A resource model of phonological working memory

Christopher R. Hepner (chepner3@jhu.edu)
Department of Neurology, Johns Hopkins University,
1629 Thames Street, Suite 350, Baltimore, MD 21231 USA

Nazbanou Nozari (nozari@jhu.edu)
Department of Neurology; Department of Cognitive Science, Johns Hopkins University,
1629 Thames Street, Suite 350, Baltimore, MD 21231 USA

Abstract

The classic Baddeley and Hitch (1974) model divides working memory into domain-specific subsystems and a shared, domain-general central executive, which plays a role in allocating resources to items stored in the subsystems. The nature of this resource—in particular, its quantization (discrete vs. continuous) and the flexibility of its allocation—has been studied extensively in the visual domain, with evidence from experiments using continuous response measures providing support for models with flexibly and continuously divisible resources. It remains unclear, however, whether similar mechanisms mediate the division of resources in phonological working memory. In this paper, we show that, despite representational differences between visual and auditory processing, continuous measures can also be employed for studying phonological working memory. Using such measures, we demonstrate that the principles of resource division in visual and phonological processing are indeed similar, providing evidence for a domain-general mechanism for allocating working memory resources.

Keywords: phonological working memory; cognitive resources; central executive; domain-generality; resource models; slot models

Introduction

In the classic model proposed by Baddeley and Hitch (1974), working memory has both domain-specific and domain-general components: separate subsystems for verbal and visual information (the phonological loop and visuospatial sketchpad, respectively), with a shared central executive. While the domain-specific properties of verbal and visual working memory have been studied extensively, the function of the central executive has remained obscure. There is some consensus that the central executive plays a role in allocating working memory resources to items stored in the domain-specific subsystems; however, the exact nature of this operation has been hotly debated: Are resources discrete or continuous? Is there an upper limit on the number of items to which resources can be allocated? Are resources divided equally between items, or can higher-priority items receive a larger share? These questions have been investigated extensively in visual working memory, leading to significant support for models with flexibly and continuously divisible resources (i.e., *resource models*) in the visual domain (Ma, Husain, & Bays, 2014). The evidence has come from experiments in which, instead of binary accuracy, the deviation of responses from the target has been measured, allowing the *quality* of the stored representations to be investigated rather than just the *quantity*. For example, instead of probing memory for colors with a choice between a limited set of discrete values (e.g., prototypical red, orange, yellow, etc.) and scoring the response as either correct or incorrect, allowing participants to select any hue on a continuous color wheel and measuring the distance between that hue and the target. The question remains: are resources divided the same way in the verbal domain? This paper investigates this issue. Specifically, we examine whether a resource model is appropriate for phonological working memory.

Division of Resources

Generally speaking, two classes of model have been proposed for the division of resources: discrete and continuous. *Slot models* (e.g., Cowan, 2001) propose that working memory resources are discrete, divided up into a fixed number of slots. Each slot can store exactly one item. When set size is less than or equal to the number of slots, all items receive slots and can be recalled with little or no error. However, when set size exceeds the number of slots, some items do not receive slots, and probing one of these items will result in a random response. Slot models thus make identical predictions for both error rates and the deviation of responses as a function of set size: minimal until set size exceeds the number of slots, then rising steeply. *Resource models* (e.g., Ma et al., 2014), on the other hand, propose that resources are continuous and can be divided between any number of items. The quality of a stored representation is dependent on the amount of these resources it receives: items receiving more resources can be recalled with greater precision, i.e., less deviation from the target response. Any increase in set size, even from 1 to 2 items (well below the capacity of any slot model), would stretch the resources a little thinner and thus reduce the quality of the stored representations. Binary accuracy measures may not be sensitive enough to detect this difference when the set sizes are small (e.g., 1 vs. 2 items), since the quality of the stored representations may still be sufficient to select the correct response. This makes binary accuracy measures a suboptimal tool for distinguishing between slot and resource models. However, a continuous measure of the deviation of the response from the target(i.e., the quality or precision of the response) can provide the necessary sensitivity. Using such measures, Ma et al. (2014) showed that, in line with the predictions of a resource model, the deviation between the target color and participants' responses on a color wheel increased monotonically as a function of the number of colors to be remembered.

Within the framework of a resource model, the allocation of resources may be either *fixed*, meaning that resources are divided equally between all items to be remembered, or *flexible*, meaning that one or more items may receive a larger share of resources than the others, e.g., due to manipulation of top-down attention. In keeping with the prediction of a flexible resource model, experiments in which attentional cues were manipulated (e.g., Gorgoraptis, Catalao, Bays, & Husain, 2011) have shown that deviation scores are significantly lower for prioritized items, and significantly higher for those that have been deprioritized, compared to a neutral baseline.

In summary, evidence from the visual domain supports a flexible version of the resource model.

Visual and Verbal Representations

In the previous section, we explained why distinguishing between slot and resource models requires a continuous measure of the deviation of a response from the target. This is easy to obtain in the visual domain, since many of the features associated with visual representations can take any value on a continuum, rather than a small set of discrete values. Importantly, observers can often imagine "in-between" values rather easily. For example, people can imagine a variety of greenish-blues and bluish-greens between the prototypical colors blue and green. Similarly, people can imagine different orientations between vertical and horizontal. Consequently, the deviation of a response from the target can be measured as the distance between the corresponding points on the continuum.

The task of identifying similar continua in the phonological domain is more complex. Though the acoustic properties of speech sounds also vary continuously, only variation that crosses category boundaries is relevant to distinguishing between phonemes. For this reason, people tend to hear a /k/-ish /g/ as either a /k/ or a /g/, but not as something in between (i.e., categorical perception; Liberman, Harris, Hoffman, & Griffith, 1957). This tendency is not absolute, though. If asked to rate a /k/-ish /g/ on a continuous scale between /k/ and /g/, participants are capable of doing so (Massaro & Cohen, 1983). This finding suggests that, despite surface differences in how people perceive visual and auditory information, both information types are perceived in enough detail to allow for fine-grained measurements of the deviation of a response from the target.

The Current Study

In the current study, we used the syllable rating task from Massaro and Cohen (1983) to obtain measurements of deviation in phonological working memory. Using this paradigm allowed us to test whether the results in support of resource models in the visual domain can be extended to auditory perception. If so, we can conclude that the same domain-general principles are at work in both visual and verbal domains at the level of the central executive. If not, domain-specific models of resource division must be proposed. In Experiment 1, we manipulated the number of syllables presented in each trial to determine the relationship between set size and the deviation of responses. In Experiment 2, we manipulated attentional cues while maintaining a constant set size in order to determine the flexibility of resource allocation in phonological working memory.

Experiment 1

Participants

Forty-eight native speakers of American English (31 females, M_{age} = 43.2, age range: 24-65 years) participated in an online experiment developed using jsPsych (de Leeuw, 2015) for payment through Amazon Mechanical Turk (AMT; https://www.mturk.com).

Stimuli and Procedures

Stimuli were 28 syllables, seven from each of four acoustic continua: /bɑ/–/dɑ/, /kɑ/–/ɡɑ/, /ɹɑ/–/lɑ/, and /sɑ/–/ʃɑ/. The syllables at the ends of the continua were recordings of a native speaker of American English. The five intermediate syllables on each continuum were created by progressively changing the acoustic properties of the initial consonant to create five equally-spaced consonants between the two recorded syllables while leaving the vowel unchanged. To minimize interference, a different distinctive feature was manipulated in each continuum: [−coronal] vs. [+coronal], [−voice] vs. [+voice], [−lateral] vs. [+lateral], and [+anterior] vs. [−anterior], respectively. Each participant completed two sessions 24–72 hours apart with the same structure but a different trial order. Each session consisted of two phases: a *baseline phase* and a *working memory phase*.

Baseline Phase The baseline phase was divided into four blocks, one for each acoustic continuum. In each block, participants first completed an orientation in which all seven syllables along the continuum were played in order. As each syllable was played, its position was shown on a slider visually representing the range between the most extreme syllables on the continuum (e.g., between the most /bɑ/-like syllable at the left end, labelled "B", and the most /dɑ/-like syllable at the right end, labelled "D"). This orientation procedure was repeated four times to give participants enough opportunities to learn the relationship between the syllables on the acoustic continuum and the corresponding positions on the visual slider.

Once the orientation was over, participants were tested on their ability to rate syllables (*baseline test*). They listened to the same syllables, presented in a random order, and indicated the position of each one on the continuum using the slider. Although there were only seven syllables in each continuum, participants could adjust the slider continuously. Once participants had adjusted the slider to their satisfaction, they pressed a "submit" button and the position was recorded on a scale from 1 to 100. Only one syllable was played in each trial and there was no deadline for responding. The baseline test in each block consisted of 14 practice trials (two per syllable) followed by 56 experimental trials (eight per syllable). Thus, across the two sessions, participants completed a total of 448 experimental baseline test trials (16 for each of the seven syllables in each of the four continua).

Working Memory Phase This phase tested the effect of set size on the deviation of responses from the target. On each trial, participants were presented with a sequence of one, two, or four syllables from different acoustic continua played at 1 s intervals. One second after the final syllable was played, the slider appeared, and participants had 1 s to rate the relevant syllable on the slider. Since two syllables from the same continuum were never played during the same trial, the labels on the slider unambiguously indicated which syllable to rate. In each session, there were 15 practice trials, followed by 12 blocks of 28 experimental trials with pseudorandomized order, such that no more than two consecutive trials had the same set size. Across the two sessions, participants completed a total of 672 experimental working memory trials (224 for each of the three set sizes). The design was fully counterbalanced, so each syllable was probed the same number of times (eight) in each set size for each participant. Within each set size, each syllable appeared the same number of times in each position.

Statistical Analyses

Dependent Variable To measure the magnitude of the error in the responses (and thus the precision), we obtained a "Deviation Score" for each response made in the working memory phase in three steps: (1) We calculated the median of the participant's 16 ratings for the same syllable in the baseline phase. (2) We then subtracted this baseline median from the response. If, for example, the median of the participant's ratings was 30, the results for the responses 33 and 29 in the working memory phase would be $33 - 30 = 3$ and $29 - 30 = -1$, respectively. (3) Finally, we took the absolute value of the number from (2) to get a Deviation Score for each response. These Deviation Scores were the dependent variable in both experiments.

Statistical Models The main analyses in this study were carried out with linear mixed-effects modeling (LMEM) using the lme4 package (version 1.1-14; Bates, Mächler, Bolker, & Walker, 2015) in R (version 3.4.2; R Core Team, 2017). We strove to include the maximal random effects structure tolerated by the model. All numeric variables were centered and scaled, and the dependent variable (the Deviation Score) was log-transformed to approximate a normal distribution. The p-values were calculated based on Satterthwaite approximations using the lmerTest package (version 2.0-33; Kuznetsova, Brockhoff, & Christensen, 2016).

Results

Baseline Test Figure 1 shows the distributions of participants' ratings for the syllables in the /kɑ/–/gɑ/ continuum in the baseline phase; the rating distributions for the other continua were similar. To determine whether participants had been able to rate the syllables continuously, rather than categorically, we analyzed the ratings using uninformed mixture modelling by means of the mclust package (version 5.3; Fraley & Raftery, 2002) in R. If participants were rating the syllables continuously, the overall distribution of the ratings should be a mixture of seven distributions centered on or near the "correct" rating for each syllable. The differences between the means of the model distributions and the correct ratings were small: the root-mean-square deviation (RMSD) was only 4.27. To formally test whether seven distributions provided a better model for the data than two distributions near the ends of the rating scale (as would be expected if participants were rating categorically), we also fitted a model with only two distributions and compared the fit of the two models using the Bayesian information criterion (BIC), which penalizes for additional parameters. The BIC for the seven-distribution model (196,329) was much lower than the BIC for the two-distribution model (201,838; a difference of 5,509), providing very strong evidence against the two-distribution model. These results indicate that the participants were able to perceive and rate the syllables continuously. Next, we tested the effect of set size on ratings for the same syllables in the working memory phase.

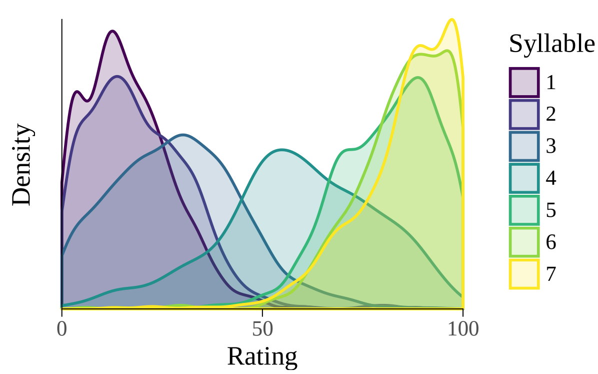

Figure 1: Distribution of ratings for syllables in the /kɑ/–/gɑ/ continuum in the baseline test phase of Experiment 1.

Working Memory Figure 2 shows the relationship between set size and the Deviation Score at each position. Before analyzing the effect of set size on deviation scores, we first established that the bow-shaped serial position effect characteristic of working memory performance was present in our data: better performance at the beginning (primacy effect) and end (recency effect) of the sequence compared to the middle. We fitted a model to data from trials with set size four (set sizes one and two are too small to allow for clear testing of position effects) to test (a) whether the canonical position effects were obtained, and (b) which covariates needed to be included in subsequent models to control for the effects of nuisance variables.

This model included position as a set of polynomial contrasts, with separate fixed effects for linear and quadratic (i.e., bow-shaped) serial position effects, along with random slopes for each of these by subject and item. The model also included fixed effects for several nuisance variables: (a) *baseline median*, which was the absolute value of the distance between the participant's baseline rating for the syllable and the center of the rating scale, to account for the reduction in variability at the ends of the scale (see the ratings for syllables 1 and 7 vs. syllable 4 in Figure 1); (b) *baseline variability*, which was the standard deviation of the participant's baseline rating for the syllable; and (c) *session*, which was coded as a contrast between the first and second sessions. We also included random intercepts for participants and items, i.e., syllables. Critically, there was a significant linear effect of position ($t = -2.70$, $p = .012$), indicating a decrease in Deviation Score for more recent syllables, and a significant quadratic effect of position ($t = -4.46$, $p < .001$), indicating higher Deviation Scores in the middle of the sequence than at the ends. There were also main effects of baseline median ($t = -10.33$, $p < .001$), corresponding to a decrease in the Deviation Scores closer to the ends of the rating scale, and baseline variability ($t = 10.17$, $p < .001$), but not session ($t = -0.66$, $p = .509$). These findings suggest that (a) Deviation Scores do indeed reflect working memory performance, and (b) both baseline median and baseline variability have significant influence on Deviation Scores. We thus included these covariates in all subsequent analyses. We also included serial position and its interaction with set size because serial position alone could potentially have created spurious set size effects.

The main prediction of the resource model investigated in this experiment—that the Deviation Score would increase as a function of set size—was tested using an LMEM with fixed effects for set size, serial position, and the interaction between the two, along with the baseline median and baseline variability as covariates. The random effect structure included random intercepts for participants and syllables and random slopes for set size, position, and the interaction between the two by participant and syllable. The model revealed a significant main effect of set size ($t = 5.32$, $p < .001$), with the Deviation Score increasing as a function of set size.

There was no interaction between set size and position; however, to confirm that the effect of set size was robust across positions, we conducted additional post-hoc analyses. Four post-hoc tests compared: (a) the initial positions of set sizes one and two, (b) the final positions of set sizes one and two, (c) the initial positions of set sizes two and four, and (d) the final positions of set sizes two and four. For each test, we compared the mean difference in Deviation Scores between the two set sizes across participants to the distribution generated by a Monte Carlo simulation with 1,000,000 permutations, resampling within participants. After Bonferroni correction for multiple comparisons, the set size effect was significant in all cases: (a) $M = 0.76$, 95% CI = [0.35, 1.18], $p < .001$, (b) $M = 0.57$, 95% CI = [0.18, 0.97], $p = .009$, (c) $M = 1.68$, 95% CI = [1.05, 2.31], $p < .001$ and (d) $M = 0.80$, 95% CI = [0.19, 1.44], $p = .018$.

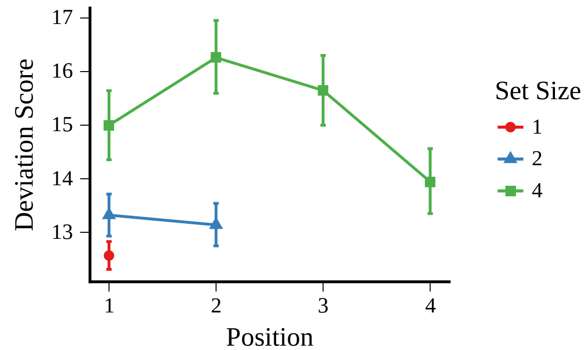

Figure 2: Deviation Score as a function of set size and serial position in Experiment 1. Error bars are 95% CIs.

Discussion

Analysis of the baseline phase confirmed that participants were able to perceive and rate the syllables continuously, as previously reported by Massaro and Cohen (1983). This finding makes these materials appropriate for testing the predictions of a resource model. If such a model is appropriate for phonological working memory, then the Deviation Scores of the syllable ratings should increase as set size increases. This increase in the Deviation Scores should be visible for any increase in set size, even from one to two syllables, which is well below the capacity limit proposed by any slot model. The results of Experiment 1 confirmed this prediction. Deviation Scores increased significantly as a function of set size. This effect was robust in post-hoc analyses which took primacy and recency effects into account by restricting comparisons between set sizes to matching positions. Thus, despite the clear differences between visual and verbal stimuli, the results of Experiment 1 closely resembled those found in the visual domain (e.g., Bays, Catalao, & Husain, 2009; Wilken & Ma, 2004), and were in full accord with the predictions of a resource model. Experiment 2 tested whether a fixed or a flexible version of the resource model is more appropriate for phonological working memory.

Experiment 2

Participants

Forty-eight native speakers of American English (27 females, $M_{age} = 36.6$, age range: 21–58 years) participated for payment through AMT.

Materials and Procedures

The materials were the same as Experiment 1. The same two-session design as Experiment 1 was used, with a similar session structure. The baseline phase was unchanged. In the working memory phase, the presence (or absence) and validity of cues appearing before the presentation of the syllables was manipulated. A fixed set size of four was used in all trials. In one third of the trials, no cue was presented. These *no-cue trials* were identical to the trials with set size four in Experiment 1. On the other two thirds of the trials, a cue (a number between 1 and 4) was presented at the beginning of the trial. This cue indicated that the syllable in the corresponding position (1 through 4) had a 50% chance of being probed (one third of all trials; *valid-cue trials*). The other syllables each has a 16.7% chance of being probed (one third of all trials; *invalid-cue trials*).

On cued trials, the cue appeared for 1 s, after which the four syllables were presented and one of them was probed in the same way as in Experiment 1. Participants completed three blocks of trials in the working memory phase: one block with *no-cue* trials and two blocks with a mixture of *valid-* and *invalid-cue* trials, in counterbalanced order. Each block consisted of 14 practice trials and 112 experimental trials, with breaks between sets of 28 trials, for a total of 224 trials in each cue condition (*no cue*, *valid-cue*, and *invalid-cue*) across both sessions. Each syllable was probed exactly once in each position in each block, resulting in a total of 8 samples for each syllable in each cue condition from each participant.

Results

Deviation Scores were calculated in the same manner as before. Figure 3 shows the Deviation Scores as a function of attentional cueing.

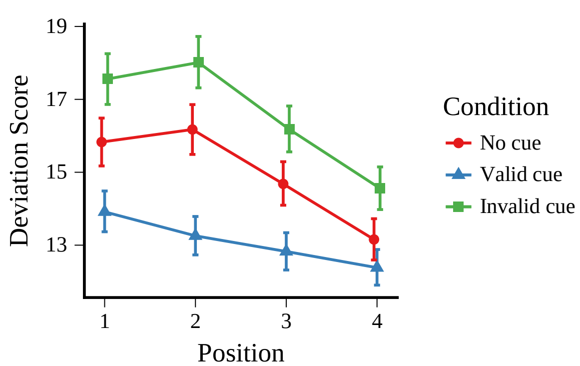

Figure 3: Deviation Score as a function of cue condition and serial position in Experiment 2. Error bars are 95% CIs.

To test for effects of cue condition, we used an LMEM with fixed effects for cue condition (contrast-coded as *valid-cue* vs. *no-cue* and *invalid-cue* vs. *no-cue*), position, the interaction between cue condition and position, and the same covariates as Experiment 1. The random effect structure included random intercepts for participants and syllables, along with random slopes for cue condition, position, and the interaction between the two by participant and syllable. *Valid-cue* trials had significantly lower Deviation Scores ($t = -3.34$, $p = .002$) and *invalid-cue* trials had significantly higher Deviation Scores ($t = 2.76$, $p = 0.008$) compared to *no-cue* trials. There was also a significant main effect of serial position, with Deviation Scores decreasing for more recent syllables ($t = -4.52$, $p < .001$). Interestingly, there was also a reliable interaction between the *valid-cue* condition and position ($t = 2.14$, $p = .037$), but no such interaction between the *invalid-cue* condition and position ($t = -0.35$, $p = .729$).

To further explore the influence of cueing on serial position effects, we fitted separate models for the *valid-cue* and *no-cue* conditions with polynomial position contrasts for position to test for both linear and quadratic (bow-shaped) serial position effects. In the *no-cue* model, there were both significant linear ($t = -5.02$, $p < .001$) and quadratic ($t = -2.38$, $p = .023$) effects of position, whereas only the linear effect was significant ($t = -3.34$, $p = .001$) in the *valid-cue* model.

Discussion

The results of Experiment 2 closely mirrored those reported in visual working memory (Gorgoraptis et al., 2011). Manipulation of attention led to a significant decrease in Deviation Scores for *valid-cue* trials (i.e., cued syllables), with a corresponding decrease in precision for *invalid-cue* trials (i.e., the uncued syllables presented in the same sequence as a cued syllable), relative to *no-cue* trials. Cueing also reduced the effect of serial position on recall, as evidenced by the significant interaction between the *valid-cue* condition and both linear and quadratic effects of position. In particular, the prototypical decrease in accuracy (or in this case precision) for items in the middle of a list, which was present in the *no-cue* condition for this experiment, was eliminated by cueing. In summary, the results of Experiment 2 showed that the distribution of resources is flexible and can be influenced by attention.

General Discussion

In the visual domain, evidence from experiments measuring the deviation of responses from the target instead of binary accuracy has provided support for a resource model of working memory. Moreover, the division of such resources has been shown to be flexible and subject to regulation through top-down attention. Through the use of similar continuous measures in a phonological working memory task, we were able to (a) verify the predictions of a resource model for phonological materials, and (b) to show that, as in the visual domain, resources can be flexibly allocated, such that items prioritized by top-down attentional cues will receive more resources at the cost of those that are deprioritized. These results are consistent with a flexible resource model of phonological working memory and, more generally, with a domain-general mechanism of resource allocation operating on both visual and verbal domains, as proposed in Baddeley and Hitch's (1974) central executive.

A key advantage of a resource model is its biological plausibility: it has been proposed that the quality of the representations stored in working memory is proportional to the *gain* (amplitude of neural activity) of the populations of neurons encoding those representations (van den Berg, Shin, Chou, George, & Ma, 2012). Due to the energy cost of maintaining high gain, there is an upper bound on the total activation across populations. Given these constraints, the optimal strategy for a resource allocation mechanism is to divide the available gain between items in proportion to their relative importance (e.g., based on attentional cues)—which would produce precisely the pattern of performance observed in the experiments described in this paper and previous experiments in the visual domain. The identification of working memory resources with neural gain also provides a natural explanation for the relationship between working memory and attention, in that attention has been shown to modulate neural gain (e.g., McAdams & Maunsell, 1999). This is consistent with the claim that working memory is the subset of long-term memory currently within the focus of attention (e.g., Cowan, 2001).

One might object that the manipulation employed in the current study does not reflect how people process language in everyday life. While it is certainly true that the processing of phonemes in the context of words and sentences involves additional operations, the main point that these experiments make is that similar principles *can* explain the division of resources in clearly separate domains of vision and auditory verbal processing. Moreover, we have demonstrated that, as in the visual domain, a continuous deviation score can be obtained and used to measure the performance of phonological working memory. Future work can take advantage of these results and further explore the degree to which low-level information is retained in phonological vs. visual working memory during processing of larger units like words and sentences.

Finally, while the current results do not speak directly to another critical debate regarding the effect of temporal delay vs. interference on working memory performance (e.g., Oberauer & Lewandowsky, 2008), the paradigm used in this study easily lends itself to manipulations of time delay and the similarity between stimuli that can help distinguish between decay and interference models.

In conclusion, the results of this work shed light on the nature of the central executive proposed by Baddeley and Hitch (1974) by specifying a clearly defined, empirically falsifiable, and biologically plausible mechanism for its operation: the central executive divides resources continuously between domain-specific representations that need to be held in working memory, and, in both visual and verbal domains, the partitioning of these resources is determined by the prioritization of items in the attentional space.

References

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7–7. https://doi.org/10.1167/9.10.7

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Fraley, C., & Raftery, A. E. (2002). Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association*, *97*(458), 611–631. https://doi.org/10.1198/016214502760047131

Gorgoraptis, N., Catalao, R. F. G., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *The Journal of Neuroscience*, *31*(23), 8502–8511. https://doi.org/10.1523/JNEUROSCI.0208-11.2011

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in linear mixed effects models (Version 2.0-33).

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*(5), 358–368. https://doi.org/10.1037/h0044417

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. https://doi.org/10.1038/nn.3655

Massaro, D. W., & Cohen, M. M. (1983). Categorical or continuous speech perception: A new test. *Speech Communication*, *2*(1), 15–35. https://doi.org/10.1016/0167-6393(83)90061-4

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, *19*(1), 431–441.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review*, *115*(3), 544–576. https://doi.org/10.1037/0033-295X.115.3.544

R Core Team. (2017). R: A language and environment for statistical computing (Version 3.4.1). Vienna, Austria: R Foundation for Statistical Computing.

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*(22), 8780–8785. https://doi.org/10.1073/pnas.1117465109

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11. https://doi.org/10.1167/4.12.11