Common constraints limit Korean and English character recognition in peripheral vision

Yingchen He
Department of Psychology, University of Minnesota,
Minneapolis, MN, USA

MiYoung Kwon
Department of Ophthalmology,
University of Alabama at Birmingham,
Birmingham, AL, USA

Gordon E. Legge
Department of Psychology, University of Minnesota,
Minneapolis, MN, USA

The *visual span* refers to the number of adjacent characters that can be recognized in a single glance. It is viewed as a sensory bottleneck in reading for both normal and clinical populations. In peripheral vision, the visual span for English characters can be enlarged after training with a letter-recognition task. Here, we examined the transfer of training from Korean to English characters for a group of bilingual Korean native speakers. In the pre- and posttests, we measured visual spans for Korean characters and English letters. Training (1.5 hours × 4 days) consisted of repetitive visual-span measurements for Korean trigrams (strings of three characters). Our training enlarged the visual spans for Korean single characters and trigrams, and the benefit transferred to untrained English symbols. The improvement was largely due to a reduction of within-character and between-character crowding in Korean recognition, as well as between-letter crowding in English recognition. We also found a negative correlation between the size of the visual span and the average pattern complexity of the symbol set. Together, our results showed that the visual span is limited by common sensory (crowding) and physical (pattern complexity) factors regardless of the language script, providing evidence that the visual span reflects a universal bottleneck for text recognition.

Introduction

Human pattern recognition is limited by available visual information within a glimpse. The visual span, referring to the number of identifiable letters within one fixation, reflects such sensory limits on reading (Legge et al., 2007). In peripheral vision, the size of the visual span for English characters can be enlarged by letter recognition training, accompanied by an improvement in reading speed of 40% or more (e.g., Chung, Legge, & Cheung, 2004). This training paradigm has potential value for reading rehabilitation for people who have lost their central vision.

But what exactly is learned through the training, and can the training benefit generalize to untrained conditions? A reduction of crowding has been shown to be the major component of the training-related enlargement of the visual span (He, Legge, & Yu, 2013). The enlargement can transfer to an untrained visual-field location (Chung et al., 2004; He et al., 2013; Yu, Legge, Park, Gage, & Chung, 2010) or to an untrained print size (Yu et al., 2010). This transfer suggests that the effect is not retinotopically specific and can tolerate some variance in the size of the stimuli. However, it is not known whether the training effect is specific to the set of trained symbols.

One possible symbol-specific mechanism of the improvement is to learn better templates for letter recognition. Gold, Sekuler, and Bennett (2004) found that human subjects used suboptimal perceptual templates for pattern discrimination, but training modified their templates and yielded better performance. If enlarged visual spans are caused by subjects learning more precise templates of the trained symbols, we would not expect transfer of training to an untrained set of symbols.

Alternatively, the underlying mechanism of the improvement could be non–symbol specific. For example, Sun, Chung, and Tjan (2010) found that training to read flanked letters may have adjusted the spatial extent of the perceptual window used for

Citation: He, Y., Kwon, M., & Legge, G. E. (2018). Common constraints limit Korean and English character recognition in peripheral vision. *Journal of Vision*, *18*(1):5, 1–15, https://doi.org/10.1167/18.1.5.

https://doi.org/10.1167/18.1.5 Received June 29, 2017; published January 11, 2018 ISSN 1534-7362

Copyright 2018 The Authors

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

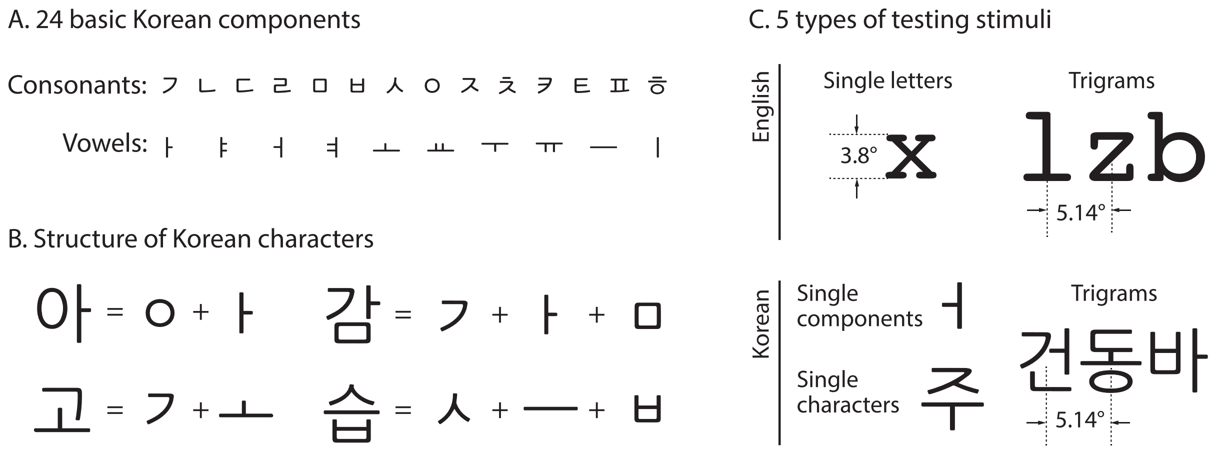
A. 24 basic Korean components
Consonants: ㄱ ㄴ ㄷ ㄹ ㅁ ㅂ ㅅ ㅇ ㅈ ㅊ ㅋ ㅌ ㅍ ㅎ
Vowels: ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ
B. Structure of Korean characters
아 = ㅇ + ㅏ
감 = ㄱ + ㅏ + ㅁ
고 = ㄱ + ㅗ
습 = ㅅ + ㅡ + ㅂ
C. 5 types of testing stimuli
English
Single letters
Trigrams
3.8°
x
lzb
5.14°
Korean
Single components
ㅓ
Trigrams
Single characters
주
건동바
5.14°

Figure 1. Testing stimuli. (A) Twenty-four basic Korean components, including 14 consonants and 10 vowels. (B) Structure of Korean characters. The top row shows two-component (left) and three-component (right) characters whose vowels contain a major vertical bar. The bottom row shows two-component (left) and three-component (right) characters whose vowels contain a major horizontal bar. (C) Five types of testing stimuli, including single English letters, English trigrams, single Korean components, single Korean characters, and Korean trigrams.

template matching. Their ideal observer analysis revealed that the effect of crowding was to increase equivalent input noise and to decrease sampling efficiency. Training either decreased noise by reducing an inappropriately large perceptual window or increased efficiency by enlarging an inappropriately small perceptual window. If enlargement of the visual span is caused by the same mechanism, training would transfer to an untrained symbol set.

Here, to test whether the underlying mechanism of the learning is symbol specific, we trained Korean-English bilingual participants to recognize Korean characters and examined whether the training transferred to better English letter recognition. Transfer would imply a non–symbol-specific learning mechanism that is not primarily due to sharpening of the character templates.

We chose Korean rather than other languages for two reasons. First, despite the differences in the symbols, both Korean and English writing are alphabetic and therefore comparable. The basic set of components that make up Korean characters contains 24 letters (14 consonants and 10 vowels; Figure 1A), close to the set size of 26 lowercase English letters.

Second, the structure of Korean characters may be informative about the role of crowding. Unlike English writing, in which letters are arranged horizontally, Korean components, each representing either a consonant or a vowel, are assembled into blocks to form characters (Figure 1B). These characters are then arranged horizontally to create words. Therefore, both within-character and between-character crowding is present in Korean reading. For English, the major sensory limit of the size of the visual span is crowding (He et al., 2013; Pelli et al., 2007), and reduced crowding is the largest contributor to training-related enlargement of the visual span (He et al., 2013). Using Korean script allows us to further tease apart the influence of within- and between-character crowding on the size of the visual span. We used a two-stage model (He et al., 2015; see the Methods section) to investigate the roles of within- and between-character crowding before and after training to better characterize the underlying mechanism of the training.

When studying crowding, an important physical property of the symbols to be considered is their pattern complexity. Pattern complexity is known to influence the magnitude of crowding and the size of the visual span (Bernard & Chung, 2011; Wang, He, & Legge, 2014; Zhang, Zhang, Xue, Liu, & Yu, 2009). Thus, we also examined the relationship between pattern complexity and the size of the visual span with the way training changes this relationship.

To summarize, our study aims to address whether or not the mechanism underlying the peripheral letter-recognition learning is symbol specific. To this end, we examined whether training to read Korean characters transferred to enlargement of the visual span for English letters. We used a two-stage model to analyze the changes in within- and between-character crowding after training. We also evaluated the effect of pattern complexity on the size of the visual span. Our results shed light on the common sensory and physical constraints limiting pattern recognition and provided further evidence for a non–symbol-specific mechanism underlying training.

Methods

Participants

Nine native Korean speakers (four men) were recruited from the University of Minnesota (mean age = 21.8 years; range = 19–24 years). All subjects were fluent in English and were able to type in both English and Korean dexterously. Participants all had normal or corrected-to-normal vision, with binocular acuity of −0.04 ± 0.01 logMAR (mean ± SEM, measured by Lighthouse Near Acuity Chart, Lighthouse Low Vision Products, Long Island City, NY). The protocol was approved by the Institutional Review Board and was in compliance with the Declaration of Helsinki. All participants gave informed consent prior to the experiment.

Stimuli and apparatus

The stimuli consisted of black text (English or Korean) on a white background (background luminance 102 cd/m^2; Weber contrast = 98%), presented at 10° in the lower visual field. We used a NEC MultiSync CRT monitor (model FP2141SB-BK, NEC, Tokyo, Japan; refresh rate = 100 Hz; spatial resolution = 0.05°/pixel) controlled by a Mac Pro Quad-Core computer (model A1186, Apple Inc., Cupertino, CA). The stimuli were generated and presented using MATLAB R2014b with Psychophysics Toolbox 3 (Brainard, 1997; Pelli, 1997). All stimuli were viewed binocularly from 30 cm in a dark room. Viewing distance was maintained using a chin rest, and the subject's central fixation was monitored using a webcam.

As shown in Figure 1C, five types of symbols were used in the experiment to measure the size of the visual span (see later for details): single English letters, English trigrams (strings of three letters), single Korean components (consonant and vowel letters), single Korean characters, and Korean trigrams. English letters were rendered in Courier, and Korean characters were rendered in Nanum Gothic. All English letters were lowercase and had an x-height of 3.8°. For trigrams, the center-to-center spacing between letters was 1.16 × x-width (standard spacing for Courier, approximately 5.14°). The size of the Korean characters was scaled so that the center-to-center spacing between adjacent characters was also 5.14° (Figure 1C). Images for Korean components were cut from images of Korean characters (see Appendix 1) so that their size and shape are representative of what commonly appears in Korean reading. Both English letter size and Korean character size exceeded their corresponding critical print size for reading at 10° in the lower visual field (Korean: 2.03° ± 0.18°, data from Baek, He, & Legge, 2016; English: 1.38° ± 0.18°, data from Chung, Mansfield, & Legge, 1998; mean ± SEM), so that the print size was not a limiting factor if we were to measure their reading speed.

Structure of Korean characters

Figure 1B illustrates different structures of Korean characters. There are at least two components in a character: a lead consonant letter and a vowel letter. For a character with two components, they are arranged left-right if the vowel has a major vertical bar (e.g., ㅣ or ㅑ) or arranged top-bottom if the vowel has a major horizontal bar (e.g., ㅡ or ㅜ). For a character with more than two components, the third component, that is, a tail consonant letter, is always placed below the other two components regardless of the shape of the vowel.

Korean character set

In a related project, we used 900 Korean sentences to compare Korean reading speed in central and peripheral vision (Baek et al., 2016). Here, we used the most frequent 279 characters from those sentences (accounting for 90% occurrences) as our testing stimuli. In our set, 114 characters are two-component (40.9%) and the rest are three-component. We acknowledge that the set size was large when compared with English (set size = 26), but since reducing the set size would compromise the resemblance of the task to real Korean reading, we kept the set of 279 characters. To minimize the influence of set size, all of the characters were printed on hard-copy paper, and we encouraged the subjects to review the list before each testing block. Subjects were not required to memorize the list, but any out-of-set response triggered a warning, and we ran replacement trials until the response was within the set. Eleven percent of all trials were out-of-set responses.

Experimental design

Our experiment consisted of three parts (Figure 2A): pretest (Day 1), training (Days 2–5), and posttest (Day 6). Each daily session lasted for approximately 1.5 to 2 hours.

In the pre- and posttests, subjects' visual span profiles for both Korean and English were measured (Figure 2A; see later for details). English visual-span profiles were measured in the order of single letters and trigrams. Korean visual-span profiles were measured in the order of single components, single characters, and trigrams. Whether to first measure Korean or English visual span was counterbalanced between subjects in

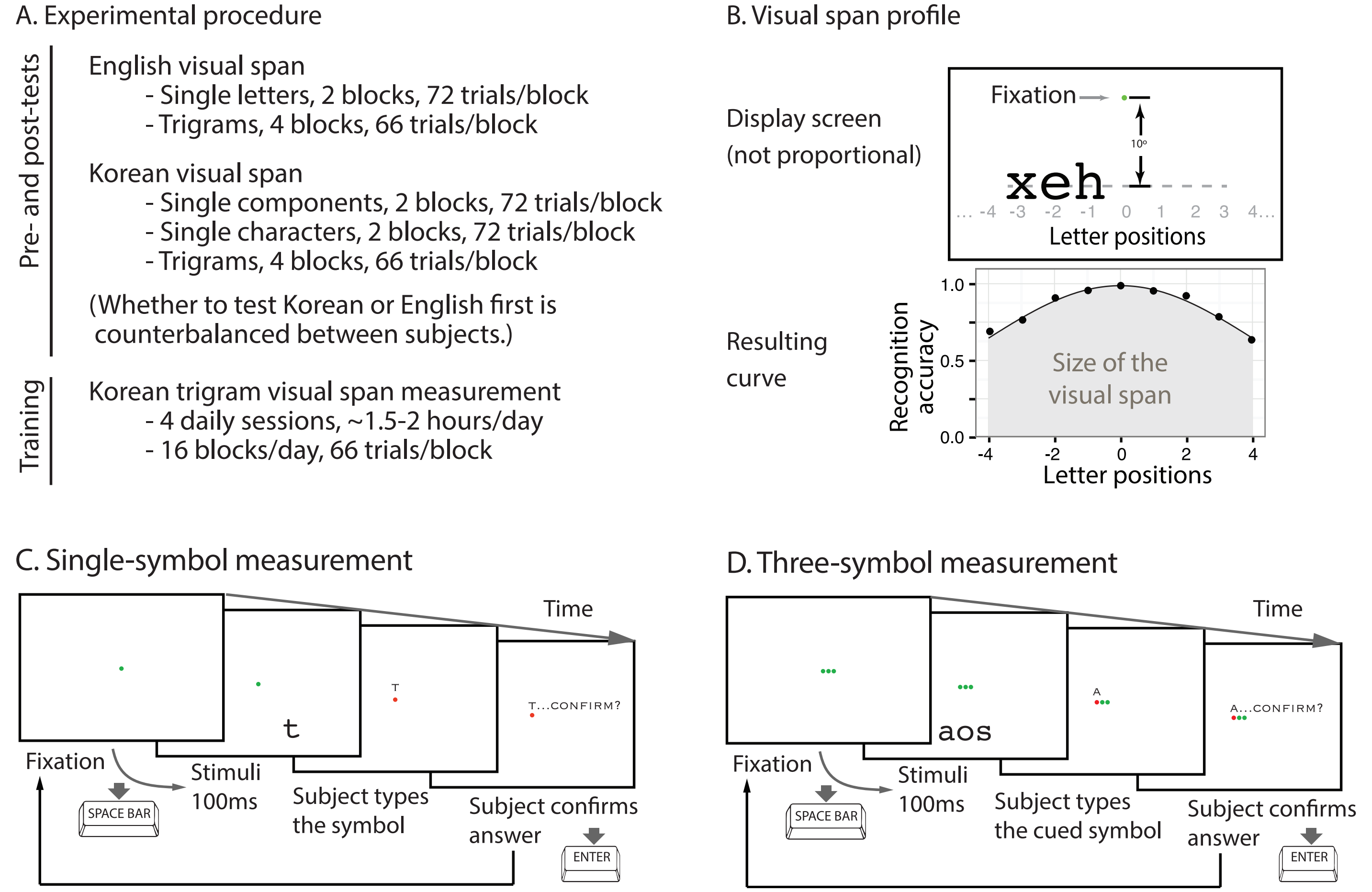
A. Experimental procedure
Pre- and post-tests
English visual span
- Single letters, 2 blocks, 72 trials/block
- Trigrams, 4 blocks, 66 trials/block
Korean visual span
- Single components, 2 blocks, 72 trials/block
- Single characters, 2 blocks, 72 trials/block
- Trigrams, 4 blocks, 66 trials/block
(Whether to test Korean or English first is counterbalanced between subjects.)
Training
Korean trigram visual span measurement
- 4 daily sessions, ~1.5-2 hours/day
- 16 blocks/day, 66 trials/block
B. Visual span profile
Display screen (not proportional)
Fixation
10°
xeh
... -4 -3 -2 -1 0 1 2 3 4...
Letter positions
Resulting curve
Recognition accuracy
1.0
0.5
0.0
-4
-2
0
2
4
Size of the visual span
Letter positions
C. Single-symbol measurement
Time
t
T
T...CONFIRM?
Fixation
SPACE BAR
Stimuli 100ms
Subject types the symbol
Subject confirms answer
ENTER
D. Three-symbol measurement
Time
aos
A
A...CONFIRM?
Fixation
SPACE BAR
Stimuli 100ms
Subject types the cued symbol
Subject confirms answer
ENTER

Figure 2. Experimental procedure and visual span measurement. (A) Experimental procedure. (B) Diagram of a visual span profile. (C) Diagram of single-symbol measurement, including English letters, Korean components, and Korean single characters. (D) Diagram of three-symbol measurement, including English and Korean trigrams.

the pretest and reversed in the posttest for each subject (for example, Korean-English in the pretest and English-Korean in the posttest). From Day 2 to Day 5, subjects participated in training sessions, each consisting of 16 blocks of visual span measurements using Korean trigrams.

Visual span measurement

To measure the size of the visual span, we performed a letter-recognition task, similar to the task in He et al. (2013). Letters were presented in predefined slots, which were horizontally arranged on an imaginary line at 10° in the lower visual field (Figure 2B). The slot on the fixation midline was labeled 0, and left and right slots were labeled with negative and positive numbers, respectively. The center-to-center spacing between adjacent slots was approximately 5.14° (1.16° × x-width, corresponding to standard spacing in the Courier font).

As described previously, there were five different types of testing stimuli. In a single-symbol trial (English letters, Korean components, or Korean characters; Figure 2C), the subject first fixated on a dot, then pressed the space bar to initiate a trial. The symbol then appeared for 100 milliseconds, and the subject was asked to type that symbol using a keyboard. To reduce typing errors, the subject's response was shown on the screen, and the subject was asked to confirm it. The visual feedback for typing was rendered in a different font (and also rendered in uppercase for English) to minimize its potential influence on recognition performance. Within a block, stimuli appeared eight times on each slot from −4 to 4 in a random order. In the pre- and posttests, single English letters, single Korean components, and single Korean characters were each measured in two blocks, respectively.

For English and Korean trigrams (Figure 2D), we used a partial report method instead of full report. In the beginning of a trial, there were three horizontally arranged green dots in the center of the screen. The subject fixated on the middle dot and pressed the space

		Pretest	Posttest
Korean component	$R_{1,\ K\_component}$ = Isolated-component accuracy (corrected for guessing)	95.5% (1.0)	97.1% (1.0)
	$R_{1,\ K\_component} \times R_{2,\ K\_component}$ = Component accuracy within isolated characters (corrected for guessing)	88.8% (1.2)	92.9% (0.6)
	$R_{2,\ K\_component}$	93.2% (1.4)	96.0% (0.8)
Korean character	$R_{1,\ K\_character}$ = Isolated-character accuracy (corrected for guessing)	74.1% (2.3)	82.3% (1.2)
	$R_{1,\ K\_character} \times R_{2,\ K\_character}$ = Character accuracy within trigrams (corrected for guessing)	35.3% (2.7)	45.1% (2.4)
	$R_{2,\ K\_character}$	46.6% (2.4)	53.9% (2.3)
English letter	$R_{1,\ E}$ = Isolated-letter accuracy (corrected for guessing)	98.3% (0.3)	99.2% (0.3)
	$R_{1,\ E} \times R_{2,\ E}$ = Letter accuracy within trigrams (corrected for guessing)	72.1% (1.9)	78.3% (2.3)
	$R_{2,\ E}$	73.3% (1.9)	78.9% (2.3)

Table 1. R_1 and R_2 values before and after training, mean (*SEM*).

bar to initiate the trial. A trigram appeared for 100 milliseconds. After the trigram disappeared, one of the three dots turned red, indicating the symbol to report. For example, if the left dot turned red, the subject needed to report the left symbol of the trigram. Within a block, stimuli were centered six times on each slot from −5 to 5 in a random order, with the left, middle, and right letters each being cued two times. Because slots ±5 and ±6 had fewer symbols than the other slots, only data from slots –4 to 4 were used in further analysis. English and Korean trigrams were each measured in four blocks for both the pre- and posttests. To obtain the visual span profiles, letter-recognition accuracy was plotted against letter positions (Figure 2B). We will report the average accuracy across the nine slots as a summary of the size of the visual span.

Two-stage model for pattern recognition

To examine the effect of crowding on the recognition of different symbol types, and to evaluate the reduction of crowding after training for both the trained and untrained symbols, we adopted a two-stage model to quantify the level of crowding (He et al., 2015).

Briefly, this model assumes that pattern recognition is a serial, independent two-stage process. At the first stage, recognition is limited by factors affecting the processing of isolated symbols. The second stage represents the additional interfering effects of nearby symbols on recognition. Each stage is characterized by its reliability, that is, the probability that the correct information is transmitted through this stage. We have built three separate models in which the first stage corresponds to recognizing single English letters, Korean components, and Korean characters. To distinguish between them, we will use subscripts *E*, *K_component*, and *K_character*, respectively when describing the corresponding reliabilities (see Table 1).

According to our model, the reliability of the first stage (R_1) is equal to the probability (corrected for guessing) of correctly identifying isolated symbols. We refer to the factors influencing this stage as an "acuity effect."

The probability of recognizing the same symbol in a crowd (also corrected for guessing) is the product of the reliabilities of the two stages ($R_1 \times R_2$). For example, if the corrected-for-guessing accuracy is 90% to recognize an isolated English letter and 70% to recognize a letter in a trigram, then

$$R_{1,\ E} = 0.9, \text{ and} \tag{1}$$

$$R_{1,\ E} \times R_{2,\ E} = 0.7. \tag{2}$$

From here, $R_{2,\ E}$ can be computed as $0.7/0.9 = 0.78$.

Korean characters are structurally different from English letters, because they are each made up of individually identifiable components. This allows us to build two separate models for Korean characters. One model treated the recognition of individual Korean components as the first stage, where the influence of other components (within the same character) is represented in the second stage. If the recognition accuracy is 90% for an isolated component and 80% for a component in an isolated character, then $R_{1,\ K\_component} = 0.9$ and $R_{2,\ K\_component} = 0.8/0.9 = 0.89$. Note that during single Korean character and Korean trigram measurements, we asked the subjects to report the identity of only the entire character, not the components. However, once we know the identity of the character, for example, 갑, we can infer that the subject thinks the three components are ㄱ, ㅏ, and ㅁ and then score them based on component accuracy.

The second model for Korean treated the recognition of individual characters as the first stage, and the influence of nearby characters is represented in the second stage. According to this model, if the recognition accuracy is 70% for an isolated character and 50%

for a character in a trigram, then $R_{1,\ K\_character} = 0.7$ and $R_{2,\ K\_character} = 0.7/0.9 = 0.78$.

The values of R_2 can be interpreted as the letter recognition accuracy when Stage 1 processing is not a limiting factor, that is, when the reliability of Stage 1 is 100% ($R_1 = 1.0$). In this way, we can remove the influence of the acuity effect from the visual span and estimate the influence of within- and between-symbol crowding. The closer to 1 R_2 is, the less crowding there is.

Statistical analysis

When comparing training effects between types of symbols, we performed a 5 × 2 repeated-measures analysis of variance (ANOVA) on the average accuracy of the visual span profiles, with two within-subject factors being symbol type (English letters/English trigrams/Korean components/Korean characters/Korean trigrams) and session type (pretest/posttest). The ANOVA was performed after the proportions were remapped to [0.025, 0.975] and logit-transformed, but the average recognition accuracy reported in the Results section is the original untransformed value. When evaluating the changes in crowding, we performed a 3 × 2 repeated-measures ANOVA on the R_2 values (logit-transformed), including $R_{2,\ E}$, $R_{2,\ K\_component}$, and $R_{2,\ K\_character}$. Two within-subject factors were symbol type (English letters/Korean components/Korean characters) and session type (pretest/posttest). If a significant interaction was found, we further analyzed the interaction using *R* with the package *phia* (Post-Hoc Interaction Analysis; Martínez, 2015). If a significant interaction was not present, we removed the interaction term from our model and ran the analysis again, and then we performed post hoc comparisons using *R* with the package *multcomp*. The reported *p* values were adjusted for multiple comparisons within each analysis.

Results

Figure 3 shows the group-averaged progress on the recognition of Korean trigrams across the training blocks. The slope of the linear fit (accuracy against training block) was significantly larger than 0 ($p < 0.001$), indicating significant improvement as training progressed. The average accuracy of letter recognition improved from 0.32 in the first block to 0.50 in the last training block, from 0.39 on the first training day to 0.46 on the last day, and from 0.36 in the pretest to 0.45 in the posttest.

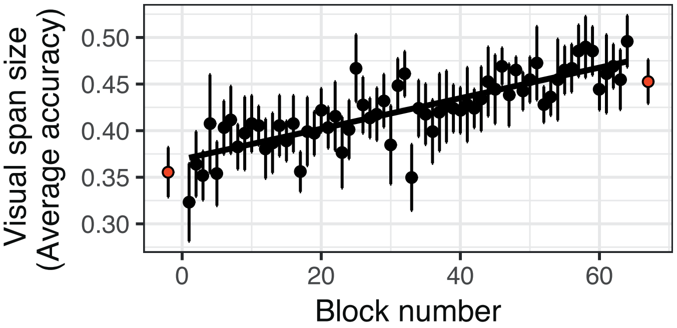

Figure 3. Training progress. Group-averaged recognition accuracy for Korean trigrams is plotted as a function of training block number. Red dots, data from pre- and posttests. Black dots, data from training blocks. Black line, linear fit of training data. Error bars: ±1 *SEM*.

In the following sections, we will first examine how training changed the size of the visual spans for the five types of symbols, then evaluate the level of crowding using our two-stage model, and lastly investigate the influence of complexity on the size of the visual span. To preview, we found that the visual span for Korean trigrams enlarged after training, and the enlargement appeared to transfer to untrained Korean and English symbols. The two-stage model revealed that training reduced the within-character and between-character crowding in Korean recognition, as well as between-letter crowding in English letter recognition. The size of visual spans negatively correlated with the pattern complexity of the symbols and ranked as English letters > Korean components > Korean characters = English trigrams > Korean trigrams.

Enlargement of visual span

We first compared the visual span profiles between the pretest and the posttest for the five types of symbols. From left to right, Figure 4 shows visual span profiles for English letters, Korean components, Korean characters, English trigrams, and Korean trigrams. After training, the visual spans for Korean characters, Korean trigrams, and English trigrams all enlarged, with their profiles moving up and broadening. For English letters and Korean components, their visual spans did not exhibit any noticeable changes due to high performance in the pretest.

Figure 5 summarizes the changes in average recognition accuracy from pre- to posttest for the five types of stimuli. A repeated ANOVA of symbol type (English letters/English trigrams/Korean components/Korean characters/Korean trigrams) × session type (pretest/posttest) was performed on the logit-transformed average recognition accuracy, which represents the size of the visual span. There was no significant interaction,

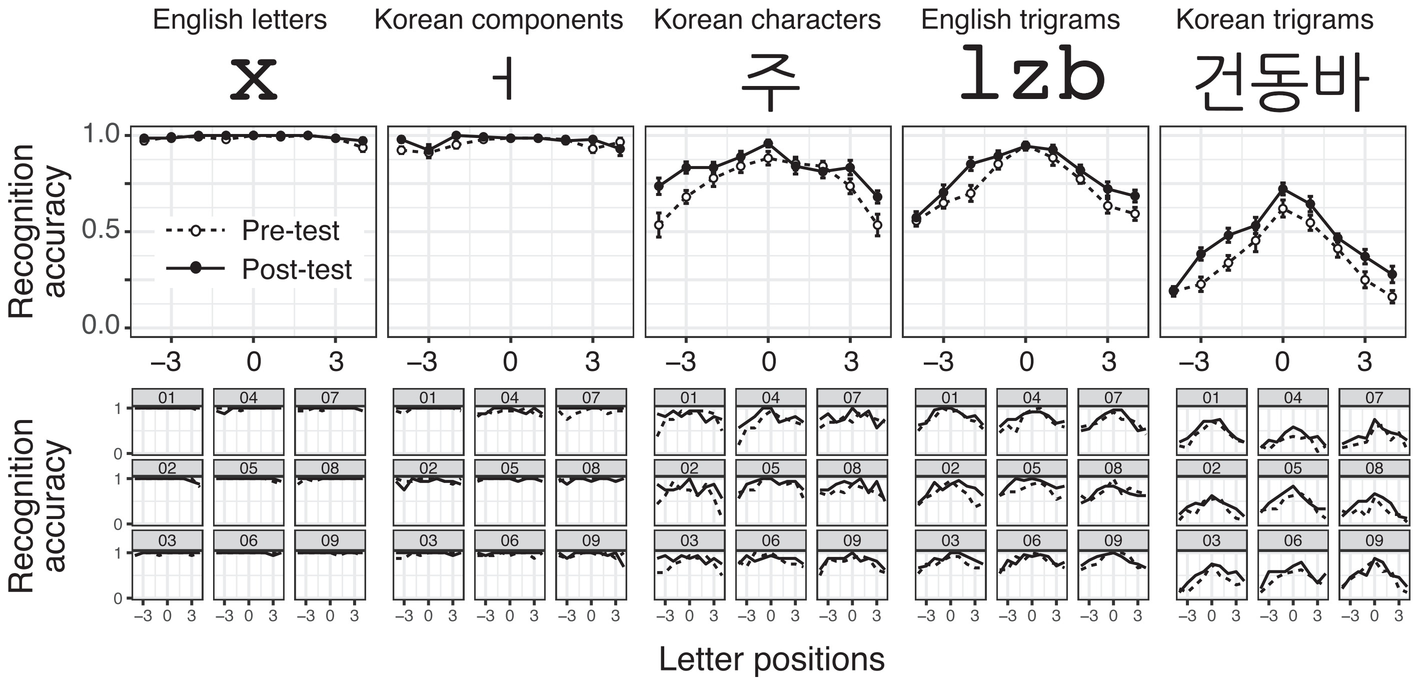

Figure 4. Korean and English visual span profiles before and after training. (Top) Average visual span profiles across all subjects. (Bottom) Individual data. Dashed lines and open symbols: pretest. Solid lines and filled symbols: posttest. Error bars: ±1 *SEM*.

so we removed the term from our model. In the updated model, we found significant main effects of symbol type, $F(4, 76) = 586.03$, $p < 0.001$, and session type, $F(1, 76) = 38.22$, $p < 0.001$. In general, the average accuracy was the highest for English letters (98.4% and 99.2% in the pre- and posttests, or 98.3% and 99.2% corrected for guessing), followed by Korean components (95.7% and 97.2%, or 95.5% and 97.1% corrected for guessing), Korean characters (74.2% and 82.4%, or 74.1% and 82.3% corrected for guessing), letters in English trigrams (73.2% and 79.1%, or 72.1% and 78.3% corrected for guessing), and characters in Korean trigrams (35.5% and 45.3%, or 35.3% and 45.1% corrected for guessing). Post hoc comparisons between the symbol types showed that the logit-transformed average accuracy (i.e., the size of visual spans) ranked as English letters > Korean components > Korean characters = letters in English trigrams > characters in Korean trigrams (">" signs mean "significantly larger than," and all of the adjusted $p < 0.001$).

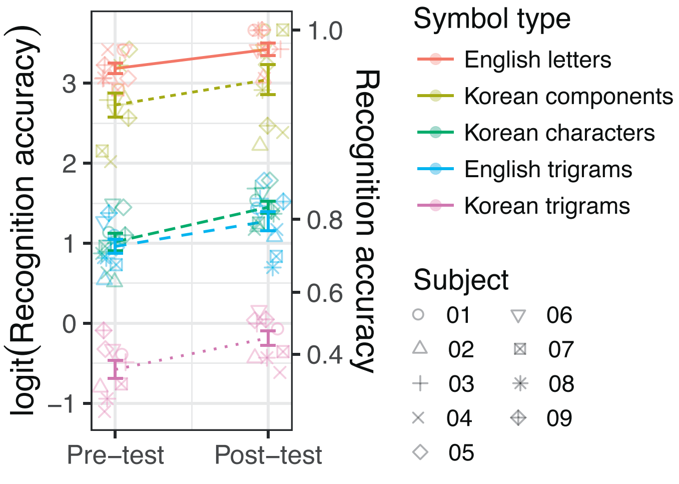

Figure 5. Changes in average recognition accuracy from pre- to posttest. From the top to the bottom: English letters (red), Korean components (olive), Korean characters (green), English trigrams (blue), and Korean trigrams (magenta). Faded points: individual data, with some jitter in the horizontal direction to increase discriminability. Error bars: ±1 *SEM*.

The parallel lines in Figure 5 and the lack of interaction between symbol type and session type indicate that training enlarged the visual span for all symbol types similarly (in the logit space). Although the exact factors underlying the transfer effect remain unclear, our results suggest a rather complete transfer of training effect from the trained symbol to all other nontrained symbols. But the absolute improvement in accuracy was the largest for Korean trigrams (i.e., the trained symbol; average accuracy +9.7%), followed by Korean single characters (+8.2%), English trigrams (+5.9%), Korean components (+1.5%), and English letters (+0.8%). The differences in the absolute value of improvement are expected considering the difference in the baseline sizes of visual spans before training.

Taken together, the results showed that our training successfully enlarged the visual span for the trained symbols. More important, this training appeared to transfer to untrained English symbols. We will discuss

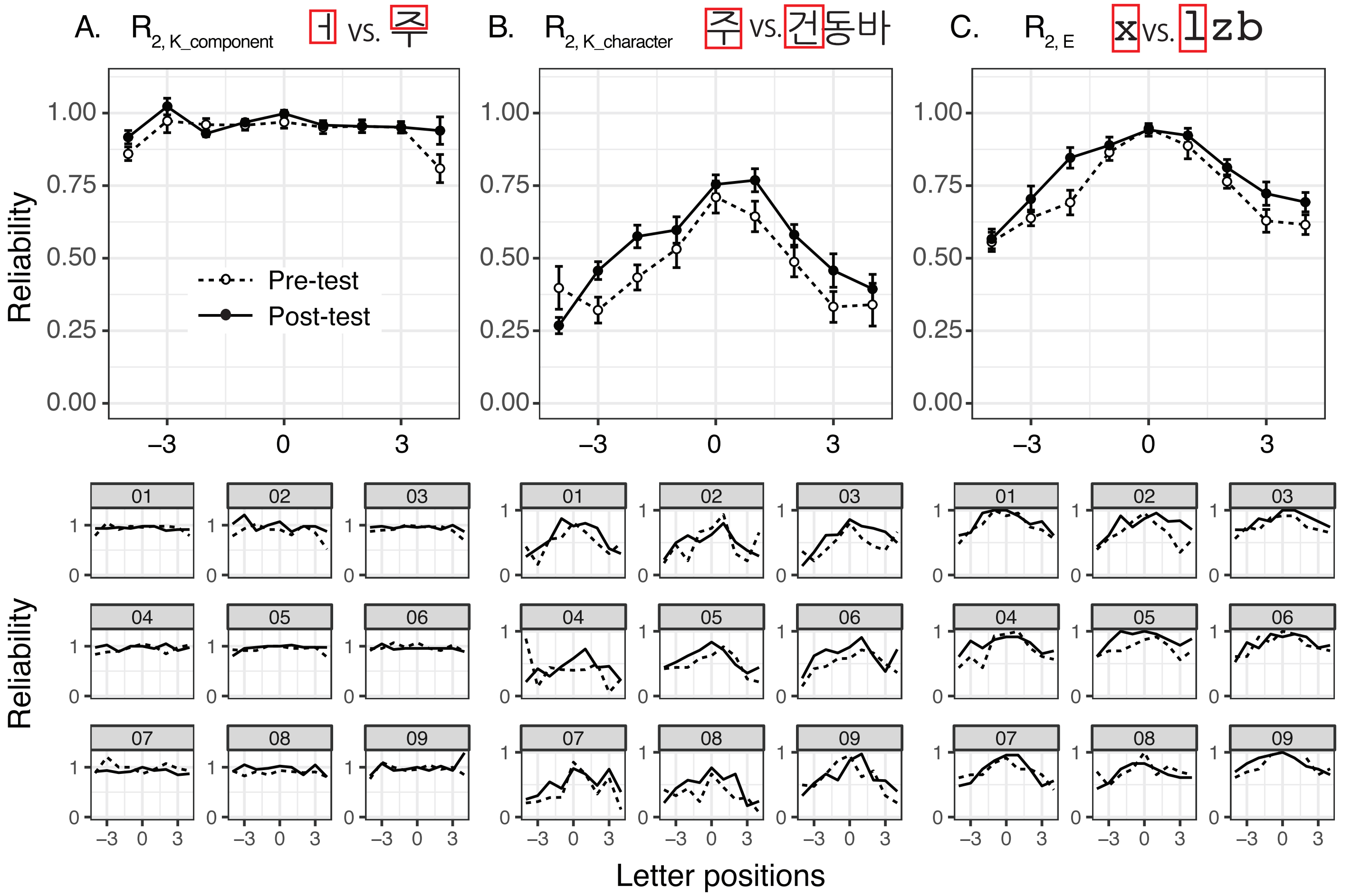

Figure 6. R_2 profiles before and after training. Each panel contains R_2 profiles, plots of R_2 reliability value versus letter position. (A) Profiles for $R_{2,\ K\_component}$, reflecting the effect of between-component, within-character crowding on Korean components. (B) Profiles for $R_{2,\ K\_character}$, reflecting the effect of between-character crowding on Korean characters. (C) Profiles for $R_{2,\ E}$, reflecting the effect of between-letter crowding on English letters. Small panels on the bottom show individual curves. Dashed lines and open symbols, pretest; solid lines and filled symbols, posttest; error bars: ±1 *SEM*.

the potential mechanisms underlying the training effect in the Discussion section.

The reduction of crowding

Next, we used a two-stage model (see the Methods section) to evaluate the influence of crowding. We estimated the reliability of the two stages, R_1 and R_2, for three types of symbols: Korean components, Korean characters, and English letters. Table 1 summarizes the calculations. R_1 was equal to the corrected-for-guessing recognition accuracy when these symbols were in isolation, which we refer to as the "acuity effect." In the pre- and posttests, $R_{1,\ K\_component} = 95.5\%$ and 97.1%, $R_{1,\ K\_character} = 74.1\%$ and 82.3%, and $R_{1,\ E} = 98.3\%$ and 99.2%, respectively. In the following paragraphs, we will focus on R_2 (which reflects the influence of crowding) and its changes after training.

Figure 6 shows R_2 profiles for Korean components, Korean characters, and English letters in the pre- (open symbols, dashed lines) and posttests (filled symbols, solid lines).

Figure 6A illustrates the effect of within-character crowding for recognizing Korean components. $R_{2,\ K\_component}$ was close to 1 in both the pretest (mean value 93.2%) and the posttest (mean value 96%). This indicates that within-character crowding was small overall, and a small yet noticeable reduction in within-character crowding occurred after training.

Figure 6B illustrates the effect of between-character crowding on recognizing Korean characters. $R_{2,\ K\_character}$ was the highest on the midline and declined rapidly as the letter position moved farther away from midline, indicating more severe crowding for more eccentric locations. In the pretest, $R_{2,\ K\_character}$ ranged from 35.5% to 59.3%, with a mean of 46.6%. In the posttest, it increased notably across almost all letter positions, ranging from 42.7% to 62.8% with a mean of

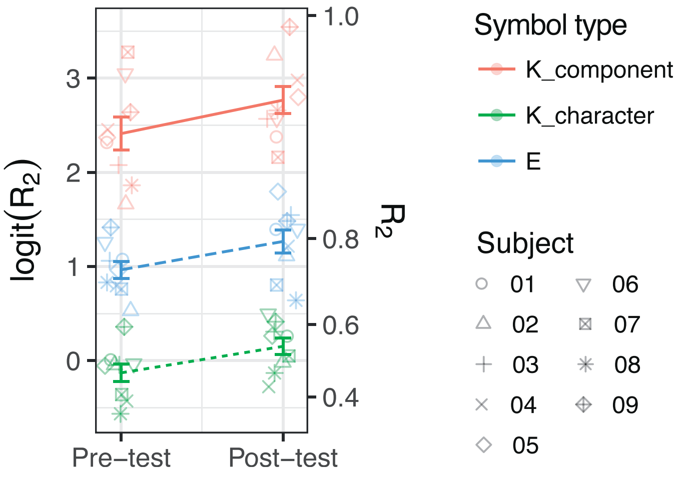

Figure 7. Changes in R_2 from pre- to posttest. From top to the bottom: $R_{2,\ \mathrm{K\_component}}$ (red), $R_{2,\ \mathrm{E}}$ (blue), and $R_{2,\ \mathrm{K\_character}}$ (green). Faded points: individual data, with some jitter in the horizontal direction to increase discriminability. Error bars: ±1 *SEM*.

53.9%. This shows that between-character crowding also reduced as a result of training, and this reduction appeared to be larger than the reduction of within-character crowding.

Figure 6C shows changes of $R_{2,\ \mathrm{E}}$, which reflects the effect of between-letter crowding on recognizing English letters. The profiles exhibit a very similar shape as the profiles for Korean characters in panel B but with larger R_2 values: In the pretest, $R_{2,\ \mathrm{E}}$ ranged from 63.6% to 82.1%, with a mean of 73.3%. In the posttest, it increased notably across almost all letter positions, ranging from 66.3% to 87.6%, with a mean of 78.9%.

Figure 7 shows the changes in the average R_2 values across letter positions from pre- to posttest for Korean components, Korean characters, and English letters. A two-factor repeated-measures ANOVA was performed on the logit-transformed R_2 values, with two within-subject factors being symbol type (Korean components/Korean characters/English letters) and session type (pretest/posttest). No significant interaction was found, so we removed the term from the model. In the updated model, we found significant main effects of both symbol type, $F(2, 42) = 310.12$, $p < 0.001$, and session type, $F(1, 42) = 13.54$, $p < 0.001$. Further post hoc comparisons showed that $R_{2,\ \mathrm{K\_component}} > R_{2,\ \mathrm{E}} > R_{2,\ \mathrm{K\_character}}$ (both of the adjusted $p < 0.001$). This indicates that crowding is the least severe between Korean components within a character, more severe for English letters within a trigram, and the most severe for Korean characters within a trigram.

After training, R_2 increased for all types of symbols (+2.8% for Korean components, +7.3% for Korean characters, and +5.6% for English letters). Despite the numerical differences in the improvement, the lack of an interaction between symbol type and session type indicates that training reduced crowding for all the symbols similarly, suggesting a successful transfer of training benefit from Korean to English. We did notice that the improvement in $R_{2,\ \mathrm{E}}$ here was smaller than in a previous study where native English speakers were trained to read English trigrams (+13.5%, analyzed from He et al., 2013). However, because the two studies differ in various aspects such as subject group, trained symbols, reporting method (full vs. partial report), and print size, further investigation is needed to resolve the discrepancy between studies.

The influence of pattern complexity

As shown in Figure 4, the visual span profiles for English letters and Korean components were very similar and close to perfect recognition, but the profile for Korean characters was narrower and had a lower peak, indicating a smaller size. The visual span for English trigrams was similar to that for Korean characters but larger than that for Korean trigrams. To understand whether the size of the visual span is limited by physical properties of the stimuli regardless of the language, we examined the relationship between the size of the visual span and pattern complexity.

We used perimetric complexity (Attneave, Arnoult, & Attneave, 1956; Pelli, Burns, Farell, & Moore-Page, 2006) to quantify pattern complexity. For a binary figure, perimetric complexity is defined as the square of the perimeter of the figure divided by the ink area of the figure:

$$\mathit{Perimetric\ Complexity} = \frac{\mathit{Perimeter}^2}{\mathit{Ink\ Area}}. \quad (3)$$

This measure of complexity "tends to capture how convoluted a character is, and is easily computed, independent of size, and additive, i.e., the perimetric complexity of two equal-area adjacent objects (considered as one) is equal to the sum of their individual complexities" (Pelli et al., 2006, p. 4648).

Figure 8 shows the relationship between average perimetric complexity of the five types of symbols and their visual span sizes before and after training. From the most simple to the most complex, the average perimetric complexity of the sets of symbols ranked as Korean components (96), English letters (102), Korean characters (243), English trigram (307), and Korean trigram (729). Their corresponding average accuracy for visual-span profiles was 95.7%, 98.4%, 74.2%, 73.2%, and 35.5% in the pretest and 97.2%, 99.2%, 82.4%, 79.1%, and 45.3% in the posttest. In most cases, more complex patterns corresponded to lower recognition accuracy, but note that there was a small reversal for Korean components and English letters: Korean components had lower complexity compared with

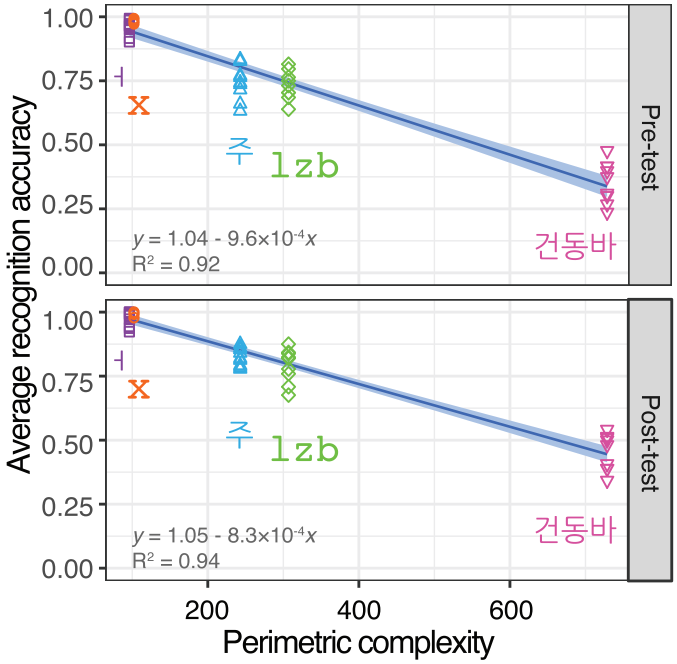

Figure 8. Perimetric complexity analysis. Each data point represents one subject's visual span size in one of the following conditions: Korean components (perimetric complexity = 96), English letters (102), Korean characters (243), English trigrams (307), and Korean trigrams (729). The straight line represents the best linear fit to the data, with the shaded area showing 95% confidence intervals.

English letters but had slightly poorer recognition performance. This reversal may be due to the higher pattern similarity between Korean components, but as we discuss in Appendix 2, similarity does not seem to be the dominant factor explaining the difference in recognition accuracy across symbol types.

Despite this reversal, there was a high negative correlation between perimetric complexity of the sets of symbols and the average accuracy in both the pretest ($r = -0.96$) and the posttest ($r = -0.97$). The slopes for the linear regressions were -9.6×10^{-4} in the pretest and -8.3×10^{-4} in the posttest, meaning that when perimetric complexity increases by 100, average recognition accuracy will decrease by 9.6% in the pretest and 8.3% in the posttest. This indicates that pattern complexity has a negative impact on recognition performance, but training can mitigate this negative impact.

Discussion

In the current study, we sought to test whether training-related enlargement of the visual span transfers to untrained symbols and whether similar sensory and physical constraints limit the size of Korean and English visual spans.

Visual span and its enlargement following training

Our first finding is that training to read Korean characters enlarged the Korean visual span and also transferred to the English visual span. We have therefore rejected the hypothesis that training sharpened the templates for the symbols used in the training task. Then, what may explain the training-related enlargement of the visual span observed in our study?

First, can general improvement in peripheral attention alone explain the enlargement of the visual span? Selective attention has been shown to enhance visual performance on various tasks (Bashinski & Bacharach, 1980; Cameron, Tai, & Carrasco, 2002; Carrasco, Ling, & Read, 2004; Carrasco & McElree, 2001; T. Liu, Pestilli, & Carrasco, 2005), including letter recognition (Talgar, Pelli, & Carrasco, 2004). Previous studies have found concurrent improvements in both peripheral attention and peripheral letter recognition after training in peripheral vision (Lee, Kwon, Legge, & Gefroh, 2010; R. Liu & Kwon, 2016). Thus, it is possible that improved peripheral attention following training might have enhanced a person's ability to recognize a target letter in the periphery. We, however, do not think it is likely to be the case for our study: Using a similar training protocol to ours, Lee et al. (2010) indeed examined whether or not better deployment of attention following peripheral vision training explained the training benefit (i.e., an enlargement of the visual span). Although they found significant improvements in both the deployment of attention and trigram letter recognition following the training, no correlation was found between the two (Lee et al., 2010). Furthermore, their no-training control group exhibited no significant enlargement of the visual span yet showed a significant improvement in the deployment of attention. Although we cannot rule out the potential contribution of attention in general, the findings of Lee et al. (2010) suggest that the deployment of attention is not likely to be the major contributor of the observed training effect. Nevertheless, future studies with carefully designed sham-training groups would help us further elucidate the factors mediating the training effect.

Some may argue that even in the absence of training, simply having subjects take the tests twice (i.e., test-retest benefits) might have led to the observed enlargement of the visual span. However, previous studies with similar training paradigms

(Bernard, Arunkumar, & Chung, 2012; Chung et al., 2004; He et al., 2013; Lee et al., 2010; Yu et al., 2010) have already shown that no-training control groups do not exhibit significant improvements from pre- to posttest (e.g., 1.1% nonsignificant improvement in average accuracy for trigram recognition; He et al., 2013), suggesting a minimal or no effect of the test-retest benefit at least for the peripheral letter-recognition studies like ours.

Although the exact mechanism underlying the training effect remains unclear, here are some potential explanations why there was a near complete transfer effect of training from Korean symbols to English symbols: One possible mechanism is that training reduced crowding by adjusting the spatial extent of the perceptual window for template matching (Sun et al., 2010). Another possibility is that the subjects became better at detecting and recognizing features that are shared across Korean and English symbols. Suchow and Pelli (2013) used carefully designed "letters" to separate feature detection and feature combination in human object recognition. They found that feature detection and combination are two independent steps for human object recognition, and both can be improved with training. When the training uses familiar patterns (such as letters and characters in our study), the improvement in recognition occurs mostly in feature detection. Therefore, instead of learning better templates for letters/characters, subjects may have learned better templates for features. Indeed, perceptual templates can be improved after training to have higher sampling efficiency and to better extract useful information for recognition in the presence of noise (Chung, Levi, & Tjan, 2005; Gold et al., 2004). If the subjects learned better templates for the most important features that overlap across symbol sets, training benefits would likely transfer between scripts as demonstrated in the current study, as well as between print sizes and visual fields as shown in previous studies (Chung et al., 2004; He et al., 2013; Yu et al., 2010). Possible shared features include line terminations and vertices, which have been shown to be crucial for English letter recognition (Fiset et al., 2008; Lanthier, Risko, Stolz, & Besner, 2009) and may also be important features for Korean recognition.

Our study examined only the transfer of training effect from Korean to English symbols, but the transfer may not be bidirectional if learning to recognize Korean and English characters involves nonoverlapping stages. Future research in the reverse direction, from English to Korean symbols, would help to better quantify the level of transfer and to further examine the relationship between task difficulty and transfer, thereby shedding light on the underlying mechanism of enlarged visual span.

Pattern complexity influences the size of the visual span

Our second finding is that the size of the visual span is influenced by the complexity of the pattern, regardless of the language. Below, we will discuss how complexity affects isolated and crowded pattern recognition, respectively. Another related influencing factor is the similarity of the symbols within a set, which we will discuss in Appendix 2.

Increased pattern complexity was associated with lower accuracy in recognition. For isolated symbols, complexity level ranked as English letter ≈ Korean component < Korean character, and the averaged visual-span accuracy ranked in the opposite order (i.e., Korean character < Korean component ≈ English letter). Korean characters are made up of multiple components. Components within a Korean character introduced within-character crowding, reducing the size of the visual span. This is consistent with the findings that patterns with higher complexity have larger acuity sizes (Zhang, Zhang, Xue, Liu, & Yu, 2007). Presumably, in more complex patterns, the features crowd each other, so that larger between-feature spacing (therefore larger character size) is needed to escape between-feature crowding.

For crowded symbols (English and Korean trigrams), the sizes of their visual spans were much smaller than isolated symbols because of between-symbol crowding. Judging from the values of $R_{2\_K,\ character}$ and R_{2_E}, it seems that more complex symbols (Korean characters) have a smaller second-stage reliability and are more influenced by between-symbol crowding. Other studies have also found that more complex patterns are susceptible to more severe crowding (Bernard & Chung, 2011; Wang et al., 2014; Zhang et al., 2009).

Our findings here suggest that complexity has a major role in determining the size of the visual span. An increase in pattern complexity is associated with a reduction in the size of the visual span, primarily because of both increased within-symbol crowding and accumulated between-symbol crowding. Whether within- and between-symbol crowding have different neural origins remains to be investigated.

Conclusion

Our conclusion is that the training to recognize scripts for one language can enlarge its visual span, and the improvement transfers to another language, possibly because of shared features between scripts. The visual span describes a presymbolic constraint on pattern recognition. Its size is mainly influenced by

physical properties of the pattern such as perimetric complexity and pattern similarity and sensory constraints such as crowding, in a similar way across languages.

Keywords: visual span, reading, pattern complexity, pattern recognition, crowding

Acknowledgments

The authors would like to thank Charles Bigelow for valuable information regarding font metrics, Jihyang Jun for her help in recruiting subjects, and Seoyoung Park for her help with data collection. This study was supported by National Institutes of Health Grant EY002934.

Commercial relationships: MiYoung Kwon, P; Gorden E. Legge, Precision Vision (R).

Corresponding author: Yingchen He.
Email: hexxx340@umn.edu.
Address: Department of Psychology, University of Minnesota, Minneapolis, MN, USA.

References

Attneave, F., Arnoult, M. D., & Attneave, F. (1956). The quantitative study of shape and pattern perception. *Psychological Bulletin*, *53*, 452–471, doi.org/10.1037/h0044049.

Baek, S. R., He, Y., & Legge, G. E. (2016). Reading speed for Korean text in central and peripheral vision. *Investigative Ophthalmology and Visual Science*, *57*, 1948–1948. [Abstract]

Bashinski, H. S., & Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception & Psychophysics*, *28*, 241–248, doi.org/10.3758/BF03204380.

Bernard, J.-B., Arunkumar, A., & Chung, S. T. L. (2012). Can reading-specific training stimuli improve the effect of perceptual learning on peripheral reading speed? *Vision Research*, *66*, 17–25, doi.org/10.1016/j.visres.2012.06.012.

Bernard, J.-B., & Chung, S. T. L. (2011). The dependence of crowding on flanker complexity and target-flanker similarity. *Journal of Vision*, *11*(8):1, 1–16, doi.org/10.1167/11.8.1. [PubMed] [Article]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. Retrieved from http://bbs.bioguider.com/images/upfile/2006-4/200641014348.pdf

Cameron, E. L., Tai, J. C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research*, *42*, 949–967, doi.org/10.1016/S0042-6989(02)00039-1.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*, 308–313, doi.org/10.1038/nn1194.

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences, USA*, *98*, 5363–5367, doi.org/10.1073/pnas.081074098.

Chung, S. T. L., Legge, G. E., & Cheung, S.-H. (2004). Letter-recognition and reading speed in peripheral vision benefit from perceptual learning. *Vision Research*, *44*, 695–709, doi.org/10.1016/j.visres.2003.09.028.

Chung, S. T. L., Levi, D. M., & Tjan, B. S. (2005). Learning letter identification in peripheral vision. *Vision Research*, *45*, 1399–1412, doi.org/10.1016/j.visres.2004.11.021.

Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research*, *38*, 2949–2962.

Fiset, D., Blais, C., Éthier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for uppercase and lowercase letter identification. *Psychological Science*, *19*, 1161–1168.

Gervais, M. J., Harvey, L. O., & Roberts, J. O. (1984). Identification confusions among letters of the alphabet. *Journal of Experimental Psychology. Human Perception and Performance*, *10*, 655–666, doi.org/10.1037//0096-1523.10.5.655.

Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, *28*, 167–207, doi.org/10.1016/j.cogsci.2003.10.005.

He, Y., Legge, G. E., & Yu, D. (2013). Sensory and cognitive influences on the training-related improvement of reading speed in peripheral vision. *Journal of Vision*, *13*(7):14, 1–14, doi.org/10.1167/13.7.14. [PubMed] [Article]

He, Y., Scholz, J. M., Gage, R., Kallie, C. S., Liu, T., & Legge, G. E. (2015). Comparing the visual spans for faces and letters. *Journal of Vision*, *15*(8): 7, 1–13, http://doi.org/10.1167/15.8.7. [PubMed] [Article]

Kwon, M., & Legge, G. E. (2013). Higher-contrast requirements for recognizing low-pass–filtered letters. *Journal of Vision*, *13*(1):13, 1–15, http://doi.org/10.1167/13.1.13. [PubMed] [Article]

Lanthier, S. N., Risko, E. F., Stolz, J. A., & Besner, D. (2009). Not all visual features are created equal: Early processing in letter and word recognition. *Psychonomic Bulletin & Review*, *16*, 67–73, doi.org/10.3758/PBR.16.1.67.

Lee, H.-W., Kwon, M., Legge, G. E., & Gefroh, J. J. (2010). Training improves reading speed in peripheral vision: Is it due to attention? *Journal of Vision*, *10*(6):18, 1–15, http://doi.org/10.1167/10.6.18. [PubMed] [Article]

Legge, G. E., Cheung, S.-H., Yu, D., Chung, S. T. L., Lee, H., & Owens, D. P. (2007). The case for the visual span as a sensory bottleneck in reading. *Journal of Vision*, *7*(2):9, 1–15, doi.org/10.1167/7.2.9. [PubMed] [Article]

Liu, R., & Kwon, M. (2016). Integrating oculomotor and perceptual training to induce a pseudofovea: A model system for studying central vision loss. *Journal of Vision*, *16*(6): 10, 1–21, doi:10.1167/16.6.10. [PubMed] [Article]

Liu, T., Pestilli, F., & Carrasco, M. (2005). Transient attention enhances perceptual performance and fMRI response in human visual cortex. *Neuron*, *45*, 469–477, doi.org/10.1016/j.neuron.2004.12.039.

Loomis, J. M. (1990). A model of character recognition and legibility. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 106–120, doi.org/10.1037/0096-1523.16.1.1066.

Martínez, H. D. R. (2015). *Analysing interactions of fitted models*. Retrieved from http://cran.wustl.edu/web/packages/phia/vignettes/phia.pdf

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. http://doi.org/10.1163/156856897X00366

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*, 4646–4674, doi.org/10.1016/j.visres.2006.04.023.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, *7*(2):20, 1–36, doi.org/10.1167/7.2.20. Modeling. [PubMed] [Article]

Suchow, J. W., & Pelli, D. G. (2013). Learning to detect and combine the features of an object. *Proceedings of the National Academy of Sciences*, *USA*, *110*, 785–790, doi.org/10.1073/pnas.1218438110.

Sun, G. J., Chung, S. T. L., & Tjan, B. S. (2010). Ideal observer analysis of crowding and the reduction of crowding through learning. *Journal of Vision*, *10*(5): 16, 1–14, doi.org/10.1167/10.5.16. [PubMed] [Article]

Talgar, C. P., Pelli, D. G., & Carrasco, M. (2004). Covert attention enhances letter identification without affecting channel tuning. *Journal of Vision*, *4*(1): 3, 1–10, doi.org/10.1167/4.1.3. [PubMed] [Article]

Wang, H., He, X., & Legge, G. E. (2014). Effect of pattern complexity on the visual span for Chinese and alphabet characters. *Journal of Vision*, *14*(8):6, 1–17, doi.org/10.1167/14.8.6.doi. [PubMed] [Article]

Yu, D., Legge, G. E., Park, H., Gage, E., & Chung, S. T. L. (2010). Development of a training protocol to improve reading performance in peripheral vision. *Vision Research*, *50*, 36–45, doi.org/10.1016/j.visres.2009.10.005.

Zhang, J.-Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2007). Legibility variations of Chinese characters and implications for visual acuity measurement in Chinese reading population. *Investigative Ophthalmology & Visual Science*, *48*, 2383–2390, doi.org/10.1167/iovs.06-1195. [PubMed] [Article]

Zhang, J.-Y., Zhang, T., Xue, F., Liu, L., & Yu, C. (2009). Legibility of Chinese characters in peripheral vision and the top-down influences on crowding. *Vision Research*, *49*, 44–53, doi.org/10.1016/j.visres.2008.09.021.

Appendix 1. Korean stimuli

Although Korean components can be directly "typed" onto the backdrop in MATLAB and cut using the default bounding box for the specified size (282 pixel [W] by 345 pixel [H]), we did not adopt this method. Our concern was that components typed in this way were generally larger than they appear in characters (for example, ㅐversus the upper-right component in the character 낼) and thus would not allow us to test the real acuity limit for Korean recognition. We therefore first generated images of Korean characters and then cut component images from them. Because of the variations in the structure of the characters (two or three components, vertical or horizontal), the components within a character could vary in size and shape. A component in a two-component character may appear bigger than the same component in a three-component character (for example, the component ㄴ in 내 versus in 낼). To make sure that we could test the acuity limit in recognizing Korean characters, we chose three-letter characters to cut our testing images of Korean letters

가 각 간 갈 감 갑 갔 강 같 개 거 건
걸 겁 것 게 겠 겨 경 계 고 곳 공 과
관 교 구 귀 그 금 기 긴 길 까 깨 께
꼭 끄 끼 나 난 날 남 내 너 넘 넣 네
노 놀 놓 누 눈 느 는 늘 니 님 다 단
달 답 당 대 더 던 데 도 돈 돌 동 돼
되 된 두 드 든 들 디 따 때 떨 라 락
란 람 랑 래 러 럽 렇 레 려 렸 로 록
루 르 른 를 름 리 린 림 마 만 많 말
맛 망 맞 매 머 먹 며 면 명 모 못 무
문 물 미 바 반 받 발 밥 방 배 버 번
보 복 부 분 불 비 빠 빨 쁜 사 산 살
상 새 색 생 서 선 성 세 셨 소 속 손
수 스 습 시 식 신 실 싫 심 싶 쓰 씨
아 안 앉 않 알 았 야 약 양 어 언 얼
엄 없 었 에 여 연 열 였 영 예 오 온
올 옷 와 왔 요 용 우 운 울 웃 워 원
위 유 으 은 을 음 의 이 인 일 입 있
자 작 잘 잠 잡 장 재 저 적 전 절 점
접 정 제 져 졌 조 종 좋 주 줄 중 지
직 진 질 집 차 참 찾 책 처 추 치 친
침 커 크 키 타 터 트 틀 파 편 프 피
하 학 한 할 함 합 항 해 했 행 혼 화
회 후 히

Figure A1-1. Full list of 279 Korean characters used in the current study.

(see Figure A1-2). When generating the characters for cutting consonants, we kept the vowel and the tail components unchanged while varying only the lead consonant. Similarly, when generating the characters for cutting vowels, we kept the lead and the tail consonants unchanged while varying only the vowel. In this way, we intended to minimize the influence of letter configuration on the shape of the components we cut.

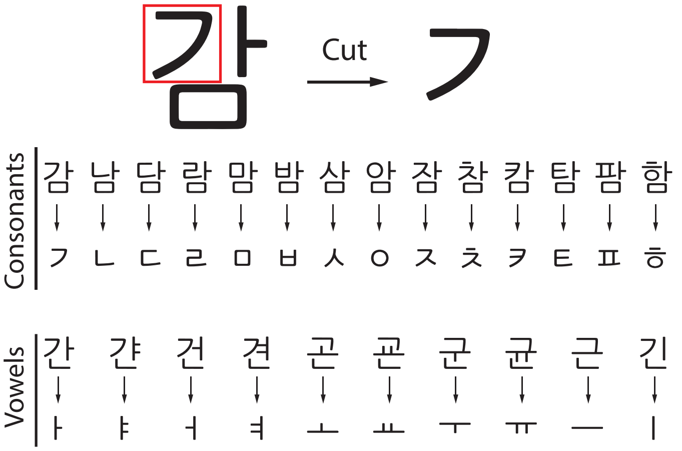
감
Cut
ㄱ
Consonants
감 남 담 람 맘 밤 삼 암 잠 참 캄 탐 팜 함
ㄱ ㄴ ㄷ ㄹ ㅁ ㅂ ㅅ ㅇ ㅈ ㅊ ㅋ ㅌ ㅍ ㅎ
Vowels
간 갼 건 견 곤 굔 군 균 근 긴
ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ

Figure A1-2. Generation of Korean component stimuli.

Appendix 2. Pattern similarity

In addition to complexity, the similarity between stimuli in the symbol sets also influences recognition. Similarity within a given symbol set is sometimes defined functionally by a confusion matrix between the items in the set. Theoretical confusion matrices derived from pairwise similarity scores can largely explain the legibility of visual or tactile symbols (Loomis, 1990) and the decrease in legibility for degraded visual symbols (Kwon & Legge, 2013). The similarity between alternatives is also a good predictor of the contrast threshold for a template-matching ideal observer (Pelli et al., 2006). For human observers recognizing crowded targets, higher target-flanker similarity results in more identification errors and mislocation (reporting the correct identity of a flanker instead of the target) errors (Bernard & Chung, 2011). It is thus important to consider the role of similarity when interpreting our results. We quantified the similarity for each symbol set by computing the average pairwise Euclidean distance in pixel space (Gervais, Harvey, & Roberts, 1984) within that set. There are more advanced models (Loomis, 1990) or empirical methods to quantify similarity, but we adopted Euclidean distance as a purely physical measure for its parallel nature with perimetric complexity.

Here we found that although Korean components had slightly lower perimetric complexity than English letters (96 vs. 102), recognition performance was slightly worse than that for English letters (averaged accuracy 95.7% vs. 98.4% before training; significantly different after transformation; see the Results section). There was a higher similarity (i.e., smaller Euclidean distance) between Korean components and English letters (Figure A2-1). For Korean component recognition, when compared with English letter recognition,

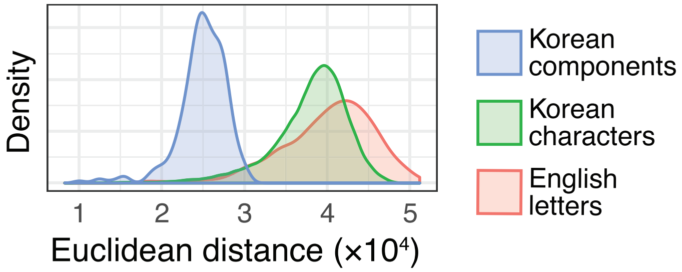

Figure A2-1. Distribution of similarity scores (pairwise Euclidean distance). Each profile shows the distribution of pairwise Euclidean distance of two symbols within a set. The larger the Euclidean distance, the less similar they are. Red, English letters. Green, Korean characters. Blue, Korean components.

the disadvantage of higher similarity seemed to offset the potential benefit of smaller perimetric complexity.

To determine the best predictors for recognition accuracy, we ran a multiple regression model in which mean perimetric complexity, pattern similarity (mean Euclidean distance), and stimulus type (isolated/trigram) were entered as predictors. Here we report just the results from the pretest data as the posttest data showed a similar pattern of results. As expected, complexity was a significant contributor ($p < 0.001$): Recognition accuracy decreased with increased complexity. But neither similarity nor being a trigram had any effect on recognition accuracy ($p > 0.5$). This point was further highlighted in Figure A2-1: The within-set similarities for English letters and Korean characters are both smaller than that for Korean components, but English letters were recognized better than Korean components. It is possible that our design lacked power to detect a potential effect of similarity, and the unequal set size makes it more difficult to interpret the results. But it appears that the average within-set pattern similarity did not play a major role in determining the recognition performance of the stimuli tested in the current study.

Although we discuss here the effect of complexity and similarity separately, they often interact in real-world situations. For example, natural scripts with higher complexity are often found to have less similarity (Pelli et al., 2006). To separate their effects, carefully controlled paradigms are needed.