was due to less deactivation relative to rest-baseline period observed for the Trained condition compared to the Novel condition. The condition regressor of this model revealed a very similar pattern to that of the contrast Trained > Novel and Novel > Trained (Supplementary Material Fig. S2). The parametric modulator of the reaction time regressor revealed that longer RT trials engaged similar areas to that observed for Novel > Trained contrast (Supplementary Material Fig. S1b). For trials that showed shorter RTs, small clusters in the bilateral inferior parietal cortex, of which the right extending to the angular gyrus was observed but neither clusters survived the cluster level threshold (left $p=.293$, right $p=.058$, Supplementary Material Fig. S1c). Thus the activation pattern observed for the condition effect can be attributed to processes that reflect over and above any between condition differences in response times.

We expected the Recombined condition to show an activation pattern that was in between Trained and Novel condition since Recombined pseudowords were recombination of characters (equivalent to syllables) used in the Trained condition. Comparison between the Trained and Recombined conditions on Day 5 showed a similar pattern to the contrast between the Trained and Novel conditions on Day 5, indicating that pronunciation of Recombined pseudowords required more visual attention, working memory and motor preparation than pronunciation of Trained pseudowords. Although smaller, the contrast between the Recombined and Novel conditions on Day 5 showed a similar pattern to the contrast between the Trained and Novel conditions (Fig. 3B, Table 5). In other words, Recombined condition showed less

deactivation relative to the rest period than the Novel condition in the right inferior parietal lobule. Decoding of Recombined pseudowords engaged the right inferior parietal lobule/angular gyrus to a higher degree than the Novel pseudowords, whereas for the Novel condition required more processes related to the bilateral occipital to inferior temporal gyrus. This suggests that the character (syllable) level representations also emerged through training on pseudowords that contained these characters.

3.2.3. Change in activation pattern from Day 1 to Day 5

To see if there were changes in the brain activation patterns with more training and time, we compared the brain patterns between the two sessions separately for each condition. Direct comparisons between Days for each of the conditions showed that for the Trained condition, more activation was observed in the left middle temporal gyrus extending to the fusiform gyrus, and another cluster in the right middle/superior frontal gyrus on Day 1 relative to Day 5. For the Recombined condition, more activation in the midline prefrontal cortex and left middle occipital clusters was observed on Day 1 relative to Day 5. For both the Trained and Recombined conditions, we did not observe any levels of activity that was greater for Day 5 compared to Day1. For the Novel condition, no clusters survived our statistical threshold for both contrasts (i.e., Day 1 > Day 5 and Day 5 > Day 1). To test if the difference between Trained and Novel increased as a function of training and time, we contrasted the difference between the Trained and Novel conditions for Day 1 and Day 5. No clusters survived our statistical threshold. Comparison of between session scans is often difficult due to low scores on test-retest reliability on higher cognitive tasks (e.g., Brandt et al. (2013)), and this may partially explain our non-significant findings in comparing between-session data.

3.2.4. Activation pattern on Day 5 predictive of behavioural performance on Day 30

To test whether specific brain activation patterns were predictive of performance after a longer delay, another model was tested. This modeled only the Day 5 data, with two regressors of interest (correct and incorrect trials, irrespective of training conditions), and two parametric modulators for the correct trials regressor (one for repetition across blocks and another accounting for the memory strength at Day 30 (calculated as mean RT of all correct pseudowords at Day30 - mean RT for each of the pseudowords across five blocks at Day30 if pronounced correctly)). Activation level increase during the pronunciation task on Day 5 in the bilateral supramarginal and angular gyrus was predictive of faster RT on Day 30, with the left middle to inferior temporal gyrus cluster showing a trend (Fig. 3C, Table 6). For those pseudowords which were pronounced slower on Day 30, the activation level in the dorsal visual attentional network was higher during the pronunciation task on Day 5. Also for this model, to ensure that the activation pattern is reflecting brain processing difference that is

Table 6.
Significant clusters for the activation pattern for correct trials that correlate with memory strength on Day 30.

Positive correlation with memory strength on Day 30							
Cluster		Peak voxel		MNI coordinates			Anatomical areas
<i>p</i> (FWE-corr)	Size	T	Z	x	y	z	
0.016	235	5.1	4.16	64	−50	30	R supramarginal gyrus
		5.01	4.1	62	−50	38	R angular gyrus
0.004	315	4.95	4.07	−54	−52	34	L angular gyrus
		4.95	4.07	−58	−52	32	L supramarginal gyrus
0.076	155	4.33	3.69	−62	−20	−16	L middle temporal gyrus
		4.33	3.69	−60	−12	−24	L inferior temporal gyrus
Negative correlation with memory strength on Day 30							
<i>p</i> (FWE-corr)	Size	T	Z	x	y	z	Anatomical areas
<.001	16,457	10.2	6.28	−14	−62	54	L precuneus
		9.83	6.17	−18	−62	52	L superior parietal lobule
		9.3	6	−28	−66	42	L middle occipital gyrus
		8.66	5.78	−30	−62	44	L inferior parietal lobule
		8.24	5.62	−24	−58	56	L superior parietal lobule
		8.15	5.59	20	−68	50	R superior parietal lobule
		8.08	5.56	20	−66	46	R superior occipital gyrus
		8.55	5.74	−8	18	42	L medial superior frontal gyrus
<.001	9860	8.5	5.72	−4	10	52	L supplementary motor area
		8.07	5.56	−4	14	44	L middle cingulate gyrus
		8.06	5.55	−20	2	56	L superior frontal gyrus
		7.97	5.52	−44	8	30	L inferior frontal gyrus pars Opercularis
		7.73	5.43	−42	24	24	L inferior frontal gyrus pars Triangularis
		7.88	5.49	−2	6	58	L supplementary motor area
		7.4	5.29	−48	0	46	L precentral gyrus
		7.27	5.23	6	18	40	R middle cingulate gyrus
0.001	446	5.31	4.28	54	6	32	R precentral gyrus
		3.97	3.45	54	0	34	R postcentral gyrus
0.007	286	4.41	3.74	40	44	30	R middle frontal gyrus
		4.21	3.61	44	30	26	R inferior frontal gyrus pars Triangularis
		4.21	3.6	42	38	22	R middle frontal gyrus

R: right, L: left, MNI: Montreal Neurological Institute, FWE-corr: Family-wise error corrected

over and above the reaction time difference across the trials on Day 5, we ran another model including both Day 5 RT and Day 30 RT as modulating regressors. After accounting for RT differences on Day 5, the Day 30 RT regressor showed areas that were predictive of better performance on Day 30 in the left angular (peak voxel [−54 −52 34], cluster size 65 voxels) and middle temporal gyrus (peak voxel [−66 −30 −10], cluster size 110 voxels), although the two clusters did not survive our cluster-level threshold. The reverse contrast showed areas in the bilateral fronto-parietal attentional network (Supplementary Material, Figs. S3a and S3b). When the parametric modulator for the RT on Day 30 was tested separately for the three conditions, none of the contrasts survived our statistical threshold. However, inspection of mean beta values extracted from the regions that predicted better performance on Day 30 (i.e., left angular, right angular and left middle/inferior temporal gyrus clusters) revealed that the effect was mainly driven by the Trained condition (Supplementary Fig. S3c).

3.2.5. Summary imaging results

In line with the study by Takashima et al. (2014b), when participants had to rely on serial decoding for pronunciation, they strongly engaged the occipito-parieto-frontal network (visual attentional network and phonological working memory) extending to motor preparatory areas in the inferior frontal gyrus, supplementary motor area, and premotor cortex. With training on a specific combination of radicals (Trained pseudowords), increased engagement of the bilateral angular/inferior parietal gyrus, left

middle temporal gyrus and left middle frontal gyrus was observed. This suggests that holistic representations of repeatedly encountered pseudowords may have emerged in these areas. Furthermore, engagement of the bilateral angular and left posterior middle temporal gyrus during test on Day 5 was predictive of performance three weeks later without further training, indicating that emergence of these representations in these areas led to stable performance over time. The newly formed representations also appeared to extend to the sub-word level, in this case the character (syllable) level, as the activation pattern observed during decoding of recombined characters (syllables) fell in between the patterns observed in the Trained and Novel conditions.

4. Discussion

In this study, we investigated the effect of training on the decoding of pseudowords in a novel script. We tested whether repeated training on pronouncing a novel orthographical pattern induces a holistic representation, such that decoding and pronunciation become more efficient after multiple exposures. Participants were trained to pronounce a set of pseudowords written in an initially unfamiliar but transparent script (Korean Hangul characters), which became familiar with repeated practice on three levels: the word-level (Trained), syllable-level (Recombined) and letter-level (Novel). Brain activation during pronunciation was measured after the initial training. We also tested after an extra

training session and a delay of four days, in order to probe extra training- and/or time-related neural changes.

A training effect was observed in behavioural data as well as in brain activation patterns. Repeatedly experienced combinations of radicals (i.e., the Trained pseudowords) were read more accurately and faster than the novel combination of the same radicals (i.e., the Novel condition). The difference between Trained and Novel pseudowords increased even more after training, with responses on Trained pseudowords speeding up on Day 5 relative to Day 1. At the brain level, less engagement of brain areas that are known to reflect visual-attention in the dorsal occipito-parietal areas, and motor preparation and phonological working memory storage areas in the posterior portion of the inferior frontal cortex for Trained relative to Novel pseudowords on Days 1 and 5 were observed. Instead, on Day 5, the Trained pseudowords recruited the bilateral angular/inferior parietal, left middle temporal and left middle frontal gyri. The areas that increased and decreased with training overlapped to a large degree with the study by Takashima et al. (2014b), in which Dutch participants were trained on speeded decoding of Greek orthography. The areas that decreased in involvement with training also overlap to a great extent for assembled phonology network reported by Mei et al. (2014). They report the bilateral inferior frontal/precentral gyri, bilateral supramarginal extending to superior parietal lobe, and bilateral inferior occipital gyri of which the left extending to the fusiform gyrus as the areas that were more engaged during assembled phonology. As for the addressed phonology, they reported not only the bilateral angular, left middle temporal gyrus, and left middle frontal gyri, similar to our findings, but also areas in the right middle frontal, anterior and posterior cingulate, right orbitofrontal and right fusiform gyri were reported. The observed pattern was also similar to the results of a study that compared pseudowords and words written in either familiar or unfamiliar fonts (Ischebeck et al., 2004), in which inferior frontal regions exhibited more activation for pseudowords and words in unfamiliar scripts, whereas words in familiar fonts elicited an activation increase in the bilateral inferior parietal lobule, left middle temporal gyrus and left middle frontal gyrus.

Better performance was also observed in the Recombined condition relative to the Novel condition. Pseudowords in the Recombined condition were comprised of the same characters that appeared in the Trained condition, and thus shared sub-lexical syllabic-level units. If syllable-level representations also emerge through repeated exposure to words in the Trained condition, conversion-and-compilation processes can be applied at the syllable-level rather than at the grapheme-level, which should lead to more efficient decoding than in the Novel condition. Since compilation of syllable units is still necessary and thus less efficient than word-level decoding, performance in the Recombined condition would, however, not be expected to reach the level of the Trained condition. Both behavioural and imaging results showed that responses to the Recombined condition indeed fell in between the Trained and Novel conditions, suggesting that intermediate sub-lexical units (syllable/character) became available for decoding.

These results confirm our first hypothesis that repeated experience with a specific combination of radicals (letters/graphemes) leads to a different decoding process. In the beginning, participants had to decipher each radical separately and compile them in order to pronounce the two-character pseudowords comprised of four radicals. This process was slower, and mistakes on the smallest unit (replacing a phoneme with another one) were most often observed. In terms of neural responses, the Novel condition elicited activation increases in occipito-parietal and frontal areas, suggesting that high levels of visual-attention, phonological working memory load, and motor preparation were

demanding for this condition. These findings indicate that during the initial stage of decoding when conversion takes place at the level of grapheme-phoneme units (given that grapheme-phoneme correspondence is transparent), more computational resources are required to pronounce the whole word because there is a demand for more units to be converted and these converted units need to be kept online for compilation.

With practice, participants became more efficient at reading aloud repeatedly exposed pseudowords. One possible explanation for the observed decrease in reaction times to the trained pseudowords could be that a more holistic decoding for these pseudowords and characters is taking place. Brain activation pattern on Day 5 support this idea, with increased levels of activation in the bilateral angular/supramarginal, left middle temporal and left frontal gyri for the Trained condition relative to the Novel condition. The posterior regions are often reported in functional imaging studies of reading (Price and Mechelli, 2005). Although with the current study we cannot dissociate between orthographic, phonological and motoric word-level representations, the areas that increased in activation for the trained pseudowords on Day 5 are related to processing of higher-order multimodal association areas. The posterior middle temporal gyrus is often reported as an audio-visual convergence zone (Gow, 2012; Hickok and Poeppel, 2007) or an area that codes for the phonological lexicon (e.g. Takashima et al. (2014a)). The angular gyrus is often referred to as an area that codes for semantic information (Binder and Desai, 2011; Graves et al., 2010), but since our stimuli did not involve semantic information, this area may also be involved in converging lexical information without semantics (Sandak et al., 2004). A study in Korean native speakers reported that the angular gyrus pathway was recruited for reading Hangul (phonogram) compared to Hanja (logogram) scripts (Cho et al., 2014), and related the function of this area to orthography-phonology conversion process.

If it were the case that the same processes are used throughout training and for all item types, namely the serial grapheme-to-phoneme conversion-and-compilation, then we would expect the training to affect all three conditions to the same degree since the radicals used in the three conditions were all the same with the same number of repetitions. At the behavioural level, we observed an interaction effect of Test day and Condition, showing faster RT to the Trained condition compared to the Recombined and the Novel condition that became more pronounced on Days 5 and 30. On the imaging level, we saw a qualitative difference, with more engagement of the bilateral angular and left middle temporal gyrus for the Trained condition on Day 5. Thus it is more likely that a qualitative change in decoding process occurred after repeated training and time to consolidate.

In his self-teaching hypothesis, Share suggests that with each successful decoding experience of an unfamiliar word, children have the opportunity to learn word-specific orthographic information (Share, 1995). Ziegler and colleagues further implemented this idea in their phonological decoding self-teaching model (Ziegler et al., 2014). Here the model assumes that the system already has knowledge about a cohort of phonological lexicons. Once the basic grapheme-phoneme rule is taught to the system, the system can decode the unfamiliar letter strings according to the grapheme-phoneme rule. If the output of the compiled phonemes matches with an existing phonological lexicon, then a direct mapping between the whole orthography and the corresponding phonology is initiated. This model fits nicely with the current experiment where the participants were taught the grapheme (radical)-phoneme rule at the start of the experiment. Initially, decoding took place serially. The compiled phonemes can be mapped to their existing phonological knowledge, and the experience of decoding the two-character words led to the emergence of holistic pseudoword-level orthographic and

phonological representations, which then could be used to decode once the representations were stable enough after some repetitions. Mapping of two radical combinations (CV) to a syllable and the link between orthography and phonology at the sub-lexical (character/syllable) level prompted faster reading time for the pseudowords in the Recombined condition relative to the Novel condition.

Behavioural data comparing the first occurrence during test (block 1) and repeated within session occurrence (blocks 2–5) gives an interesting view. Within-session repetition of naming is known to reduce RTs, especially for lower-frequency words relative to high-frequency words (Katz et al., 2005). All three conditions showed a benefit of repetition within the same test session. This can be explained by the possibility that retrieval of the episodic memory of the previous trial was enough to execute the production in the repeated trials rather than converting grapheme/orthography to phoneme/phonology for reading aloud for each repeated occurrence. Indeed some participants made the same mistake on the same word (usually mispronunciation of a single radical), even though they had the chance to correct the mispronounced radical through exposure to other pseudowords with this mispronounced radical. Although this explanation is probable, it does not explain the interaction effect only found for Trained and Recombined conditions on Days 5 and 30. The difference between the first and the repeated occurrence within session did not change across Test days (between sessions) for the Novel condition, but speeded up for the Trained and Recombined conditions. Thus repetition benefit is not merely reflecting the episodic retrieval of the previous trial. A more likely explanation therefore is that (sub) lexical representations for characters/words in these two conditions emerged after training, but access to these codes may have weakened after the training had stopped. The first re-exposure to the trained character/pseudoword will rekindle the link between the visual input and the emerged orthographic and phonological lexical representations, leading to faster pronunciation on consecutive trials. This idea is very likely since speeding up from the first to the consecutive trials was most obvious on Day 30, where participants had not been exposed to the pseudowords for three weeks after Day 5. It may take some time to access stored (sub) lexical information initially if it is not used frequently, but once the link is reconfigured, then the presence of the (sub) lexical representation aids speeded decoding.

Correlation of behavioural performance (both accuracy and reaction times) across days revealed that good learners were good performers. Both measures were highly correlated between Day 1, Day 5 and Day 30. This finding is interesting in the light of developmental studies in reading (Ehri, 2014; Share, 1995), in which reading readiness at kindergarten was shown to be predictive of later reading skills in children (Duncan et al., 2007; Grissmer et al., 2010). No between-participants correlation on the behavioural performance and the degree of neural activation was found. Nevertheless, the pattern that was stronger on Day 5 for Trained pseudowords in the bilateral angular and left posterior middle temporal gyri was predictive of better performance three weeks after the training had finished (Day 30). What functional roles could these areas serve? A study that investigated later-literates found more grey matter in bilateral dorsal occipital areas, left supramarginal and superior temporal areas, and the angular and posterior middle temporal gyri in this population than in illiterates (Carreiras et al., 2009). This later-literate study furthermore reported stronger functional connectivity between the bilateral angular gyri during reading relative to object-naming, suggesting an important role for the angular gyrus in orthography-phonology correspondence. Thus, in our study, the level of activity of the angular and posterior middle temporal gyri during decoding on Day 5 may reflect the stability of the lexical representation that

directly associated orthography to phonology. If the level of activation in these areas reflected the strength of the lexical representation(s) for a given script input on Day 5, then those representations are probably still available at delayed test on Day 30, contributing to efficient reading.

We observed recruitment of the visual word form area (VWFA), an area in the left ventral occipitotemporal cortex/fusiform gyrus that is often reported to be engaged in visual perception of words and letters (Dehaene and Cohen, 2011; Dehaene et al., 2005; McCandliss et al., 2003). The activity level was greater for the Novel compared to the Trained condition (the area was a part of the huge clusters that entailed a large part of the posterior ventral cortex bilaterally) for both Days 5 and 30 (Supplementary Fig. S4a and b). This area has been reported to show a posterior to anterior gradient with increasing complex orthographic units (Vinckier et al., 2007) or familiarity with a specific sequence of letters (Binder et al., 2006). This area has also been shown to respond to auditorily presented words when the task is focused on the orthographic aspects of the word (Ludersdorfer et al., 2016). Some researchers also advocate the function of this area not as an orthographic representation area but rather as an interface to associated semantic and phonological representations (Price and Devlin, 2011). Over all, the area seems to code for representation of scripts, but the activity level is often reported to be higher for less familiar strings, such as pseudowords, low frequent words, unfamiliar as compared to trained logographs, etc (e.g. Graves et al. (2010), Kronbichler et al. (2007), Mechelli et al. (2003), Xue et al. (2006)). In a recent meta-analysis, Taylor et al. proposed an engagement-effort relationship correspondence that reflects an inverted U-shape of the brain signal (Taylor et al., 2013). They argued that if we do not have a mental representation of the input stimulus, the area will not respond, however, depending on the amount of effort needed to process the incoming information, the degree of response will differ, with higher signal change for more demanding stimulus such as pseudowords, infrequent words, and unfamiliar letter combinations compared to a more familiar input. Our data is in line with this proposal. The pseudowords in the Novel condition were comprised of Trained radicals, but the combination of radicals was new. This unfamiliarity to the combination could have led to increased engagement of this area.

With this study, in which participants actively pronounced the pseudowords during test, we cannot dissociate whether it is the orthographic-phonological mapping or the motor-related code that contributes most to efficient reading of frequently occurring words. We can only speculate from the activation increase in the posterior areas known to process audio-visual integration, that orthographic-phonological mapping is the more likely explanation. Future research using overt and covert reading manipulations, for instance, may lead to a better understanding of emergent holistic representations in different stages of overt reading (i.e., visual perception, phonological conversion, motor preparation, and motor execution). Furthermore, for the sake of less complexity, we left the semantic component out of our experimental manipulations. Although pseudowords have been observed to show patterns similar to real words if experienced multiple times, especially in the visual processing stream (Fisher et al., 2011), the existence of semantic representation may have a top-down effect on reading (Taylor et al., 2015; Twomey et al., 2011). One study reported that direct orthography-to-phonology conversion recruited the left supramarginal, posterior middle temporal, and fusiform gyri, whereas semantic processing involved the left middle/inferior temporal gyri, bilateral angular and precuneus/posterior cingulate gyri (Graves, et al., 2010). However, another study showed elevated activation levels in the posterior middle temporal gyrus for meaningful words (Simos et al., 2002). Thus, it is not yet clear to what extent semantic processing plays a role in

activating these areas during reading. Investigating behavioural and neural changes as a consequence of training on reading novel meaningful scripts would be an interesting next step for future research.

5. Conclusions

Repeated exposure influences learning to read novel scripts. We have shown that repeated exposures to a specific combination of radicals (letters) led to faster and more accurate pronunciation through the possible use of holistic decoding processes, and this was related to increased engagement of the bilateral angular/supramarginal and left middle temporal gyri in the brain, and decreased processing demands in the visual attentional, cognitive control, phonological working memory, and motor preparation areas. Activation level in the bilateral angular and left middle temporal gyri, which could be coding for holistic unit (both on word- and syllable-levels) representations, was predictive of efficiency in reading aloud the trained unfamiliar script after a no-training period of at least three weeks.

Acknowledgments

This work was supported by a grant from the Netherlands Organization for Scientific Research (NWO: 056-33-014). The authors would like to thank Paul Gaalman and Uriël Plönes for their technical support, Jiyoun Choi for pronunciation video clips, and Iske Bakker for useful suggestions on the manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.neuropsychologia.2016.05.014>.

References

- Ans, B., Carbonnel, S., Valdois, S., 1998. A connectionist multiple-trace memory model for polysyllabic word reading. *Psychol. Rev.* 105 (4), 678–723. <http://dx.doi.org/10.1037/0033-295x.105.4.678-723>.
- Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. *Trends Cogn. Sci.* 15 (11), 527–536. <http://dx.doi.org/10.1016/j.tics.2011.10.001>.
- Binder, J.R., McKiernan, K.A., Parsons, M.E., Westbury, C.F., Possing, E.T., Kaufman, J. N., Buchanan, L., 2003. Neural correlates of lexical access during visual word recognition. *J. Cogn. Neurosci.* 15 (3), 372–393. <http://dx.doi.org/10.1162/089892903321593108>.
- Binder, J.R., Medler, D.A., Westbury, C.F., Liebenthal, E., Buchanan, L., 2006. Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33 (2), 739–748. <http://dx.doi.org/10.1016/j.neuroimage.2006.06.053>.
- Brandt, D.J., Sommer, J., Krach, S., Bedenbender, J., Kircher, T., Paulus, F.M., Jansen, A., 2013. Test-retest reliability of fMRI brain activity during memory-encoding. *Front. Psychiatry* 4. <http://dx.doi.org/10.3389/fpsy.2013.00163> (Original Research).
- van den Boer, M., de Jong, P.F., 2015. Parallel and serial reading processes in children's word and nonword reading. *J. Educ. Psychol.* 107 (1), 141–151. <http://dx.doi.org/10.1037/a0037101>.
- Carreiras, M., Seghier, M.L., Baquero, S., Estevez, A., Lozano, A., Devlin, J.T., Price, C.J., 2009. An anatomical signature for literacy. *Nature* 461 (7266), 983–986. <http://dx.doi.org/10.1038/nature08461> http://www.nature.com/nature/journal/v461/n7266/supinfo/nature08461_S1.html.
- Cattinelli, I., Borghese, N.A., Gallucci, M., Paulesu, E., 2013. Reading the reading brain: a new meta-analysis of functional imaging data on reading. *J. Neurolinguistics* 26 (1), 214–238. <http://dx.doi.org/10.1016/j.jneuroling.2012.08.001>.
- Cho, Z.-H., Kim, N., Bae, S., Chi, J.-G., Park, C.-W., Ogawa, S., Kim, Y.-B., 2014. Neural substrates of Hanja (Logogram) and Hangul (Phonogram) character readings by functional magnetic resonance imaging. *J. Korean Med. Sci.* 29 (10), 1416–1424. <http://dx.doi.org/10.3346/jkms.2014.29.10.1416>.
- Coltheart, M., 2005. Modeling reading: the dual-route approach. In: Snowling, M., Hulme, C. (Eds.), *The Science of Reading: A Handbook*. Blackwell, Oxford.
- Coltheart, M., Curtis, B., Atkins, P., Haller, M., 1993. Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychol. Rev.* 100 (4), 589–608.
- Daselaar, S.M., Prince, S.E., Dennis, N.A., Hayes, S.M., Kim, H., Cabeza, R., 2009. Posterior midline and ventral parietal activity is associated with retrieval success and encoding failure. *Front. Hum. Neurosci.* 3. <http://dx.doi.org/10.3389/fnhum.09.013.2009> (Original Research).
- Davis, M.H., Di Betta, A.M., Macdonald, M.J.E., Gaskell, M.G., 2009. Learning and consolidation of novel spoken words. *J. Cogn. Neurosci.* 21 (4), 803–820. <http://dx.doi.org/10.1162/jocn.2009.21059>.
- Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. *Trends Cogn. Sci.* 15 (6), 254–262. <http://dx.doi.org/10.1016/j.tics.2011.04.003>.
- Dehaene, S., Cohen, L., Sigman, M., Vinckier, F., 2005. The neural code for written words: a proposal. *Trends Cognit. Sci.* 9 (7), 335–341. <http://dx.doi.org/10.1016/j.tics.2005.05.004>.
- Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Japel, C., 2007. School readiness and later achievement. *Dev. Psychol.* 43 (6), 1428–1446. <http://dx.doi.org/10.1037/0012-1649.43.6.1428>.
- Ehri, L.C., 2014. Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18 (1), 5–21. <http://dx.doi.org/10.1080/1088438.2013.819356>.
- Fisher, J.E., Cortes, C.R., Griego, J.A., Tagamets, M.A., 2011. Repetition of letter strings leads to activation of and connectivity with word-related regions. *Neuroimage*, 2839–2849. <http://dx.doi.org/10.1016/j.neuroimage.2011.09.047>.
- Forster, K.I., Chambers, S.M., 1973. Lexical access and naming time. *J. Verbal Learn. Verbal Behav.* 12 (6), 627–635. [http://dx.doi.org/10.1016/S0022-5371\(73\)80042-8](http://dx.doi.org/10.1016/S0022-5371(73)80042-8).
- Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11 (2), 127–138. <http://dx.doi.org/10.1038/nrn2787> http://www.nature.com/nrn/journal/v11/n2/supinfo/nrn2787_S1.html.
- Gow, D.W., 2012. The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. *Brain Lang.* 121 (3), 273–288. <http://dx.doi.org/10.1016/j.bandl.2012.03.005>.
- Graves, W.W., Desai, R., Humphries, C., Seidenberg, M.S., Binder, J.R., 2010. Neural systems for reading aloud: a multiparametric approach. *Cereb. Cortex* 20 (8), 1799–1815. <http://dx.doi.org/10.1093/cercor/bhp245>.
- Grissmer, D., Grimm, K.J., Aiyer, S.M., Murrain, W.M., Steele, J.S., 2010. Fine motor skills and early comprehension of the world: two new school readiness indicators. *Dev. Psychol.* 46 (5), 1008–1017. <http://dx.doi.org/10.1037/a0020104>.
- Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8 (5), 393–402.
- Indefrey, P., Levelt, W.J.M., 2004. The spatial and temporal signatures of word production components. *Cognition* 92 (1–2), 101–144. <http://dx.doi.org/10.1016/j.cognition.2002.06.001>.
- Ischebeck, A., Indefrey, P., Usui, N., Nose, I., Hellwig, F., Taira, M., 2004. Reading in a regular orthography: an fMRI study investigating the role of visual familiarity. *J. Cogn. Neurosci.* 16 (5), 727–741. <http://dx.doi.org/10.1162/089892904970708>.
- Jobard, G., Crivello, F., Tzourio-Mazoyer, N., 2003. Evaluation of the dual route theory of reading: a meta-analysis of 35 neuroimaging studies. *Neuroimage* 20 (2), 693–712. [http://dx.doi.org/10.1016/s1053-8119\(03\)00343-4](http://dx.doi.org/10.1016/s1053-8119(03)00343-4).
- Kastner, S., Ungerleider, L. G., 2000. Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* 23 (1), 315–341. <http://dx.doi.org/10.1146/annurev.neuro.23.1.315>.
- Katz, L., Lee, C.H., Tabor, W., Frost, S.J., Mencl, W.E., Sandak, R., Pugh, K.R., 2005. Behavioral and neurobiological effects of printed word repetition in lexical decision and naming. *Neuropsychologia* 43 (14), 2068–2083. <http://dx.doi.org/10.1016/j.neuropsychologia.2005.03.022>.
- Keene, O.N., 1995. The log transformation is special. *Stat. Med.* 14 (8), 811–819.
- Kronbichler, M., Bergmann, J., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., Wimmer, H., 2007. Taxi vs. Taksi: on orthographic word recognition in the left ventral occipitotemporal cortex. *J. Cogn. Neurosci.* 19 (10), 1584–1594. <http://dx.doi.org/10.1162/jocn.2007.19.10.1584>.
- Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* 9 (12), 920–933. <http://dx.doi.org/10.1038/nrn2532>.
- Ludersdorfer, P., Wimmer, H., Richlan, F., Schurz, M., Hutzler, F., Kronbichler, M., 2016. Left ventral occipitotemporal activation during orthographic and semantic processing of auditory words. *Neuroimage* 124, 834–842. <http://dx.doi.org/10.1016/j.neuroimage.2015.09.039>.
- Maloney, E., Risko, E.F., O'Malley, S., Besner, D., 2009. Tracking the transition from sublexical to lexical processing: On the creation of orthographic and phonological lexical representations. *Q. J. Exp. Psychol.* 62 (5), 858–867. <http://dx.doi.org/10.1080/17470210802578385>.
- McCandliss, B.D., Cohen, L., Dehaene, S., 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7 (7), 293–299. [http://dx.doi.org/10.1016/s1364-6613\(03\)00134-7](http://dx.doi.org/10.1016/s1364-6613(03)00134-7).
- Mechelli, A., Gorno-Tempini, M.L., Price, C.J., 2003. Neuroimaging Studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. *J. Cogn. Neurosci.* 15 (2), 260–271. <http://dx.doi.org/10.1162/089892903321208196>.
- Mei, L., Xue, G., Lu, Z.-L., He, Q., Zhang, M., Wei, M., Dong, Q., 2014. Artificial language training reveals the neural substrates underlying addressed and assembled phonologies. *PLoS One* 9 (3), e93548. <http://dx.doi.org/10.1371/journal.pone.0093548>.

- Milner, A.D., Goodale, M.A., 2008. Two visual systems re-viewed. *Neuropsychologia* 46 (3), 774–785. <http://dx.doi.org/10.1016/j.neuropsychologia.2007.10.005>.
- Mukherjee, C., White, H., Wuyts, M., 1998. *Econometrics and Data Analysis for Developing Countries*. Routledge.
- Paulesu, E., Frith, C.D., Frackowiak, R.S.J., 1993. The neural correlates of the verbal component of working memory. *Nature* 362 (6418), 342–345. <http://dx.doi.org/10.1038/362342a0>.
- Penny, W.D., Holmes, A.P., Friston, K.J., 2003. Random effects analysis. In: Frackowiak, R.S.J., Friston, K.J., Frith, C., Dolan, R., Price, C.J., Zeki, S., Ashburner, J., Penny, W.D. (Eds.), *Human Brain Function*, 2nd ed. Academic Press.
- Plaut, D.C., McClelland, J.L., Seidenberg, M.S., Patterson, K., 1996. Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychol. Rev.* 103 (1), 56–115. <http://dx.doi.org/10.1037/0033-295x.103.1.56>.
- Poser, B.A., Versluis, M.J., Hoogduin, J.M., Norris, D.G., 2006. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magn. Reson. Med.* 55 (6), 1227–1235. <http://dx.doi.org/10.1002/mrm.20900>.
- Price, C.J., 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62 (2), 816–847. <http://dx.doi.org/10.1016/j.neuroimage.2012.04.062>.
- Price, C.J., Devlin, J.T., 2011. The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci.* 15 (6), 246–253. <http://dx.doi.org/10.1016/j.tics.2011.04.001>.
- Price, C.J., Mechelli, A., 2005. Reading and reading disturbance. *Curr. Opin. Neurobiol.* 15 (2), 231–238. <http://dx.doi.org/10.1016/j.conb.2005.03.003>.
- Ratcliff, R., 1993. Methods for dealing with reaction time outliers. *Psychol. Bull.* 114 (3), 510–532. <http://dx.doi.org/10.1037/0033-2909.114.3.510>.
- Sandak, R., Mencl, W.E., Frost, S.J., Pugh, K.R., 2004. The neurobiological basis of skilled and impaired reading: recent findings and new directions. *Sci. Stud. Read.* 8 (3), 273–292. http://dx.doi.org/10.1207/s1532799xssr0803_6.
- Seidenberg, M.S., 2007. Connectionist models of word reading. In: Gaskell, M.G. (Ed.), *Oxford Handbook of Psycholinguistics*. Oxford University Press, Oxford, pp. 235–250.
- Seidenberg, M.S., McClelland, J.L., 1989. A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96 (4), 523–568.
- Share, D.L., 1995. Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition* 55 (2), 151–218. [http://dx.doi.org/10.1016/0010-0277\(94\)00645-2](http://dx.doi.org/10.1016/0010-0277(94)00645-2).
- Share, D.L., 2008. On the Anglocentricity of current reading research and practice: the perils of overreliance on an "outlier" orthography. *Psychol. Bull.* 134 (4), 584–615. <http://dx.doi.org/10.1037/0033-2909.134.4.584>.
- Simos, P.G., Breier, J.I., Fletcher, J.M., Foorman, B.R., Castillo, E.M., Papanicolaou, A.C., 2002. Brain mechanisms for reading words and pseudowords: an integrated approach. *Cereb. Cortex* 12 (3), 297–305. <http://dx.doi.org/10.1093/cercor/12.3.297>.
- Smith, E.E., Jonides, J., 1999. Storage and executive processes in the frontal lobes. *Science* 283 (5408), 1657–1661. <http://dx.doi.org/10.1126/science.283.5408.1657>.
- Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., Zoccolotti, P., 2005. Length effect in word naming in reading: role of reading experience and reading deficit in Italian readers. *Dev. Neuropsychol.* 27 (2), 217–235. http://dx.doi.org/10.1207/s15326942dn2702_2.
- Takashima, A., Bakker, I., van Hell, J.G., Janzen, G., McQueen, J.M., 2014a. Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *Neuroimage* 84, 265–278.
- Takashima, A., Nieuwenhuis, I.L.C., Jensen, O., Talamini, L.M., Rijpkema, M., Fernández, G., 2009. Shift from hippocampal to neocortical centered retrieval network with consolidation. *J. Neurosci.* 29 (32), 10087–10093.
- Takashima, A., Wagensveld, B., van Turenout, M., Zwitserlood, P., Hagoort, P., Verhoeven, L., 2014b. Training-induced neural plasticity in visual-word decoding and the role of syllables. *Neuropsychologia* 61 (0), 299–314. <http://dx.doi.org/10.1016/j.neuropsychologia.2014.06.017>.
- Taylor, J.S.H., Duff, F.J., Woollams, A.M., Monaghan, P., Ricketts, J., 2015. How word meaning influences word reading. *Curr. Dir. Psychol. Sci.* 24 (4), 322–328. <http://dx.doi.org/10.1177/0963721415574980>.
- Taylor, J.S.H., Rastle, K., Davis, M.H., 2013. Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychol. Bull.* 139 (4), 766–791. <http://dx.doi.org/10.1037/a0030266>.
- Taylor, J.S.H., Rastle, K., Davis, M.H., 2014. Interpreting response time effects in functional imaging studies. *Neuroimage* 99 (100), 419–433. <http://dx.doi.org/10.1016/j.neuroimage.2014.05.073>.
- Thielen, J.-W., Takashima, A., Rutters, F., Tendolcar, I., Fernández, G., 2015. Transient relay function of midline thalamic nuclei during long-term memory consolidation in humans. *Learn. Mem.* 22, 527–531.
- Twomey, T., Kawabata Duncan, K.J., Price, C.J., Devlin, J.T., 2011. Top-down modulation of ventral occipito-temporal responses during visual word recognition. *Neuroimage* 55 (3–3), 1242–1251. <http://dx.doi.org/10.1016/j.neuroimage.2011.01.001>.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J.P., Sigman, M., Cohen, L., 2007. Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55 (1), 143–156. <http://dx.doi.org/10.1016/j.neuron.2007.05.031>.
- Waters, G.S., Seidenberg, M.S., 1985. Spelling-sound effects in reading: time-course and decision criteria. *Mem. Cogn.* 13 (6), 557–572.
- Xue, G., Chen, C., Jin, Z., Dong, Q., 2006. Language experience shapes fusiform activation when processing a logographic artificial language: an fMRI training study. *Neuroimage* 31 (3), 1315–1326. <http://dx.doi.org/10.1016/j.neuroimage.2005.11.055>.
- Ziegler, J.C., Perry, C., Zorzi, M., 2014. Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philos. Trans. R. Soc. B: Biol. Sci.* 369 (1634), 20120397. <http://dx.doi.org/10.1098/rstb.2012.0397>.
- Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., Spinelli, D., 2005. Word length effect in early reading and in developmental dyslexia. *Brain Lang.* 93 (3), 369–373. <http://dx.doi.org/10.1016/j.bandl.2004.10.010>.