Perceptual learning of contrast discrimination under roving: The role of semantic sequence in stimulus tagging

Lin-Juan Cong
Department of Psychology, Peking University, Beijing, China

Jun-Yun Zhang
Department of Psychology, Peking University, Beijing, China

Perceptual learning may occur when multiple contrasts are practiced in a fixed, but not in a roving (random), temporal sequence. However, learning may escape roving disruption when each contrast is assigned a letter tag (i.e., A, B, C, D). Because these letter tags carry not only stimulus identity information, but also semantic sequence information, here we investigated whether the semantic sequence information is necessary for learning of tagged contrasts under the roving condition. We found that assigning number tags (i.e., 1, 2, 3, 4), which also contained both identity and semantic sequence information, to four roving contrasts enabled significant learning of discrimination of each contrast, confirming previous data. However, learning became insignificant when the contrast tags were replaced with Greek letters that were familiar to our Chinese observers except their sequence or Chinese characters that carried no sequence information. In addition, assigning orientation tags, which carried no sequence information either, to roving contrasts was ineffective as well because learning occurred only with sequenced but not roving contrasts. These results suggest that semantic sequence information is necessary for stimulus tagging to effectively enable perceptual learning of multiple contrast discrimination under roving.

Introduction

Perceptual learning refers to improvement in discrimination of visual and other sensory features through training. In a typical perceptual learning study, only a single stimulus level (e.g., a specific contrast or tone) is practiced, and the performance is usually improved after a few sessions of training. However, if multiple levels of stimuli are being learned simultaneously, learning of several visual and auditory stimuli is found to occur when the stimuli are presented in a fixed temporal sequence (Kuai, Zhang, Klein, Levi, & Yu, 2005; Nahum, Nelken, & Ahissar, 2010; Zhang et al., 2008) but not in a roving order (Adini, Wilkonsky, Haspel, Tsodyks, & Sagi, 2004; Nahum et al., 2010; Yu, Klein, & Levi, 2004).

However, several studies have shown that perceptual learning can escape roving disruption when the roving stimuli are sufficiently different (Tartaglia, Aberg, & Herzog, 2009; Zhang et al., 2008) or are presented in different retinal locations (Otto, Herzog, Fahle, & Zhaoping, 2006), probably to reduce the overlap of corresponding neural populations. Long-lasting training is also shown to help learning under stimulus roving (Parkosadze, Otto, Malania, Kezeli, & Herzog, 2008). Specific to the current study, our previous study also shows that learning can escape roving disruption when each roving stimulus is given a letter tag (Zhang et al., 2008). The observers are unable to learn the discrimination of four roving contrasts (i.e., 0.20, 0.30, 0.47, and 0.63). But significant learning becomes evident when a distinct letter tag (i.e., A, B, C, or D) is assigned to each contrast and is shown before a relevant trial. Additional evidence shows that contrast thresholds are unaffected by stimulus sequencing or roving at the beginning of training (Adini et al., 2004; Zhang et al., 2008) or after learning is established through sequenced stimulus presentations (Kuai et al., 2005). Therefore, it is not the encoding and retrieval stages, but the consolidation stage, of information processing that are more affected by contrast roving. We suggest that the letter tags may help the brain to attend to the responses of proper sets of neurons for learning consolidation (Zhang et al., 2008).

One issue that remains unsolved is whether the above letter tagging acts like stimulus sequences. The A, B, C, D letter tags carry both stimulus identity and semantic information.
sequence information. Is it the identity tagging or semantic sequence tagging that really matters in learning of multiple roving contrasts? In this study, we separated the roles of these two factors, and our results indicate that it is the semantic sequence information that plays a critical role.

**Methods**

**Observers and apparatus**

A total of 36 observers (undergraduate students in their early 20s) with normal or corrected-to-normal vision participated in different experiments of this study. All were inexperienced to psychophysical experiments and were unaware of the purposes of the study. Informed written consent with IRB approval was obtained from each observer before data collection.

The stimuli were generated by a Matlab-based WinVis program (Neurometrics Institute, Oakland, CA) and presented on a 21-in. Dell P1130 CRT monitor (1024 pixel × 768 pixel resolution, 0.38 mm × 0.38 mm pixel size, 75 Hz frame rate). The mean luminance of the monitor was 53.7 cd/m². The luminance of the monitor was linearized by an eight-bit lookup table. A chin-and-head rest helped stabilize the head of the observer. Experiments were run in a dimly lit room. Viewing was binocular at a distance of 3 m.

**Stimuli and procedure**

The stimuli (Figure 1a) were foveal Gabor patches (Gaussian enveloped sinusoidal grating) presented on a mean luminance screen background. The Gaussian envelope of the Gabor patches had a standard deviation of 0.17°. The sinusoidal grating carrier had a spatial frequency at 6 c/° and a phase at 90°. The four reference contrasts were 0.20, 0.30, 0.47, and 0.63 (actually 0.202, 0.295, 0.473, and 0.632 due to the eight-bit contrast resolution). The Gabors were vertical except in Experiment 4 in which they had various orientations serving as orientation tags.

In a contrast discrimination trial, the test and reference stimuli were separately presented in two 92-ms stimulus intervals in a random order separated by a 500-ms interstimulus interval. In Experiments 1–3, a number, Greek letter, or Chinese character tag of 300 ms preceded the first stimulus interval with a 200-ms
time gap. In Experiment 4, the stimulus orientations per se were used as stimulus tags. A fixation cross preceded each trial by 300 ms and disappeared 250 ms before the onset of the first stimulus interval. The observers’ task was to judge which stimulus interval contained a higher contrast Gabor. Auditory feedback was given on incorrect responses.

The contrast discrimination thresholds were estimated with a temporal two-alternative forced choice (2AFC) staircase procedure. The staircases followed a three-down-one-up staircase rule that resulted in a 79.4% convergence rate. The initial contrast difference was sufficiently large so that the observers could always make correct discrimination. Each staircase consisted of four preliminary reversals and six experimental reversals (approximately 50 trials). The step size of the staircase was 0.05 log units. The geometric mean of the experimental reversals was taken as the threshold. Because of the limited contrast resolution (eight bit), there were up to approximately 0.05% error of the posttraining thresholds at the lowest pedestal contrast (0.20), and up to approximately ±8% error of the posttraining thresholds at the highest pedestal contrast (0.63).

During testing, the staircases for all reference contrasts were randomly interleaved trial by trial unless otherwise stated. The intertrial interval (ITI) was 2000 ms in Experiments 1–3, which included a 1000-ms (randomized from 0 ms to 2000 ms in Experiment 3) delay after an observer pushed a button to respond to a previous trial, a 300-ms presentation of the fixation cross, a 200-ms time gap between the fixation and the onset of the 300-ms stimulus tag, and a 200-ms time gap between the tag and the onset of the first stimulus interval of the next trial. The ITI in Experiment 4 was 1550 ms, which included a 1000-ms delay after an observer pushed a button to respond to a previous trial, a 300-ms presentation of the fixation cross, and a 250-ms time gap between the fixation and the onset of the first stimulus interval of the next trial. The ITIs did not include the observer’s response time, which could add another few hundred milliseconds. When a specific staircase ended, the stimuli would still be presented at the last step value until all interleaving staircases were completed. An observer typically completed five repeats of interleaved staircase runs in a 2-hr training session.

Results

Experiment 1: Perceptual learning of multiple roving contrasts with number tags

Previously, we showed that multistimulus perceptual learning can escape roving disruption when four roving contrasts were each assigned an English letter tag: A, B, C, and D, respectively (Zhang et al., 2008). To replicate this tagging effect, we repeated the same multicontrast discrimination experiment but assigned four roving contrasts each a number tag: 1, 2, 3, and 4 (Figure 1). The same four reference contrasts (0.2, 0.3, 0.47, and 0.63) were interleaved with constant 2-s ITIs. Five daily sessions of practice led to significant learning. The post/pretraining threshold ratios (PPRs) were 0.77 ± 0.07, 0.86 ± 0.08, 0.79 ± 0.05, and 0.88 ± 0.10 for 0.2, 0.3, 0.47, and 0.63 contrasts, respectively (PPR < 1 indicates reduced thresholds after practice). The overall PPR was 0.82 ± 0.04, which was significantly different from a no-learning PPR = 1, F(1, 5) = 14.6, p = 0.012 (repeated measures ANOVA), similar to the overall PPR (0.77 ± 0.04) in perceptual learning of the same four contrasts tagged with English letters in our previous study (Zhang et al., 2008).

Experiment 2: Perceptual learning of multiple roving contrasts with Greek letter and Chinese character tags

Letter and number tags carry both identity and sequence information. To investigate which information in stimulus tagging really helped learning of roving contrasts, we changed the stimulus tags to Greek letters (i.e., δ, ω, λ, and θ) or Chinese characters (i.e., 各, 未, 什, and 四). Our observers were Chinese college students. They were all familiar with these Greek letters from math and physics courses, but they typically had no knowledge of Greek letters’ alphabetic sequence. The four Chinese characters were common family names that carried no sequence information and were used here to double check the learning results with Greek letter tags.

Two groups of new observers, six observers per group, practiced discrimination of the same four roving contrasts with tags of Greek letters or Chinese characters for five sessions, respectively. However, this time, learning was insignificant. For the Greek letter tag group (Figure 2a, b), the PPRs were 0.95 ± 0.08, 0.89 ± 0.07, 1.04 ± 0.11, and 0.84 ± 0.06 for 0.2, 0.3, 0.47, and 0.63 contrasts, respectively. The overall PPR was 0.93 ± 0.04, not significantly different from a no-learning PPR = 1, F(1, 5) = 2.51, p = 0.17. For the Chinese character tag group (Figure 2c, d), the PPRs were 1.23 ± 0.15, 0.86 ± 0.06, 0.82 ± 0.05, and 0.91 ± 0.09 for 0.2, 0.3, 0.47, and 0.63 contrasts, respectively. The overall PPR was 0.96 ± 0.05, not significantly different from PPR = 1 either, F(1, 5) = 2.88, p = 0.15. These results indicate that stimulus tags carrying no sequence information are unable to induce significant contrast learning under stimulus roving.

We performed a comprehensive statistical analysis to examine the role of semantic sequence in stimulus...
tagging. The PPR data from the previous letter tag condition (Zhang et al., 2008) and number tag condition (Figure 1) were pooled into a semantic sequence group (SG), and those from the Greek letter tag and Chinese character tag conditions were pooled into a no semantic sequence group (NSG). A mixed-design ANOVA with sequence (SG vs. NSG) as a between-subjects factor and contrast (0.2, 0.3, 0.47, and 0.63) as a within-subjects factor showed a significant main effect of sequence, $F(1, 22) = 16.03, p = 0.001$, and an insignificant main effect of contrast, $F(3, 66) = 1.21, p = 0.31$. There was no significant interaction between sequence and contrast, $F(3, 66) = 2.27, p = 0.089$. The sequence main effect confirmed the impact of semantic sequence in stimulus tagging on learning of multiple roving contrasts.

**Experiment 3: The impact of stimulus rhythm on number tagging–enabled learning of multiple roving contrasts**

Previously, we demonstrated that a temporal rhythm is necessary for stimulus sequence to enable
perceptual learning of multiple roving contrasts. When the ITI was jitters, learning was disabled even if the stimulus sequence was unchanged (Zhang et al., 2008). Would semantic sequence be strong enough to evade this disruption? We repeated the first experiment (number-tagged roving contrasts) except that the ITIs were jittered from 1 s to 3 s (mean = 2 s). The jittered ITIs interrupted the stimulus rhythm but preserved the stimulus sequence information. The training results showed that the learning was unaffected. The PPRs were 0.77 ± 0.06, 0.81 ± 0.09, 0.95 ± 0.11, and 0.73 ± 0.04 for 0.2, 0.3, 0.47, and 0.63 contrasts, respectively. The overall PPR was 0.81 ± 0.04, $F(1, 5) = 12.9, p = 0.016$ (Figure 3a), in contrast to previous results that jittering ITIs interrupted learning (PPR = 0.99 ± 0.05; Zhang et al., 2008). To directly compare PPRs from these two experiments, we performed a mixed-design ANOVA with sequence (semantic tags or sequenced contrasts) as a between-subjects factor and contrast (0.2, 0.3, 0.47, and 0.63) as a within-subjects factor. The analysis showed a significant main effect of sequence, $F(1, 10) = 6.07, p = 0.033$, but there was no significant main effect of contrast (0.2, 0.3, 0.47, and 0.63), $F(3, 30 = 0.185, p = 0.91$, or interaction between sequence and contrast, $F(3, 30) = 1.04, p = 0.39$. Therefore, unlike stimulus temporal sequence, semantic sequence tags are effective to induce learning even with ITI jittering.

**Experiment 4: Perceptual learning of multiple roving contrasts with orientation tags**

The multiple contrasts can also be tagged with orientation information. In this experiment, the same four roving contrasts were each tagged by one distinct orientation (i.e., the Gabors were oriented at 36°, 81°, 126°, or 171°, each corresponding to one specific contrast). However, like the Chinese characters we used previously, these orientation tags carried no sequence information. Contrast learning with orientation tags was disabled regardless of five sessions of training (PPR = 1.00 ± 0.06), $F(1, 5) = 0.27, p = 0.62$ (Figure 4a, b). It was only after the four contrasts were presented in a specific sequence (i.e., 0.20, 0.30, 0.47, 0.63) that contrast learning became evident after the same amount of practice (PPR = 0.72 ± 0.05), $F(1, 5) = 11.43, p = 0.020$ (Figure 4c, d). The different learning results were confirmed with a mixed-design ANOVA with sequence (with or without sequence) as a between-subjects factor and contrast (0.2, 0.3, 0.47, and 0.63) as a within-subjects factor. There was a significant main effect of sequence, $F(1, 10) = 5.14, p = 0.047$, and an insignificant main effect of contrast, $F(3, 30) = 0.454, p = 0.72$. There was no significant interaction between sequence and contrast, $F(3, 30) = 0.029, p = 0.99$. 

---

**Figure 3.** The impact of ITI jittering on number tagging with multicontrast learning. (a) The post- versus pretraining contrast discrimination thresholds at four trained contrasts with number tags and jittered ITI for all observers. (b) Individual session-by-session changes of contrast thresholds.
Our results indicate that it is the semantic sequence information in letter or number tagging that facilitates perceptual learning of roving contrasts. Unlike stimulus temporal sequence, the semantic sequence is strong enough to escape the disruption of ITI jittering. These and our previous results together demonstrate that sequence is necessary for stimulus tagging to enable perceptual learning of roving stimuli.

Previously, we have suggested that stimulus roving may have a larger impact on the consolidation stage of information processing in perceptual learning (Zhang et al., 2008). The stimulus temporal sequence may guide proper attention to the responses of different sets of neurons, each set responding to one specific roving stimulus, to facilitate learning consolidation (Zhang et al., 2008). Alternatively, it may align the rewards to the corresponding stimuli in reward-based perceptual learning (Herzog, Aberg, Fremaux, Gerstner, & Sprekeler, 2012). Moreover, the stimulus temporal
sequence per se may facilitate learning consolidation. There is evidence that memory consolidation requires reactivation of neuronal responses (O’Neill, Senior, Allen, Huxter, & Csicsvari, 2008; Wilson & McNaughton, 1994), and such reactivation is found to be associated with learned stimulus temporal sequences (Eagleman & Dragoi, 2012).

The novel finding in the current study is that the semantic sequence information implied in stimulus tags is responsible for learning of roving stimuli. Stimulus tags containing no known sequence information are ineffective. The semantic sequence may affect learning of roving stimuli through top-down influences in a manner similar to what the stimulus sequence does, which may make it easy to associate the stimuli and corresponding learning processes. For example, it could direct the system to accurately switch attention to the responses of proper set of neurons or align the rewards to the proper responses to a specific roving stimulus. It is interesting to know whether semantic sequence information could guide the reactivation of neuronal responses for learning consolidation after training. These top-down influences are important assets of human memory and learning.

Semantic sequence is useful when to-be-learned stimuli cannot be rearranged in a specific temporal order. Roving interference is also evident in auditory learning (Amitay, Hawkey, & Moore, 2005; Nahum et al., 2010). For example, learning of complex speech stimuli is affected by stimulus variability during training (Nahum et al., 2010). However, such interference can also be avoided when multiple stimuli are presented in a temporal sequence (Nahum et al., 2010). In motor learning, humans cannot simultaneously learn opposing force fields or opposing visuomotor rotations (Karniel & Mussa-Ivaldi, 2002; Tong, Wolpert, & Flanagan, 2002). It is interesting to know in these cases whether semantic sequences are equally effective in enabling auditory and motor learning as they are in visual learning.

Keywords: perceptual learning, roving, sequence, contrast discrimination

Acknowledgments

This research was supported by Natural Science Foundation of China Grants 31470975 and 31230030.

Commercial relationships: none.
Corresponding author: Jun-Yun Zhang.
Email: zhangjunyun@pku.edu.cn.
Address: Department of Psychology, Peking University, Beijing, China.

References


