Research report

# Alphabetical knowledge from whole words training: effects of explicit instruction and implicit experience on learning script segmentation

T. Bitan[a,*], A. Karni[a,b]

[a]Department of Neurobiology, Brain Res., The Weizmann Institute of Science, Rehovot, 76100 Israel
[b]Brain-Behavior Research Center, Haifa University, Mt. Carmel, Haifa, 31905 Israel

## Abstract

We investigated the possibility that pattern segmentation skills, specifically, phonological decoding, evolve implicitly in adult readers given training in an artificial script. In this Morse-like script each phoneme was represented by 2–3 discrete symbols. Subjects were trained in five consecutive sessions, on reading six nonsense words using a forced choice task that required translating symbol strings to sound patterns written in Latin letters. Three training conditions were compared within subject in terms of the time-course of learning and the ability to generalize the acquired knowledge (transfer): alphabetical whole words with letter decoding instruction (Explicit); alphabetical whole words (Implicit), and non-alphabetical whole words (Arbitrary). In separate blocks in each training session, a visual-matching task was administered using the same stimuli. Our results show: (a) that while all three training conditions were equally effective in terms of magnitude and time-course of learning accurate translation, each training condition resulted in a different type of knowledge (i.e. differential transfer). (b) Declarative knowledge of letters evolved from training on whole words only in subjects with previous experience in Explicit training. However, even with declarative knowledge of the specific letters subjects did not develop general letter segmentation skills. (c) Contrary to the robust transfer of learning gains to different stimuli within a given task, there was no significant transfer across tasks indicating that the locus of learning was task dependent. Altogether our results suggest that even given explicit letter instruction, training on word decoding may result in letter recognition rather than in alphabetic segmentation skills.
© 2002 Elsevier Science B.V. All rights reserved.

## 1. Introduction

A tenet of most reading theories is that in the process of reading alphabetical scripts, visually presented patterns can be segmented and decoded into letters and sounds, i.e. the acquisition of reading skills involves learning of alphabetical rules [11,21,26,57]. The acquired alphabetical knowledge is partly declarative in nature with a conscious knowledge of the way each letter is mapped to the corresponding sound (grapheme–phoneme correspondence) in all skilled readers. On the other hand, this knowledge is acquired in a procedural manner, namely as a function of repeated experience, in a time course charac-

teristic of rote learning and skill acquisition [35,59]. Thus, learning to read is a gradual process that requires numerous repetitions to achieve skillful performance. A critical question is whether alphabetical knowledge, namely grapheme decoding can be acquired implicitly, i.e. from training on whole word recognition. As the alphabetical rules are rather complex and synthetic it is not unreasonable to expect that explicit instruction on grapheme–phoneme correspondence is essential to learning. Alternatively, given that whole word training results in improved reading [4,23,63,64], training on the correspondence of whole word patterns to sound patterns may be sufficient for extracting alphabetical knowledge in an implicit (procedural) manner. It is not clear, therefore, what is being learnt through training on whole words: implicit knowledge of the alphabetical rules or improved

*Corresponding author. Tel.: +972-54-490-012.
E-mail address: tali.bitan@weizmann.ac.il (T. Bitan).

recognition of the patterns of specific words. A related question of much practical importance concerns the relative effectiveness of implicit vs. explicit training modes.

Declarative knowledge (of facts and events) is typically distinguished from procedural knowledge by being accessible to awareness, being often acquired through a single experience and involving medio-temporal brain regions. Procedural skill learning, on the other hand, is evident by improvement in the performance of a given task; it is not necessarily conscious, requires multiple repetitions and is subserved by different cortical areas independent of medio-temporal processing [3,35,45,50,59–61]. While the terms 'declarative' and 'procedural' are theory based and refer to distinct learning mechanisms, the terms 'explicit' and 'implicit' learning can be operationally defined, and thus will be used here, although the dichotomies are not identical [17]. Implicit learning is contingent on repeated experience, but the resulting knowledge may be different from the intended goal of the learning experience, although strongly constrained by task and specific parameters of the training experience [35]. The nature of the resulting knowledge can be inferred only indirectly via performance changes.

A number of paradigms provide data suggesting that complex rules can be learned implicitly (e.g. Artificial Grammar Learning [50], Serial Reaction Time [45], dynamic control of computerized input–output systems [3], and predictive judgments of events according to a probabilistic rule [52]). Implicit learning of complex rules was shown not only in normal healthy individuals but also in amnesic patients in whom declarative memory is disrupted [40,41,60]. In both groups the data indicated dissociation between (procedural) improvement in the performance of the task, presumably related to implicit knowledge of the rules, and explicit knowledge of the underlying rules. Thus, the data from these nonlinguistic paradigms provide indirect support for the hypotheses that knowledge of alphabetical rules can evolve implicitly from training on whole word patterns.

In Artificial Grammar Learning (AGL) subjects are presented with letter strings and required to memorize them. Subsequently, during the test phase, subjects are told that the strings were constructed according to specific grammatical rules, and are required to judge the grammaticality of novel strings. Normal subjects and amnesic patients perform the grammaticality judgement task above chance level, even when the items are composed of different letters (preserving the grammar rules), while they are unable to explicitly describe the grammar rules that underlie their judgements. This finding was taken as an indication for implicit learning of the abstract grammar rules [40,41,50]. However, recent studies have suggested that learning of surface features, rather than rule abstraction processes, can account for the improved performance in these tasks [8,15,47,53].

An influential model on reading instruction within the education context is the dual route model for visual word recognition ([18,19] but see [30,48,55] for an alternative connectionist model). The dual route model assumes two distinct routes from the visual form of the word to its meaning and sound: a lexical route, constituting a direct link from the orthographic visual form of the word to its meaning and sound, and a phonological route that requires graphemic and phonemic segmentation and grapheme–phoneme conversion. Others, applying a developmental approach to reading, argue that in skilled readers lexical and phonological processes form a continuum, wherein the phonological decoding mechanism becomes sensitive to the context [57,65,46]. Thus, the size of the relevant units for decoding becomes larger than one-grapheme-to-one-phoneme However, these models have contradicting assumptions on the role that phonological and lexical processes play in the different stages of development, assumptions that give rise to different prediction on the appropriate and most effective reading instruction. For example, Share [57] argues that explicit phonological instruction is essential for the beginning reader to gain alphabetical knowledge and to acquire skilled reading. This prediction is supported by studies that show that exposure to alphabetical orthography does not spontaneously induce the discovery of the alphabetic principle in children [9,11,13,22,32,34,56], and by studies showing that explicit instruction on phonological decoding enhanced reading acquisition ([2,4,12,24,25,67] and see [57] for review). In contrast, the models of Perfetti [46] and Van Orden [65] suggest that explicit phonemic knowledge is not a prerequisite of reading acquisition, and that the phonemic knowledge and rule-like behavior implicitly emerge from the correlational structure of trained words. This view is supported by studies suggesting that grapheme–phoneme correspondences were learned by young beginning readers from training on whole words [23,63,64]. However, in natural settings additional factors may have critically contributed to the children's acquired knowledge, e.g. knowledge of letter names, spelling exercises and explicit alphabetical instructions outside the classroom [63].

It is not clear, therefore, what is learned from whole word training. Whole word training may result in the formation of word-specific (orthographic) representations, or it may induce procedural learning of segmentation and graphemes to phonemes conversion. The acquired procedures of segmentation and conversion may be implicit (e.g. AGL [40,41,50]), or it may subsequently evolve into declarative knowledge of alphabetical rules as, for example, in the bottom-up model proposed by Sun et al. [62]. Declarative knowledge may be advantageous as it may speed up the extraction of conversion rules, and facilitate transfer by making the knowledge more flexible, i.e. applicable in different novel conditions [10,62]. Byrne and Carroll [10] found that adult subjects did not learn the mapping of phonetic regularities onto artificial letters without explicit instruction of the mapping. On the other hand, declarative knowledge, which requires extensive

working memory resources, may interfere with the process of proceduralization and automatization [51,62], thus reducing the benefits of training in terms of speed and accuracy [1,35]. As suggested by the reading stages model of Frith [26], children receiving little instruction in letter–sound correspondence are expected to skip the alphabetical reading stage, and leap directly to the application of word-specific orthographic representations, which is supposedly an advanced stage of skilled reading.

Evidence from studies of nonlinguistic learning suggests that implicit learning has an advantage over explicit learning of very complex rules [43]. In some studies general explicit instructions, such as encouraging subjects to discover underlying rules, were found to be helpful only when the rules were salient, and could be easily discovered [51]. For complex rules general instructions had either no effect or were found to be detrimental to learning as reflected in gains in task performance [3]. Other studies, however, suggest that some complex rules may only be learnt explicitly [16]. Dominey et al. [20] found that although a sequence of movements was learned in an explicit as well as an implicit manner, only subjects that received explicit instructions learned the abstract rule underlying the sequence, and transferred learning gains to sequences composed of different surface components.

The relative efficacy of letter versus whole word training was studied in a number of studies using artificial orthographies or novel script [4,32,6,7]. Bishop [4] trained adult subjects in reading Arabic script and tested their transfer to novel words. Although letter training was more effective than word training for the transfer to novel words, word training did have some value for transfer compared to a control group. Jeffrey and Samuels [32] trained preliterate children on reading either single letters or two-letter words written in an artificial orthography and found that letter training was better than word training both in the initial encounter with the transfer words, and in the number of repetitions required to learn them. Contrary to Bishop's [4] results the word training group was not different from the control group given no training on the artificial script. There was no difference between the groups in the number of repetitions required for learning the words or letters during the training phase. However, different units were used for training in the two groups making the comparison of learning curves difficult. Moreover, both groups were given prior practice on target phonemes. Brooks and Miller [7] trained adults to read non-words written in an artificial script, and found that the reading of non-words was slower after explicit instruction on the individual component letters, compared to reading without such instruction, suggesting a disadvantage for explicit training. However, it is important to note that in the Brooks and Miller study, the pattern of transfer results to new words composed of the trained letters, did not indicate that subjects learned the individual letters in any of the training conditions.

One methodological problem with studying reading acquisition in adult subjects is that their long reading experience with alphabetical systems may predispose them to apply their word segmentation skills to the novel orthography. In the current study we attempted to minimize this effect by using a Morse-like script in which a sequence of more than one symbol represents one letter. Thus, none of the standard spatial cues for word segmentation into letters were retained. The learning of the alphabetic code would, therefore, entail the segmentation of the symbol string into letters as well as the mapping of letters to sounds. By introducing this and additional changes to the basic paradigm of Brooks and Miller we here show that learning of letters does not necessarily evolve implicitly from training on whole words, even in normal adult readers with extensive experience in alphabetical systems, and that explicit letter instruction may be crucial. On the other hand, our results show that explicit knowledge of letters may result in effective, but highly specific, letter recognition rather than in a general word segmentation ability.

## 2. Materials and methods

### 2.1. Subjects

Nine adult volunteers, aged between 17 and 28, with normal linguistic and reading skills participated in the experiment and were paid for their time. The group consisted of five males and four females, eight right handed and one left handed. Each subject participated in all three training conditions.

### 2.2. Stimuli

The training stimuli constituted an artificial Morse-like script, in which each grapheme was represented by a sequence of 2–3 symbols, and three symbols in different orders were used to compose all graphemes (i.e. P: |^* L: *^| N: ^|^ O: *| etc.) (Table 1). The graphemes were composed according to the following rules: (1) A consonant was represented by three symbols and a vowel was represented by two symbols. (2) A given symbol could not appear in two adjacent locations within a grapheme (e.g. ^|^ was legitimate, while ^^| was not). Three sets of symbols were used, one for each set of trained non-words, with symbol sets balanced across training condition. The number of graphemes with symmetrical patterns (e.g. ^|^,<)<, etc.) was equal in all sets of graphemes (Table 1).

The training stimuli consisted of three sets of six non-words. The non-words were composed of two consonants (C) and one vowel (V), and each of the three training sets contained all phonological patterns: CVC, VCC and CCV. Four consonants and two vowels were used to compose all non-words in a given set, with each element repeating three times (e.g.: PON, LOP, LAT, ONT, PTA and NAL). Each non-word was represented in the novel script using

Table 1
Correspondence of phonemes (in Latin letters) to graphemes using the Morse-like artificial script

| Set 1 | | Set 2 | | Set 3 | |
|---|---|---|---|---|---|
| Phonemes | Graphemes | Phonemes | Graphemes | Phonemes | Graphemes |
| l | *^\| | d | )<- | b | ~>~ |
| p | \|^* | f | )-< | r | ~>\| |
| n | ^\|^ | s | -<- | m | >~\| |
| t | *^* | g | )<) | k | >\|> |
| a | ^* | i | -) | u | ~\| |
| o | *\| | o | )- | i | >\| |

The three sets of phonemes were each composed of three unique symbols with a string of three symbols per consonant and two symbols per vowel. Each of the three phoneme–grapheme sets was used in one training condition, with stimulus sets balanced across conditions. In the arbitrary condition there was no consistent correspondence between the symbol string grapheme and a given phoneme, across words.

two different transformations: an alphabetical transformation, in which each phoneme consistently corresponded to a grapheme (e.g. PON=|^**|^|^; LOP=*^|*||^*), and an arbitrary transformation, in which phoneme to grapheme correspondence differed across words (e.g.: PON= |^**|^|^; LOP=^|^|^**|). Thus, the symbol strings in the arbitrary condition could only be read as pictographs (in similarity to Japanese Kanji).

### 2.3. Apparatus

The stimuli were presented on a 17-inch 60 Hz. PC screen, with each item subtending 1° viewing angle, from a viewing distance of 60 cm.. Stimulus presentation as well as the recording of responses (using a standard three button mouse), was controlled by 'Psy', a psychophysical measurements program, operating on LINUX environment (Y. Bonneh, 1998).

### 2.4. Experimental procedure

Each subject was trained in three training conditions, similar to those used in the Brooks and Miller [7] study: 'Implicit'—training on alphabetical non-words; 'Explicit'—training on alphabetical non-words following instruction on the grapheme–phoneme correspondence; and 'Arbitrary'—training on non-alphabetical non-words with no consistent mapping of graphemes to phonemes (pictographs). In contrast to the Brooks and Miller study, training conditions were administered serially rather than in parallel, with the order of conditions balanced across subjects, in order to have better control for interactions between conditions. Each subject was trained with a different set of stimuli (non-words and symbols) in each condition, with the sets of non-words and symbols balanced across training conditions.

Training was given on two tasks, using the same stimulus set in both tasks. In the first session of each training condition the Translation task was presented with a 'whole-word instruction' block. In this block the subject was presented with each target non-word in novel script with its corresponding translation to Latin letters below

(Fig. 1a). Each stimulus was presented for 2000 ms and subjects were instructed to read it aloud and memorize the association. The non-words appeared in a fixed order that repeated for five times (total of 30 trials). A 'letter-instruction' block was used only in the explicit training condition, and consisted of 30 trials in which the individual letter patterns in the new script were presented together with their corresponding Latin letter translation, each pair for 2000 ms. Subjects were required to pronounce the related phoneme and memorize the association. The letters appeared in a fixed order that repeated for five times.

In the training blocks of the translation task, each target non-word with a translation to Latin letters presented below appeared for 800 ms (Fig. 1a). The subject's task was to indicate, for each test item, whether the translation to Latin was correct or not, by pressing one of two keys (two alternative forced choice). Auditory feedback was given for errors. Each block consisted of 60 trials.

The second training task was a Visual matching task, in which two symbol strings (from the same set of non-words used in the translation task) were briefly presented in sequence, each followed by a patterned mask. Fig. 1b depicts the structure of each trial in the visual matching task: stimulus 1 (80 ms), blank screen (150 ms), mask (80 ms), blank screen (100 ms), stimulus 2 (80 ms), blank screen (150 ms), mask (80 ms). A key press was used to indicate whether the two strings were identical or different (two alternative forced choice). Auditory feedback was given for errors. Each block consisted of 60 trials. Only six subjects participated in this task. In the instruction blocks as well as in both training tasks the onset of each trial was determined by the viewer.

In each training condition subjects were given training on five daily session, spaced 1–3 days apart (Fig. 2). In the first session of each training condition subjects were first tested on the visual-matching task (one block). In the explicit, but not in the implicit or arbitrary training conditions, a 'letter-instruction' block was given next. The following order of tasks was identical for all conditions: one block of 'whole-word instruction' for the translation task, a second block of the visual-matching task, five
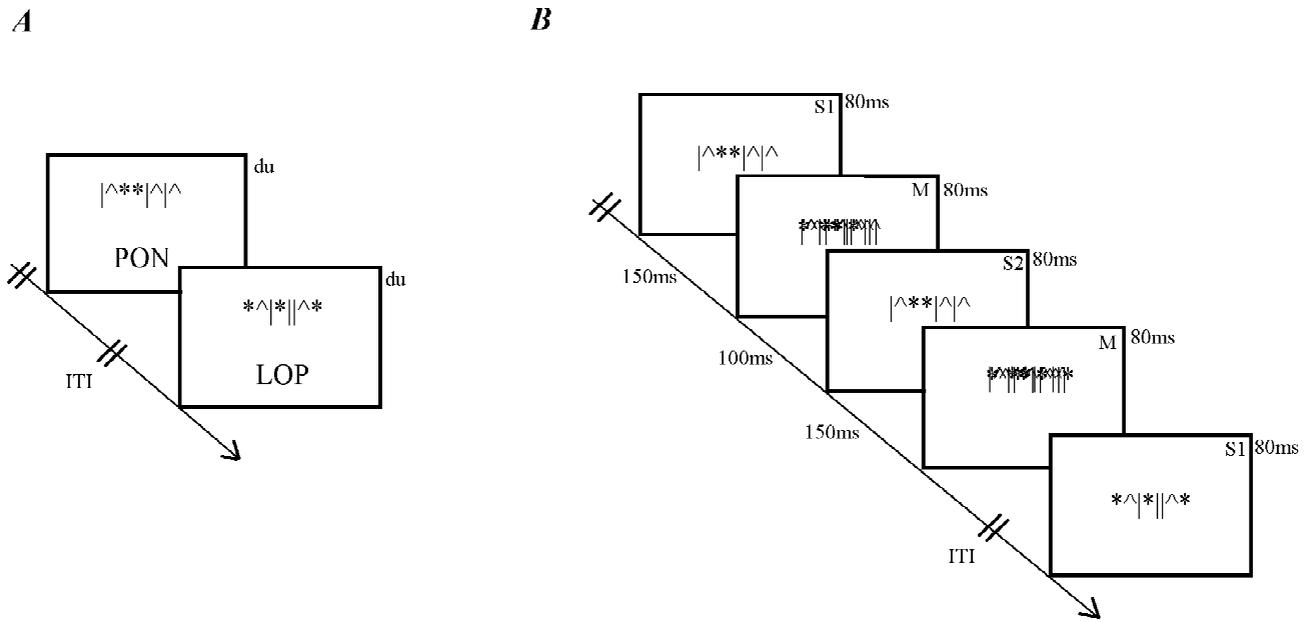
**A**

**B**



Fig. 1. Training tasks: temporal sequence of displays for the translation (A) and visual matching (B) tasks. (A) Two consecutive trials in the translation task. In the instruction blocks du=2000 ms, and in the training blocks du=800 ms. (B) Each trial in the visual matching task consists of two stimulus (S) and mask (M) pairs. ITI (Intertrial interval) in both tasks was controlled by the participant (by using a key press to initiate each successive trial).

blocks of the translation task, and finally a third block of the visual-matching task. Thus the explicit training condition differed from the implicit and arbitrary training conditions only in the structure of the first training session. The structure of training sessions 2–5 was identical in all conditions. Subjects first received two blocks of the visual-matching task, followed by five blocks of training in the translation task, and finally two additional blocks of the visual-matching task.

At the end of the five training sessions in each training condition the transfer of learning gains to novel stimuli was tested, in order to probe the level of representation at which learning occurred [36] (Fig. 2). Three transfer tests were administered, six non-words in each test. The 'word transfer' test consisted of new non-words composed of the original letters, and written with the same set of symbols; (i.e. after training on: PON=$|^{\wedge}**|^{\wedge}|^{\wedge}$ testing the transfer to: NOP=$^{\wedge}|^{\wedge}*||^{\wedge}*$). The 'letter transfer' test consisted of new non-words composed of new letters written with the same set of symbols; (i.e. after training on: LOP=$*^{\wedge}|*||^{\wedge}*$ testing the transfer to: KIB=$|^{\wedge}||^{\wedge}*|*$). A comparison of 'word transfer' to 'letter transfer' was planned to enable

the differentiation between learning gains that were due to learning of the letters per se, and learning gains related to the familiarity with the elementary symbols and general features of the task. A third transfer test was the 'symbol transfer' test in which the original non-words were written using a new set of symbols, with consistent mapping between the sets of symbols. Thus, the pattern of symbol repetitions and internal symmetries within each string was preserved (i.e. after training on LOP=$*^{\wedge}|*||^{\wedge}*$ testing the transfer to LOP=$\sim|>\sim>>|\sim$). This condition was designed to test whether the pattern of symmetries and repetitions in the symbol sequence were learned.

Each of the three transfer tests was administered in a separate session with the order of transfer tests fixed for all subjects ('word transfer'; 'symbol transfer'; 'letter transfer'). In each of the three transfer sessions subjects first performed a single block of the visual matching task, followed by three blocks of the translation task using the originally trained non-words. The level of performance with the trained stimuli served as the reference for calculating the transfer of performance gains to the transfer stimuli. Subjects then performed a visual-matching task
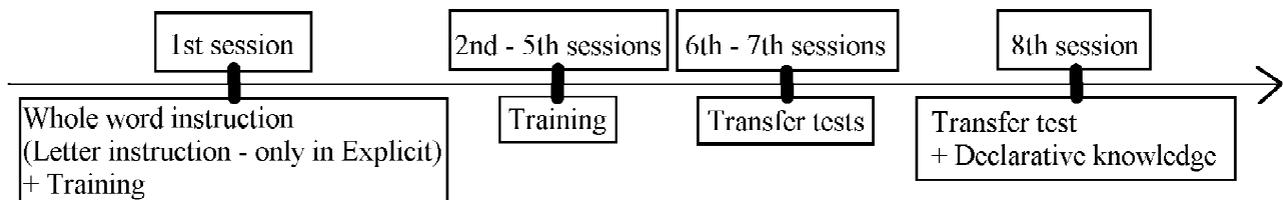


Fig. 2. Structure of one training condition. In the overall study design, for each participant the eight consecutive sessions were administered in each of the three training conditions.

block with the transfer stimuli, and a 'whole-word instruction' block in which the transfer stimuli and their Latin letter equivalents were presented. No 'letter-instruction' was given during the transfer sessions. A second block of the visual-matching task, followed by five blocks of the translation task and a final block of the visual-matching task were then administered, all using the transfer stimuli.

A declerative knowledge test (in a pen and paper format) was administered at the last (8th) session of each training condition. Subjects were required to write the appropriate translation of symbol strings to Latin letters. The symbol strings included in the test were: (a) the six trained non-words; (b) the six component letters of the trained non-words; (c) six novel non-words composed of the original letters.

The Masked translation task, was a time-limited version of the translation task that was administered using all three training sets in a separate session at the end of all three training conditions. In this version of the translation task the Latin letter translation was presented after, rather than simultaneously with the target symbol string. Subjects were required to indicate whether the translation was correct by pressing one of two keys (two alternative forced choice). Eleven test blocks, of 60 trials each, were given for each stimulus set, with decreasing duration of exposure: 800, 800, 600, 400, 200, 150, 100, 80, 60, 40 and 20 ms. Each target symbol string was immediately followed by a Latin letter string (800 ms), which served also as a mask. All three sets of trained stimuli were given serially in the same session. For each subject the order of stimulus sets was identical to the order in which they were administered during training.

Two subjects were tested 8 months after the termination of training for long-term retention of the training effects. Both subjects were retested on the stimulus sets used in the explicit and the arbitrary conditions.

## 3. Results

### 3.1. Learning curves

Clear learning effects were evident in the translation task for all subjects, in all three training methods. Subjects improved from an average performance of 60% to 93% of correct trials (Fig. 3). In all conditions $d'$ has increased from an average of 1.14 in the first session to 3.31 on the fifth session, indicating this was a true learning process rather than a change in the criterion of response [28]. Learning was gradual and incremental over the sessions. The learning curves were steepest initially, and shallower on late sessions before reaching an asymptote, and hence had a good fit to power functions ($R^2 = 0.90$–$0.94$ in the different conditions). A $3 \times 5 \times 5$ repeated measures ANOVA, with training condition, session and block as within subject variables, and accuracy as the dependent variable revealed significant main effects for session [$F(4,4) = 133.81$, $P < 0.001$], and for block [$F(4,4) = 21.91$ $P < 0.01$], but no effect of training condition [$F(2,6) = 1.21$]. These results indicate that there was significant learning within sessions (between blocks) and across sessions, but no difference between training conditions. A comparison of the slopes of the power functions fitted to the learning curves in a one-way ANOVA showed no effect of training conditions [$F(2,6) < 1$], further indicating there was no difference between the learning curves of the three training conditions.

Although subjects were not instructed to respond as fast as possible, an analysis of reaction time showed a trend of improvement in all three conditions. Individual subjects' performance speeds either remained constant or showed monotonic improvement, with, however, no speed accuracy tradeoff in any training condition, indicating a skill learning process [29]. A one-way ANOVA on the linear
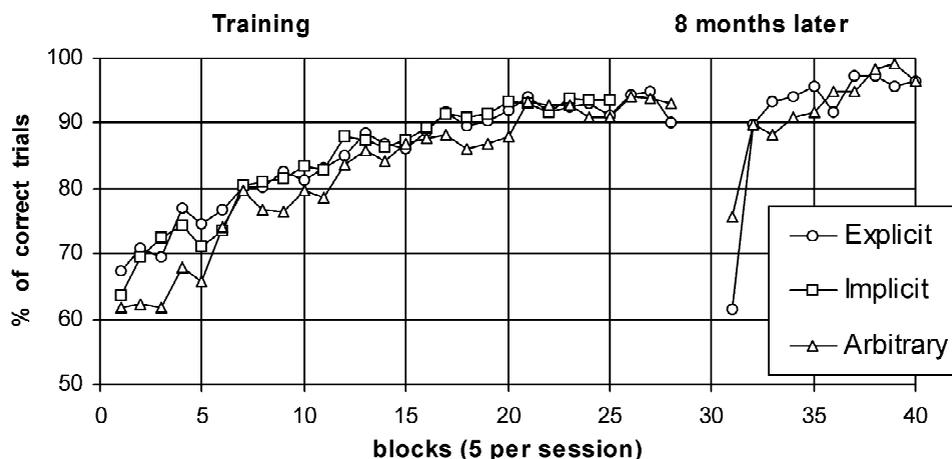


Fig. 3. Learning curves of the three training conditions of the translation task (average across subjects). Performance accuracy in the five training sessions (five blocks each) (9 Ss) and after 8 months (2 Ss) is shown. In all three conditions the standard errors were less than 10%.

slope coefficients of reaction times in the three training conditions showed no significant effect of training condition [$F(2,6)=2.053$].

The effects of training for both the explicit and the arbitrary training conditions were preserved for a long time, as indicated by the performance of two subjects 8 months after training. A single block was sufficient to regain the same level of performance that was attained by the end of training (Fig. 3).

### 3.2. Transfer results

Although there was no difference between the learning curves of the different training conditions, a differential pattern of transfer was found, indicating that learning occurred at different levels of representations in the three conditions (Fig. 4a). Transfer was calculated, for each transfer condition, as the gain in performance in the first transfer block compared to the first training block, as a proportion of the gain in performance resulting from training. Following training in the explicit condition about 50% of the gains transferred to novel words composed of the original letters ('word transfer'), while the transfer to novel words composed of novel letters ('letter transfer') and to the trained words composed of novel symbols ('symbol transfer') were negative. In the arbitrary and in the implicit training conditions, however, 'symbol transfer' was above 50%, while both 'word' and 'letter transfer' were lower (10–20%) (Fig. 4a).

In a 3×3 ANOVA on the above measure for transfer, with training condition and transfer condition as within subject factors, none of the effects were significant. One subject was excluded from analyses of transfer results due to missing data, and another due to outlying values. One-tailed paired *t*-tests performed between individual conditions in order to test planned comparisons showed that 'word transfer' was significantly higher than 'letter transfer' [$t(6)=2.32$, $P<0.05$] following training in the explicit condition, while in the arbitrary and in the implicit training conditions, there was no difference between 'word' and 'letter transfer' [$t(6)<1$]. This indicates that only in the explicit condition, but not in the arbitrary and implicit training conditions, learning involved processing of letter representations. In addition, 'word transfer' after training in the explicit condition was significantly higher than after training in the arbitrary condition [$t(6)=2.47$, $P<0.05$], reflecting an advantage of alphabetical training over whole-word training in encountering new words composed of trained phonemes, which in the explicit condition correspond to trained letters.

The transfer effects, however, were dynamic even within the limited time frame of a single transfer session, with block by block improvement. A significant advantage for 'word transfer' over 'letter transfer' was found in the explicit training condition [$t(7)=2.5$, $P<0.05$] but not in
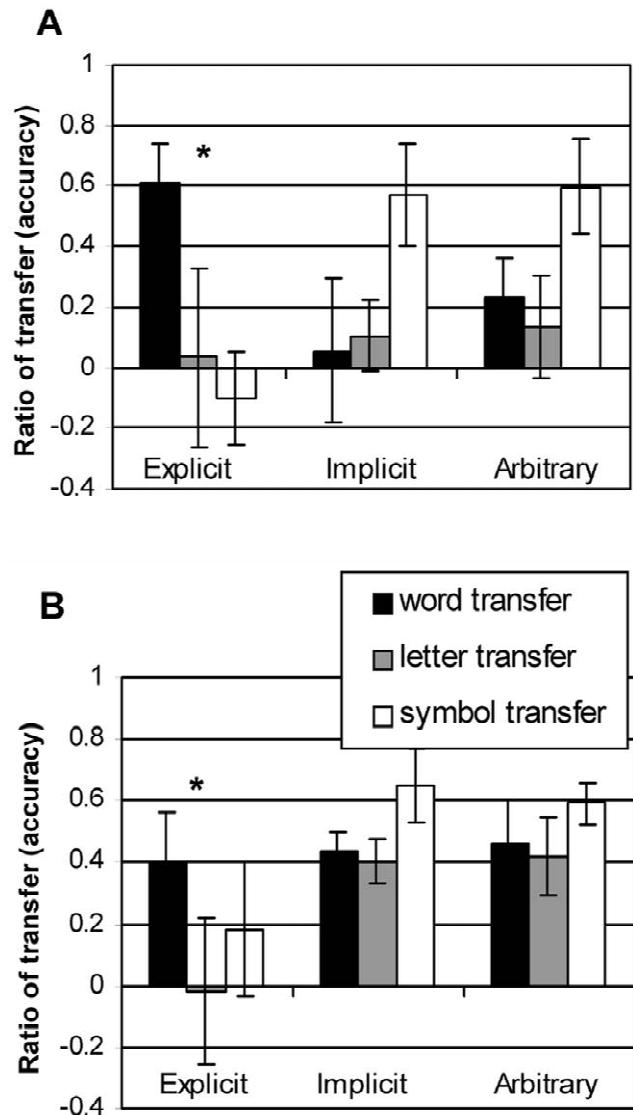


Fig. 4. Transfer results in the translation task (7 Ss). The measure for transfer was calculated as: (transfer–beginning of training)/(end of training–beginning of training), with transfer ratio of 1.0 indicating full transfer. (A) Transfer ratios for the first transfer block, in each transfer condition. (B) Transfer ratios for the mean performance in the session (across blocks). * Indicates significant Word-transfer Letter-transfer difference.

the arbitrary or in the implicit training conditions, even when transfer was calculated as the average over the entire session (Fig. 4b). In contrast to the results of the first transfer block, however, the mean of the entire transfer session did not show an advantage for the explicit compared to the arbitrary training condition in terms of the amount of 'word transfer' (both are around 40%). Moreover, as can be seen in Fig. 4b, the advantage of 'word' over 'letter transfer' in the explicit condition was the result of a clear reduction in 'letter transfer' in the explicit condition. Thus our results suggest that learning the

specific individual letters significantly impeded the learning of new letter systems.

The slope of the power function fitted to the 'word transfer' session was not significantly different in the explicit compared to the arbitrary training condition, again indicating that the time course of learning may be similar, although presumably different levels of representations were involved in the two conditions.

### 3.3. Effect of the position of training condition in the sequence

The sequence in which the training conditions were given to each individual, regardless of training method, had a highly significant effect on the time course of learning (Fig. 5). The average learning curve of the first, second and third training experiences across subjects had a good fit to power function ($R^2$ between 0.90 and 0.94). A one-way ANOVA on the intercepts of these functions revealed a significant effect of the sequence of training conditions [$F_{(2,7)}=8.53$, $P<0.05$], indicating that the starting point of learning was higher as a function of previous experience in other training conditions. Furthermore, a one-way ANOVA on the slopes of these functions also revealed a significant effect of sequence of training conditions [$F_{(2,7)}=7.4$, $P<0.05$], with the slopes becoming less steep with experience in other training conditions. (The contribution of the sequence of training conditions could not be analyzed together with the effect of training method because the design of the experiment resulted in six groups of different sequences with only nine subjects altogether). These findings indicate a considerable amount of transfer between conditions irrespective of the differences in the stimuli and training methods used in each training condition.

### 3.4. Visual matching task

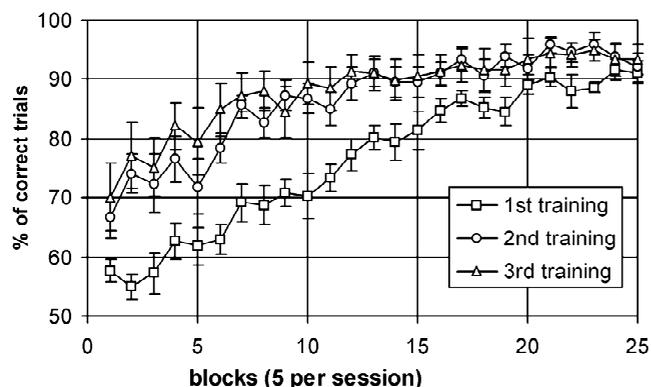Learning in the visual matching task was less effective than learning in the translation task in terms of gains in accuracy. The maximum level of performance, at the end of the fifth session was 82% (average across training conditions), compared to 93% in the translation task. In contrast to the translation task, the learning curves for the visual matching task did not fit a power function ($R^2=0.23$–0.32) and were better fitted by a linear function ($R^2=0.52$–0.68) (Fig. 6a). In contrast to the translation task, in the visual matching task there was no effect of the sequence in which the training conditions were administered. In a GLM analysis on the accuracy of performance, with the sequence of training conditions and the blocks as independent variables, the effect of the sequence was not significant [$F_{(2,10)}<1$]. The two tasks also differed in terms of the transfer of learning gains to other stimuli after training. In the visual matching task, the 'symbol transfer' condition showed the lowest amount of transfer in all training conditions (Fig. 6b) suggesting that learning in the visual matching task involved representations of specific
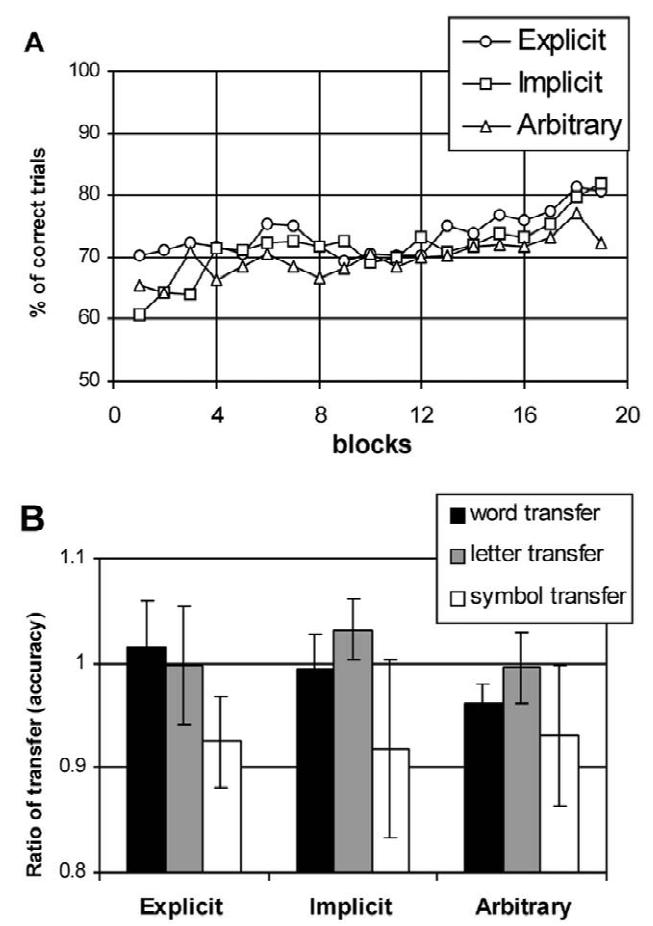


Fig. 6. Visual matching task (6 Ss). (A) Learning curves of the three training conditions, over five training sessions (three blocks were given in the initial session and four blocks in sessions 2–5). (B) Transfer results. The measure for transfer was the proportion of the average performance in the transfer session (across blocks), from the average performance in the last training session. A transfer measure that combines both the first and last training sessions could not be used because of the small learning gains during training.



Fig. 5. Learning curves in the translation task averaged by the order of training experience (irrespective of training method).

symbols. A significant effect of transfer condition [$F(2,4)=7.97$, $P<0.05$], was found in a 3×3 repeated measures ANOVA, with training condition and transfer condition as within subject factors and the ratio between performance in the transfer session and performance in the last training sessions as a dependent variable. (Transfer ratio was not normalized to the gains of training, as in the translation task, since learning gains in the visual matching task were in some cases negative.) Thus, the difference in the effect of the sequence of training conditions on the accuracy of the translation and the visual-matching task, together with the difference in the transfer pattern in these tasks suggest that there was no transfer between tasks even though the same stimuli were used.

### 3.5. Declarative data

The test for declarative knowledge showed that significant alphabetical knowledge of the trained stimuli was gained in the explicit, and to a much lesser degree, in the implicit conditions (Fig. 7a). In the explicit condition the average score for the original letters was 4.6 of 6 items [$t(7)=6.33$, $P<0.001$], and the score for the new words composed of the original letters was 3.4 of 6 [$t(5)=3.48$, $P<0.05$]. In the implicit condition an average of 2.4 out of 6 letters were recognized, significantly above zero [$t(9)=3.44$, $P<0.01$], but also significantly less than in the explicit condition [in one tailed paired $t$-test $t(6)=2.17$ $P<0.05$]. The sequence of training conditions had no significant effect on the performance of the 'Original Words Declarative test' [$F(2,12)=2.26$, $P=0.15$].

However, there was an important contribution of experience in the explicit training condition on the evolution of declarative letter knowledge in the implicit training condition. Subjects that received training in the implicit condition after completing training in the explicit condition had significantly higher declarative letter knowledge in the implicit condition compared to subjects trained in the implicit condition before the explicit condition (Fig. 7b). One-tailed $t$-test on the scores of the original letters testing showed a significant difference between the two groups [$t(7)=(-2.06)$, $P<0.05$]. Only in subjects who were given implicit training before the explicit training, significantly less letters were recognized in the implicit compared to the explicit conditions [$t(5)=2.12$, $P<0.05$]. For comparison, recognition of the original letters in the declarative test after training in the implicit condition was not dependent upon previous experience in the arbitrary condition. No significant difference was found in the scores of the original letters testing between subjects that received training in the implicit condition after completing training in the arbitrary condition, and subjects receiving training in the implicit condition before training in the arbitrary condition [$t(4.9)=1.2$; n.s.]. Thus, it is unlikely that an order effect per se accounted for the gains in the implicit following explicit training. It may therefore be concluded
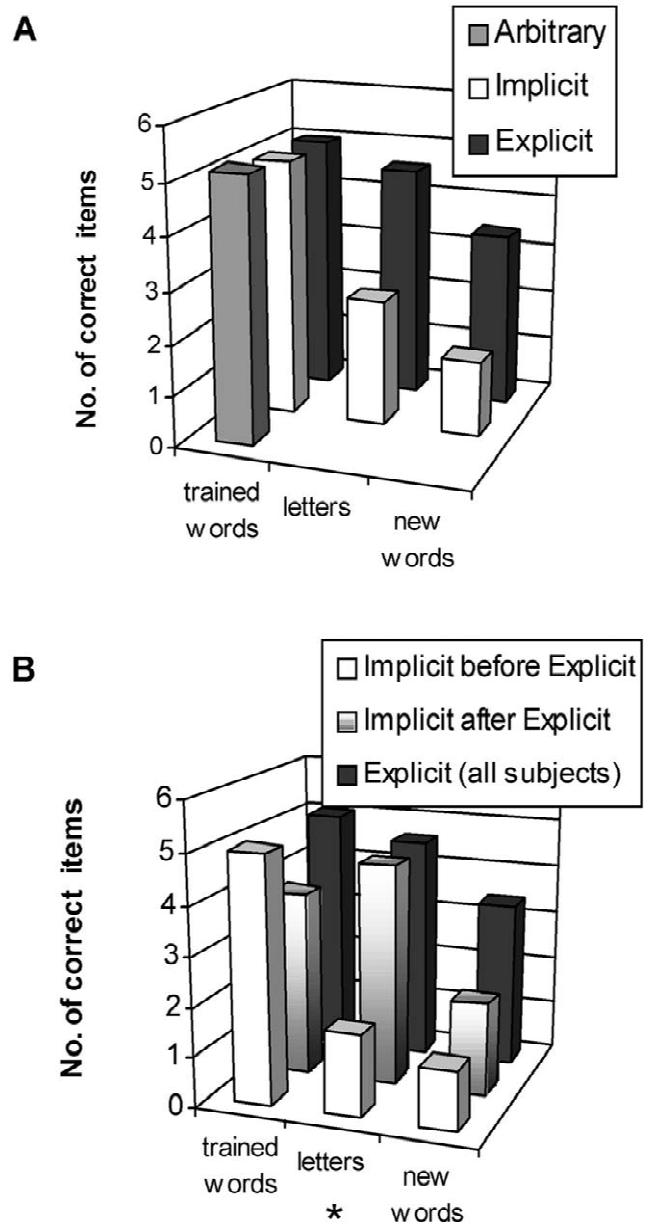


Fig. 7. Results of the declarative knowledge test in the three training conditions (9 Ss). (A) Number of correct items translated. In the arbitrary condition, 'letters' and 'new words' translation could not be tested. (B) Partially shaded and white bars designate the results in the implicit condition split according to whether experience with explicit training preceded implicit training (white bars) or given after implicit training (shaded bars).

that previous experience with explicit training was necessary for gaining declarative knowledge of letters during implicit training on whole-words.

### 3.6. Groups in the implicit condition

Because declarative knowledge in the implicit training condition was found to be dependent on whether explicit experience was available before training, the data on

accuracy and transfer in the implicit condition was reanalyzed according to the same grouping criterion. Previous experience with explicit training did not have an effect on the learning curves in the implicit condition, or on the pattern of transfer results following training in the implicit condition. A GLM analysis on the accuracy in the implicit condition, with session and block as within subject variables, and previous experience in the explicit condition as between subject variable, revealed no significant main effect or interaction with previous experience in explicit training [$(F 1,7)<1$]. A similar analysis on the transfer tests ratios following training in the implicit condition revealed no significant main effect or interaction with previous experience in explicit training [$F(1,7)<1)$]. Similar results were obtained for previous experience in arbitrary training which had no significant effect on the learning curves of the implicit condition [$F(1,6)<1$], and no significant effect on the transfer tests ratios following training in the implicit condition [$F(1,6)=2.6$; n.s.].

## 4. Discussion

Taken together the results of the present study show that very effective learning occurs in both explicit and implicit training, as well as in training to read non-alphabetical word patterns (pictographs). Moreover, within a given task, only the amount of experience, rather than the training mode had a significant effect on the time course of learning. However, the differential pattern of transfer results suggests that although the learning curves were similar in all three conditions, training in the translation task resulted in changes in different levels of brain representation, as a function of training condition. Our results also show that the ability to segment the trained whole words did not evolve spontaneously or implicitly. Extraction of the letters from whole words occurred only when participating individuals have had previous experience in explicit training, suggesting a critical contribution from declarative knowledge mechanisms to the development of letter knowledge. Thus, explicit training, by inducing the learning of letters, had an advantage over whole word training when encountering new words composed of the same letters. However, this type of training was found to be disadvantageous relative to whole word learning with respect to the ability to transfer the effects of training (generalization) to a new alphabetic system. Finally, our results show that a considerable part of the performance gains in all training conditions were specific to the requirements (constraints) of the task, but transferable across stimuli and training conditions.

Our results show that the learning curves of the different training conditions were not significantly different. Whole words (pictographs) and alphabetical reading evolved at an identical rate, although given the differential pattern of transfer results learning related changes presumably occurred at different levels of neural representations [36]. Given a reductionist approach our results suggest the possibility that similar, basic learning mechanisms may underlay practice related changes in different neural populations subserving script decoding [27,36]. The similarity in the time course of learning in all three conditions, does not indicate that an identical number of task repetitions is required to induce a given amount of improvement. Rather, it suggests that a threshold amount of training required to induce optimal improvement was exceeded in the current training protocol [36–38].

The learning curves (in all training conditions) show some of the critical characteristics of procedural learning [36]. In similarity to learning of sensory-perceptual and motor skills, improvement in task performance in the current study required time, multiple sessions and numerous repetitions [35]. The learning curves in all three training conditions showed a good fit to power functions (a leading characteristic of skill learning [1,39,44] but see [31,42]), and learning gains were preserved for a long time (months) after training [8,14,36]. In addition, data to be reported elsewhere has shown that, in similarity to the time course of simple perceptual and motor skill learning [29,38,42] fast, within session improvements, occurred in the initial sessions, while slow, inter-session gains, continued throughout the whole training process [5].

Our data indicate that there was also a considerable involvement of declarative knowledge in the performance of the translation task. In all training conditions the subjects initially received an instruction block in which they were explicitly informed of the correct translation for each target symbol string. This declarative knowledge may account for the initial above chance level of performance in the translation task (about 60% accuracy). Moreover, the results of the declarative tests at the end of training showed good declarative knowledge of the trained words in all training conditions.

### 4.1. What subjects learn in the different training conditions

In the explicit condition the transfer to novel words composed of the original letters was significantly better than the transfer to novel words composed of novel letters, indicating that learning involved the representation of specific letters. In the arbitrary training condition, there was no advantage for new words composed of the original letters over new words composed of novel letters, indicating that learning occurred at the level of word-specific orthographic pattern representation[1].

---

[1]Despite the usage of terms employed in the Dual Route Model (whole word vs. letter representations), our results cannot be considered as a support for either model of visual word recognition [14,15,23,38,43,46,57,65].

One should note, however, that the transfer pattern was dynamic even within the limited timeframe of the transfer session (Fig. 4). In the initial phase of testing the transfer to new words composed of the original phonemes, explicit alphabetical instruction resulted in an advantage compared to training on whole words. This advantage for alphabetical knowledge diminished, however, during the limited training that occurred during the transfer session, again indicating that both training methods resulted in effective learning. These results are partly in accord with the results of Jeffrey and Samuels [32] that found an advantage for letter training compared to word training in the transfer to new words, both in the first encounter and in the repetitions required for learning them. The simplicity of the stimuli in Jeffrey and Samuels' study resulted in a very short training process before participants reached maximum accuracy (ceiling). The reason for the disappearance of the advantage for the explicit training condition in the current study is not clear. This may be the result of a switch from phonological decoding processes to a direct retrieval of word-specific representation in which explicit training has no advantage over arbitrary training. Such a switch from phonological to orthographic reading is suggested by the 'self-teaching hypothesis' to occur after a certain amount of familiarity with the target words [57], and supported by some findings in children [24,67].

Explicit training showed a clear disadvantage compared to whole word training (in both the implicit and the arbitrary conditions) in terms of the ability to transfer knowledge to words in a new alphabet ('letter transfer'). A high 'letter transfer' ratio may be the result of familiarity with the elementary symbols (used in the trained condition), and learning of other task features that are common to all conditions. Additionally, in the alphabetical conditions a high 'letter transfer' ratio may reflect learning of the segmentation rules in various degrees of elaboration. For example: each phoneme corresponds to more than a single symbol; the number of symbols per phoneme is variable (2–3); each consonant corresponds to a string of three symbols while each vowel corresponds to a string of two symbols. However, when segmentation knowledge is not acquired in an alphabetical condition, the usage of recombination of familiar symbols may cause interference. For example, after training on: LOP=*^|*||^* the 'letter transfer' item KIB=|^||^*|* consists of the trained letter P in the 4th, 5th and 6th positions. The poor performance on the 'letter transfer' test in the explicit condition compared to the arbitrary condition suggests that explicit training resulted in knowledge of specific letters, rather than in a more general ability to segment a string of symbols and recognize letters in a new alphabetical system. This would presumably enhance the interference from trained letters on translation of words composed of new letters. However, in the whole word training conditions (in addition to learning specific words) subjects became more efficient in recognizing patterns and extracting salient features from a sequence of symbols. This knowledge could then be applied to other strings with a similar structure.

There is however an apparent contradiction between this interpretation and the declarative knowledge data, showing that previous experience in the explicit training condition enhanced letter recognition in subsequent implicit training conditions. This contradiction can be resolved by assuming that translation performance in the explicit condition was based on the recognition of specific letters, rather than on knowledge of the segmentation rules (i.e. the number of symbols per grapheme). Thus the segmentation rules were inferred and applied to the new alphabet used in the implicit condition, only in the declarative knowledge test. Thus, previous experience in the explicit condition improved declarative knowledge of segmentation in the subsequent implicit condition, but did not improve segmentation in the performance of the 'letter transfer' test.

Another indication that learning may have occurred at a different level of representation in the different training conditions is given by the pattern of results in the 'symbol transfer' tests. In both the arbitrary and the implicit training conditions the largest transfer effects were found in the 'symbol transfer' test, while the explicit training condition resulted in the least amount of transfer in the 'symbol transfer' test. The stimuli in the 'symbol transfer' test preserved both the phonological pattern of the trained words and the internal regularities of the symbols in the trained words (as symmetries and repetitions of symbols in the sequence). It is unlikely that preserving the phonological pattern of the trained words would have a differential effect on the explicit compared to the arbitrary and the implicit conditions (although this interpretation cannot be ruled out by the current results). However, symmetries and repetitions of symbols are presumably crucial in learning specific word patterns (as in the arbitrary training condition), but less relevant for word recognition when the word is segmented into the individual letters (as in the explicit training condition). Thus, the differential pattern of transfer in the 'symbol transfer' test supports the notion that letter specific knowledge resulted from explicit training, and ability to process structural features resulted from implicit and arbitrary training.

Transfer of learning gains to the 'symbol transfer' test may result from learning different sizes of symbol patterns (2–8 symbols) within the symbol string, in different positions within the string, in a continuous or even a discontinuous manner (e.g. ABB_ _ _ BA). The high degree of 'symbol transfer' in the arbitrary condition adds to the evidence that word-specific representations may contain detailed orthographic information of parts (or perhaps even the complete) specific sequence of symbols rather than a global visual configuration of the word [46,49,54,57,58].

The nature of the 'symbol transfer' test is, in many aspects, similar to the type of transfer test used in the Artificial Grammar Learning task [50]. In the AGL task, strings composed of a new set of letters but with the same

abstract structural rules (so that internal symmetries and element repetitions were preserved) were often used as a transfer condition. The classical interpretation for the ability of individuals, both amnesic and neurologically intact, to perform this transfer condition above chance level, was that learning had occurred at the level of the abstract rules that underlie the composition of the letter strings (deep structure) [40,41]. Alternative interpretations, however, suggest that the performance in the transfer test could be the result of feature based learning [53]. Thus, during the test phase subjects may be able to map 'transfer' letters to the trained letters, given that they have learned to recognize surface recurring fragments (of two to three letters) during the learning phase (e.g. xax and bxx into: ycy and dyy respectively). Our results, therefore, lend support to the notion that symmetries and patterns of repetitions of symbols in a given complex pattern may constitute important surface markers for recognition learning. The notion that surface recurring fragments of the whole pattern are learned, rather than (deep) abstract segmentation rules, is compatible with our findings of a high degree of 'symbol transfer' (>50%) following training in the arbitrary condition, in which there were no underlying rules for the structure of the words.

### 4.2. No spontaneous (implicit) segmentation

Word segmentation skills are an important component in the acquisition of reading in alphabetical systems such as English or French, in which a grapheme (the written representation of a phoneme) could be composed of more than a single letter (e.g. 'ch') [18,33,64]. Joubert and Lecours [33] found that non-words that were composed of two-letter graphemes were read slower than non-words composed of one-letter graphemes, indicating that the process of graphemic segmentation requires time in reading of alphabetical languages even in experienced readers [33,66]. In addition, in deep orthographies, such as English, the effect of context on phonological decoding suggests that the relevant unit for decoding is larger than one-phoneme-to-one-grapheme [46,57], therefore a process of segmentation and identification of the relevant units is necessary prior to decoding.

Our results show that subjects in the implicit condition did not employ knowledge of letters when performing the translation task, as reflected in the similar amount of transfer to new words composed of the original letters compared to new words composed of new letters. This finding indicates that, in the implicit condition, adult readers, experienced in word to letter segmentation in their native alphabet, processed the Morse-like script stimuli as pictographic units, without segmenting the words into their component letters. The data from the declarative knowledge test indicate that only after subjects were given explicit training on letters in one Morse-like script, they were able to segment whole words written in a second

Mores-like script, trained in the implicit condition. When learning to read, children have to acquire the alphabetical principle, which, several studies suggest, includes segmentation rules, the concept of letter to sound correspondence and specific associations between letters and sounds [18,33,64]. Given that in the explicit training condition, subjects received explicit instruction on the specific associations between the Morse-like ''letters'' and their corresponding sounds, it is likely that this procedure resulted in declarative knowledge of the letter to sound correspondence and the principle that the Morse-like script is segmented to letters. Our results clearly show that this declarative knowledge was essential for the extraction of specific letter to sound associations in the implicit condition (i.e. in training on whole-word patterns). Moreover, as we have shown, even with explicit training subjects may have acquired only specific letter to sound associations rather than the ''abstract'' segmentation rules.

Our finding that segmentation did not occur spontaneously in training on whole words is in accord with the evidence that word-specific representations are rapidly produced in beginning readers, and that beginning readers tend to rely on graphemic processing rather than on phonological decoding for word recognition [49,54,65]. Our results are also in accord with the studies showing that exposure alone to alphabetical words does not spontaneously induce the discovery of the alphabetic principle in children and that explicit instruction on symbol–sound correspondence is required [9,11,13,22,32,34,56]. For example, Byrne and Fielding-Barnsley [11] showed that children required both training on phonemic awareness and training on letter–sound associations in order to be able to decode novel words, and that even extensive exposure to whole-words could leave a child ignorant of the alphabetic code [11]. The results of the current study extend these conclusions to adult experienced readers acquiring a new alphabet. Our results do not support the claim that implicit learning is sufficient for acquiring phonological decoding [23,46,63–65]. The lack of spontaneous segmentation together with the current findings in the 'symbol transfer' test, are compatible with studies that showed that subjects learn global resemblance [6] or permissible ''surface'' fragments [47] rather than abstract grammar rules in the AGL paradigm.

The participants in the current study were trained to associate symbol strings with sound patterns that were represented orthographically by Latin letters. Although subjects were instructed to articulate the words (read aloud) during the instruction block, in order to enhance the association of symbol string to phonological representation, subjects may have learned to associate the symbol string patterns with the orthographic patterns of the Latin letters. At the very least this procedure, made the notion of letters in a grapheme string more salient. The finding that segmentation did not occur spontaneously even under these facilitating conditions further supports the notion that

segmentation does not necessarily evolve spontaneously from extensive whole word training. Nevertheless it is not clear to what extent the lack of segmentation in the implicit condition was the result of the small number of words used in each training set or the result of the complexity of segmentation rules [10,47,57]. The effects of larger sets of words and less complicated segmentation rules on the induction of segmentation skills in implicit training conditions is under current investigation.

### 4.3. No advantage for implicit training

Brooks and Miller [7] reported an advantage for the arbitrary over the explicit training conditions, and an advantage of implicit over arbitrary training conditions in learning to associate symbol patterns with phonological patterns, similar to the task used in the current study. Several methodological differences may account for the lack of such advantages in our results. Brooks and Miller had their subjects practice to maximal accuracy before measuring their improvement in terms of reaction time, thus testing a late phase of learning. The current study was designed to address initial as well as more advanced stages of learning, and the main parameter for improvement was reading accuracy. However, our RT data show that when maximal accuracy was reached, reading in the arbitrary condition was faster than in the explicit condition, similar to the Brooks and Miller [7] results.

Brooks and Miller [7] did not find an advantage for either of the alphabetical conditions (explicit and implicit) compared to the arbitrary condition in the transfer to new words composed of the trained letters. This indicates that their subjects may not have had an effective knowledge of the letters in the alphabetical conditions. Ineffective alphabetical knowledge may account for the slow reading in the explicit condition compared to the effective whole word reading in the arbitrary condition. Support for this interpretation is provided by a second experiment [7] in which the explicit condition became faster than the arbitrary condition after longer training. These transfer results indicate that the speed advantage gained in implicit over explicit training cannot be taken as indication for a substantial advantage for the former mode of training, particularly in light of the current study's finding that letter learning is dependent on explicit knowledge.

### 4.4. Stimulus independent and task specific learning

Our results show a clear effect of the sequence of training conditions, although different alphabetical systems were used in each subsequent condition. This transfer between training conditions and sets of stimuli indicates that in addition to learning that was specific to each training condition, substantial learning related changes also occurred at a level of representation that is shared by all training conditions and all our Morse-like stimulus sets, presumably reflecting the shared requirements of the task. This notion is supported by the finding that although learning gains transferred across different stimuli, there was no transfer of the training dependent gains between the translation and the visual matching tasks using the same stimuli. Our results show that training in the visual matching task was less effective in terms of gains in accuracy across the five training sessions (Fig. 6a) compared to training in the translation task. Moreover, performance in the translation task improved with the experience in previous conditions, while performance in the visual matching task was not affected by previous experience or the order of conditions. In addition, the low 'symbol transfer' across all training conditions in the visual matching task suggests that learning occurred at the level of representations of the individual symbols independent of the training method in contrast to learning in the translation task. Altogether these differences indicate that although identical symbol strings were used in the two tasks they were processed differently and learnt independently to meet the requirements and constraints of each task. Our results, thus, supports the conjecture that the locus of learning is determined not only by the nature of the stimuli but also by the requirements of the task[2] [27,35].

## 5. Conclusions

Our findings suggest that learning in different levels of representation may be subserved by similar learning mechanisms. In addition, our results indicate that in some conditions, segmentation skills may not develop spontaneously from extensive exposure to whole patterns, and that declarative knowledge of segmentability and of the segmentation rules may be essential. However, even the process of explicit training on the individual segments of a given pattern, may induce segment-specific recognition skills, rather than a general segmentation skill, thus reducing the ability to transfer this knowledge to patterns composed of different segments. Finally, our results indicate that task constraints, rather than the nature of the input per se, can determine what is learned from repeated experience.

---

[2]In a control 'masked translation' task (in which the task requirements were the same as in the translation task but with timing constraints and masking as in the visual-matching task) subjects were only 70% accurate by the end of training, compared to more than 90% accuracy in the original translation task. This reduction in performance indicates that the limited exposure duration can account, at least partially, for the lack of transfer between tasks.

# References

[1] J.R. Anderson, Acquisition of cognitive skill, Psychol. Rev. 89 (1982) 369–406.

[2] N.J. Benson, M.W. Lovett, C.L. Kroeber, Training and transfer of learning effects in disabled and normal readers: evidence of specific deficits, J. Exp. Child Psychol. 64 (1997) 343–366.

[3] D.C. Berry, D.E. Broadbent, Interactive tasks and the implicit–explicit distinction, Br. J. Psychol. 79 (1988) 251–272.

[4] C. Bishop, Transfer effects of word and letter training in reading, J. Verb Learn Verb Behav. 3 (1964) 214–221.

[5] T. Bitan, A. Karni, Training on whole words does not necessarily result in alphabetical knowledge, Soc. Neurosci. Abstr., 27, Program No. 949.4, (2001).

[6] D.N. Brooks, Nonanalytic concept formation and memory for instances, in: E. Rosch, B.B. Lloyd (Eds.), Cognition and Categorization, Erlbaum, Hillsdale, NJ, 1978, pp. 169–215.

[7] L. Brooks, A. Miller, A comparison of explicit and implicit knowledge of an alphabet, in: P.A. Kolers, M.E. Wrolstad, H. Bouma (Eds.), Processing of Visible Language, Vol. 1, Plenum Press, New York, 1979, pp. 391–401.

[8] L.T. Butler, D.C. Berry, Implicit memory: intention and awareness revisited, Trends Cog. Sci. 5 (2001) 192–197.

[9] B. Byrne, Studies in the acquisition procedure: rationale, hypotheses and data, in: P.B. Gough, L.C. Ehri, R. Treiman (Eds.), Reading Acquisition, 1992, pp. 1–34.

[10] B. Byrne, M. Carroll, Learning artificial orthographies: further evidence of a non-analytic acquisition procedure, Mem. Cog. 17 (1989) 311–317.

[11] B. Byrne, R. Fielding-Barnsley, Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle, J. Edu. Psychol. 81 (1989) 313–321.

[12] B. Byrne, R. Fielding-Barnsley, L. Ashley, Effects of preschool phoneme identity training after six years: outcome level distinguished from rate of response, J. Edu. Psychol. 92 (2000) 659–667.

[13] D. Carnine, Phonics vs. look–say: transfer to new words, Reading Teacher 30 (1977) 636–640.

[14] D.M. Clawson, K.A. Ericsson, A.F. Healy, L.E. Bourne, Retention and transfer of Morse-code reception skill by novices: part-whole training, JEXP–Appl 7 (2001) 129–142.

[15] A. Cleeremans, A. Destrebecqz, M. Boyer, Implicit learning: news from the front, Trends Cog. Sci. 2 (1998) 406–416.

[16] A. Cohen, R.I. Ivry, S.W. Keele, Attention and structure in sequence learning, J. Exp. Psychol: Learn. Mem. Cog. 16 (1990) 17–30.

[17] N.J. Cohen, R.A. Poldrack, H. Eichenbaum, Memory for items and memory for relations in the procedural/declarative memory framework, Memory 5 (1997) 131–178.

[18] M. Coltheart, B. Curtis, P. Atkins, M. Haller, Models of reading aloud: dual-route and parallel-distributed-processing approaches, Psychol. Rev. 100 (1993) 589–608.

[19] M. Coltheart, K. Rastle, C. Perry, R. Langdon, J. Ziegler, DRC: a dual route cascaded model of visual word recognition and reading aloud, Psychol. Rev. 108 (2001) 204–256.

[20] P.F. Dominey, T. Lelekov, J. Ventre-Dominey, M. Jeannerod, Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences, J. Cog. Neurosci. 10 (1998) 734–751.

[21] L.C. Ehri, Reconceptualizing the development of sight word reading and its relationship to recoding, in: P.B. Gough, L.C. Ehri, R. Treiman (Eds.), Reading Acquisition, Erlbaum, Hillsdale, NJ, 1992, pp. 107–143.

[22] L.C. Ehri, J. Sweet, Fingerpoint-reading of memorized text: what enables beginners to process the print, Read Res. Quart. 24 (1991) 442–462.

[23] L.C. Ehri, L.S. Wilce, Movement into reading: is the first stage of printed word learning visual or phometic?, Read. Res. Quart. 20 (1983) 163–179.

[24] B.R. Foorman, D.J. Francis, J.M. Fletcher, C. Schatschneider, P. Mehta, The role of instruction in learning to read: preventing reading failure in at risk children, J. Edu. Psychol. 90 (1998) 37–55.

[25] B.R. Foorman, D.J. Francis, D.M. Novy, D. Lieberman, How letter–sound instruction mediates progress in first-grade reading and spelling, J. Edu. Psychol 83 (1991) 456–469.

[26] U. Frith, Beneath the surface of developmental dyslexia, in: K.E. Patterson, J.C. Marshall, M. Coltheart (Eds.), Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading, Earlbaum, London, 1985, pp. 301–330.

[27] C.D. Gilbert, M. Sigman, R.E. Crist, The neural basis of perceptual learning, Neuron 31 (2001) 681–697.

[28] D.M. Green, J.A. Swets, Signal Detection Theory and Psychophysics, Wiley, New York, 1966.

[29] M. Hallet, A. Pascual-Leone, H. Topeka, in: J. Bloedel, T. Ebner, S.P. Wise (Eds.), Acquisition of Motor Behavior in Vertebrates, MIT Press, Cambridge, MA, 1996, pp. 289–301.

[30] M.W. Harm, M.S. Seidenberg, Phonology, reading acquisition, and dyslexia: insights from connectionist models, Psychol Rev. 106 (1999) 491–528.

[31] B. Hauptmann, M.J.E. Richardson, T. Flash, A. Karni, Motor sequence learning and the co-articulation of movements in adults. Soc. Neurosci. Abstr., 26, Program No. 63.8 (2000).

[32] W.E. Jeffrey, S.J. Samuels, Effects of method of reading training on initial learning and transfer, J. Verb. Learn. Verb. Behav. 6 (1967) 354–358.

[33] S.A. Joubert, A.R. Lecours, The role of sublexical graphemic processing in reading, Brain Lang. 72 (2000) 1–13.

[34] C. Juel, The development and use of mediated word identification, Read Res. Quart. 18 (1983) 306–327.

[35] A. Karni, The acquisition of perceptual and motor skills: a memory system in the adult human cortex, Cog. Brain Res. 5 (1996) 39–48.

[36] A. Karni, G. Bertini, Learning perceptual skills: behavioral probes into adult cortical plasticity, Cur. Opin. Neurobiol. 7 (1997) 530–535.

[37] A. Karni, G. Meyer, C. Rey-Hipolito, P. Jezzard, M.M. Adams, R. Turner, L.G. Ungerleider, The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex, Proc. Natl. Acad. Sci. 95 (1998) 861–868.

[38] A. Karni, D. Sagi, The time course of learning a visual skill, Nature 365 (1993) 250–252.

[39] K. Kirsner, C. Speelman, P. Schofield, Implicit memory and skill acquisition: Is synthesis possible? (1993) 119–139.

[40] B.J. Knowlton, L.R. Squire, The information acquired during artificial grammar learning, J. Exp. Psychol: Learn. Mem. Cog. 20 (1994) 79–91.

[41] B.J. Knowlton, L.R. Squire, Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information, J. Exp. Psychol: Learn Mem. Cog. 22 (1996) 169–181.

[42] M. Korman, N. Raz, T. Flash, A. Karni, Practice makes transfer imperfect: evidence for stages in motor skill acquisition, Soc. Neurosci. Abstr., 26, Program No. 691.11 (2000).

[43] Y. Lee, D.A. Vakoch, Transfer and retention of implicit and explicit learning, Br. J. Psychol. 87 (1996) 637–651.

[44] G. Logan, Toward an instance theory of automatization, Psychol. Rev. 95 (1988) 492–527.

[45] M.J. Nissen, P. Bullemer, Attentional requirement of learning: evidence from performance measures, Cog. Psychol. 19 (1987) 1–32.

[46] C.A. Perfetti, The representation problem in reading acquisition, in: P.B. Gough, L.C. Ehri, R. Treiman (Eds.), Reading Acquisition, Erlbaum, Hillsdale, NJ, 1992, pp. 145–174.

[47] P. Perruchet, C. Pacteau, Synthetic grammar learning: implicit rule abstraction or explicit fragmentary knowledge?, JEXP: Gen. 119 (1990) 264–275.

[48] D.C. Plaut, J.L. McClelland, M.S. Seidenberg, K. Patterson, Understanding normal and impaired word reading: computational principles in quasi-regular domains, Psychol. Rev. 103 (1996) 56–115.

[49] K. Rayner, Word recognition cues in children: the relative use of graphemic cues, orthographic dues and grapheme–phoneme correspondence rules, J. Edu. Psychol. 80 (1988) 473–479.

[50] A.S. Reber, Implicit learning of artificial grammars, J. Verb. Learn. Verb. Behav. 6 (1967) 855–863.

[51] A.S. Reber, Implicit learning and tacit knowledge, J. Exp. Psychol: Gen. 118 (1989) 219–235.

[52] P.J. Reber, B.J. Knowlton, L.R. Squire, Dissociable properties of memory systems: differences in the flexibility of declarative and non-declarative knowledge, Behav. Neurosci. 110 (1996) 861–871.

[53] M. Redington, N. Chater, Transfer in artificial grammar learning: a reevaluation, JEXP: Gen. 119 (1996) 123–138.

[54] P. Reitsma, Printed word learning in beginning readers, J. Exp. Child Psychol. 36 (1983) 321–339.

[55] M.S. Seidenberg, D.C. Plaut, A.S. Petersen, J.L. McClelland, K. McRae, Nonword pronunciation and models of word recognition, JEXP:HPP 20 (1994) 1177–1196.

[56] P.H.K. Seymour, L. Elder, Beginning reading without phonology, Cog. Neuropsychol. 1 (1986) 43–82.

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

[58] J. Shimron, D. Navon, The dependence on graphemes and on their translation to phonemes is reading: a developmental perspective, Read. Res. Quart. 17 (1982) 210–228.

[59] L.R. Squire, Memory and Brain, Oxford University Press, New York, 1987.

[60] L.R. Squire, B.J. Knowlton, Learning about categories in the absence of memory, Proc. Natl. Acad. Sci. 29 (1995) 12470–12474.

[61] L.R. Squire, S.M. Zola, Structure and function of declarative and nondeclarative memory systems, Proc. Natl. Acad. Sci. 93 (1996) 13515–13522.

[62] R. Sun, E. Merill, T. Peterson, From implicit skills to explicit knowledge: a bottom-up model of skill learning, Cog. Sci. 25 (2001) 203–244.

[63] G.B. Thompson, D.S. Cottrel, C.M. Fletcher-Flinn, Sublexical orthographic-phonological relation early in the acquisition of reading: the knowledge sources account, J. Exp. Child Psychol. 62 (1996) 190–222.

[64] G.B. Thompson, C.M. Fletcher-Flinn, D.S. Cottrel, Learning correspondence between letters and phonemes without explicit instructions, Appl. Psychol. 20 (1999) 21–50.

[65] G.C. Van Orden, B.F. Pennington, G.O. Stone, Word identification in reading and the promise of subsymbolic psycholinguistics, Psychol. Rev. 97 (1990) 488–522.

[66] N. Walter, S. Cliche, S. Joubert, M. Beauregard, Y. Joanette, Grapheme–phoneme processing of single words, Brain Cog. (2001) 295–9.

[67] R. Wessling, P. Reitsma, The transfer role of explicit phonological recoding for reading acquisition, Read Writ. 13 (2000) 313–336.