

Visual serial dependence in an audiovisual stimulus

Wee K. Lau

Psychology Programme, School of Social Sciences,
Nanyang Technological University, Singapore



Gerrit W. Maus

Psychology Programme, School of Social Sciences,
Nanyang Technological University, Singapore

Serial dependence is a phenomenon that biases the perception of features or objects systematically toward sensory input from the recent past (Fischer & Whitney, 2014). There is an active debate whether this effect is rooted directly in perception or reflects biases in decision making. We investigated serial dependence across three experiments by manipulating the decision made on each trial. A multimodal audiovisual stimulus comprising a Gabor and a vowel sound was presented repeatedly. On each trial, participants reported either the Gabor orientation or the vowel sound. Participants either ignored one modality (Experiment 1) or attended to both modalities (Experiments 2 and 3). In Experiments 2 and 3, the response task was randomized to prevent anticipating which modality to respond to until the response phase. In Experiment 3, no-response trials were additionally interleaved. Results across the three experiments demonstrated serial dependence only when participants reported the visual modality. Serial dependence was also present in visual reports when participants completed auditory reports or made no reports on previous trials. The previous stimulus alone was enough to elicit an effect. Serial dependence is unlikely to be an effect of the previous decision on the stimulus, but rather an effect of perceiving the previous stimulus.

Introduction

Our perception of the world is continuous, although the input to various senses is fraught with disruptions, discontinuities, and ambiguities. In vision, for instance, eyeblinks, saccades, and frequent occlusions cause gaps in the stream of images from the retina, but we continue to perceive an object as stable. The auditory system also needs to deal with problems of continuity in its input. When we listen to music, we do not hear the individual auditory elements produced by the instruments. Instead, we hear the overall structure of the musical phrase (Krumhansl, 1985). How does

perception remain stable despite these discontinuities in sensory information?

A recent proposal of how perceptual continuity might be achieved introduced the concept of *serial dependence* (Fischer & Whitney, 2014), although similar phenomena—such as sequential dependency (Lages & Treisman, 1998; Treisman & Williams, 1984) and serial effects (Deese & Kaufman, 1957)—have been described previously. Serial dependence arises from spatially and temporally tuned *continuity fields*. Objects which are spatially similar and temporally closer together appear to be more similar than they are. For example, a Gabor differing by $\sim 10^\circ$ in orientation from a previous Gabor presented in the same location is perceived to be more similar, whereas larger orientation differences of 45° or more, or different presentation locations, do not lead to this bias.

Serial dependence is shown in a wide array of studies, including ones of category learning (Jones, Love, & Maddox, 2006; Petzold & Haubensak, 2001), recognition memory (Malmberg & Annis, 2012), discrimination of auditory features including loudness (Jesteadt, Luce, & Green, 1977) and pitch (Alais, Orchard-Mills, & Van der Burg, 2015; Arzounian, de Kerangal, & de Cheveigné, 2017), discrimination of visual features like spatial frequency (Lages & Treisman, 1998), motion (Alais, Leung, & Van der Burg, 2017), and higher level features like numerosity (Cicchini, Anobile, & Burr, 2014), facial expression (Liberman, Fischer, & Whitney, 2014; Liberman, Manassi, & Whitney, 2018), and attractiveness (Tauber, Van der Burg, & Alais, 2016).

Serial dependence may be associated with working memory during spatial judgments. Bliss, Sun, and D'Esposito (2017) manipulated the delay before the response phase and had participants remember the cue location while fixating. Participants then matched the cue location by moving a cursor. Results revealed increasing serial dependence for longer delays. The bias then decreased for delays longer than 10 s. More importantly, when the researchers manipulated the

Citation: Lau, W. K., & Maus, G. W. (2019). Visual serial dependence in an audiovisual stimulus. *Journal of Vision*, 19(13):20, 1–15, <https://doi.org/10.1167/19.13.20>.

<https://doi.org/10.1167/19.13.20>

Received June 6, 2019; published November 26, 2019

ISSN 1534-7362 Copyright 2019 The Authors



intertrial interval (ITI) and kept the delay time constant, they showed that varying the ITI also varied serial dependence. Thus, they concluded that working memory played a role in this bias. This experiment was partially replicated in a different study. Manassi, Liberman, Kosovicheva, Zhang, and Whitney (2018) did not find serial dependence when they eliminated the response-time delay, but serial dependence was present when the stimulus contrast was drastically reduced. They concluded that while working memory plays a role in serial dependence, the uncertainty of the stimulus (i.e., noise) is equally important.

There is also a recent debate on whether serial dependence occurs at the perceptual level or the decision level. Fritsche, Mostert, and de Lange (2017) presented evidence which suggests that serial dependence is related to postperceptual decisions. In their experiment, participants had to sustain the Gabor orientation information in working memory for either 300 or 3,750 ms before reproducing the orientation during the response. Results indicated that the positive attraction of serial dependence was larger in the longer delay. In addition, participants who experienced the greatest serial-dependence bias experienced the largest repulsion when completing the task using a two-alternative forced-choice paradigm instead of a method-of-adjustment task. These results suggest that the type of decision made during a trial affects whether positive or negative serial dependence is observed, and led those authors to attribute serial dependence to postperceptual decision processes rather than perception itself. Similarly, Pascucci et al. (2019) have suggested a multistage model that combines perceptual aftereffects and attractive decision biases to explain serial-dependence effects.

Other researchers argue that serial dependence is a perceptual effect. For example, Cicchini, Mikellidou, and Burr (2017) reasoned that serial dependence occurs at the perceptual level. First, they replicated the results of Fritsche et al. (2017) by showing that negative effects (i.e., repulsion) occurred when successive stimuli were largely different ($> 20^\circ$). They further analyzed the response errors for stimuli with smaller differences ($< 10^\circ$) and found positive serial dependence. In a separate experiment, they provided more support for perceptual origins of serial dependence by showing that when successive stimuli were similar, there was strong serial dependence even when participants had to respond with an orthogonal orientation. When successive responses were similar and successive stimuli were orthogonal, there was no serial dependence. Further, serial dependence was shown to be tuned to specific stimulus locations (Fischer & Whitney, 2014), which seems in contrast with a high-level decision account.

Participants demonstrate serial-dependence bias even when recall memory for the previous stimulus is poor (Fischer & Whitney, 2014). Visually evoked potentials during passive viewing of random-dot arrays have been shown to be biased by the dot quantity from the previous trial (Fornaciai & Park, 2018). More recently, Fornaciai and Park (2019) have shown that the numerosity of sequentially presented flashes can elicit attractive serial dependence when participants judge the numerosity of a subsequently presented dot array. They conclude that serial dependence is likely perceptual, since this bias was observed for various visual presentation formats of numerosity, although it was absent when visual flashes were replaced with auditory tones.

While studies show compelling arguments for serial dependence at the perceptual level and at the decision level, it is important to consider the role of expectations on serial dependence and the actual instance a decision was made. Task expectations might contribute to serial biases. Expectation may facilitate visual processing by allowing us to more readily detect or recognize a stimulus (Summerfield & Egner, 2009). In these serial-dependence experiments, participants are presented with the same kind of information throughout the experiment—for example, always judging Gabor orientations (Fischer & Whitney, 2014; Fritsche et al., 2017) or faces (Liberman et al., 2014; Liberman et al., 2018). Participants in such tasks become accustomed to encoding the same stimulus feature and performing the same adjustment task throughout the experiment. Therefore, they expect to see and do the same thing on every trial. It is possible that this expectation influences judgment of the stimulus in the current trial, since even the encoding of the stimulus is done under the expectation to perform a subsequent decision on it. Evidence of serial dependence when such expectations on the task to be performed are eliminated would add compelling evidence to the perceptual/decision debate.

The goal of this study was to determine the impact of such task expectations on serial dependence. We presented a multimodal audiovisual stimulus on every trial and varied the task expectations across three experiments. In Experiment 1, we first tested serial dependence for the multimodal stimulus in the typical way. Participants were asked to ignore one modality and report the other repeatedly throughout the experiment. We found strong evidence for visual serial dependence, whereas our auditory stimulus and task did not elicit serial dependence.

In Experiment 2, we randomized the task on every trial to abolish strong task expectations. Participants had to pay attention to both modalities, as they did not know which modality to report until the response phase. We again found serial dependence for reports on the visual modality, even when participants reported

the auditory modality on the previous trial and could not anticipate which modality was task relevant.

In Experiment 3, we further investigated expectation effects by additionally including no-response trials. Participants had to switch between reporting different modalities and, occasionally, making no responses. Again, the previous stimulus always influenced subsequent visual reports, but not auditory reports, regardless of whether a report was made in the previous trial.

Methods

Participants

Participants ($N = 19$, $M = 23.74$ years old, $SD = 3.89$; 12 women, seven men) were recruited from undergraduate courses at Nanyang Technological University for Experiment 1. They were unaware of the purpose of the experiment, except for one. Eight participants from Experiment 1 ($M = 25.75$ years old, $SD = 4.71$; five women, three men) proceeded to complete Experiment 2. Data from one participant in Experiment 2 were excluded for failing to complete both sessions (resulting in $N = 7$, $M = 26.0$ years old, $SD = 5.03$; four women, three men). A different group of participants ($N = 10$, $M = 22.10$ years old, $SD = 3.21$; six women, four men) was recruited for Experiment 3. Participants had normal or corrected-to-normal visual acuity, had no hearing problems, and gave consent prior to the start of the experiment. Upon completion, participants were either rewarded with course credits or paid up to S\$10/hr. The experiments were approved by Nanyang Technological University's Institutional Review Board.

Stimuli and apparatus

The experiment was programmed using MATLAB R2015a (MathWorks, Natick, MA) with Psychtoolbox 3 (Brainard, 1997; Kleiner et al., 2007). Angular statistics were calculated using the circStat toolbox (Berens, 2009). A pair of over-ear Creative Aruvana Live! headphones (frequency response: 10 Hz–30 kHz) was used for both experiments. Visual stimuli were presented on a 20-in. Sun Microsystems CRT monitor at a resolution of $1,152 \times 864$ and a refresh rate of 100 Hz. A chin rest was positioned 57 cm away from the monitor.

A white fixation dot (diameter = 0.5°) was always shown at the center of a gray screen (gray = 64.19 cd/m^2 , white = 132 cd/m^2 , black = 0.75 cd/m^2). The multimodal stimulus comprising a Gabor patch and a vowel sound appeared for 500 ms. The Gabor orientation and the specific vowel were selected

randomly. There was a systematic relationship between the Gabor and the vowel (Figure 1a). A 0° (vertical) Gabor corresponded to /ɜ:/, a 60° Gabor corresponded to /u:/, and a 120° Gabor corresponded to /ɑ:/. Orientations in between were related to in-between vowel morphs. For instance, a 15° Gabor was mapped to morph step 15 between /ɜ:/ and /u:/ (see later for morphing procedure). We use *orientations*, *clockwise*, and *counterclockwise* to refer to both visual and auditory stimuli. Participants were initially unaware of the relationship, except for one.

Multimodal noise consisted of white-noise pixels smoothed with a Gaussian kernel and a scrambled vowel morph continuum (see below), and was presented for 1,000 ms immediately after the stimulus. Then there was an interstimulus interval of 250 ms before the response phase. No time limit was imposed for the response phase. After each response, there was an ITI of 2,000 ms before the next trial (Figure 1b).

Parameters of the Gabor were like those used by Fischer and Whitney (2014). Each Gabor was drawn with a fixed spatial frequency of 0.33 c/° at 25% Michelson contrast, using a Gaussian envelope of 1.5° SD . Noise was added to the Gabor every trial to increase task difficulty. The noise was generated by smoothing white noise with a Gaussian kernel ($SD = 0.91^\circ$) and then combining with the Gabor at a ratio of 3:1. Thus, a 25% contrast Gabor was added to 75% noise. All possible Gabor orientations were presented. Another set of noise was generated using the same parameters to mask visual aftereffects after Gabor offset. The visual stimulus was presented 6.5° to the right of fixation at the center of the screen.

Vowels were synthesized and morphed along a circular continuum with the sounds /u:/, /ɜ:/, and /ɑ:/ as anchors (Figure 1a) using a modified *vowel* function from the MATLAB audio toolkit (Smith, 2011). The morphed stimulus space was analogous to a morphed continuum of three distinct faces used in previous studies (Lieberman et al., 2014; Lieberman et al., 2018). Vowels were synthesized with a fundamental frequency of either 122 Hz (Experiments 1 and 2) or 120 Hz (Experiment 3), at a sampling frequency of 44.1 kHz. Each vowel was synthesized using three formant frequencies. The details of the morphing method and the audio files for the morphed vowels are described in Supplementary File S1. The term *stimulus vowel* denotes vowels heard at the onset of the multimodal stimulus, and *response vowel* denotes vowels heard during the response phase. In Experiments 1 and 2, the stimulus vowels were 2 Hz higher than the response vowels, whereas in Experiment 3 both the stimulus and response vowels had the same pitch.

Auditory noise was generated by temporally scrambling the morphed continuum derived from the three canonical vowels (Ellis, 2010). The auditory noise was

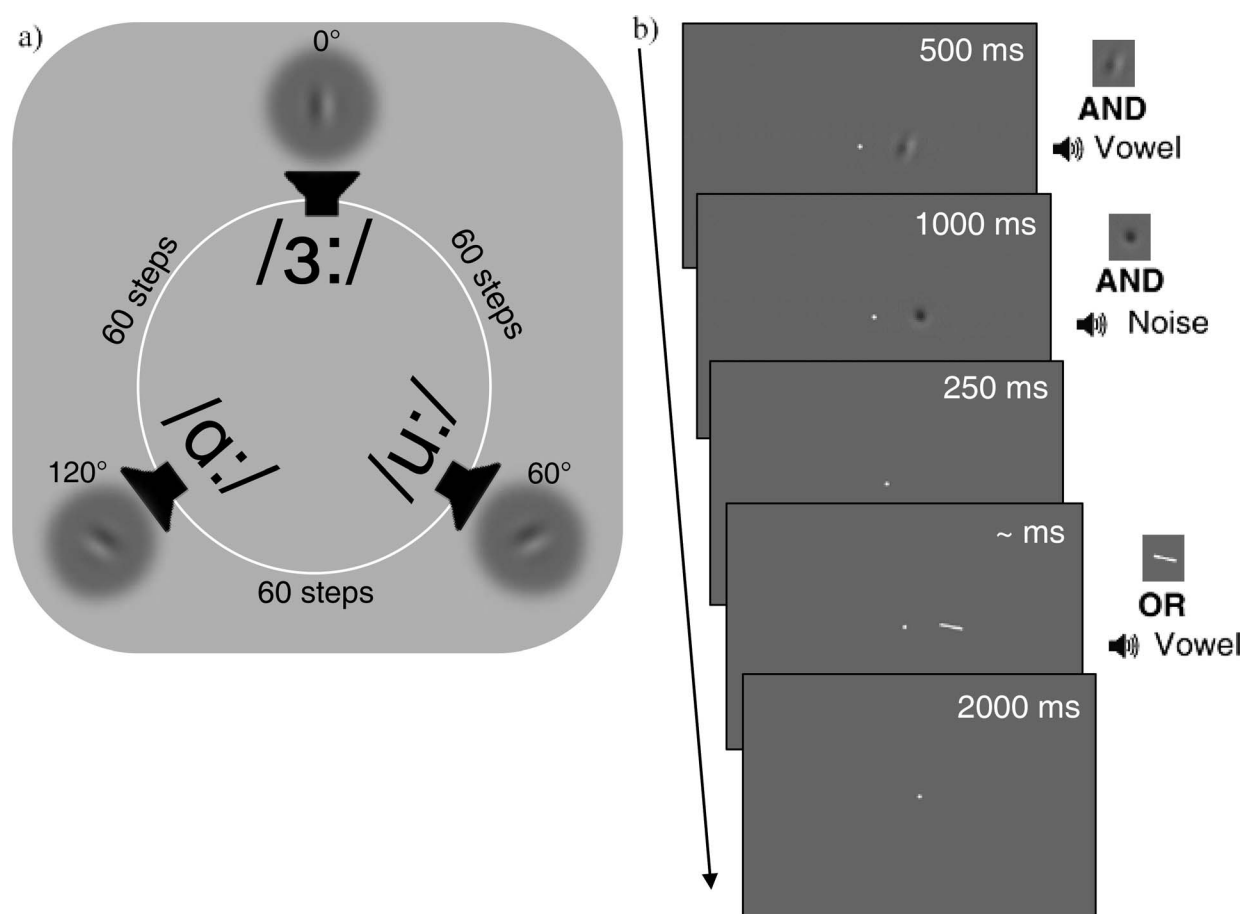


Figure 1. Gabor orientation, vowel continuum, and trial sequence. (a) Examples of a Gabor presented at various orientations, and the Gabor's relationship with the vowels /u:/, /3:/, and /ɑ:/, morphed along a circular continuum. There were 60 morph steps between each vowel pair. (b) Sequence of a single trial in all experiments. Participants fixated the dot and the multimodal stimulus (Gabor + vowel) was presented for 500 ms. Noise was presented for 1,000 ms after the offset of the stimulus. There was an interstimulus interval of 250 ms before the response phase. In Experiment 1, participants only saw the adjustment bar in condition V and only heard the response vowel in condition A. In Experiments 2 and 3, participants either saw the adjustment bar or heard the response vowel, prompting the appropriate response: the Gabor orientation if they saw the adjustment bar or the stimulus vowel if they heard the response vowel. After a response was given by a mouse click, there was an intertrial interval of 2,000 ms before the start of the next trial. In Experiment 3, 20% of the trials were no-response trials, in which neither the adjustment bar nor the adjustment vowel was presented. A video example of the trial sequence can be found in Supplementary Movie S1.

used to mask auditory aftereffects. Temporal scrambling preserved the feature space of the auditory stimuli without making it intelligible: The noise sounded *like* the vowels without containing information about any one vowel. Scrambled masks are effective at masking the stimulus information (Brungart, 2001). We faded in and out to the first and final milliseconds of the auditory stimuli to eliminate clicking artifacts. A sample of the individual vowels, auditory noise mask, and morphed continuum are provided in Supplementary Movie S2.

During the response phase, participants had to respond to either the Gabor orientation or the vowel morph. Visual responses to orientation were made by using a computer mouse to adjust the orientation of a white bar (width = 0.61°, length = 1.5°) at the same location as the previously presented Gabor. Auditory

responses to vowels were made by adjusting a response vowel which morphed in real time when the computer mouse was moved. The white bar and the response vowels were never presented concurrently. The initial bar orientation or response-vowel morphed step was randomized on every trial.

In Experiment 3, participants did not make responses during no-response trials. The duration until the next trial in no-response trials was determined by the participant's average response time since the start of the block.

Procedure: Experiment 1

The experiment was conducted in a dimly lit room. Participants rested their heads on the chin rest and

wore headphones. They were instructed to maintain fixation throughout the experiment. There were two conditions, visual (V) and auditory (A), corresponding to the response modalities. Participants had to match either the orientation of the Gabor using the adjustment bar (V) or the response vowel to the stimulus vowel (A). However, the stimulus was always multi-modal. Participants were randomly assigned to start with either condition and completed both conditions over two separate visits. They were instructed to respond as quickly and as accurately as possible.

Conditions V and A were divided into four blocks of 104 trials per condition (416 trials in total). Participants were allocated 5 min of rest time after completing each block, and were encouraged to take longer breaks when necessary before proceeding with subsequent blocks. Each condition lasted roughly 90 min with breaks.

Before the start of the experiment, participants received 40 trials of training. Participants selected either condition for practice by pressing a key. Feedback was provided at the end of each training trial. Training data were excluded from the analysis.

Procedure: Experiment 2

The procedure for Experiment 2 was identical to that of Experiment 1 except for the following: Participants were instructed to pay attention to both stimulus modalities. The response modality was randomized on every trial using a uniform distribution. On any given trial, participants did not know which modality to report until the response phase. During the response phase, they were to respond to the stimulus orientation if they saw the adjustment bar or to the stimulus vowel if they heard the response vowel. Participants completed Experiment 1 before Experiment 2, so they were familiar with the task and the stimulus. Experiment 2 comprised eight blocks of 104 trials, for a total of 832 trials per participant. On average, there were 52 (50%) visual trials and 52 (50%) auditory trials per block.

Procedure: Experiment 3

The procedure for Experiment 3 was identical to that of Experiment 2 except for the following: Participants familiarized themselves with the stimulus and task by completing 30 training trials (discarded from the analysis). They were then briefed about no-response trials prior to the start of the experiment. During no-response trials, the adjustment bar or the adjustment vowel was not presented. Participants completed 10 blocks of 102 trials (1,020 trials total per participant). On average, there were 20 (20%) no-response trials, 41 (40%) visual trials, and 41 (40%) auditory trials.

Analysis

To investigate serial dependence for reports of each modality in Experiment 1, we looked at whether stimulus history influenced the response errors. The analysis followed that used by Fischer and Whitney (2014). Response errors were calculated by subtracting response orientations from the stimulus orientations. Positive errors meant that participants responded more clockwise to the stimulus. Errors larger than 90° were remapped in the opposite direction. Trials with response times longer than 15 s (for visual responses) and 20 s (for auditory responses) were discarded from the analysis. We conducted two bias-subtraction methods. In the first method, we centered trial response errors by subtracting them from the mean response error in each condition (Experiment 1) or from the mean response errors of the same response modality (Experiments 2 and 3). In the second method, data were centered by subtracting biases from the raw responses for each stimulus value. This procedure removes biases such as the tendency to avoid cardinal directions or respond with unmorphed vowels. The methods and results for this analysis are presented in Supplementary File S1. Using either bias-subtraction method did not significantly affect the results.

N-back histories were calculated by subtracting the current stimulus orientation from the stimulus orientation of the *n*th previous trial. We computed histories for up to *n* = 5 for Experiment 1, but only analyzed 1-back histories for Experiments 2 and 3. Responses were pulled toward the direction of the *n*th stimulus if both the *n*th stimulus and response errors were in the same direction (e.g., clockwise). We visualized serial dependence by binning the response errors using a rolling averaging window (bin width = 15°).

The magnitude of serial dependence was quantified by the amplitude parameter of a fitted function. Following Fischer and Whitney (2014), we fitted the rolling means with a derivative of Gaussian (DoG) of the form $y = xawce^{-(wx)^2}$, where *x* is the distance between the *n*th previous stimulus and the current stimulus, *a* is the amplitude of the function fit, *w* is the width of the curve, and *c* is the constant $\sqrt{2} \div e^{-0.5}$. The constant scales the parameter *a* so that it corresponds to the maximum amplitude and thus the strength of serial dependence. A positive amplitude represents an attractive serial-dependence bias.

For Experiments 2 and 3, we analyzed serial dependence separately depending on the modalities reported in the current and previous trials. All the analyses were abbreviated accordingly. For instance, in Experiment 2, VA refers to pairs where the current trial was A and the previous trial V. In Experiment 3, we use N for no-responses trials; the notation NV denotes

pairs where the current trial was V and the previous trial was a no-response trial.

We conducted fixed-effects analyses for both experiments to measure serial dependence at the group level (Fritsche et al., 2017), because at the individual level it can be a small and subtle effect. We pooled the raw responses and stimuli from all participants in the respective conditions and experiments.

Significance testing was carried out on the group data using a permutation test. We shuffled the order of the stimuli and raw responses without replacement, so that each response was still associated with the same stimulus. Shuffling of data removes all forms of serial-order effects by randomizing the stimulus histories, and therefore controls for other potential artifacts that could have induced a bias like serial dependence, such as biases to respond closer to cardinal directions or unmorphed vowels. The separate bias-subtraction method which removed such biases before serial-dependence analysis also did not produce noteworthy changes in the results (see Supplementary File S1). Data were shuffled over 10,000 iterations. For each iteration, we computed the group response errors and *n*-backs and centered the response errors based on the corresponding conditions (Experiment 1) or response modalities (Experiments 2 and 3). For each shuffled iteration, the empirical widths of condition V and condition A in Experiment 1 were used to constrain the *w* parameter of DoG fits for each respective modality in Experiments 2 and 3 to perform bootstrap significance testing at the group level. The resulting distribution of fitted amplitude values was used to construct 95% confidence intervals (Efron & Tibshirani, 1986).

Significant empirical amplitudes were detected if values fell outside this confidence interval. We also performed Bonferroni correction to account for multiple comparisons in both experiments (Abdi, 2007). With an alpha value of 0.05, the cutoffs for each experiment were Experiment 1: $p < 0.01$ per condition for comparisons up to 5-backs; Experiment 2, $p < 0.0125$ (VV, AA, VA, and AV); and Experiment 3, $p < 0.0083$ (VV, AV, NV, AA, VA, and NA).

Results

We visualized response accuracy for all experiments using a scatterplot of the raw responses against stimuli (Figure 2). Responses were noisy but centered around the unity line (i.e., response = stimulus), indicating that, in general, participants were accurate at performing the task. There were no obvious biases in visual responses. For auditory responses, there were three clusters (Figure 2b, 2d, and 2f) representing the three canonical vowels, meaning that subjects were biased to respond

closer to a canonical vowel than to the morphed stimulus. However, this bias did not affect the analysis of serial dependencies, because the permutation analysis creates empirical null distributions that include the same canonical bias. If the clustering around canonical values could somehow cause a serial-dependence bias, it would also be reflected in permuted iterations. We also subtracted systematic biases from the responses before analyzing for serial dependence in an additional analysis. This alternative analysis did not produce qualitatively different results (Supplementary File S1; Supplementary Figures S1–S4).

Trials with response times (RTs) exceeding 15 s (for visual responses) and 20 s (for auditory responses) were removed. In total, 15 (0.19%) trials from Experiment 1 and 11 (0.19%) trials from Experiment 2 were removed. In Experiment 1, participants were significantly quicker at responding to V (mean RT = 3.02 s, $SD = 0.70$) than A (mean RT = 4.67 s, $SD = 2.10$), $t(18) = 44.62$, $SD = 0.16$, $p = 6.94 \times 10^{-20}$. Participants also made statistically smaller errors when responding to V (mean magnitude = 14.10, $SD = 4.49$) than to A (mean magnitude = 32.93, $SD = 5.34$), $t(18) = -10.59$, $SD = 7.75$, $p = 3.66 \times 10^{-9}$.

Reaction times for Experiment 2 were similarly shorter when participants responded to V (mean RT = 3.30 s, $SD = 0.63$) than to A (mean RT = 4.36 s, $SD = 0.61$), $t(6) = -3.01$, $SD = 0.92$, $p = 0.024$. Participants made significantly smaller errors when responding to V (mean magnitude = 14.51, $SD = 2.60$) than to A (mean magnitude = 32.78, $SD = 6.92$), $t(6) = -5.85$, $SD = 8.26$, $p = 0.001$.

In Experiment 3, participants were also statistically faster at responding to V (mean RT = 3.18 s, $SD = 0.28$) than to A (mean RT = 4.54 s, $SD = 0.77$), $t(9) = -7.59$, $SD = 0.57$, $p = 3.35 \times 10^{-5}$. Similarly, participants made substantially smaller response errors to V (mean magnitude = 15.33, $SD = 4.25$) than to A (mean magnitude = 35.97, $SD = 6.67$), $t(9) = -14.14$, $SD = 4.61$, $p = 1.88 \times 10^{-7}$.

Experiment 1: Visual serial dependence and no auditory serial dependence

To test for serial dependence, we examined whether response errors were attracted toward stimuli from previous trials. When the orientation from the previous trial was more clockwise than in the current trial, serial dependence would predict a clockwise response error, and vice versa. We visualized the data by plotting response errors as a function of the change of stimulus orientation. The amplitude parameter of a fitted DoG model quantified serial dependence.

In Experiment 1, we found serial dependence in condition V (Figure 3). For each *n*-back history, the

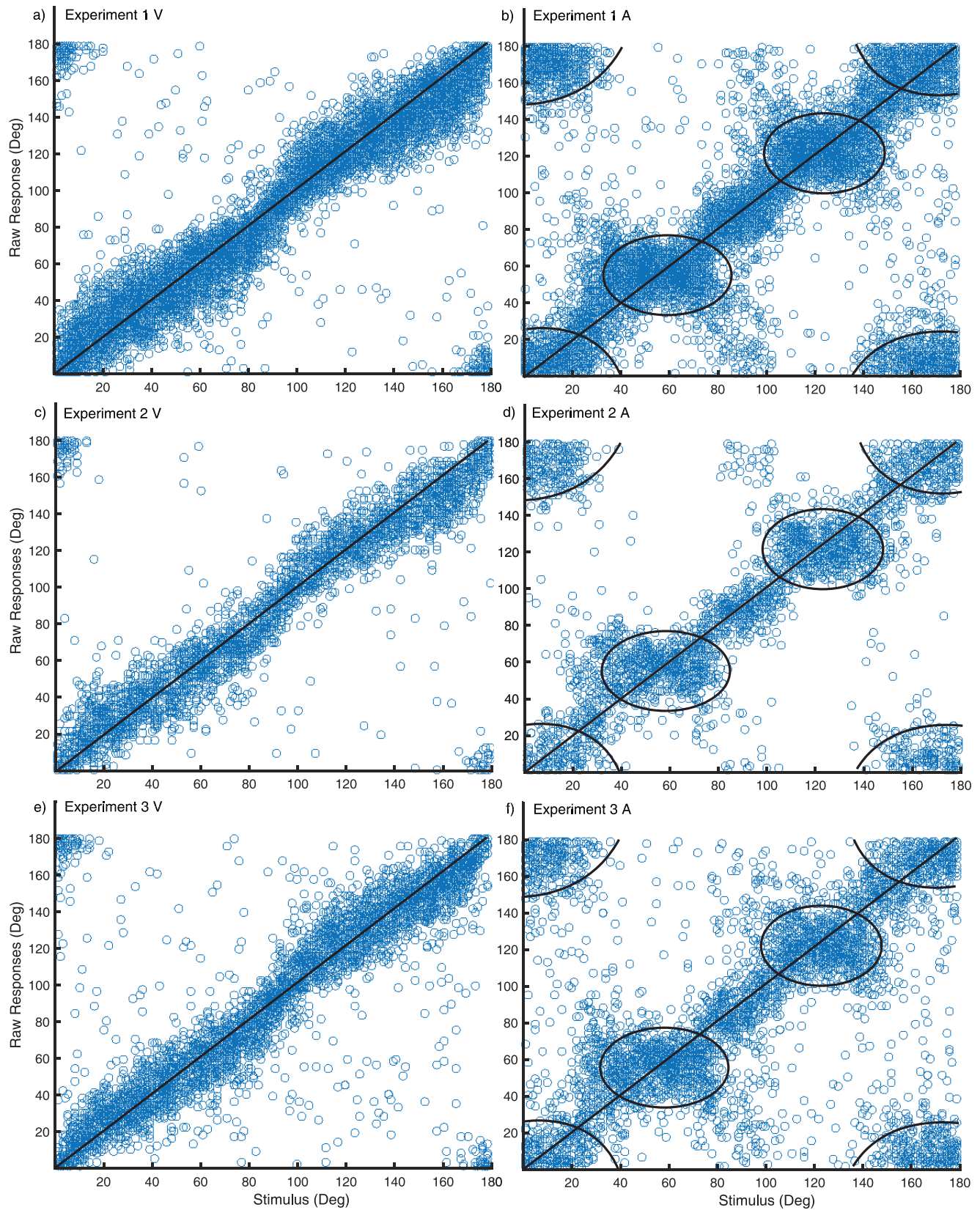


Figure 2. Plot of raw responses against stimulus for all experiments: (a, c, e) Visual responses and (b, d, f) auditory responses. The diagonal line represents the unity line. In (b, d, f), the ovals mark the three clusters corresponding to the three canonical vowels /u:/, /ɜ:/, and /ɑ:/.

shape of the rolling means resembled the typical serial-dependence function (Fischer & Whitney, 2014). Group responses were pulled toward the previous stimulus by 1.53° , $R^2 = 0.74$, root mean square error (RMSE) = 0.010, for 1-back histories (Figure 3a); 0.85° , $R^2 = 0.61$, RMSE = 0.009, for 2-back (Figure 3b); and 1.39° , $R^2 = 0.85$, RMSE = 0.007, for 3-back (Figure 3e). The DoG curve was flat and a bad fit for 4-back, $R^2 = 0.02$, RMSE = 0.011, and 5-back histories, $R^2 = 0.0004$, RMSE = 0.013. Permutation tests revealed that serial dependence was significant for 1-back and 3-back histories, but not 2-back: $p_{1\text{-back}} = 0.0002$, $p_{2\text{-back}} = 0.041$, $p_{3\text{-back}} = 0.0005$, Bonferroni-corrected $\alpha = 0.01$.

For condition A, auditory response errors as a function of stimulus history revealed no evidence for serial dependence (Figure 3c through 3e). Group responses were pulled toward stimulus history by 0.75° , $R^2 = 0.30$, RMSE = 0.013, at 2-back histories (Figure 3d). Group responses were repelled by 0.27° , $R^2 = 0.06$, RMSE = 0.012, at 4-back and by 0.18° , $R^2 = 0.001$, RMSE = 0.014, at 5-back. The DoG curve was irregular for 1-back, $R^2 = 0.03$, RMSE = 0.017 (Figure 3d) and 3-back, $R^2 = 0.12$, RMSE = 0.012. Permutation tests revealed no significant auditory serial dependence: $p_{1\text{-back}} = 0.67$, $p_{2\text{-back}} = 0.69$, $p_{3\text{-back}} = 0.49$.

From Experiment 1, we found significant serial dependence for the visual task for 1-back and 3-back histories. Responses in the auditory task were much more varied and not well captured by a DoG curve indicating serial dependence.

Experiment 2: Visual serial dependence within and across modalities

Experiment 2 was designed to test whether serial dependence occurred even when participants did not know in advance which modality to report from trial to trial. Participants had to pay attention to both stimulus modalities during each trial because the task modality was only known during the response phase.

We first determined whether there was serial dependence for successive trials of the same modalities (VV & AA). Only 1-back histories were tested. Group results revealed the presence of serial dependence. V responses were pulled toward the stimulus of the previous trial by 2.43° , $R^2 = 0.79$, RMSE = 0.014 (Figure 4a). For auditory responses, the DoG curve was flat and poorly fitted, $R^2 = 0.17$, RMSE = 0.042 (Figure 4b). Permutation tests revealed evidence of serial dependence for VV: $p_{\text{VV}} = 0.0107$, Bonferroni-corrected $\alpha = 0.0125$.

We then examined whether serial dependence was present when participants responded to different modalities for successive trials (VA and AV). Group results revealed the presence of serial dependence for V

responses only. V responses on the current trial were pulled toward stimulus history by 2.91° , $R^2 = 0.84$, RMSE = 0.014 (Figure 4c). The DoG curve was flat and poorly fitted for VA, $R^2 = 0.18$, RMSE = 0.033 (Figure 4d). Permutation tests indicated significant serial dependence for AV: $p_{\text{AV}} = 0.0021$, Bonferroni-corrected $\alpha = 0.0125$.

For Experiment 2, we found significant serial dependence when participants made visual responses. The effect was present regardless of the response modality in the previous trial. There was no serial dependence when participants made auditory responses.

Experiment 3: Serial dependence without responses

In Experiment 3, we further examined the task dependence of serial dependence by also including 20% no-response trials. We compared whether the previous stimulus also influenced subsequent reports if participants made no responses in the 1-back trial (NV and NA).

Visual responses showed positive attraction toward stimulus history when participants did not respond in the previous trial (NV). Visual responses were pulled by 4.33° , $R^2 = 0.53$, RMSE = 0.038 (Figure 5a). The DoG curve was flat and poorly fitted for NA, $R^2 = 0.07$, RMSE = 0.056 (Figure 5b). Permutation tests found significant serial dependence for visual responses: $p_{\text{NV}} = 0.0009$, Bonferroni-correct $\alpha = 0.0083$.

We further examined if there was serial dependence when participants responded to the same modalities on successive trials (VV and AA), as in the previous two experiments. Visual responses were pulled toward stimulus histories by 2.48° , $R^2 = 0.36$, RMSE = 0.031 (Figure 5c). The DoG curve was flat and poorly fitted for AA, $R^2 = 0.33$, RMSE = 0.047 (Figure 5d). Permutation tests confirmed serial dependence in VV: $p_{\text{VV}} = 0.004$, Bonferroni-corrected $\alpha = 0.0083$.

We further analyzed serial dependence for responses from different modalities across successive trials (AV and VA). Group responses indicated evidence of serial dependence for AV. V responses on the current trial were attracted by 1.32° toward stimulus histories, although the fit was poor, $R^2 = 0.24$, RMSE = 0.026 (Figure 5e). The DoG curve was flat and poorly fitted for VA, $R^2 = 0.25$, RMSE = 0.06 (Figure 5f). Permutation tests did not reveal any significant effect: $p_{\text{AV}} = 0.13$.

From Experiment 3, we found significant serial dependence for visual reports in the current trial when participants did not respond in the previous trial. We also found the effect only when participants responded to the visual modalities on successive trials. There was no effect for auditory reports in the current trial if

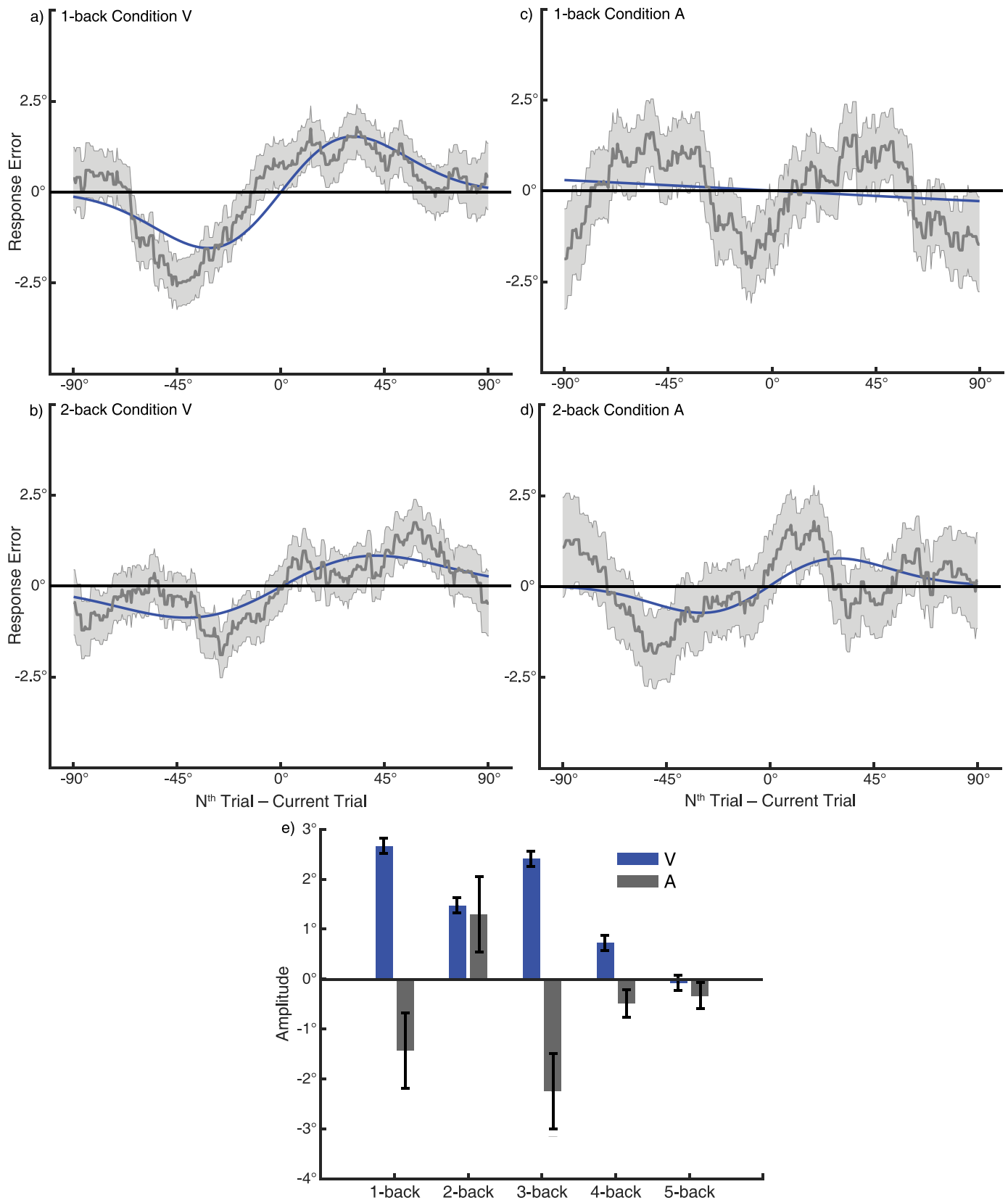


Figure 3. Serial-dependence results of Experiment 1: Response errors as a function of stimulus history for 1-back (a) and 2-back (b) in the visual (V) condition and (c–d) in the auditory (A) condition. The x-axis represents the difference between the stimulus orientation in the n th previous trial and the stimulus orientation in the current trial. The average was calculated using a 15° binning window. The gray areas demarcate the standard error of the mean. The blue lines indicate a fitted derivative-of-Gaussian curve. (e) Amplitudes of fitted derivative-of-Gaussian curves for up to 5-back across both conditions. Error bars represent 1 standard error of the mean of the permutation distribution.

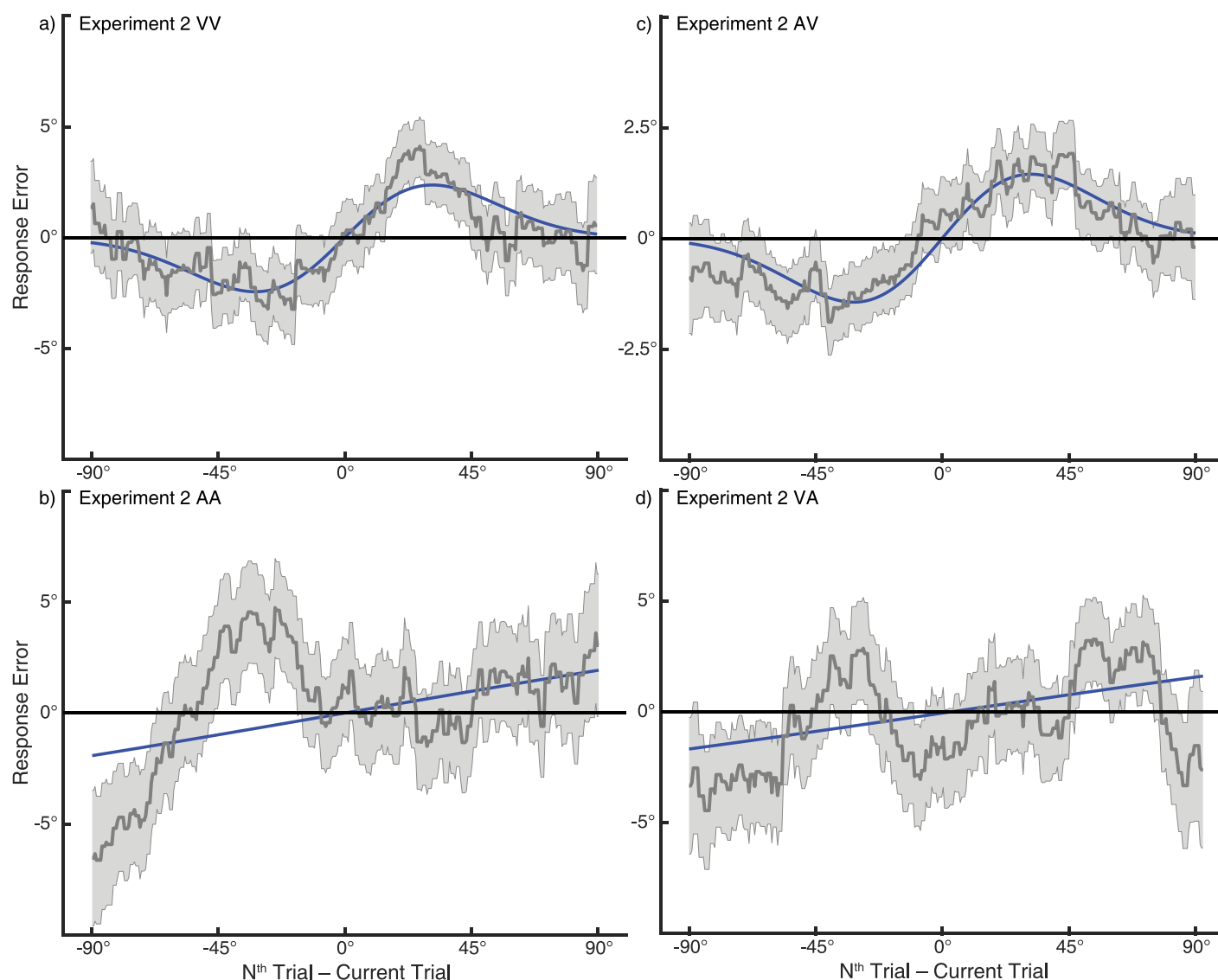


Figure 4. Serial-dependence results of Experiment 2. (a–b) Response errors as a function of stimulus history for 1-back trials with the same response modalities (VV and AA). (c–d) Response errors for 1-back trials with different response modalities.

participants did not respond, reported auditory modalities, or reported visual modalities on the previous trial.

Discussion

Recent debate surrounding serial dependence has revolved around whether this effect occurs at the perceptual level or at a decision level. We further investigated this by using a multimodal stimulus which contained more information to provide sensory estimates of the stimulus features and multiple ways to report them. In our stimuli, visual orientation of a grating was always consistently paired with a vowel sound from a morph wheel, providing redundant

information to execute the task. In Experiment 1, we found compelling evidence for serial dependence when participants had to report the visual orientation. We did not find evidence for serial dependence when the auditory feature was reported.

We then evaluated, in Experiment 2, whether task expectations affected serial dependence. Participants attended to both modalities but reported only one of them. The report modality was unknown to the participants because it was randomized on every trial, and only announced in the response phase of each trial. We found serial dependence when participants reported two successive trials in the visual domain. This bias was also present on the current visual report when participants made auditory reports in the trial before, but not vice versa.

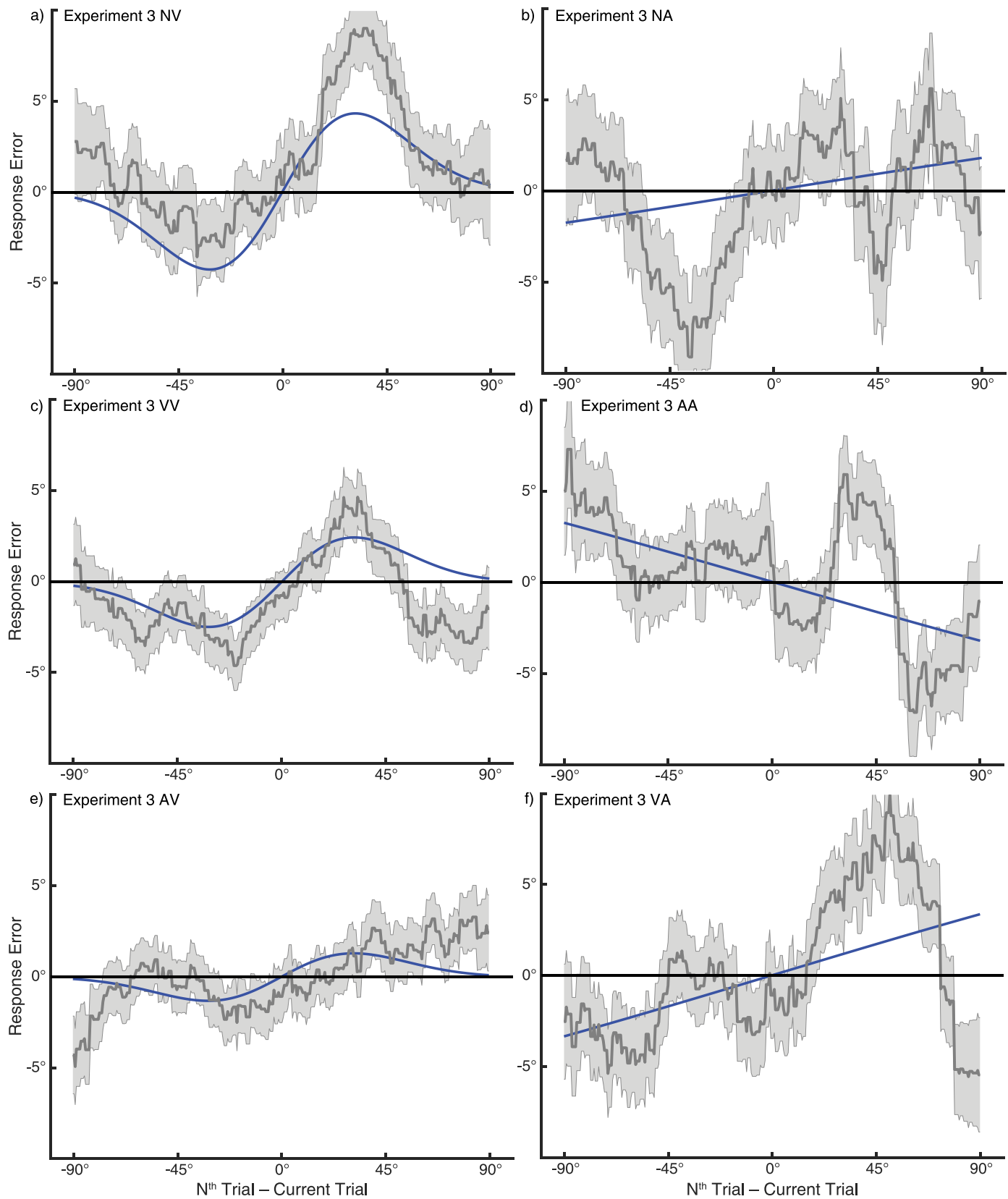


Figure 5. Serial-dependence results of Experiment 3: (a, c, e) Visual response errors and (b, d, f) auditory response errors, as a function of 1-back histories.

Finally, we investigated in Experiment 3 whether serial dependence was evident when no-response trials were included, in addition to randomizing tasks. There was serial dependence on the current visual-response trial even if the trial was preceded by a no-response trial. Again, no serial-dependence effect was found for auditory response trials.

Serial dependence: An effect of perceiving the previous stimulus

Fritsche et al. (2017) argue that serial dependence appears at the decision level. The analysis and interpretation of serial-dependence experiments are complicated by the meaning of *decision*. Serial dependence could be an effect of making decisions about the previous stimulus which subsequently influences how the current stimulus is reported, or it could be an effect on making decisions about the current stimulus.

We showed that serial dependence was evident in visual responses even when participants made auditory reports in the preceding trial (AV). Furthermore, our stimulus was always multimodal and contained both features across our experiments. If serial dependence were an effect of making decisions about the previous trial, we should observe a strong effect when the task remained the same on the current trial (i.e., VV or AA). A weaker effect might be present, albeit less likely, if the task were different on the current trial (i.e., VA or AV). However, we observed serial dependence only when participants reported visual orientation, regardless of the response modality on the previous trial.

Previous studies have argued for a perceptual basis of serial dependence based on finding an effect after no-response trials (Fischer & Whitney, 2014; Liberman et al., 2014). No-response trials cannot fully justify whether serial dependence is perceptual or decisional. Participants in those studies expect to perform a specific task. They might have already judged, internally, the stimuli immediately after presentation, regardless whether they were subsequently asked to respond or not.

In our Experiment 2, participants did not have a specific expectation to perform one task. Arguably, they only decided on a response once the response modality was revealed to them. Our Experiment 3 included additional no-response trials in the context of an unpredictable task. Despite these additions, we showed that whether the previous trial included a response (and decision) of the same, different, or no modality at all was irrelevant to the serial-dependence bias. Only visual responses were subjected to serial dependence. This finding is also consistent with evidence for serial dependence in the absence of any tasks, using visually evoked potentials as dependent

measures (Fornaciai & Park, 2018). This means that future experiments do not require the use of no-response trials, as serial-dependence bias, when found, continues to be present in the current trial.

When taken together, our evidence suggests that serial dependence is better characterized as an effect of perceiving the previous stimulus, since no-response trials and trials with responses to a different modality produce the same serial-dependence biases on subsequent reports. Capacity constraints make it unlikely that participants internally decide on two possible responses simultaneously. Whether the bias occurs at a perceptual or a decisional level warrants further investigation.

Serial dependence and multimodal integration

Let us now consider whether the response bias is determined at a modality-specific level or after a multimodal integration stage. If serial dependence arises because a decision was made after a multisensory integration stage, then reports toward a multimodal stimulus should be biased regardless of the response modality on the previous or current trial. However, we did not find serial dependence of our multimodal stimulus when auditory information was used for reports. Instead, the effect was apparent only when participants reported visual information.

Does serial-dependence bias occur in a multimodal stimulus? Our stimuli might not have led to strong multisensory integration because the mapping between vowel sounds and orientation was arbitrary. A recent report also failed to find serial dependence across modalities: Auditory tones did not influence visual numerosity judgment in a serial-dependence task (Fornaciai & Park, 2019). Other stimuli with a stronger correspondence between different modalities (e.g., pitch and size) might be used in future experiments to investigate multisensory integration of serial dependence.

Lack of auditory serial dependence

Similarity between stimuli is important in eliciting serial dependence. Cicchini et al. (2017) showed evidence of positive serial dependence only for participants comparing similar stimuli. Likewise, Liberman et al. (2018) confirmed that serial dependence was only present when participants reported facial expressions within the same identity (i.e., same genders). In our Experiments 1 and 2, there was a subtle difference between the pitch of the stimulus vowels and the response vowels, of 2 Hz. It is possible that the vowels were too different to influence serial dependence. One

could therefore argue that the vowels did not share the same identity. However, in Experiment 3 both stimulus and response vowels had the same pitch, and we still did not find serial dependence for auditory reports.

The lack of auditory serial dependence in our study may be surprising. Recent studies have shown serial dependence in auditory feature judgments (Alais et al., 2015; Arzounian et al., 2017; Roseboom, 2019). For instance, Alais et al. presented a short gliding pitch to participants on every trial and asked them to indicate whether the sweep was an upward or downward glide. Participants were likely to report subsequent trials as unchanged when the previous pitch glided in the same direction.

Our stimuli employed only one specific auditory feature and adjustment space, a circular vowel morph wheel, which may have been a suboptimal choice in eliciting serial dependence. However, in pilot experiments for the current study we failed to find serial dependence for simple reports of pitch using the method of adjustment. Further, we aimed to use a circular adjustment space to be able to use a one-to-one mapping between the visual and auditory modalities. We are not aware of previous studies reporting evidence for auditory serial dependence using the method of adjustment or circular stimulus spaces. The lack of auditory serial dependence in our study may indicate that serial dependence is a strongly visual phenomenon.

We note that our auditory stimulus caused a strong bias to report canonical vowels; morphed vowels in between the original vowels used in creating the morphed stimulus space were less likely to be reported. While this bias might have increased the difficulty in finding auditory serial dependence, by adding extra noise to the signal in our analysis, it should be independent of any additional serial-dependence bias.

Another commonality of most previous auditory serial-dependence experiments is in the timing of stimuli and the time between responses. In the experiment by Alais et al. (2015), the duration from stimulus onset to the start of the response phase was only 100 ms; the duration of an entire trial was less than 2000 ms. In other serial-dependence experiments involving auditory stimuli, the interval between stimulus onset and response has been roughly 500 ms (Arzounian et al., 2017) or 1,900 ms (Roseboom, 2019).

Timing plays an important role in serial dependence. Fischer and Whitney (2014) have demonstrated that positive or negative serial dependence can occur depending on the stimulus duration and the interval between two stimuli. In our experiments, the interval between two multimodal stimuli was greater than 4,000 ms. This was perhaps too long for auditory serial dependence. Since serial dependence for pitch was found when the stimulus duration and interval between

the two stimuli were short, our current paradigm may not be sensitive enough to capture auditory serial dependence. Further examinations with shorter intervals between trials, without compromising serial dependence in vision, may be required.

An alternate explanation for the lack of auditory serial dependence in our current experiments could be related to working memory. Some literature indicates that auditory sensory memory lasts no more than 10 s (Sams, Hari, Rif, & Knuutila, 1993). Auditory recognition memory is also shorter than visual recognition memory for the same stimulus length (Cohen, Horowitz, & Wolfe, 2009). Bliss et al. (2017) manipulated the ITI and showed that serial dependence varied depending on the response time between the previous and the current trial. In their second experiment, they showed that an ITI of 10 s resulted in a repulsive effect between the previous and the current response error. Thus, it is possible that the lack of auditory serial dependence in our experiment could be related to working memory.

Serial dependence and the natural environment

The attractive forces under serial dependence can be beneficial for interpreting the visual world. Objects in the real world do not change abruptly. Therefore, visual biases of serial dependence enhance the permanence of objects in the natural setting. However, this does not apply equally to all visual features. Taubert, Alais, and Burr (2016) conducted an experiment testing the extent to which faces and emotional expressions are influenced by serial dependence: Participants had to judge the gender of a face (male/female) and its emotion (happy/sad). They showed that the responses to gender were subject to positive serial dependence (i.e., attractive responses), but responses to emotion were repelled from the previous response.

In that experiment, the visual stimulus provided both gender and emotional information. However, only the gender information responded to serial dependence. The researchers argue that the attractive visual bias may not always be beneficial. The ability to detect a sudden change of emotion is important, as emotions convey critical information about the current state of an individual. Moreover, it is unlikely that an individual abruptly changes their physical appearance during social interaction. This could explain why gender, but not emotion, was subject to serial-dependence bias.

The lack of an auditory effect in our results may be explained by a similar logic. Vowel sounds are constantly changing in human speech. When a word is articulated in English, the vowels are sometimes modified by the preceding consonants (De Jong, 1991).

Detecting changes in vowels is crucial for understanding speech. An auditory bias which assimilates the current vowel to the previous vowel would be unfavorable and would not provide much benefit to the intelligibility of the speech.

Conclusion

Recent debates on serial dependence circle about whether the effect is a perceptual or a decision-level bias. We conducted three experiments with a multi-modal audiovisual stimulus to further test the origins of the effect. We found serial dependence only when participants responded to the visual information across all the experiments. This effect was present when task expectations were eliminated and when participants made no reports on the preceding trial. The presence of a previous visual stimulus alone was enough to elicit serial dependence; whether it was subject to attention or to a perceptual decision in a previous trial was irrelevant. Our results suggest that serial dependence is an effect of perceiving the previous stimulus.

Keywords: serial dependence, perceptual, decision, sequential, multimodal

Acknowledgments

This work was supported by a Nanyang Technological University Research Scholarship to W. K. L., and a Nanyang Assistant Professorship start-up award to G. W. M. We would like to thank Jason Fischer for his help in interpreting the results, and Alina Liberman for her suggestion with randomizing response modalities.

Commercial relationships: none.

Corresponding author: Wee K. Lau.

Email: wlau010@e.ntu.edu.sg.

Address: School of Social Sciences/Psychology Program, Nanyang Technological University, Singapore.

References

- Abdi, H. (2007). Bonferroni test. In N. J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics* (pp. 103–107). Thousand Oaks, CA: Sage Publications, Inc. <https://doi.org/10.4135/9781412952644>.
- Alais, D., Leung, J., & Van der Burg, E. (2017). Linear summation of repulsive and attractive serial dependencies: Orientation and motion dependencies sum in motion perception. *The Journal of Neuroscience*, 37(16), 4381–4390.
- Alais, D., Orchard-Mills, E., & Van der Burg, E. (2015). Auditory frequency perception adapts rapidly to the immediate past. *Attention, Perception, & Psychophysics*, 77(3), 896–906.
- Arzounian, D., de Kerangal, M., & de Cheveigné, A. (2017). Sequential dependencies in pitch judgments. *The Journal of the Acoustical Society of America*, 142(5), 3047–3057.
- Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, 31(10), 1–21. <https://doi.org/10.18637/jss.v031.i10>.
- Bliss, D. P., Sun, J. J., & D’Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, 7(1), 14739.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109(3), 1101–1109.
- Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences, USA*, 111(21), 7867–7872.
- Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, 17(14):6, 1–9. <https://doi.org/10.1167/17.14.6>. [PubMed] [Article]
- Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences, USA*, 106(14), 6008–6010.
- De Jong, K. (1991). An articulatory study of consonant-induced vowel duration changes in English. *Phonetica*, 48(1), 1–17.
- Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology*, 54(3), 180–187.
- Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1(1), 54–75.
- Ellis, D. P. W. (2010). *Time-Domain Scrambling of Audio Signals in Matlab*. Retrieved from www.ee.columbia.edu/~dpwe/resources/matlab/scramble/
- Fischer, J., & Whitney, D. (2014). Serial dependence in

- visual perception. *Nature Neuroscience*, 17(5), 738–743, <https://doi.org/10.1038/nn.3689>.
- Fornaciai, M., & Park, J. (2018). Attractive serial dependence in the absence of an explicit task. *Psychological Science*, 29(3), 437–446.
- Fornaciai, M., & Park, J. (2019). Serial dependence generalizes across different stimulus formats, but not different sensory modalities. *Vision Research*, 160, 108–115, <https://doi.org/10.1016/j.visres.2019.04.011>.
- Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, 27(4), 590–595.
- Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 3(1), 92–104.
- Jones, M., Love, B. C., & Maddox, W. T. (2006). Recency effects as a window to generalization: Separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 316–332.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1–16.
- Krumhansl, C. L. (1985). Perceiving tonal structure in music: The complex mental activity by which listeners distinguish subtle relations among tones, chords, and keys in Western tonal music offers new territory for cognitive psychology. *American Scientist*, 73(4), 371–378.
- Lages, M., & Treisman, M. (1998). Spatial frequency discrimination: Visual long-term memory or criterion setting? *Vision Research*, 38(4), 557–572.
- Liberman, A., Fischer, J., & Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology*, 24(21), 2569–2574, <https://doi.org/10.1016/j.cub.2014.09.025>.
- Liberman, A., Manassi, M., & Whitney, D. (2018). Serial dependence promotes the stability of perceived emotional expression depending on face similarity. *Attention, Perception, & Psychophysics*, 80(6), 1461–1473, <https://doi.org/10.3758/s13414-018-1533-8>.
- Malmberg, K. J., & Annis, J. (2012). On the relationship between memory and perception: Sequential dependencies in recognition memory testing. *Journal of Experimental Psychology: General*, 141(2), 233–259.
- Manassi, M., Liberman, A., Kosovicheva, A., Zhang, K., & Whitney, D. (2018). Serial dependence in position occurs at the time of perception. *Psychonomic Bulletin & Review*, 25(6), 2245–2253.
- Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., & Chelazzi, L. (2019). Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. *PLoS Biology*, 17(3), e3000144.
- Petzold, P., & Haubensak, G. (2001). Higher order sequential effects in psychophysical judgments. *Perception & Psychophysics*, 63(6), 969–978.
- Roseboom, W. (2019). Serial dependence in timing perception. *Journal of Experimental Psychology: Human Perception and Performance*, 45(1), 100–110. <https://doi.org/10.1037/xhp0000591>.
- Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence. *Journal of Cognitive Neuroscience*, 5(3), 363–370.
- Smith, S. (2011). *MATLAB Audio Toolkit*. Retrieved from www.cs.uml.edu/~stu/
- Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, 13(9), 403–409.
- Taubert, J., Alais, D., & Burr, D. (2016). Different coding strategies for the perception of stable and changeable facial attributes. *Scientific Reports*, 6, 32239.
- Taubert, J., Van der Burg, E., & Alais, D. (2016). Love at second sight: Sequential dependence of facial attractiveness in an on-line dating paradigm. *Scientific Reports*, 6, 22740.
- Treisman, M., & Williams, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review*, 91(1), 68–111. <https://doi.org/10.1037/0033-295X.91.1.68>.

Supplementary material

Supplementary Movie S1. Sample of 10 trials from Experiment 3. The first trial was an auditory-response trial. The second trial was a visual-response trial. The third trial was a no-response trial. Response trials during the experiment were fully randomized.

Supplementary Movie S2. Auditory stimulus. The individual vowels, noise mask, and three repetitions of the morphed continuum are heard.