Neural Dynamics of Grapheme-Color Synesthesia Extend to Sequential Learning of Music via Simulated Hippocampal Mechanisms of Associative Memory

Jay Gopal and Jeremy Fleming

Cognitive, Linguistics, and Psychological Sciences Department, Brown University

CLPS 1492: Computational Cognitive Neuroscience

December 12th, 2023

Introduction:

Associative memory is defined as remembering the relationship between two usually unrelated things (Suzuki, 2008). One example includes associating someone's name with their face. This ability is central to the human condition and the applications are numerous in nearly every modern profession. While the hippocampus has also been shown to do the opposite of association, using sparse representations in order to distinguish between episodic memories (O'Reilly et al., 2012), it has been known for nearly half a century that the hippocampus plays a pivotal role in forming associative memories (Levy and Steward, 1979). The ability of the hippocampus to learn associative memories even extends to associating events separated by time (Ahmed et al., 2020). And importantly, the hippocampus is able to associate memories across different stimuli (Borders et al., 2017).

However, the hippocampus's ability to quickly associate concepts can also be a disadvantage. For instance, as shown in the AB-AC task, learning new associations during the AC trials leads to a reduced ability to remember AB trial associations (O'Reilly et al., 2012). Although humans never reach the levels of catastrophic interference that artificial neural networks and models do, this could still pose a problem. For example, one may be grocery shopping and decide to memorize a list of food items to buy. Before shopping, they may also take a look at a recipe and memorize the list of ingredients. When shopping and deciding what to buy, thinking about the shopping list may become difficult: the new association between the recipe and list of ingredients interferes with the old association between groceries and list of items to buy. As a result, the shopper could end up buying the wrong items (Kliegl and Bäuml, 2021).

Another example of a strong association leading to undesired results is the scenario that inspired the present study. Oftentimes during exams, one finds themselves humming or otherwise following along to a song they listened to while studying. This can even lead to remembering the song one was listening to instead of the answers to a question, at least in our personal experience. We theorize that the hippocampus's ability to associate information across modalities is what leads to this phenomenon, as the association of learned information with auditory stimuli seems to yield this unwanted behavior. To test this hypothesis, we adapted a model of the hippocampus to learn letters (and "study for an exam") while simultaneously listening to music and saw if the model would be able to spontaneously recover the music when "taking" the exam by being prompted with letters without music. The network we adapted models yet another scenario where the brain has an over-association of stimuli: synesthesia.

Synesthesia is a phenomenon in which one stimuli is involuntarily linked to another stimuli, such as a number being linked to a color (Day, 2005) or a date being linked to a portion of space (Mann et al., 2009). While there are many different manifestations of this condition, the common thread is that a trigger is strongly linked to a synesthetic concurrent in a way that may not be expected. Synesthesia may occur across multiple sensory modalities, or be confined to just one sense (Neckar and Bob, 2014). Studying synesthesia may unlock the secrets of the brain, as it has been linked to a variety of other conditions and may connect to several fascinating biological phenomena. For example, a recent study found a potential link between synesthesia and autism (Hoadley and Hughes, 2017). Researchers have even argued that studying synesthesia could illuminate the secrets of consciousness itself (van Leeuwen et al., 2015).

One variant of the condition is grapheme-color synesthesia, which involves an individual seeing either a letter or number, but also seeing a corresponding color. This subtype of synesthesia is well-studied across a variety of domains. Studies have examined the cognitive processing mechanisms of grapheme-color synesthetes, and have found that they may have distinct neural patterns during motor execution and inhibition (Auki et al., 2023). Synesthetes exhibit diverse experiences, with some describing visually perceiving the color (referred to as projector synesthesia), while others have a cognitive recognition of the color's presence without a direct visual experience (known as associator synesthesia) (Dixon et al., 2004).

The model we adapted belongs to a project examining how plausible two possible different neural explanations of grapheme-color synesthesia are when implemented computationally. The first theory is that synesthetes have increased top-down feedback between traditionally distinct higher-order processing areas (Lungu et al., 2021). The second is that synesthesia involves heightened cross-activation between pathways at lower levels of processing (Barnett et al., 2008, Brang et al., 2012, Brang et al., 2013). These two theories may be somewhat concordant due to the distinction between projector and associator synesthesia. From a structural neuroscience perspective, researchers have reported increased connectivity of the inferior temporal cortex in synesthetes with projector synesthesia when compared to associator synesthesia, and that both subtypes of grapheme-color synesthesia display greater connectivity in the parietal and frontal cortices when compared to controls (Rouw and Scholte, 2007). This may support the theory that projector synesthesia is more closely tied to cross-activation than the other. Along the same lines, a study used dynamic causal modeling for functional magnetic resonance imaging (fMRI) and found that area V4 cross-activation was more closely tied to the top-down pathway for associator synesthesia, while it was tied to bottom-up activation for projector synesthesia (van Leeuwen et al., 2011). More broadly, the difference between the projector and associator forms of grapheme-color synesthesia is well-supported, with a study finding that the two subtypes could be reliably distinguished based on neural oscillations alone (Cohen et al., 2015).

In the present study, we examine how and why a grapheme stimulus can lead to involuntary recall of a sequential stimulus, in a similar fashion to how grapheme-color synesthesia leads to an involuntary linking of numbers or letters and colors. Given this clear link, we decided that adapting the synesthesia model was suitable. We will now discuss the steps we took to modify this model, and the experiments we ran to explore associative learning in an environment with music.

Methods:

Datasets:

To create our datasets, we adjusted the .dat files from the synesthesia model. These .dat files mapped two sets of 7x5 inputs to two sets of 5x2 outputs. The 7x5 inputs were able to show 10 different letters (representing a grapheme) and 10 different numbers (representing a color) respectively. The outputs then correspond to the 10 different correct classifications for both grapheme and color.

In our task, we kept the inputs and outputs for graphemes to represent how one might be studying 10 different questions for an exam. However, since music is sequential and often repeats chords or melodies throughout a song, we decided to create a new set of inputs to reflect this. Taking inspiration from the Simple Recurrent Network (SRN) model of the midterm, we aimed to compose our own sequence to represent music one would be listening to when studying for an exam. To represent this music input and output, we decided to change the numbers to instead be a 5x1 input and 5x1 output, as shown in Figure 1. These represent different notes of a scale that can be connected to make our composition. The 5x1 output in turn represents the model's prediction of the next note in the sequence. Our final "song" was the sequence ABCDBCDECB repeating, so for our first training example, we paired a first neuron input (A) with a second neuron output (B) to represent how the model should learn that the next note will be B (see Figure 1). The second training example, therefore, had an input of B with an output of C, and so on.

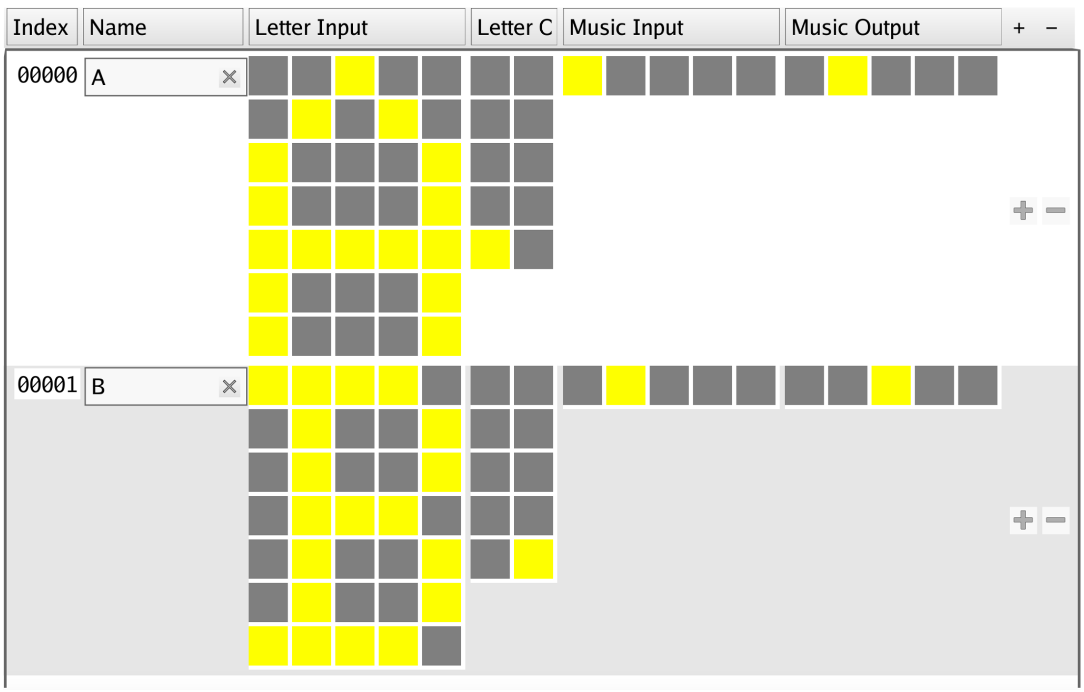

Figure 1: Examples from the training set are shown. Both inputs and outputs are included for letters and music. Importantly, the music input is provided, so it is possible to predict the music output solely from the music input.

The synesthesia project tested the model by comparing congruent and incongruent trials in a Stroop task and seeing if response times were affected. Since we instead tried to model spontaneous recall of music, our test dataset ended up diverging from the Stroop task test dataset. We simply removed all of the music inputs while keeping the grapheme inputs and outputs, as well as the music outputs, the same, as shown in Figure 2. The goal for the model was to be cued by the grapheme inputs and spontaneously begin thinking about the song it had associated with the graphemes. Therefore, to test our model, we looked at how many letters were correctly recovered when running through the test trials (and after our model was trained).

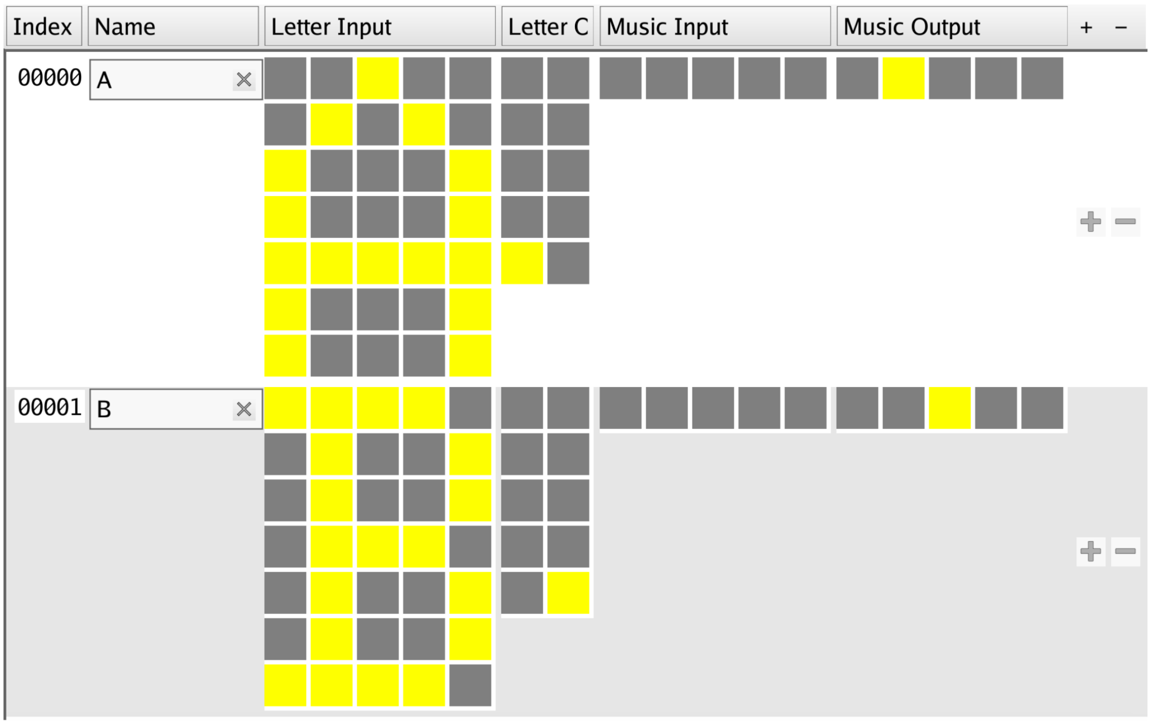

Figure 2: Examples from the test set are shown. The inputs and outputs for letters remain unchanged from the training set. However, no music inputs are included, so the prediction of music output relies on effectively connecting letters to music.

<u>Modeling Overview</u>:

The synesthesia model we were given consisted of two parallel networks: one aiming to learn the classification of letters and the other aiming to learn the classifications of colors. The two networks were connected in two ways that could be chosen by the user: an associator layer that connects each network via their output layers in a top-down manner, inspired by associator synesthesia, and a set of connections that connect the letter and color hidden layers, similar to what studies have noted in projector synesthesia.

We first created a control model in which the letter and music tasks were kept completely separate. After Jay had our control network working, Jeremy and Jay worked individually to augment this model's capabilities. We created two models, one for each of cross-activation (Shared Hidden) and top-down feedback (Top-Down). Jeremy worked to enhance the projector synesthesia version of the model and Jay worked to enhance the associator synesthesia version of the model in order to find the model of associative learning that would most accurately spontaneously recall music. We then combined the Shared Hidden and Top-Down models into a hybrid model. Jeremy worked on a hybrid network bringing together both of our modifications, evaluating whether the combined mechanisms would yield improved performance on the music task. We continued with this equal worksplit to evaluate our modified networks.

<u>Control Model (Jay)</u>:

My first major modification was changing the color processing portion of the network into a sequence processing model. First, I renamed the color processing network neurons and changed the size of its inputs and outputs to reflect the 5x1 inputs and outputs of our dataset. Then, again taking inspiration from the SRN midterm, I added a hidden context layer next to our music hidden layer. Importantly, this had the ability to copy previous iterations' music hidden layers activities exactly. I defined a parameter called **`FmHid`** to modulate what percentage of the activity of our custom hidden context layer came from the Music Hidden layer. I defined a second parameter called **`FmPrv`** to control the percentage of context activity that came from the previous time step's Music Hidden Context activity. Since our custom musical sequence, ABCDBCDECB, has first-order dependencies, the hidden context was necessary to ensure the music model could successfully learn the sequence. We were given this insight by Dr. Frank after our final project presentation to the class.

Incorporating the Music Hidden Context layer into our model was an exciting challenge for me, as I needed to add a completely new component to a model that had several existing parts. To be sure that I had implemented this correctly, I tested whether the control network could solve both tasks, achieving a percent error of zero reliably, across several random seeds. Those dozens of training plots aren't included in order to save space in the present text. As a final sanity check, I examined the receiving and sending weights of the Music Hidden and Music Hidden Context layers manually. Example sending weights are included in Figure 3, which clearly shows the desired connectivity. The Music Hidden Context layer only connects to Music Hidden (Figure 3A), while the Music Hidden layer sends weights to both the Music Hidden Context and the Music Output (Figure 3B).

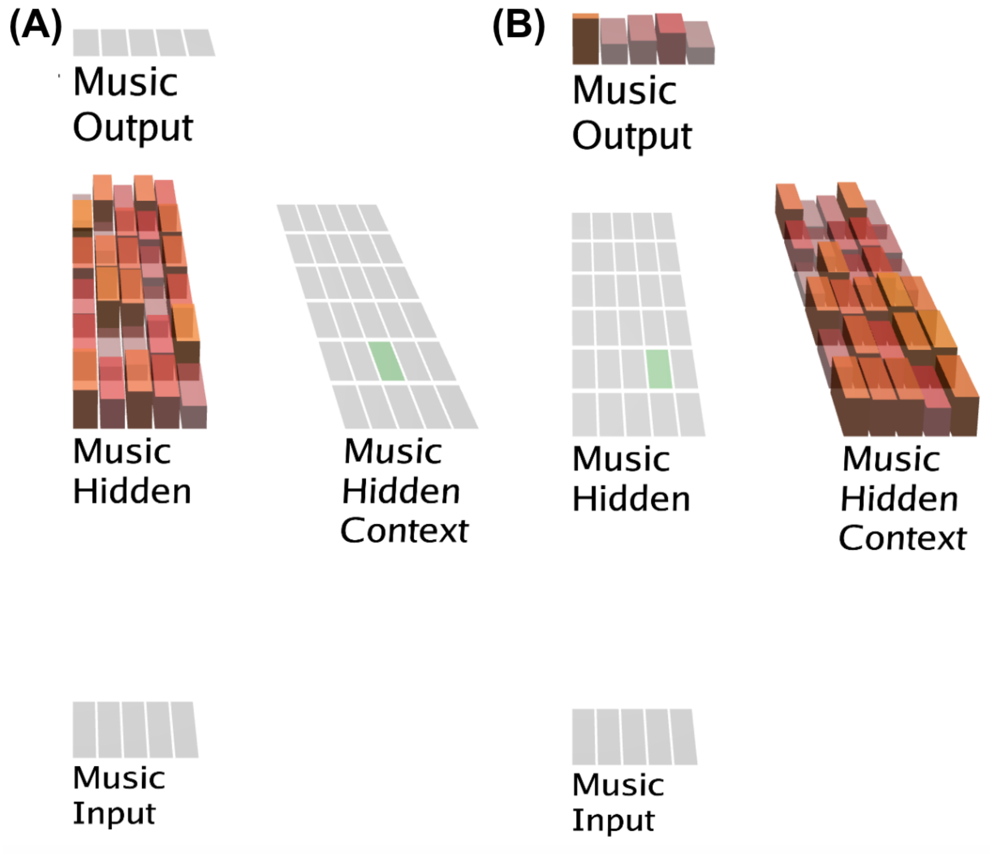

Figure 3: Sending weights for example units in (A) Music Hidden Context and (B) Music Hidden layers. Importantly, the former only projects to Music Hidden, while the latter projects to both Music Hidden Context and Music Output.

As a negative control, I ensured that the test accuracy was at random chance for the music task. The network needed to have the ability to generalize for the letter task, but not for the music task, as the latter should be completely impossible for the control network to achieve high performance on due to the lack of inputs.

<u>Shared Hidden Model (Jeremy)</u>:

My main change to the model was to include a shared hidden layer between the letter and music hidden layers. This created an even stronger connection between the hidden layers and represents a more connected version of projector synesthesia. I chose to add this layer to represent how the visual and episodic memory portion of the brain studying graphemes may be communicating with the auditory portion of the brain during our task.

Importantly, my Shared Hidden layer was bidirectionally connected to the other hidden layers, allowing them to organically learn a combined lower-level representation that aids in both tasks. My bidirectional connections added four hyperparameters, all of which I pulled into the GUI: **`SharedHiddenToHid`** (modulating connections from Shared Hidden to Letter Hidden), **`SharedHiddenToHid2`** (Shared Hidden to Music Hidden), **`HidToSharedHidden`** (Letter Hidden to Shared Hidden), and **`Hid2ToSharedHidden`** (Music Hidden to Shared Hidden). Jay and I hypothesized that these bidirectional connections would increase the testing accuracy on the music task due to the incorporation of information from the letter task, which had a surjective mapping onto the music outputs.

<u>Top-Down Model (Jay)</u>:

I modeled associator synesthesia by using an associator to connect the Music Output layer with the Letter Output layer. Echoing the work by Isabella and Peter in 2021, I started off trying bidirectional connections between the Associator and both of the Output layers. However, this proved to be incompatible with the simulation, resulting in crashes once “Init” and “Train” were hit. Critically, this was the same behavior reported for the model we were adapting. Thus, I replicated the feedforward connections of the original network, going from Color Output (converted to Music Output) to the Associator, and then going from Associator to Letter Output.

Once this was working, I reversed the direction of the connections to allow the letter task to exhibit top-down feedback onto the music task. My final network had a connection from the Letter Output to the Associator, and another connection from the Associator to the Music Output. Jeremy and I hypothesized this would allow for the network to learn to infer the Music Output solely from the Letter Output, similar to how those with projector synesthesia may be able to see a color when only presented with a letter or number.

I created two hyperparameters that could be set manually, and added these to the GUI to allow for faster iteration. The first hyperparameter, **`OutToAssoc`**, controlled the relative weights of the connections from Letter Output to Associator. The second was called **`AssocToOut2`** and modulated the connections from Associator to Music Output. Adding hyperparameters meant that we needed to rigorously search for the best combination of both. I defined "First Zero" as the first epoch during training when the network reached zero percent training error. I then modified RunLogs to also include all of the newly added hyperparameters by both Jeremy and myself, such that after several runs, the data could be extracted as a CSV to interpolate the ideal hyperparameters for any given model.

Jeremy and I worked together on hyperparameter interpolation for all models via a custom Python script that plotted all of the relevant hyperparameters on a heatmap and used the SciPy library to interpolate the ideal combination. I experimented with various interpolation methods, including linear, cubic, and nearest. Once the parameters minimizing First Zero were found, they were used in the simulation to yield the final results. I did not optimize hyperparameters specifically for music accuracy on the test set because we did not have a separate validation set for hyperparameter tuning. Thus, Jeremy and I determined that using the test set to find the ideal hyperparameters would lead to invalid results, and would leave us unable to determine the generalizability of our network.

Hybrid Model (Jeremy):

After adding our components to the model individually, we realized that we could combine the features we created in a new kind of model. I took the lead on combining our networks. Fortunately, the way we implemented our two main modifications allowed them to be easily compatible, as we both had moved the parameters to modulate our new components' weights into the GUI.

This model has both strong hidden layer connectivity, as well as strong top-down feedback from the associator layer. We hypothesized that the additional connectivity afforded by this combination of features would yield a more accurate model. After creating our two individual models, a control model, and this combined model, we decided to split up the models between the two of us when training and testing. Jay examined the control and the top-down feedback (associator) models and I examined the shared hidden (projector) and combined (hybrid) models.

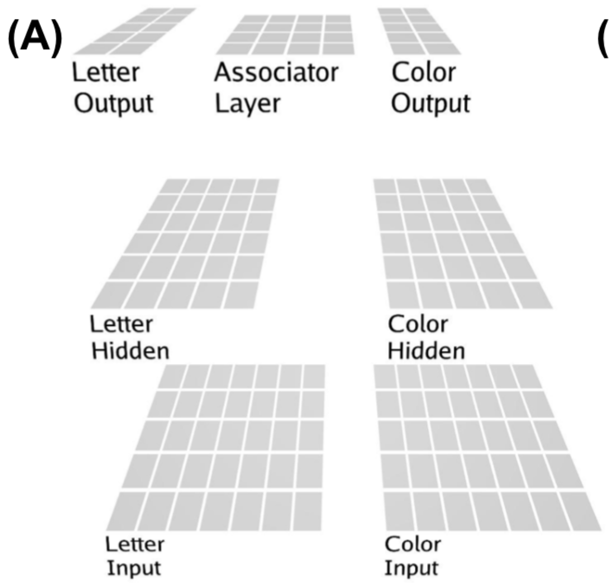

Figure 4: Overview of network structure for (A) the starter model that we adapted and (B) our final model. We replaced color with music, added a shared hidden layer, added a hidden context layer for music. Not pictured: we reversed the direction of top-down feedback.

Figure 4 above shows the comparison between the starter model from Isabella and Peter (Figure 4A) and the final hybrid model that we created (Figure 4B). The main additional components are clear: the Shared Hidden and Music Hidden Context layers.

Results:

Control (Jay):

I trained the control network with the Music Hidden Context layer enabled. Both hyperparameters modulating the bidirectional connections to this layer were set to 1 after they were tuned based on the epoch at which the first zero was reached. **`FmHid`** and **`FmPrv`** were tuned, and the ideal combination was found to be 0.3 and 0.7, respectively.

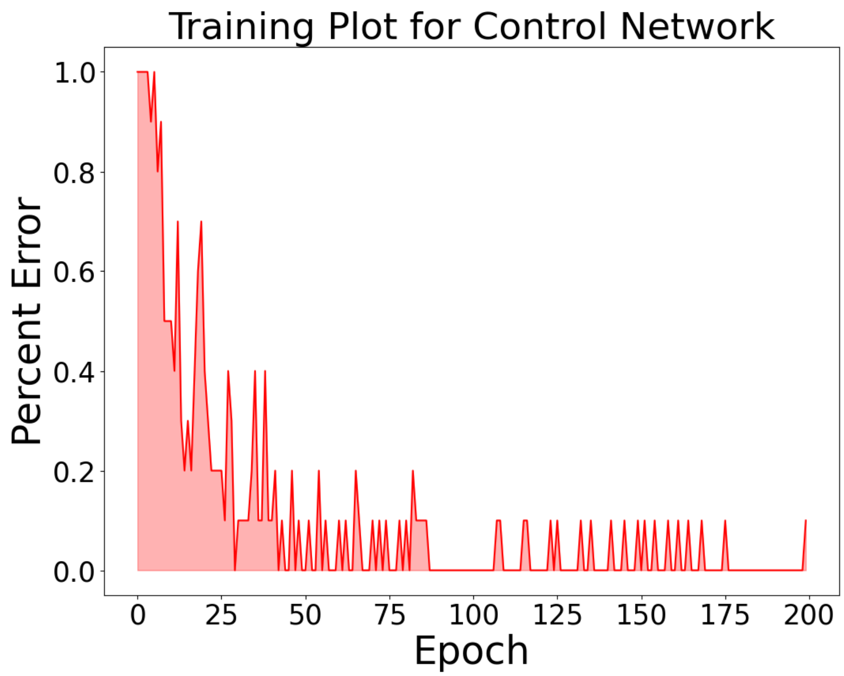

Figure 5: Training plot (percent error vs. epoch) for the control network, in which there was no connection between the music and letter tasks.

As shown in Figure 5, the percent error drops to zero during training, but fluctuates rather than staying at zero. I observed this trend across a wide variety of hyperparameter settings, and manually double-checked sending and receiving weights to ensure they were correct. While the fluctuation may not be fully understood, it's clear that the network reaches zero percent error for many training epochs.

Test set accuracy was measured for both tasks. Importantly, the network had 100% accuracy on the letters task. This was fully expected, as the music and letter components were not connected at all. Additionally, I found the network had 10% accuracy on the test set for music. This was also completely reasonable. Given no inputs, the network was simply guessing which note it should output at any given time.

Shared Hidden (Jeremy):

I trained many variants of the Shared Hidden network, fully exploring the hyperparameter space. For each hyperparameter, I logged the First Zero via RunLogs, and also visualized the training plot to observe how consistently the network remained at zero percent error. The results from the best combination of hyperparameters is shown in Figure 6. Like the control network, my network does not stay at zero for many epochs in a row. Instead, it fluctuates between zero and a small percent error. Interestingly, I show that my network is capable of achieving zero percent error for 14 epochs in a row (between epochs 168 and 182), but then starts fluctuating yet again. I saw many of these instances and chose a set of hyperparameters with the longest such streak of zeros, shown in Figure 6.

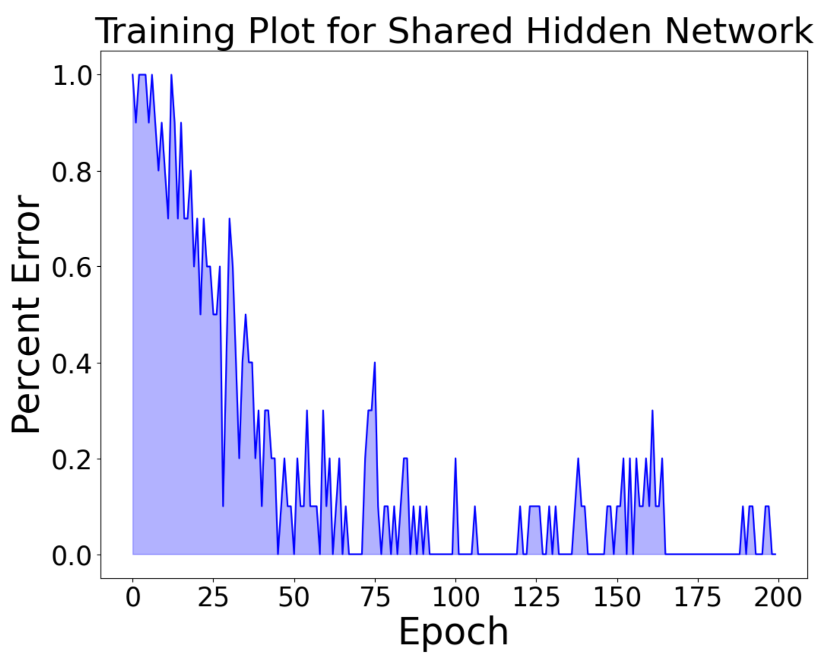

Figure 6: Training plot for the network with the Shared Hidden layer. This was bidirectionally connected to both the Music Hidden and the Letter Hidden layers, serving as a bridge between the two tasks.

My network achieved test set accuracy of 100% on the letter task, as expected. I found it fascinating that with some sets of hyperparameters (results not pictured), the Shared Hidden layer was actually detrimental to this task, and accuracy dropped to 80% or 90%. On the music task, the Shared Hidden network achieved an accuracy of 30%. To be sure this was not above random chance, I tested across some random seeds, and saw that it was often above 20%. While this was somewhat above random chance, the improvement was marginal, suggesting that cross-activation (modeling associator synesthesia) is largely insufficient to fully replay a music sequence at test time when no input is given.

Top-Down Feedback (Jay):

I started by manually experimenting with a few hyperparameter combinations and visualizing the training plot for my Top-Down network. Notably, I saw that unlike the control network and Jeremy's Shared Hidden network, my network was able to reliably reach and stay at zero percent error for over 30 consecutive epochs, as shown in Figure 7. This suggests that the associator is able to create more stable learned representations of the inputs when training.

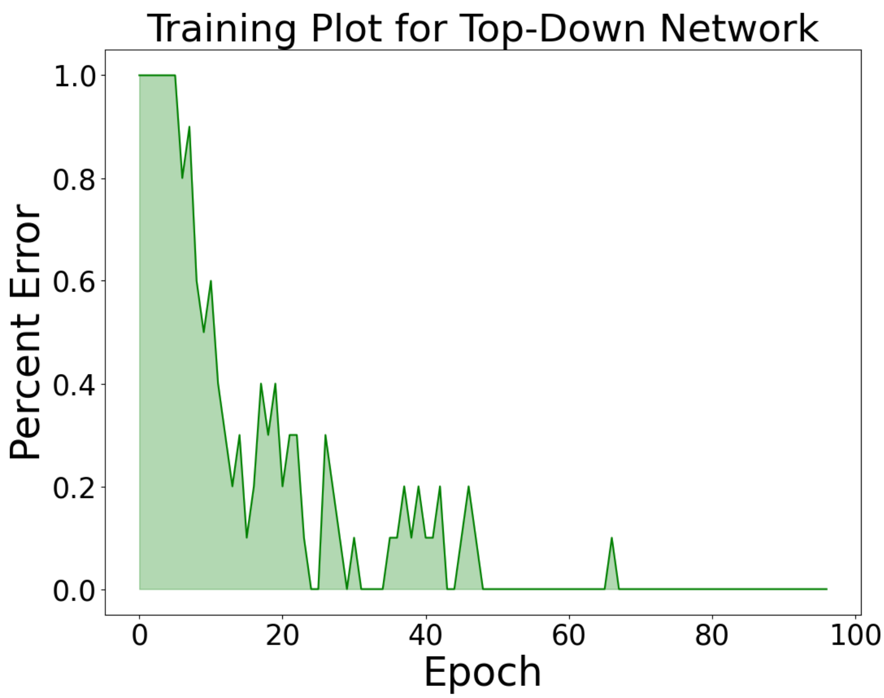

Figure 7: Training plot for the network with top-down feedback. An associator layer was added, taking input from Letter Output and feeding into Music Output, allowing one task to influence the other.

In order to identify the ideal Top-Down network, shown in Figure 7, I needed to do a thorough hyperparameter search beyond my manual exploration. The heatmap created using SciPy's interpolate function is shown in Figure 8. Cubic interpolation produced the most reliably and visually reasonable results. I ran over 50 simulations with different hyperparameter combinations. First Zero was the target parameter to minimize, and the ideal combination was **`AssocToOut2`** = 1.414 and **`OutToAssoc`** = 1.919.

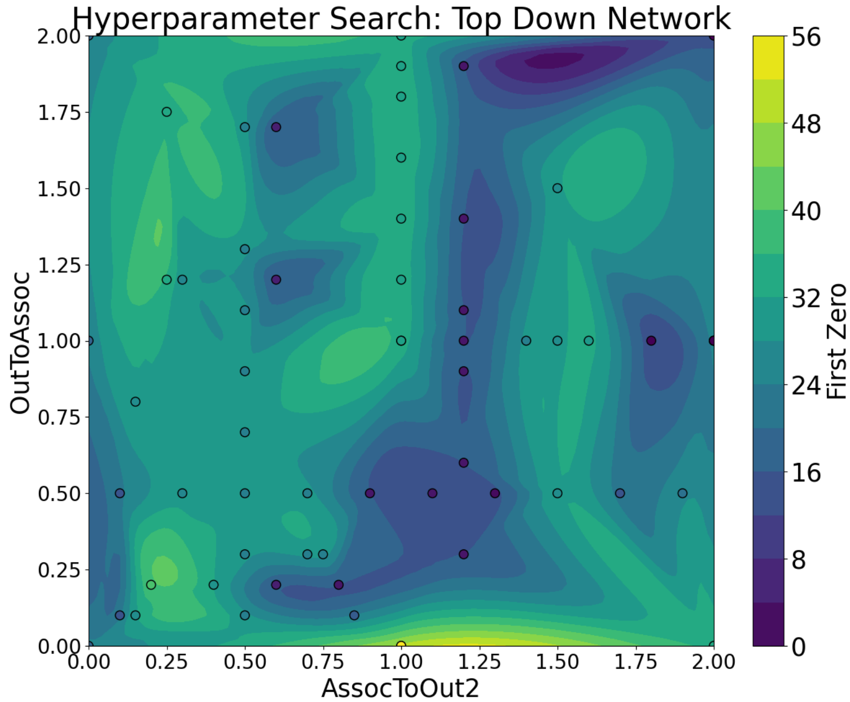

Figure 8: Heatmap showing a rigorous hyperparameter search for the relative weights of the connections controlling top-down feedback. The epoch at which the network first reaches a percent error of zero (First Zero) is minimized. The two hyperparameters are OutToAssoc (connecting Letter Output to Associator) and AssocToOut2 (controlling Associator to Music Output). Each parameter combination that was manually simulated is shown as a circular black point.

The test set accuracy on the letters was 100% for this model as well. This was consistent across nearly all hyperparameter combinations tested. On the music task, the best Associator network achieved an impressive 90% accuracy! Unfortunately, this was not consistent across random seeds. Additionally, this was not observed across many different hyperparameter combinations. That being said, we still found reasonably high accuracy (≥70%) on a regular basis for this model architecture, suggesting that it is capable of replaying music when given no musical input.

Hybrid Model (Jeremy):

When training and testing the combined model, I knew that finding the optimal set of hyperparameters would be a difficult task, as I had 6 to optimize: **`SharedHiddenToHid`**, **`SharedHiddenToHid2`**, **`HidToSharedHidden`**, **`Hid2ToSharedHidden`**, **`OutToAssoc`**, and **`AssocToOut2`**. I started off by manually exploring the hyperparameter space, and found that it was possible to make this network reach zero percent error for 30 epochs in a row.

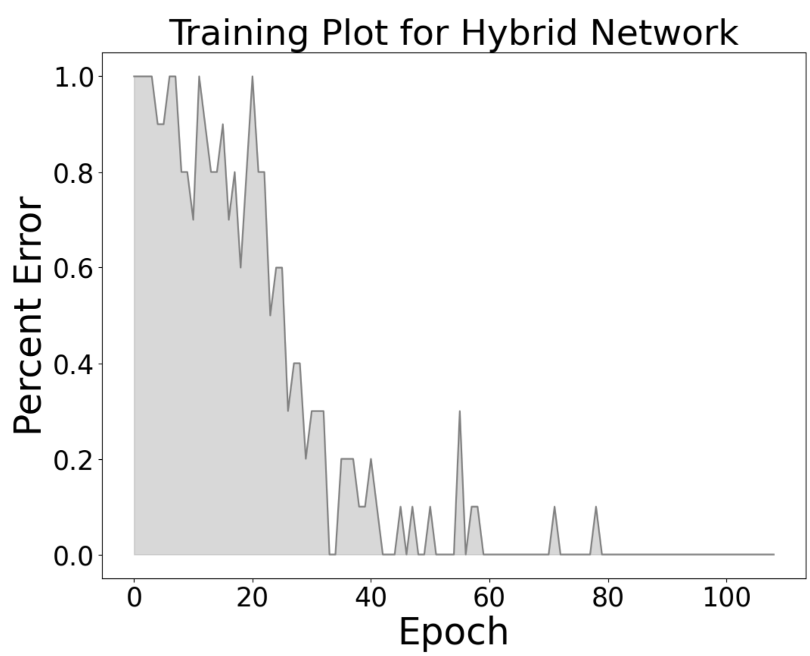

Figure 9: Training plot for the hybrid network with both the Associator and the Shared Hidden layers.

The training plot for the best set of hyperparameters I tried is shown in Figure 9 (**`AssocToOut2`** = 0.5, **`OutToAssoc`** = 1, **`SharedHiddenToHid`** = 1, **`SharedHiddenToHid2`** = 1, **`HidToSharedHidden`** = 1, and **`Hid2ToSharedHidden`** = 1). I tried to interpolate the best set of hyperparameters using SciPy, but the 6D interpolation proved to be too big a challenge for our timeframe. The model, like all of the others, achieved perfect test set accuracy on the letters. However, its accuracy was 50% on the test set for music, suggesting limited generalizability.

I evaluated the model's sending and receiving weights to confirm whether the model was configured properly, and this confirmed that the issue was likely with the hyperparameters. To qualitatively evaluate the model, I visualized a few test trials. An example is shown in Figure 10. Even when the network was correct, I noticed that the associator layer was barely activated. This added credence to my results, as it validated the importance of the associator for high accuracy.

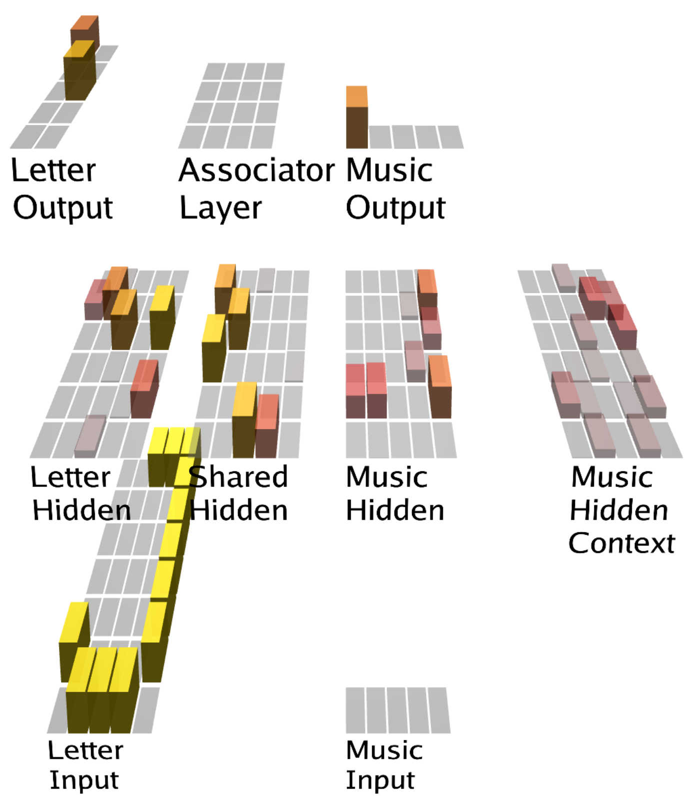

Figure 10: Example test step for our hybrid network. The input consists of the letter "J" and no music. Activities of all hidden and output layers are shown, with increased height correlating to increased magnitude. In this case, the network produces the correct letter and music outputs.

Jay and I suspect the hybrid model is failing because it is not using the associator layer enough. If we had more time, we would've tried more hyperparameter combinations or added some lateral inhibition to force the use of the associator.

<u>Modeling: Results Overview</u>

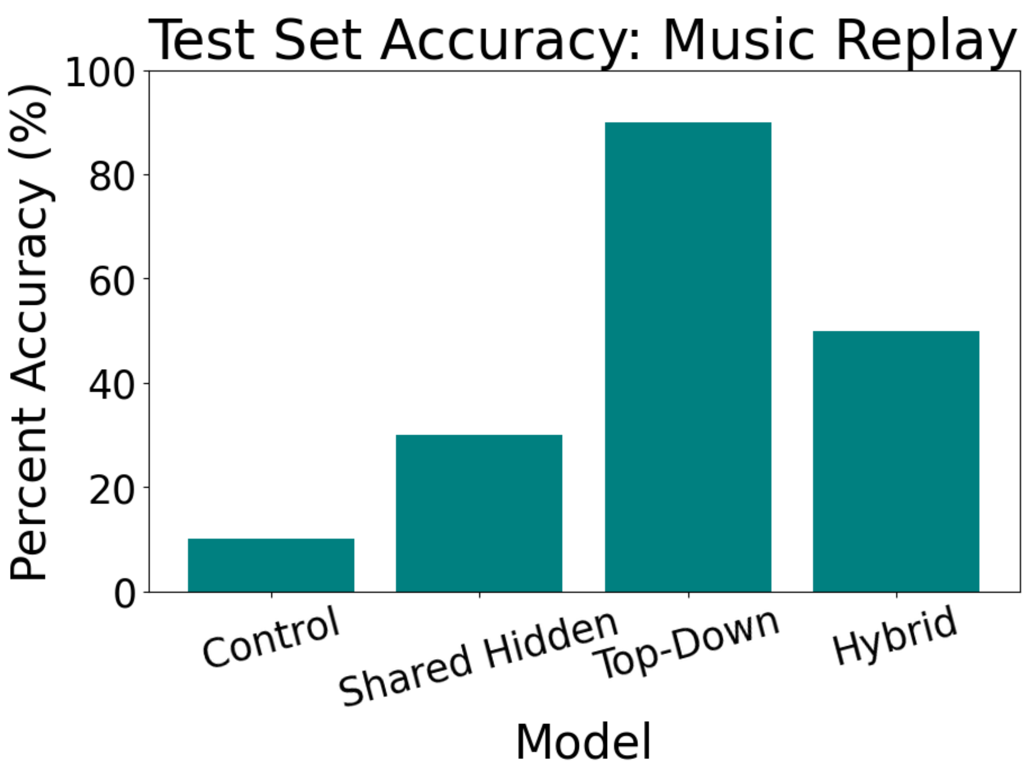

Figure 11: Test set accuracies of the four models compared in the present study. We find that the model with top-down feedback (Top-Down) is able to most accurately replay the musical melody it heard during training without being given any input at test time.

In conclusion, all four of our models were able to achieve 100% test set accuracy for letters. Test set accuracy for music, however, varied between our networks. The accuracies are shown in Figure 11. The top-down model, which is the augmented and reversed version of the model inspired by associator synesthesia, yielded the highest music test set accuracy of 90%, outperforming the rest of our models.

Discussion:

In the present study, we demonstrated that associative learning in the hippocampus can successfully lead to spontaneous recall of linked stimuli. The top-down model, drawing on connectivity found in associator synesthesia, ended up being the best network to represent our phenomenon, reaching a test accuracy of 90% compared to the 10% accuracy of the control.

A similar, but distinct, area of work focuses on state-dependent memory. Significant work has been done to show music can enhance memory, specifically in therapeutic contexts such as language recovery after stroke (Sihvonen et al. 2020). Music therapy is an area of work that relies on these findings, and it has been proposed for a variety of memory-related conditions, including Alzheimer's disease (Flo et al. 2022). Notably, our work explores the opposite direction of connectivity, whether a non-music task can effectively elicit music in a mechanism similar to grapheme-color synesthesia, and to our knowledge, this has not been evaluated in the literature. Thus, the present study provides a novel approach for understanding a unique aspect of associative learning.

In addition to having varying testing accuracies, another important finding from our models has to do with their training errors. The Shared Hidden model, inspired by projector synesthesia, was unable to consistently remain at 0% training error across our trials, whereas the two networks that contained the Associator layer could. This points to the associator layer of our network being a critical component in enabling replay of music.

Another finding of note has to do with a difference between the top-down and hybrid models. Both models reached 0% training error relatively quickly and within a similar timeframe during our experimentation. However, the top-down model consistently outperformed the hybrid model in terms of test accuracy. This demonstrates that the hybrid model has limited generalizability when moving from training to testing, and suggests that the addition of the shared hidden layer in addition to an associator layer detracts from a model's ability to perform the task. Also, as discussed in our results, the associator layer was largely unused in the hybrid model, providing additional evidence for our hypothesis about the importance of top-down feedback to spontaneously recall sequential information.

Additional work could be conducted to further understand our model. Given more time, we would have pursued the following goals: performing additional hyperparameter tuning, exploring unidirectional vs. bidirectional connections within our models, modifying model architecture, and experimenting with higher-order dependent sequences as inputs.

After the top-down model of associative learning proved promising, we successfully tuned its hyperparameters to make it yield more consistent and accurate results. However, we did not get a chance to fully explore rigorous hyperparameters tuning of the shared hidden layer and combined models, as the 4D and 6D interpolation, respectively, proved to be challenging. This choice could also present a key limitation to our study: we may have only found that the top-down model yielded the best results because we experimented with it the most. Further research would help demystify this aspect of our findings, demonstrating whether or not other models can compete with the top-down model when put on an equal playing field.

Isabella and Peter found that unidirectional connections between the associator layer and output layers was the best option for yielding promising results, and that bidirectional connections caused the simulation to crash. As a result, we also ended up deciding to use unidirectional connections for our associator layer. However, we think that bidirectional connections where both music and grapheme output layers communicate with the associator layer and thus each other make more sense as a biologically plausible system. Additional work could explore alternative simulation software to investigate whether this type of bidirectional connectivity would improve the accuracy and consistency of the top-down and combined models.

Model architecture could also be experimented with generally. Both the grapheme and music networks have one hidden layer, which may be limiting their capacity to learn the task. Additional hidden layers or other experimentation could be attempted in order to improve the model accuracy.

Lastly, higher-order dependent sequences could be used for our music input. Right now, we've represented music input as a ten note long series with first-order dependencies. Music is oftentimes longer and more complex than ten notes in reality, however. Longer sequences, and especially higher-order dependent sequences, could be used to test the robustness of our model's ability to learn complex musical patterns.

References:

Ahmed, M.S., Priestley, J.B., Castro, A., Stefanini, F., Canales, A., Balough, E.M., Lavoie, E., Mazzucato, L., Fusi, S., Losonczy, A. (2020). *Hippocampal Network Reorganization Underlies the Formation of a Temporal Association Memory*. Neuron. https://doi.org/10.1016/j.neuron.2020.04.013

Aoki, Y., Shibasaki, M. and Nakata, H. (2023). *Synesthesia has specific cognitive processing during Go/No-go paradigms*. Sci Rep 13, 6172. https://doi.org/10.1038/s41598-023-32389-8

Barnett, K. J., Foxe, J. J., Molholm, S., Kelly, S. P., Shalgi, S., Mitchell, K. J., & Newell, F. N. (2008). *Differences in early sensory-perceptual processing in synesthesia: a visual evoked potential study*. NeuroImage, 43(3), 605–613. https://doi.org/10.1016/j.neuroimage.2008.07.028

Borders, A.A., Aly, M., Parks, C.M., Yonelinas, A.P. (2017). *The hippocampus is particularly important for building associations across stimulus domains*, Neuropsychologia, Volume 99, 2017, Pages 335-342, ISSN 0028-3932, https://doi.org/10.1016/j.neuropsychologia.2017.03.032

Brang, D., Ghiam, M., and Ramachandran, V. S. (2013). *Impaired acquisition of novel grapheme-color correspondences in synesthesia*. Frontiers in human neuroscience, 7, 717. https://doi.org/10.3389/fnhum.2013.00717

Brang, D., Williams, L. E., and Ramachandran, V. S. (2012). *Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities*. Cortex; a journal devoted to the study of the nervous system and behavior, 48(5), 630–637. https://doi.org/10.1016/j.cortex.2011.06.008

Cohen, M. X., Weidacker, K., Tankink, J., Scholte, H. S., and Rouw, R. (2015). *Grapheme-color synesthesia subtypes: Stable individual differences reflected in posterior alpha-band oscillations*. Cognitive neuroscience, 6(2-3), 56–67. https://doi.org/10.1080/17588928.2015.1017450

Day, S. (2005). *Some Demographic and Socio-cultural Aspects of Synesthesia.* In L. C. Robertson & N. Sagiv (Eds.), Synesthesia: Perspectives from cognitive neuroscience (pp. 11–33). Oxford University Press.

Dixon, M. J., Smilek, D., and Merikle, P. M. (2004). *Not all synaesthetes are created equal: projector versus associator synaesthetes*. Cognitive, affective & behavioral neuroscience, 4(3), 335–343. https://doi.org/10.3758/cabn.4.3.335

Flo, B. K., Matziorinis, A. M., Skouras, S., Sudmann, T. T., Gold, C., and; Koelsch, S. (2022). *Study protocol for the alzheimer and music therapy study: An RCT to compare the efficacy of music therapy and physical activity on brain plasticity, depressive symptoms, and cognitive decline, in a population with and at risk for alzheimer's disease*. PLOS ONE, 17(6). https://doi.org/10.1371/journal.pone.0270682

Kliegl, O., Bäuml, K. T. (2021). *The Mechanisms Underlying Interference and Inhibition: A Review of Current Behavioral and Neuroimaging Research*. Brain sciences, 11(9), 1246. https://doi.org/10.3390/brainsci11091246

Levy, W.B., Steward, O. (1979). *Synapses as associative memory elements in the hippocampal formation*, Brain Research, Volume 175, Issue 2, Pages 233-245, ISSN 0006-8993, https://doi.org/10.1016/0006-8993(79)91003-5

Lungu, L., Rothen, N., and Terhune, D. B. (2021). *The time course of synaesthetic colour perception*. Cortex; a journal devoted to the study of the nervous system and behavior, 141, 322–330. https://doi.org/10.1016/j.cortex.2021.04.018

Mann H., Korzenko J., Carriere J.S.A., Dixon M. J. (2009). *Time–space synaesthesia – A cognitive advantage?*, Consciousness and Cognition, Volume 18, Issue 3, 2009, Pages 619-627, ISSN 1053-8100, https://doi.org/10.1016/j.concog.2009.06.005

Neckar, M., Bob, P. (2014). *Neuroscience of synesthesia and cross-modal associations*. Reviews in the neurosciences, 25(6), 833–840. https://doi.org/10.1515/revneuro-2014-0033

O'Reilly, R. C., Munakata, Y., Frank, M.J., Hazy, T.E. and Contributors (2012). *Computational Cognitive Neuroscience*. Wiki Book, 4th Edition (2020), URL: https://github.com/CompCogNeuro/ed4

Rouw, R., Scholte, H. *Increased structural connectivity in grapheme-color synesthesia*. Nat Neurosci 10, 792–797 (2007). https://doi.org/10.1038/nn1906

Sihvonen, A. J., Leo, V., Ripollés, P., Lehtovaara, T., Ylönen, A., Rajanaro, P., Laitinen, S., Forsblom, A., Saunavaara, J., Autti, T., Laine, M., Rodríguez‐Fornells, A., Tervaniemi, M., Soinila, S., Särkämö, T. (2020). *Vocal music enhances memory and language recovery after stroke: Pooled results from two RCTS*. Annals of Clinical and Translational Neurology, 7(11), 2272–2287. https://doi.org/10.1002/acn3.51217

Suzuki, W. A. (2008). *Associative learning signals in the brain*. Progress in brain research, 169, 305–320. https://doi.org/10.1016/S0079-6123(07)00019-2

Ward, J., Hoadley, C., Hughes, J. et al. (2017). *Atypical sensory sensitivity as a shared feature between synaesthesia and autism*. Sci Rep 7, 41155. https://doi.org/10.1038/srep41155 van Leeuwen, T. M., den Ouden, H. E., and Hagoort, P. (2011). *Effective connectivity determines the nature of subjective experience in grapheme-color synesthesia*. The Journal of neuroscience : the official journal of the Society for Neuroscience, 31(27), 9879–9884. https://doi.org/10.1523/JNEUROSCI.0569-11.2011

van Leeuwen, T. M., Singer, W., and Nikolić, D. (2015). *The Merit of Synesthesia for Consciousness Research*. Frontiers in psychology, 6, 1850. https://doi.org/10.3389/fpsyg.2015.01850