Sequential study effects of free time and word frequency in long-term memory: A critical test of four theories using a large-scale publicly available dataset

Authors: Vencislav Popov[1,2], Michael Griffin[1], Lynne M. Reder[1,3]

[1] Department of Psychology, Carnegie Mellon University

[2] Department of Psychology, University of Zurich

[3] Center for the Neural Basis of Cognition, Pittsburgh, PA

Corresponding author: Vencislav Popov, vencislav.popov@gmail.com

Authors' Note:

Vencislav Popov is now in the Department of Psychology, University of Zurich

The data and analysis code are openly available at https://github.com/venpopov/sequential-effects-of-free-time-and-frequency

Abstract

A recent theory proposes that storing information in long-term memory depletes a limited resource which recovers gradually over time. Here, we provide a critical test of this theory by showing that free recall is worse if a word is preceded during study by a low-frequency word and by a short inter-stimulus-interval (ISI). These effects interacted – the preceding study word's frequency effect decreased as the pre-ISI increased. The duration of subsequent ISI had no effect. The resource-depletion-and-recovery model explains these results by positing that the amount of resources required for encoding in long-term memory is an inverse function of the existing representation's strength, and that these resources recover more when the pre-ISI is longer. The theory's predictions are contrasted with three alternatives – temporal distinctiveness, selective rehearsal and refreshing, and short-term consolidation – none of which fit the data. We compare each set of predictions using the PEERS dataset (Healey & Kahana, 2016). The resource-depletion-and-recovery theory's predictions provided the best match to the data.

Keywords: memory resources, long-term-memory, encoding

Sequential study effects of free time and word frequency in long-term memory: A critical test of four theories using a large-scale publicly available dataset

Despite the seemingly unlimited capacity of long-term memory (LTM; Brady et al., 2008; Konkle et al., 2010; Standing, 1973), there is a limit on how much information can be added to LTM at any given time. As an example of this limit, consider that both short-term memory (STM) and LTM performance consistently degrades when the “free time” between study items is reduced (Brown et al., 2006; Criss & McClelland, 2006; Malmberg & Nelson, 2003; Mizrak & Oberauer, 2021; Ricker & Hardman, 2017). This limitation could arise through various mechanisms – it could be due to reduced temporal distinctiveness (Brown et al., 2006), due to reduced opportunity for articulatory rehearsal or attentional refreshing (Camos et al., 2009), or due to the interruption of a consolidation process (Ricker & Hardman, 2017). In addition to these three, now standard, explanations for the benefit of extra study time, a recent theory developed in our lab proposes that storing information in LTM requires the use of a limited encoding resource that *recovers gradually over time* (Popov & Reder, 2020, in press; Reder et al., 2007)[1]. This gradual resource recovery assumption was developed to account for the mixed-list word frequency paradox (Popov & Reder, 2020; Reder et al., 2007), but it could be extended to account for the free time benefit. There is little consensus as to why human memory improves with more free time during study, and the goal of this paper is to provide a critical test of the four theoretical accounts mentioned above.

[1] This theory is an evolution of the Source of Activation Confusion (SAC) model originally developed by Reder and Schunn (1996), Schunn et al. (1997) and Reder et al. (2000) and builds upon ideas about the low-frequency encoding disadvantage presented in Diana and Reder (2006) and Reder et al. (2007).

One promising area in which to test these theories is at the trial level: 1) what are the effects of free time when it varies either before or after a presentation (also see Mizrak & Oberauer, 2021); and 2) does the size of the free time benefit interact with stimulus difficulty? To test these, we turn to the Penn Electrophysiology of Encoding and Retrieval Study (PEERS) dataset, a highly powered public list-learning free recall dataset that, critically, varies the inter-stimulus-interval (ISI) between words in each list on a trial by trial basis (Figure 1a). When the ISI varies in duration from trial to trial, it is possible to distinguish between *proactive* and *retroactive* effects of free time – a longer preceding ISI (pre-ISI) or a longer following ISI (post-ISI) could affect memory differently. This is valuable, because it provides a particularly diagnostic test between the above theories: to explain the benefit of free time, each theory makes very different predictions concerning the pre- vs post-ISI effects.

We begin with the resource-depletion-and-recovery theory of SAC. This theory assumes that 1) storing information in LTM requires binding item and context information together, that 2) the amount of resources required for this binding is an inverse function of the strength of the item's representation, that 3) the strength of these bindings will be weaker if you do not (or cannot) allocate sufficient resources, and that 4) these resources recover gradually over time. Within this account, free time helps memory by allowing the encoding resources to recover or replenish to a greater extent before the next study item appears.

Support for the resource-depletion-and-recovery theory comes primarily from the recent discovery of "sequential study effects" – that memory for one study item depends on how difficult it was to process the items that preceded it during study. For example, memory for an item in study position N is better, if the items in positions N-1, N-2, etc. are (a) high- rather than low-frequency words, (b) more frequently (as opposed to less frequently) experienced pseudo-words during the training that preceded studying the critical list, (c) words presented more frequently during the study session itself, or (d) words followed by instructions to forget rather than to remember them (Popov et al., 2019; Popov & Reder, 2020). These sequential study effects accumulate, such that a longer sequence of difficult-to-process items hurts memory for subsequent items even more. Importantly, these results cannot be due to other mechanisms such as selective rehearsal, distinctiveness, or attentional borrowing (Popov et al., 2019). The key insight behind these results is that some items (low frequency words, to-be-remembered items, etc.) deplete more resources during encoding and they leave fewer resources available for processing additional items.

Since the PEERS dataset also naturally varies word frequency, this theory makes two critical predictions concerning the effect of free time between items and its interaction with the sequential frequency effect. SAC's postulation of resource depletion and recovery predicts that increasing the temporal gap *before* a study item (i.e., the pre-ISI) would allow resources to recover to a greater degree, thereby increasing memory performance for the item that follows the longer interval. In addition, since increasing the pre-ISI leads to greater resource recovery, a longer pre-ISI should also diminish the sequential study effects, described above, concerning the roll of frequency of preceding study items. Specifically, the frequency of the preceding study items should matter less if there is more time for resources to recover before studying the next item in the list.

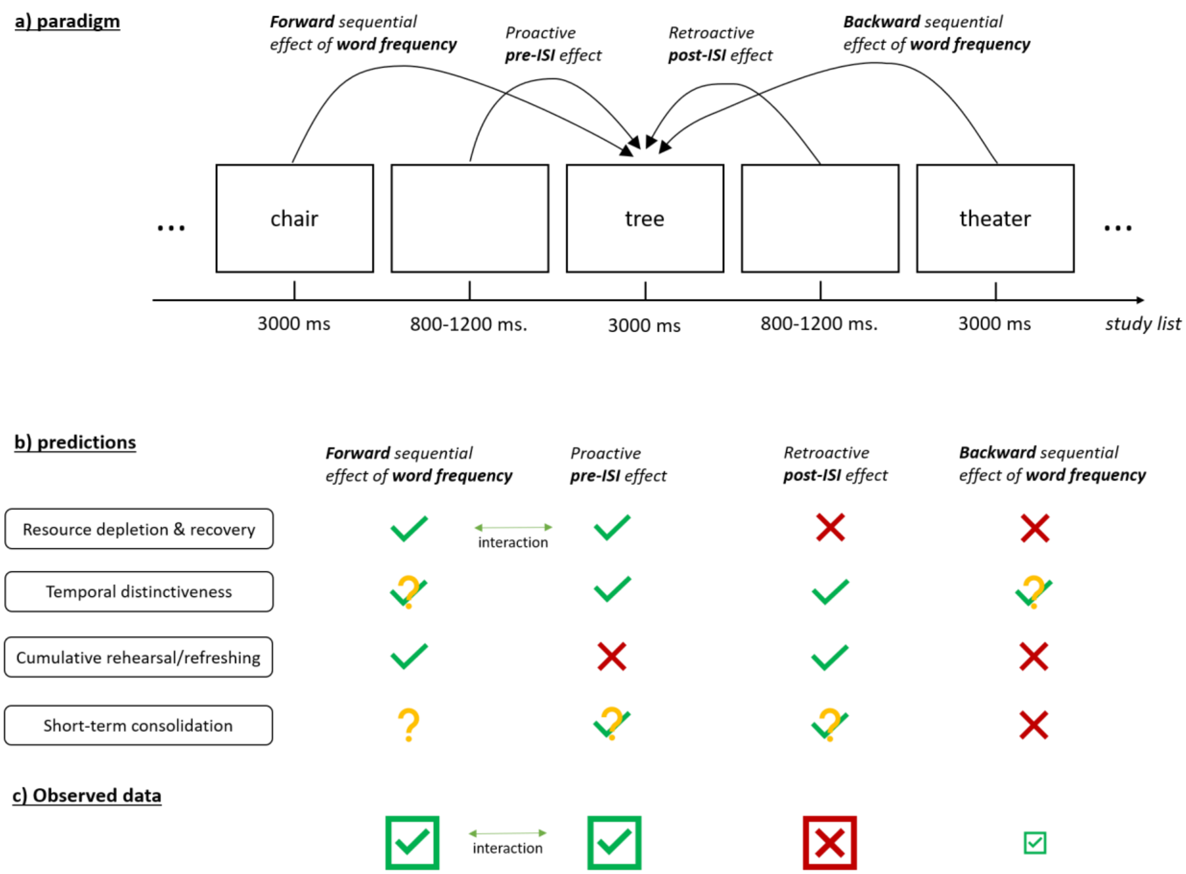

Figure 1. *a) An illustration of the study procedure in PEERS. Each word was presented for 3000ms and the ISI was uniformly distributed between 800ms and 1200 ms; free recall performance for the word "tree" could depend on the frequency of the preceding study word (forward sequential effect of word frequency), on the frequency of the subsequent word (backward sequential effect of word frequency), on the duration of the preceding interval (proactive pre-ISI effect), or on the duration of the subsequent interval (retroactive post-ISI effect). b) Predictions about the four different sequential effects for the resource-depletion-and-recovery theory and the temporal distinctiveness theory. Green checkmarks mean that the effects is predicted by the theory, red X marks mean that the theory predicts no effect, and the yellow question marks mean that the prediction from the theory is unclear because it depends on additional assumptions. c) Observed data pattern (a smaller checkmark reflects a smaller effect).*

In contrast, the other existing explanations for the benefit of free time make very different predictions concerning the pre- vs post-ISI effects and the sequential study effects (see Figure 1b). Temporal distinctiveness theory assumes that items are stored along a temporal dimension and that retrieval is a function of how temporally distinct items are from each other (Brown et al., 2006, 2007; Morin et al., 2010). Thus, adding free time on either side of a study item should benefit memory equally. When it comes to the frequency of the surrounding items, they should have no effect, unless one assumes that LF items are more distinct, which might cause interference. In this case, the effects should again be both proactive and retroactive. It is not clear whether this account predicts an interaction between the two factors.

Two maintenance mechanisms could also explain the benefit of free time and the benefit of preceding high-frequency words – verbal rehearsal and/or attentional refreshing (Barrouillet et al., 2011; Camos et al., 2009). For free time, this account predicts that the effect should be retroactive but not proactive – more time following the word presentation allows more opportunities for rehearsal or refreshing. If we assume that low-frequency items are more likely to be forgotten, this account also predicts that people could spend more time rehearsing prior low-frequency items, thus hurting memory for the current item.

A final possibility is that words take a certain amount of time to be consolidated, and that more time during the ISI allows the consolidation process to complete (Jolicœur & Dell'Acqua, 1998; Ricker et al., 2018; Ricker & Hardman, 2017). Thus, the benefit should be both retroactive and proactive – a longer post-ISI allows consolidation of the prior item to finish, while a longer pre-ISI reduces the possibility that consolidation of the prior item will interrupt encoding of the current item. When it comes to the sequential frequency effect, the prediction is unclear – one could assume that low-frequency words take longer to consolidate, which would impair words for the subsequent items. However, the typical consolidation process in STM is assumed to take between 500-1000ms (Ricker & Hardman, 2017), which is a significantly shorter time than the intervals used in this study (3800-4200ms per item, see Figure 1a).

To contrast these four mechanisms, and to provide the most stringent test to date of SAC's resource-depletion-and-recovery assumption, we reanalyzed data from the PEERS dataset. This existing dataset is suitable for testing the predictions discussed above for several reasons. First, as mentioned above, the ISI in the PEERS study varied randomly on each trial from 800ms to 1200ms (see Figure 1a). Second, while temporal isolation effects are often unreliable in serial recall tasks (Brown & Lewandowsky, 2005; Lewandowsky et al., 2006; Peteranderl & Oberauer, 2018), they have been repeatedly observed in long-term free recall and recognition tasks (Brown et al., 2006; Morin et al., 2010). Third, this dataset involves several hundred participants who completed multiple hour-long sessions – as such, it has sufficient power to detect even small effects (previous null effects in the literature could be attributed to lack of power, see Morin et al., 2010, for this discussion).

A previous reanalysis of the PEERS dataset has demonstrated that memory for word X was better when the preceding item during study X-1 was of higher frequency (Popov & Reder, 2020). The critical questions in this paper are whether there would be a beneficial effect of the pre-ISI and post-ISI on free recall performance, and whether a longer pre-ISI would diminish the sequential effects of word frequency. As Figure 1b demonstrates, the combination of these predictions provides a highly diagnostic test of the four possible explanations.

Materials and Methods

These methods are described in detail in Healey & Kahana (2016). An abbreviated version is included here to facilitate comprehension of the new information reported herein. All analyses and results are novel and have not been previously reported.

Participants

Our analyses focus on the 230 participants who completed Experiment 1 of PEERS. The sample included 192 college students (age range: 18-30 years) and 38 older adults (age range: 61-85 years). For full demographic information, please consult Healey & Kahana (2016).

Procedure and materials

In Experiment 1 of PEERS, participants performed a free recall task that consisted of 7 sessions. During each session, participants were presented with 16 lists of 16 words each. During study, words were displayed one at a time in the middle of a computer screen. Each word was shown for 3000ms, followed by a jittered, uniformly distributed ISI of 800-1250ms (see Figure 1a). Immediately after all 16 words within a list were presented, participants were given 75 seconds to recall as many words from the list as possible. For each participant, each word was randomly selected from a pool of 1638 words. These words varied randomly in their natural frequency.

Data analysis

We analyzed the free recall performance using logistic mixed-effects regression models (Baayen et al., 2008). Random effects were determined through restricted likelihood ratio tests and all final models included varying intercepts for subjects and individual words/word pairs (i.e., subjects and items differ in their overall accuracy). We excluded the last four words of each list from the analysis because performance for them could be based on short-term rather than LTM. The final dataset included 247,914 free recall observations. We inferred the significance of each effect based on likelihood ratio tests and AIC comparisons of the regression models that contained the effect in question with identical models that lacked this contrast.

Predictors in the model included the duration of the pre-ISI and the post-ISI, the log frequency of the preceding study word, the current study word and the subsequent study word (as measured by the SUBTLEX norms; van Heuven et al., 2014). In addition, the model included the interaction between the pre-ISI and the preceding word's frequency, as well as the interaction between the post-ISI and the subsequent word's frequency. All predictors used in the models were continuous; however, for illustration purposes we binned each variable. For the pre-ISI, we created four bins – 800-900ms, 900-1000ms, 1000-1100ms and 1100-1200ms. Since ISI was uniformly distributed, each bin contained approximately an equal number of observations. The log word frequency was instead normally distributed. To avoid having fewer observations for higher and lower frequencies, we sorted words by frequency and then created 20 bins such that each bin contained an equal number of observations.

Results

We found that as the pre-ISI increased, so did the recall probability for the word that followed it (see Table 1 for a summary of the regression results). As can be seen from Figure 2 (left), the effect was small – recall probability increased by ~1% as the pre-ISI increased from 800-900ms to 1100-1200ms. This is not surprising, since the average difference in duration between trials these two bins is just ~300ms (small relative to a total trial duration of 3800-4200ms). This effect size is comparable to that obtained by Brown et al (2006) – in their study the ISI varied from 0-3500ms, and memory increase by ~10% as the ISI increased from 0 to 3500ms. Consistent with the resource-depletion-and-recovery theory, there was no retroactive benefit of the post-ISI length (Figure 2, right).

As described in Popov and Reder (2020), recall probability also increased as the frequency of the preceding study word increased (Figure 3, left). Importantly, we found that the pre-ISI and the preceding word frequency factors interacted significantly. As can be seen from Figure 3, the effect of the preceding study word frequency got smaller as the pre-ISI duration got longer (as indicated by the differences in the slope of the lines for the different pre-ISI durations). The left panel shows the raw data; to illustrate the differences in the slope for the preceding word frequency as a function of preceding ISI duration, the right panel of Figure 3 shows the data after the mean main effect of preceding ISI was subtracted from each condition (see Figure A1 in the Appendix for the uncorrected raw data).

Surprisingly, there was also a smaller backward effect of the frequency of the subsequent word (Table 1) – words followed by higher frequency words tended to be recalled more easily. This effect was significantly smaller than the effect of the preceding word's frequency, as confirmed by a z-test comparison of the regression parameters (Table 1). Finally, in contrast to the interaction between the preceding item's frequency and the pre-ISI duration, the frequency of the following item did not interact with the post-ISI.

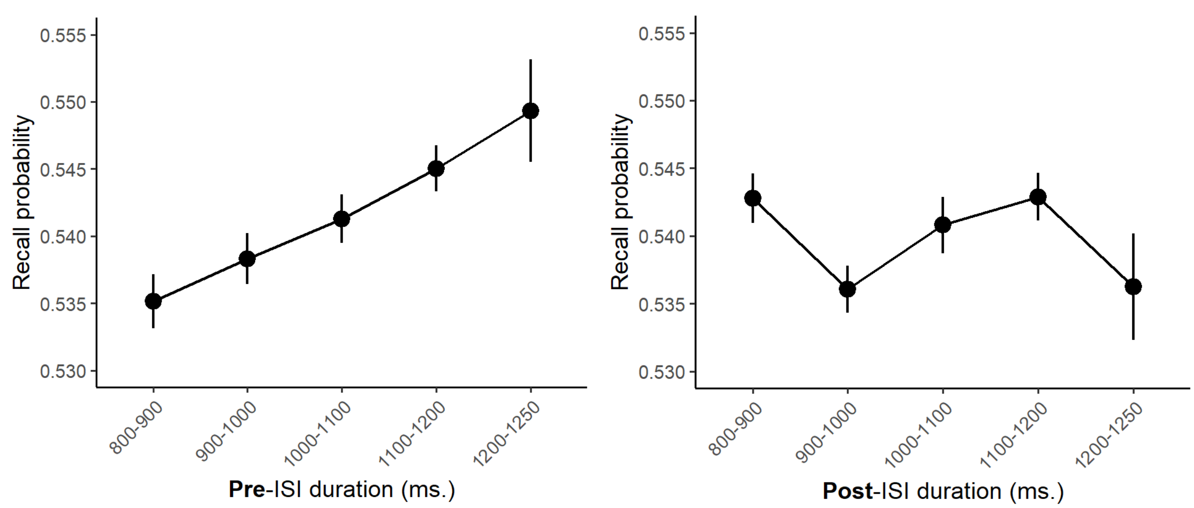

Figure 2. *Recall probability depending on the Pre-ISI interval (left) or the Post-ISI interval (right)*

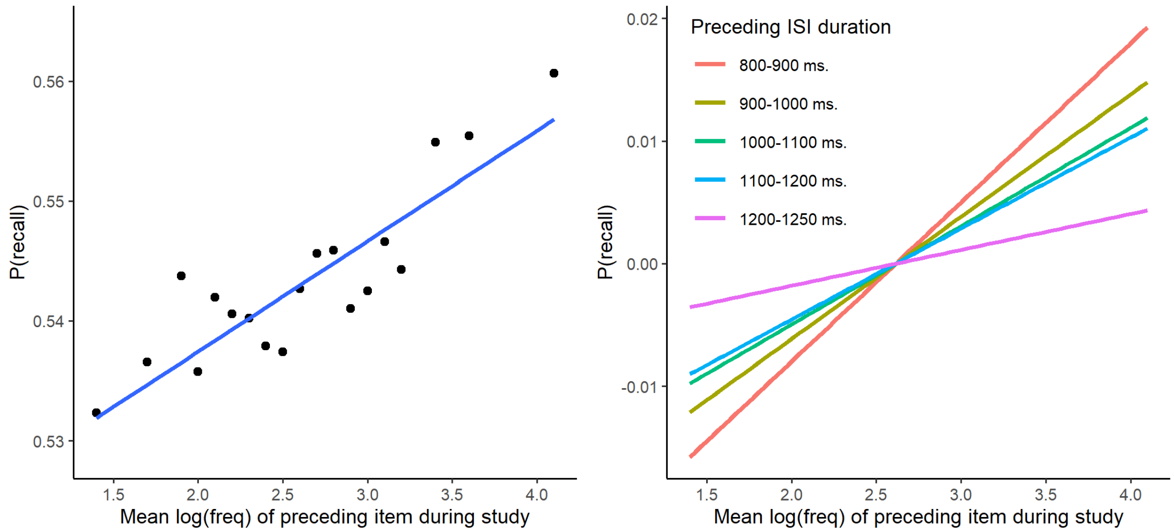

Figure 3. *Left – recall probability as a function of the frequency of the items that preceded the current. Right – changes in the preceding frequency effect slope as a function of the preceding ISI. The colored lines show the best fitting regression lines The right panel shows the regression fit after the main effect of preceding ISI (intercept) has been subtracted so that the slopes to be compared more easily (see figure A1 for the uncorrected raw data).*

Discussion

Out of the four theoretical mechanisms we discussed in the introduction, the SAC resource-depletion-and-recovery theory's predictions come closest to the data (Figure 1c). It is the only account that predicts the presence of a proactive pre-ISI benefit together with the absence of a retroactive post-ISI effect, as well as the interaction between the prior item's frequency and the pre-ISI duration. None of the other theories can account for the lack of a post-ISI effect and they cannot explain why the effect of the preceding word frequency decreases as the pre-ISI gets longer. In contrast, this prediction comes naturally from the resource depletion account – preceding LF words deplete more resources and leave fewer resources available for processing the current item but a longer pre-ISI allows these resources to recover more, diminishing the effect.

Table 1 Parameter estimates for the mixed-effects logistic regression on predicting recall performance

Main fixed-effects	**β**	**β (std. err.)**	**z**	***p***	***ΔAIC****
Intercept	-0.510	0.080			
Current word's frequency	0.067	0.015	4.447	< .001	
Preceding word's frequency	0.071	0.014	5.249	< .001	-45
Following word's frequency	0.051	0.014	3.835	< .001	-30
Pre-ISI duration	0.494	0.157	3.147	.002	-13
Post-ISI duration	0.261	0.157	1.668	.095	-1
Interactions between fixed-effects	**β**	**β (std. err.)**	**z**	***p***	***ΔAIC****
Pre-word frequency X pre-ISI duration	-0.133	0.059	-2.273	.023	-3
Post-word frequency X post-ISI duration	-0.076	0.059	-1.296	.195	0
Random-effects	**σ**				
Subject intercept	0.49				
TBR instructions for the item at lag1	0.33				
Parameter comparisons	**z**	***p***			
Pre-word freq > Post-word freq	2.307	.021			

Note: *Aikake Information Criterion (AIC) evidence for the full model relative to a model that lacks this effect. Smaller values reflect a better fit of the full model. ΔAIC < -2 reflects that including the relevant factors significantly improves the regression fit.

One surprising finding was that the subsequent item's frequency also affects memory for the current item. This prediction is the only one that is inconsistent with the resource-depletion-and-recovery theory. Only the distinctiveness account predicts both a forward and a backward sequential frequency affect, if we assume that low-frequency words are more distinct. However, the distinctiveness account would predict that the two effects should be of similar magnitude, and that there should be a post-ISI effect as well, which is not what we found. It is possible that all the other effects are generated by the resource-depletion-and-recovery-mechanism, while the smaller effect of the subsequent item's frequency reflects some form of attentional capture that is greater for low-frequency words (Murdock, 1998). The attentional-likelihood theory proposes that low-frequency words attract more attention, thus presenting a lower-frequency word after the current item might be more likely to detract attention away from it.

The effects reported here provide a critical test of different theories that attempt to explain how free time affects memory. Any model of these results should be able to account for 1) the proactive benefit of having high frequency preceding words, 2) the proactive benefit of having a longer pre-ISI, 3) the diminishing benefit of the preceding item's frequency as the pre-ISI increases, 4) the absence of a retroactive post-ISI benefit. The idea that people utilize the free time to rehearse or attentionally refresh previous items provides the worst match to the data. Furthermore, prior research has shown that suppressing rehearsal or dividing attention does not attenuate sequential study effects, as would be predicted by the differential rehearsal or refreshing accounts (Popov et al., 2019). Thus, the resource-depletion-and-recovery theory is currently the only to provide a near complete account of the data.

These results are consistent with a similar recent finding in a working memory serial recall task (Mizrak & Oberauer, 2021). Mizrak and Oberauer found that if one of the intervals in a list of seven words is longer (2500ms) rather than shorter (500ms), this leads to better memory for all subsequent items, but not for the preceding items. They also argue that these findings are not consistent with maintenance-based mechanisms such as selective rehearsal or refreshing of preceding items, nor with temporal distinctiveness accounts. They conclude that the results could be explained by either the depletion of a gradually recovering resource or by assuming that longer intervals allow the preceding items to be chunked and offloaded to LTM which, in effect, reduces the set size maintained in WM. While their findings cannot distinguish between the two explanations, the results presented here cannot be explained by offloading some parts of the list to LTM since this is already a LTM task. As such, the resource-depletion-and-recovery theory provides a more parsimonious account of both findings.

Open Practices Statement

The predictions for this experiment were not preregistered. The PEERS dataset is freely available at http://memory.psych.upenn.edu/Data_Archive, courtesy of M. Kahana. The analysis code used for the reanalysis is available at https://github.com/venpopov/sequential-effects-of-free-time-and-frequency

Appendix A:

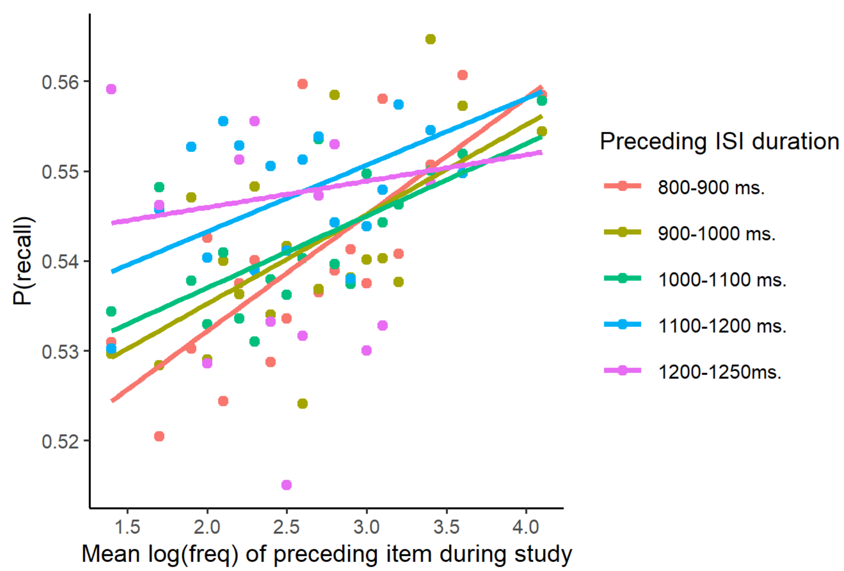

Figure A1: The raw interaction between preceding study item frequency and pre-ISI duration. In contrast to Figure 3, the main effect of pre-ISI has not been removed.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, *118*(2), 175–192.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brown, G. D., & Lewandowsky, S. (2005). Serial recall and presentation schedule: A micro-analysis of local distinctiveness. *Memory*, *13*(3–4), 283–292.

Brown, G. D., Morin, C., & Lewandowsky, S. (2006). Evidence for time-based models of free recall. *Psychonomic Bulletin & Review*, *13*(4), 717–723.

Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576.

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*(3), 457–469. https://doi.org/10.1016/j.jml.2009.06.002

Criss, A. H., & McClelland, J. L. (2006). Differentiating the differentiation models: A comparison of the retrieving effectively from memory model (REM) and the subjective likelihood model (SLiM). *Journal of Memory and Language*, *55*(4), 447–460. https://doi.org/10.1016/j.jml.2006.06.003

Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 805–815.

Healey, M. K., & Kahana, M. J. (2016). A Four–Component Model of Age–Related Memory Change. *Psychological Review*, *123*(1), 23–69. https://doi.org/10.1037/rev0000015

Jolicœur, P., & Dell'Acqua, R. (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology*, *36*(2), 138–202. https://doi.org/10.1006/cogp.1998.0684

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene Memory Is More Detailed Than You Think: The Role of Categories in Visual Long-Term Memory. *Psychological Science*, *21*(11), 1551–1556. https://doi.org/10.1177/0956797610385359

Lewandowsky, S., Brown, G. D., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language*, *54*(1), 20–38.

Malmberg, K. J., & Nelson, T. O. (2003). The word frequency effect for recognition memory and the elevated-attention hypothesis. *Memory & Cognition*, *31*(1), 35–43.

Mizrak, E., & Oberauer, K. (2021). What is time good for in working memory? In *Psychological Science*. https://doi.org/10.31234/osf.io/ahqwj

Morin, C., Brown, G. D. A., & Lewandowsky, S. (2010). Temporal isolation effects in recognition and serial recall. *Memory & Cognition*, *38*(7), 849–859. https://doi.org/10.3758/MC.38.7.849

Murdock, B. B. (1998). The mirror effect and attention-likelihood theory: A reflective analysis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *24*(2), 524–534.

Peteranderl, S., & Oberauer, K. (2018). Serial recall of colors: Two models of memory for serial order applied to continuous visual stimuli. *Memory & Cognition*, *46*(1), 1–16. https://doi.org/10.3758/s13421-017-0741-0

Popov, V., Marevic, I., Rummel, J., & Reder, L. M. (2019). Forgetting Is a Feature, Not a Bug: Intentionally Forgetting Some Things Helps Us Remember Others by Freeing Up Working Memory Resources. *Psychological Science*, *30*(9), 1303–1317. https://doi.org/10.1177/0956797619859531

Popov, V., & Reder, L. M. (2020). Frequency effects on memory: A resource-limited theory. *Psychological Review*, *127*(1), 1–46. https://doi.org/10.1037/rev0000161

Popov, V., & Reder, L. M. (in press). Frequency effects in recognition and recall. In M. J. Kahana & A. D. Wagner (Eds.), *The Oxford Handbook of Human Memory*. Oxford University Press. https://doi.org/10.31234/osf.io/xb8es

Reder, L.M. & Schunn, C.D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In Reder, L.M., (Ed.) *Implicit Memory and Metacognition.* (pp. 45-77). Psychology Press.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 294–320.

Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J., & Dickison, D. (2007). Experience is a Double-Edged Sword: A Computational Model of The Encoding/Retrieval Trade-Off With Familiarity. In *Psychology of Learning and Motivation* (Vol. 48, pp. 271–312). Elsevier. http://linkinghub.elsevier.com/retrieve/pii/S0079742107480070

Ricker, T. J., & Hardman, K. O. (2017). The nature of short-term consolidation in visual working memory. *Journal of Experimental Psychology: General*, *146*(11), 1551–1573. https://doi.org/10.1037/xge0000346

Ricker, T. J., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: Insights from studies on attention and working memory: An overview of working memory consolidation. *Annals of the New York Academy of Sciences*, *1424*(1), 8–18. https://doi.org/10.1111/nyas.13633

Schunn, C. D., Reder, L. M., Nhouyvanisvong, A., Richards, D. R., & Stroffolino, P. J. (1997). To calculate or not to calculate: A source activation confusion model of problem familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(1), 3–29.

Standing, L. (1973). Learning 10,000 pictures. *The Quarterly Journal of Experimental Psychology*, *25*(2), 207–222. https://doi.org/10.1080/14640747308400340

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521