

The limited role of hippocampal declarative memory in transient semantic activation
during online language processing

Sarah Brown-Schmidt¹, Sun-Joo Cho¹, Nazbanou Nozari², Nathaniel Klooster³, and
Melissa Duff⁴

1. Vanderbilt University, Department of Psychology and Human Development
2. Carnegie Mellon University, Department of Psychology
3. University of Pennsylvania, Department of Neurology
4. Vanderbilt University Medical Center, Department of Hearing and Speech Science

Corresponding Author:

Sarah Brown-Schmidt

sarah.brown-schmidt@vanderbilt.edu

Keywords: Hippocampus, Amnesia, Word Recognition, Eye-Tracking, Semantic,
Prediction

Abstract

Recent findings point to a role for hippocampus in the moment-by-moment processing of language, including the use and generation of semantic features in certain contexts. What role the hippocampus might play in the processing of semantic relations in spoken language comprehension, however, is unknown. Here we test patients with bilateral hippocampal damage and dense amnesia in order to examine the necessity of hippocampus for lexico-semantic mapping processes in spoken language understanding. In two visual-world eye-tracking experiments, we monitor eye movements to images that are semantically related to spoken words and sentences. We find no impairment in amnesia, relative to matched healthy comparison participants. These findings suggest, at least for close semantic links and simple language comprehension tasks, a lack of necessity for hippocampus in lexico-semantic mapping between spoken words and simple pictures.

The human hippocampus has long been known for its critical role in the encoding and use of enduring memories of experience. More recently, we have argued that hippocampus also makes contributions to the moment-by-moment processing of language (Duff & Brown-Schmidt, 2017; 2012; Brown-Schmidt & Duff, 2016). This role for hippocampus in language processing may relate to its capacity for relational binding (Eichenbaum & Cohen, 2001) and prediction (Bonhage et al., 2015; Chen et al., 2013; Gluck, Meeter & Meyers, 2003). Here we explore the role of hippocampus in lexico-semantic mapping processes during online language comprehension.

Hippocampus supports new semantic learning, computing the relational binding of the arbitrarily-related phonological, conceptual, and orthographic components of semantic knowledge (Gabrieli et al., 1998; Manns et al., 2003; see Duff et al., 2020 for review). The traditional view of hippocampal contributions to semantics, however, is that its role is limited to acquisition, and that semantic knowledge and processing become independent of hippocampus over time through neocortical consolidation (McClelland et al., 1995; O'Reily & Rudy, 2000). Consistent with hippocampal independence of the semantic network are findings that one patient with unilateral left-hippocampal damage and memory impairment shows intact cumulative semantic interference effects (Oppenheim et al., 2015) – an effect in which the time to name a picture successively slows when pictures from the same category are named (e.g., *pig*, *cow*, *horse*), reflecting accumulated processing of semantic relations within the category.

However, emerging findings point to hippocampal contributions to lexico-semantic processing beyond initial acquisition. There is good reason to think semantic processing should invoke hippocampus, as semantic knowledge is grounded in the

experiences that give rise to that meaning (Glenberg, 1997; Glenberg & Robertson, 2000). Hippocampus plays a critical role in the retrieval, simulation, and imagination of personal experience (Schacter et al., 2012; Maguire, 2001; Buckner, 2010; Verfaellie et al., 2014). New evidence indicates that semantic knowledge and processing do not, in fact, become fully hippocampal independent. Atrophy in hippocampus is correlated with deficits in a semantic association task (Butler et al., 2009), and fMRI evidence reveals left-hippocampal engagement during a semantic interference naming paradigm (deZubicaray et al., 2014). Recordings from depth electrodes find hippocampal theta oscillations are related to semantic distances between words in a word recall task (Solomon, Lega, Sperling, & Kahana, 2019). Further, fMRI evidence from a task where participants learned associations between objects and abstract stimuli shows a key role for hippocampus in the encoding of distances in an abstract, multidimensional space (Theves, Fernandez, & Doeller, 2019). These data are striking as they suggest a role for the hippocampus in tracking and representing the relations among words (and other abstract relations) in semantic memory in a manner that is similar to how the hippocampus tracks and represents relations in physical space and events in episodic memory.

Hippocampal contributions to Language

Studies of amnesic patient HM's lexico-semantic knowledge have revealed mixed findings, some of which can be attributed to the depth of the knowledge being tested. When HM's lexico-semantic knowledge was tested with tasks designed to diagnosis aphasia or semantic dementia, or that capture simple associations (e.g., naming, matching, or providing definitions of high frequency words) HM's performance

did not differ from controls (e.g., Gabrieli et al., 1998; Kensinger et al., 2001; Schmolck et al., 2002). Yet, HM did present with deficits relative to controls in lexical decision and definition tasks for low-frequency words (James & MacKay, 2001), and had a “semantic-level production deficit” when task-related discourse was analyzed relative to discourses produced by controls (MacKay, Burke, & Stewart, 1998; also see MacKay, Stewart, & Burke 1998 for a semantic-level binding account of observed deficits in language comprehension).

Subsequent behavioral studies of groups of patients with bilateral hippocampal damage point to the possibility that some but not all aspects of language may be impaired following damage to hippocampus. In prior work, we have examined the impact of bilateral hippocampal damage on the online processing of language in rich discourse contexts using the visual-world eye-tracking paradigm (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). We have generally found that this patient group is able to successfully participate in these tasks, despite a pronounced memory impairment. For example, in a dialogue task, patients with amnesia and healthy comparison participants were equally likely to look at a picture of an elephant in the shared visual world when given the instruction *"Look at the elephant"* (Rubin, et al. 2011). In some cases, looks to the mentioned object were somewhat lower in the patient group, as compared to healthy matched comparison participants, and may relate to the use of obscure word-picture pairings that required pre-training (Trude, Duff, & Brown-Schmidt, 2014), or sentences containing multiple named characters (Kurczek, et al. 2013).

In contrast, studies that involved linking referents over time in sentences have revealed pronounced deficits. A key feature of conditions in which deficits were pronounced is the need to link information across time. One clear example is narratives where two or more sentences introduced and then referred back to candidate referents in the visual display, as in “*Mouse was bringing some mail to Duck as a rainstorm was beginning, he’s carrying an umbrella, and it looks like it’s about to rain*”. In cases like this, participants with hippocampal damage struggled to use the information contained in the first part of the sentence to resolve subsequent linguistic ambiguity, i.e., in who “he” referred to (Kurczek et al., 2013; Covington et al., 2020; Rubin et al., 2011). By contrast, healthy comparison participants and young adults were significantly more likely to use prior discourse information (such as which character was most prominent in the prior discourse). These deficits were observed in cases where meaning was built up over time within a discourse; an open question then, is whether similar deficits would extend to the online processing of short and simple sentences.

Other findings point to a role for hippocampus in the use of semantic representations. For example, when provided with a known word such as “*menu*”, the ability to generate semantic features such as “usually got a plastic cover”, is markedly impaired in amnesia (Klooster & Duff, 2015; Klooster, Tranel, & Duff, 2020). Further, in this task, the features that the hippocampal patients produced tended to be closer in semantic space to the target word compared to healthy comparison participants (Cutler, Duff, & Polyn, 2019). These patients also show impaired knowledge of the meaning of collocates such as “run a bath” and “save the date” (Covington & Duff, 2017), adding to the evidence that hippocampus plays a role in lexical associations beyond its initial

acquisition. In a study of picture naming, Hilverman and Duff (in revision) tested patients with bilateral hippocampal damage and amnesia on 1458 items from the Bank of Standardized Stimuli (BOSS) database (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010; Brodeur, Guérard, & Bouras, 2014) that varied across a range of word features such as imageability, frequency, and familiarity. Hilverman and Duff (in revision) found that patients with amnesia were less likely than comparison participants to correctly name the objects that they viewed. This finding is in contrast to prior work which reported no impairment, but only tested fewer than 100 images (e.g., Kensinger et al., 2001). In a complement to these findings, in a study of picture naming, Hamamé et al (2014) found a relationship between hippocampal activity (measured using electrodes implanted in hippocampus in pre-surgical epilepsy patients) and preparation of a picture name. The activity was related to picture naming latency, pointing to a role for hippocampus in retrieving the arbitrary associations between objects and their names. Indeed, findings of hippocampal contributions to maintenance of relational information over even very brief time-scales (e.g. from one trial to the next, Hannula et al., 2006) and to the updating of previously acquired information through reconsolidation (McKenzie & Eichenbaum, 2011) further suggest that even long after acquisition of semantic knowledge, hippocampus may play a long term role in maintaining and tuning that information over time.

Hippocampal contributions to prediction in language?

Semantic representations play a critical role in the predictions that listeners make during real-time language processing. As listeners interpret the word "candle" in "Click on the candle", numerous candidate words are activated. These temporary activations

include words that sound like the intended word (e.g., *candy*, *cannery*; Allopenna et al., 1998), as well as words that are semantically related to the intended word (e.g. *lightbulb*, *matchstick*; Yee & Sedivy, 2006; Huettig & Altmann, 2005). Moreover, in sentences like *The boy will eat the cake*, semantic information conveyed by the unfolding sentence, “*The boy will eat...*” provides clues about the upcoming direct object, “*cake*”, shaping predictions about how the sentence will continue, and processing of subsequent words in the sentence (Altmann & Kamide, 1999; Altmann & Kamide, 2007; Kamide, Altmann, & Haywood, 2003; Federmeier & Kutas, 1999). This technique has been used in prior neuropsychological studies to uncover the role of specific brain regions in lexico-semantic mapping processing, such as the role of the VLPFC (Nozari, Mirman, & Thompson-Schill, 2016).

Given the observation of deficits in the ability to generate semantic features (Klooster & Duff, 2015; Klooster, Tranel, & Duff, 2020), an open question, then, is the degree to which the use of semantic representations to generate upcoming predictions would be similarly impaired following hippocampal damage. Hippocampal contributions to predictive processes have been observed in a variety of tasks. For example, in a study of language production using intracranial recordings, Jafarpour, Piai, Lin, and Knight (2017) examined patterns of hippocampal activity, specifically hippocampal high frequency band (HFB) power, as participants were about to name a picture. Greater HFB power was observed when the unfolding sentence was highly predictive of the upcoming to-be-named picture, suggesting pre-activation of the expected semantic representation (for related findings see Piai et al., 2016; Wang, Hagoort, & Jensen, 2018). Other work points to a role for hippocampus in prediction of upcoming word

forms in reading (Bonhage et al., 2015) and in the calculation of prediction errors in viewing picture sequences (Chen et al., 2013; see related discussion in Henson & Gagnepain, 2010).

On a number of theoretical proposals regarding the role of hippocampus in learning, hippocampus supports learning of predictive relationships in the world (Stachenfeld, Botvinick, & Gershman, 2017; Gruber & Ranganath, 2019; Davachi & DuBrow, 2015). For example, Gluck and Myers (1993) present a computational theory of cortico-hippocampal interactions in discrimination learning. On that view, hippocampus acts as a predictive autoencoder, which receives input and reconstructs it as output; in the process, the information is compressed to reduce redundancy while preserving bits that predict reinforcement. This new representation then acts as the “desired output” for the long term memory system, and helps learn new associations in the neocortex. Over time, the learning loop between hippocampus and long term memory leads to the development of cortical representations that are linear combinations of those developed in the hippocampus. In the face of hippocampal damage the theory predicts that other brain regions may step in to learn new associations based on previously established fixed connections. If true, such an account may predict that individuals with hippocampal damage should retain the ability to map sound to meaning. Furthermore, since the associations have been formed before hippocampal damage, such mappings could well extend to activating related words, i.e., to predictive processing (also see Gluck, Myers, & Meeter, 2005). On such a view, then, we may expect to see preservation of linguistic prediction in hippocampal amnesia.

The present research

Taken together, these findings point to a continuing role for hippocampus in the use of semantic representations to generate information in naming and recall tasks. What role the hippocampus might play in lexico-semantic mapping processes in language comprehension, however, is unknown. In two Experiments, we test the hypothesis that hippocampus provides critical support to lexico-semantic mapping processes during online sentence comprehension. The well-established role of hippocampus relational binding (Eichenbaum & Cohen, 2001) and prediction (Bonhage et al., 2015; Chen et al., 2013) points to a potential role for hippocampus in mapping spoken words to candidate referential meanings (and upcoming meanings) during spoken language understanding. Hippocampus is a likely candidate for involvement in the online processing of semantics due to its role in supporting episodic retrieval, simulation, imagination, and prediction. Sentence processing deficits in amnesia may be expected, then, particularly when interpretation requires linking spoken words and referents to generate lexico-semantic mappings, and to generate predictions about upcoming material. Alternatively, if lexico-semantic mapping processes draw on previously learned semantic associations, use of these representations to guide language processing and prediction may remain intact in the face of hippocampal damage (Gluck & Myers, 1993).

Experiment 1

Experiment 1 examines semantic processing in spoken word recognition. We use the visual world paradigm (Tanenhaus et al., 1995) to measure the activation of semantic competitors during perception of individual spoken words. If hippocampal damage impairs lexico-semantic mapping processes, activation of these competitors

should be attenuated in patients with bilateral hippocampal damage and dense amnesia.

Method

Participants

An initial sample of 18 healthy young adults from the student community at the University of Illinois Urbana-Champaign participated in this study in exchange for \$8/hour payment or partial course credit. This young adult sample was tested in order to vet the paradigm and demonstrate the effect typically seen in convenience samples with our materials. One additional young adult participant completed the study but was not included in the analysis due to experimenter error (the eye-tracker had been moved to the wrong location on the desktop). The young adult sample was tested in the first author's laboratory at the University of Illinois.

The focus of the research was on participants with bilateral hippocampal damage ($N=5$) and demographically matched healthy comparison participants ($N=5$). All healthy comparison participants were matched individually to each patient on age, sex, education, handedness, and ethnicity. The patients and comparison participants were tested either in the last author's lab at the University of Iowa, or at a location convenient to the patient (e.g., in a private conference room at a hotel near their home). Patients and healthy comparison participants were compensated \$15 per hour of participation.

The patients and healthy comparison participants were tested between late 2013 and early 2015. Participants were five (one female, four male) individuals with bilateral hippocampal damage and severe declarative memory impairment and 5 healthy comparison individuals (NC) who demographically matched each patient with amnesia

for age (+/- 5 years), sex, education (+/- 2 years), ethnicity, and handedness. This is the same group of patients with amnesia that were tested in Klooster and Duff (2015). At the time of data collection, the participants with hippocampal damage (HC) were in the chronic epoch of amnesia and were 57.6 years old. Etiologies included anoxia/hypoxia ($n=3$) resulting in bilateral hippocampal damage, and herpes simplex encephalitis (HSE; $n=2$), resulting in more extensive bilateral medial temporal lobe damage affecting the hippocampus, amygdala, and surrounding cortices. Structural MRI examinations completed on 4 of the 5 patients confirmed bilateral hippocampal damage and volumetric analyses revealed significantly reduced hippocampal volumes. Participant 2563 wears a pacemaker could not undergo MRI examination; computerized tomography confirmed that damage was confined to the hippocampus bilaterally. Anoxic participants had no visible damage to the lateral temporal lobes or anterior temporal lobes.

The patients' neuropsychological assessment results indicated severe impairment in declarative memory ($M=59$; Wechsler Memory Scale–III General Memory Index) compared to other cognitive domains (verbal IQ, vocabulary, and semantic knowledge), in which the patients tested within normal limits (Table 1). Scores within normal limits suggest that the patients with amnesia do not have deficits in general semantic knowledge and access. Further, standardized neuropsychological testing and interviews with a certified speech language pathologist confirmed that the patients do not have language deficits such as aphasia, anomia, or semantic dementia.

Table 1. Demographic, Neuroanatomical, and Neuropsychological Characteristics of Patients with Hippocampal Amnesia

Demographics						Neuroanatomical		Neuropsychological			
Patient	Sex	Age	Hand.	Ed	Etiology	Damage	HC Volume	WAIS III VIQ	WAIS III Vocab	WAIS III Info	WMS III GMI
1846	F	46	R	14	Anoxia	Bilateral HC	-4.23	88	8	8	57
1951	M	61	R	16	HSE	Bilateral HC + MTL	-8.10	107	10	11	57
2308	M	53	L	16	HSE	Bilateral HC + MTL	N/A	95	11	8	45
2363	M	53	R	18	Anoxia	Bilateral HC	-2.64	112	12	13	73
2563	M	54	L	16	Anoxia	Bilateral HC	N/A	91	9	12	75
Patient Means		57.6		16				98.8	10	10.4	59

Key: HC = hippocampus (participants with hippocampal amnesia); F = female. M = male. R = right handed. L = left handed. Ed. = years of education. HSE = Herpes Simplex Encephalitis. + MTL = damage extending into the greater medial temporal lobes. HC Volume = hippocampal volumetric z-scores as measured through high resolution volumetric MRI and compared to a matched healthy comparison group (Allen et al., 2006). WAIS-III VIQ = Wechsler Adult Intelligence Scale-III Verbal Intelligence Quotient, with Vocabulary and Information sub-scores reported. WMS-III GMI = Wechsler Memory Scale-III General Memory Index. **Bolded** scores are impaired as defined as 2 or more standard deviations below normative data.

Procedure

Participants were tested individually. After signing a consent form, the participant was seated at a computer with a desktop-mounted Eyelink 1000 eye-tracker.

Participants were instructed that on each trial, they would see four pictures and hear a man say the name of one of the pictures. When they heard the picture name, they were asked to click on it as quickly as they can. Following the instructions, participants were calibrated on the eye-tracker and the experiment began. Participants were allowed to take breaks between trials as needed. All participants (including those with memory impairment) understood and had no difficulty performing the task. The experiment was composed of a total of 612 trials and lasted approximately 90 minutes. Young adults completed the entire set of 612 trials in a single session. For patients and their matched comparisons, we included short breaks every ~100 trials, and due to scheduling conflicts, in some cases the experiment was conducted over the course of several days. For the amnesia group only, in order to have sufficient data to characterize individual participant performance, each participant with amnesia completed the full task twice.

Each trial began with a drift-check of the eye-tracker. If the drift-check failed, the eye-tracker was re-calibrated. Next, the 4 pictures appeared on the screen, followed by a 1-second delay, and auditory presentation (over the computer speakers) of the critical word, e.g., *candle*. The 4 pictures were randomly assigned to one of 4 locations on the screen (Figure 1), and remained on the screen until the end of the trial. The participant's task was to click on the picture corresponding to the spoken word. Once the computer recorded the participant's click response, there was another 1-second delay before the drift-check for the next trial.

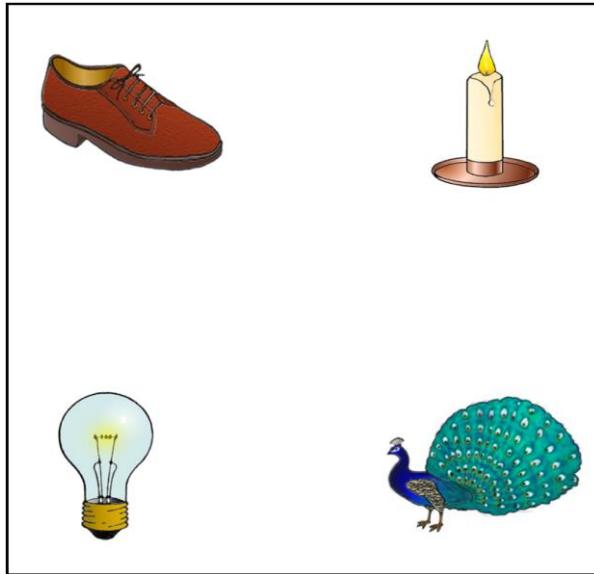


Figure 1. Experiment 1: Example experimental display.

Materials

The materials were selected on the basis of a series of norming studies (see Appendix for details on the norming studies and the complete list of materials). We created 51 stimulus sets of 6 easy-to-name pictures. The pictures were colorized versions (Rossion and Pourtois, 2004) of a set of normed images from Snodgrass & Vanderwart (1980), and similar clip art pictures.

Each image set was designed to maximize the semantic relatedness of the words corresponding to the critical image pair (e.g., *candle-lightbulb*) while minimizing the semantic relationships with the 3 other words. Pairwise similarity was calculated using LSA (see Appendix for details). The average similarity across the 51 pairs of critical semantically related pairs was .361 (SD=.16), vs. .076 (SD=.06) for the relationship between the critical words and the unrelated and filler words. We also avoided initial

phonological overlap between the words in the set (i.e., so that the words were not cohort-competitors, Allopenna et al., 1998).

Each set was designed to generate 12 trials, created by systematically grouping 4 pictures from each set of 6, and varying the target item, for a total of $51 \times 12 = 612$ total trials. Trials were in one of four conditions, designed to test for the activation of semantically related words during interpretation of spoken words, along with filler trials (see Table 2); see below for full predictions for each condition.

- On Semantic-competition trials (2 per item set), the scene contained the target (e.g. candle), a semantic competitor (e.g. lightbulb), and two unrelated items (e.g. peacock, shoe). On Semantic-competition trials we measure fixations to the semantic competitor (e.g. lightbulb) during the processing of the spoken target word (e.g. candle).
- On Unrelated-target trials with a semantic competitor in-scene (2 per item set), the scene contained the target (e.g. peacock), an unrelated item (e.g. shoe), and a pair of semantically related items that were unrelated to the target (e.g. candle, lightbulb). On these trials, we measure fixations to a non-competitor (e.g. lightbulb) that was presented in the scene along with a semantically-related item.
- On Unrelated-target trials without a semantic competitor in-scene (2 per item set), the scene contained the target (e.g. mailbox), an unrelated item (e.g. lightbulb), and a pair of semantically unrelated items (e.g. peacock, bear). On these trials, we measure fixations to a non-competitor (e.g. lightbulb) that was unrelated to the other items in the scene.

- On Filler trials (6 per item set), targets (e.g. lightbulb) were presented with 3 unrelated items. The purpose of filler trials was to make the semantic competition manipulation less noticeable, and to ensure that participants could not guess which of the 4 pictures would be the target on any given trial. We did not analyze eye-gaze from filler trials.

Trial	Condition	Target	Critical object	Other	Other
1	Semantic-competition	candle	lightbulb	peacock	shoe
2	Semantic-competition	lightbulb	candle	peacock	shoe
3	Unrelated target (competitors present)	peacock	candle	lightbulb	shoe
4	Unrelated target (competitors present)	shoe	lightbulb	candle	peacock
5	Unrelated target (competitors absent)	mailbox	lightbulb	peacock	bear
6	filler	lightbulb	mailbox	bear	peacock
7	filler	peacock	bear	mailbox	lightbulb
8	filler	bear	peacock	lightbulb	mailbox
9	Unrelated target (competitors absent)	bear	candle	mailbox	shoe
10	filler	candle	bear	shoe	mailbox
11	filler	mailbox	shoe	bear	candle
12	filler	shoe	mailbox	candle	bear

Table 2. Illustration of how a single item set was used to create 12 trials. Shaded cells indicate the objects that were the focus of the eye-tracking analyses. Filler trials were not analyzed.

Each target item (e.g. candle) appeared in two different item sets, in one set as a critical target or competitor, and in the other as an unrelated item. Altogether participants saw a total of 204 unique pictures during the task. The pictures were easily nameable full-color drawings from Rossion and Pourtois (2004) and similar clip-art pictures. The auditory stimuli were individual spoken words, e.g. “*candle*”, “*lightbulb*” which were recorded in isolation in a sound-proof booth by a male talker with a mid-western regional accent of North American English.

Predictions

The aim of this study was to examine the impact of bilateral hippocampal damage, and resulting memory impairment, on lexico-semantic mapping during spoken word recognition.

Findings from the healthy young adult sample are expected to replicate findings from related paradigms in the literature (e.g., Yee & Sedivy, 2006; Huettig & Altmann, 2005), such that listeners will make more fixations to the critical object (e.g., *lightbulb*) on Semantic competition trials (e.g., target word = *candle*), compared to Unrelated-target trials (e.g., target word = *peacock*). We did not anticipate differences between Unrelated-target trials with vs. without a competitor in-scene. The use of two different types of Unrelated-target trials was simply to control for the presence of a semantic competitor in-scene.

Prior evidence points to links between use of semantic representations and hippocampal function in generation and recall tasks (Klooster & Duff, 2015; Butler et al., 2009; deZubicaray et al., 2014; Solomon et al., 2019; Jafarpour et al. 2017; Piai et al., 2016; Hilverman & Duff, in revision), and in generating lexical associations (Covington &

Duff, 2017). If hippocampus plays a critical role in the lexico-semantic mapping processes that occur during interpretation of individual spoken words, we would expect the competition effects on Semantic-competition trials to be attenuated in the Amnesia group, compared to the matched healthy comparison participants.

Alternatively, processing of individual spoken words, and activation of semantically related concepts as those words unfold in time, may be a hippocampal-independent process for previously-learned lexico-semantic mappings (Gluck & Meyers, 1993), or one that becomes hippocampal independent over time (McClelland et al., 1995). If so, amnesic patients would show the same pattern of semantic competition as their healthy comparisons. Such a result would help circumscribe the locus of deficits that are observed in patients with hippocampal amnesia, potentially pointing to impairments in more distant semantic relations or low frequency concepts (Klooster & Duff, 2015; Hilverman & Duff, in revision), or in the integration or combination of distinct concepts (Covington & Duff, 2017).

Analysis and Results

Accuracy in clicking on the picture corresponding to the target word was > 95% for all participant groups. Table 3 presents the click accuracy data and the number of critical trials that were entered into statistical analyses. Note that participants with amnesia completed the task twice across sessions, and in a few cases, trials were accidentally repeated 3 times across sessions, resulting in a total of 10 extra critical trials.

Participant Group	Accuracy	N	Trials in Analysis
Young adults	99.6 (SD=.45)	18	5508

Amnesia patients	97.1 (SD=3.7)	5	3070
Amnesia comparisons	99.7 (SD=.43)	5	1530

Table 3. Accuracy by participant group.

Eye-gaze analyses focus on the saccades and fixations made as participants interpreted the critical word¹. The dependent measure for the analysis is the fixations made to the *critical object*: the semantic competitor (in the Semantic-competition condition) or to the non-competitor (in the Unrelated-target conditions). Note that the experiment was designed such that across trials, the same critical objects (e.g., *candle*, *lightbulb*, see Table 2) served both as competitor and non-competitor, thus controlling for any particular visual or acoustic features of these items. The time-course of fixations for each of the three groups is presented in Figure 2.

¹ An initial analysis of the data from Experiments 1 and 2 were presented in Klooster (2016).

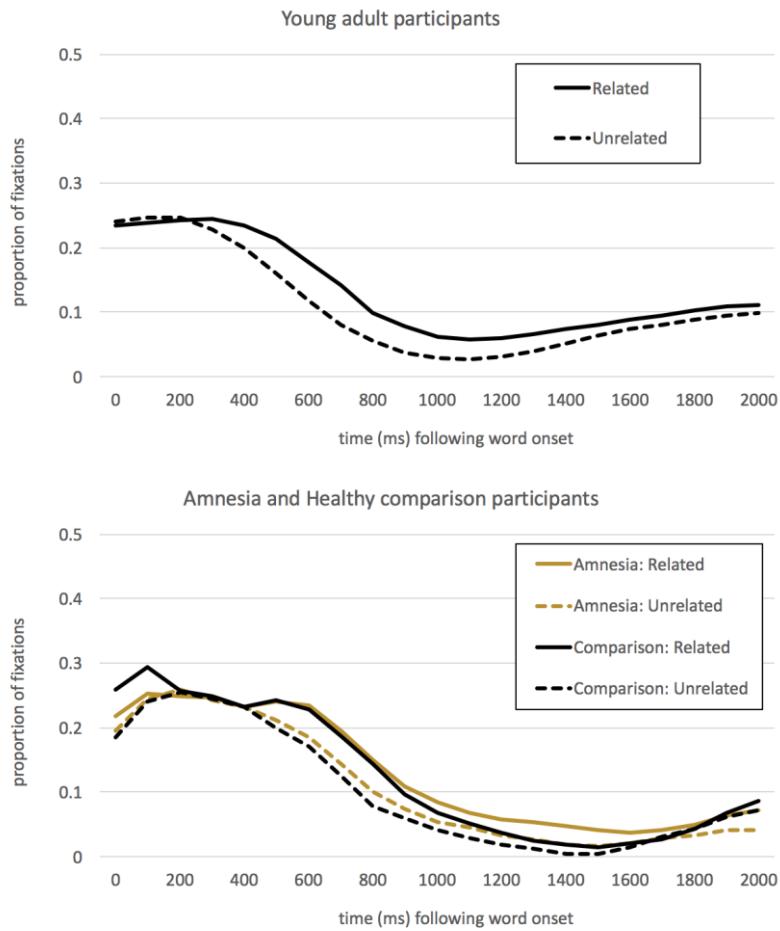


Figure 2. Experiment 1: Time-course of fixations to the critical object following critical word onset for healthy young adults (top panel); for patients with amnesia and healthy matched comparison participants (bottom panel).

The time-course data were analyzed using a binary measure of fixations to the critical object. This dependent measure indicated whether (1) or not (0) the participant was fixating the critical object in each of a series of 10ms time-bins over a period of time from 180-1300ms following the onset of the target word (e.g., *lightbulb*). Gaze data were modeled using a *dynamic GLMM* (Cho, Brown-Schmidt, & Lee, 2018), which models a binary fixation measure over a series of consecutive 10ms time-bins, while taking into account dependencies in fixations(i.e., the first-order autocorrelation, AR(1)) and trend across time points, as well as dependencies due to repeated testing of participants and items). This model was fit using the glmer function in the lme4 (Bates, Mächler, Bolker, & Walker, 2015) package in R (R Core Team, 2016). Fixations are assumed to be delayed by 200ms due to the time needed to program and launch an eye movement (Hallett, 1986). We use a 20ms baseline (180-200ms) in order to define the beginning of the AR(1) process. A fixed effect of AR(1) and random effects of AR(1) varied across participants and items (i.e., random slopes) were included in a dynamic GLMM. A trend (time) effect captured the tendency to fixate the competitor *less* over time as activation of the target increases; time was scaled (time(ms)/100) and centered. In addition, we modeled the experimental condition manipulation using Helmert contrasts. The first condition contrast compared the Semantic-Competition condition (-.66) to the two Unrelated-target conditions (+.33 and +.33). The second condition contrast directly compared the two Unrelated-target conditions (with competitor in scene = .5, without = -.5).

Analyses are presented below for (1) young adults; (2) amnesia patients and their matched comparisons. For each analysis, models were specified by first exploring

trend and AR(1) effects using descriptive statistics such as autocorrelations and partial autocorrelations, and then identifying a set of random effects that were appropriate for the data structure (see Cho et al. [2018] for details of the model-building steps). The random effects (i.e., random intercepts) included for items (the 51 stimulus sets that trials were designed around), trials (reflecting each unique trial participants completed), and participants. As noted in Bates, et al. (2018) convergence warnings (not errors) can occur when modeling large datasets. Given the large number of time-points analyzed here, convergence warnings were expected. When they occurred, following recommendations of Bates, et al. (2018) we re-fit the models using the `allFit` function to check for consistency in findings across multiple optimizers. The warning messages were considered to be false positives when estimates for the fixed condition effects (and interactions with group) were consistent out to three decimal places across the optimizers. The final selected models are presented below including estimates for the fixed and random effects.

Young Adults

Consistent with many findings in the literature, young adult participants made more fixations to the semantic competitor than unrelated objects (first Condition contrast, $\beta = -0.29$, $z = -7.13$). As predicted, the difference between the two types of unrelated trials (second Condition contrast) was not significant ($\beta = -0.05$, $z = -0.96$). A significant effect of the fixed AR(1) term ($\beta = 9.05$, $z = 140.93$) reflects serial dependency from time-point to time-point in whether or not the participant fixated the target. A significant time (trend) effect ($\beta = -0.12$, $z = -21.52$) is due to a decrease in

competitor fixations over time within the trial, consistent with participants locating and clicking the target.

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-5.981	0.029	-204.802	<.0001
AR(1)	9.049	0.064	140.930	<.0001
time	-0.118	0.005	-21.518	<.0001
Condition 1 (competition = -0.66, unrelated with competitor = 0.33, unrelated without competitor = 0.33)	-0.290	0.041	-7.125	<.0001
Condition 2 (competition = 0, unrelated with competitor = -0.5, unrelated without competitor = 0.5)	-0.046	0.048	-0.955	0.34
Random effects	Variance	SD		
trial (intercept)	0.034	0.185		
item (intercept)	0.000	0.001		
AR slope by participant	0.053	0.231		

Table 4. Results of dynamic GLMM for young adult participants ($N=18$), 306 trials, 51 items and 611,388 observations. Bolded values indicate significance at $p\text{-value}<.05$.

Participants with Amnesia and Matched Comparison Participants

Like healthy young adults, participants with amnesia made more fixations to the semantic competitor than unrelated objects. A similar pattern was observed in the healthy comparison participants matched to the participants with amnesia. Participant group was dummy coded with participants with amnesia as the reference level, allowing the fixed condition effect to be interpreted as the effect in the patient group; the condition by group interactions test whether the condition effects are larger in the healthy comparison group. A significant effect of the first Condition contrast was due to more fixations to the semantic competitor than non-competitors ($\beta = -0.27$, $z = -6.24$). While this effect did not significantly interact with participant group ($\beta = 0.43$, $z = .56$),

there was a main effect of group, due to overall fewer fixations in the healthy comparison group. A supplemental analysis that treated Healthy comparison participants as the reference level revealed a similar condition effect in the Comparison participants ($\beta = -0.23$, $z = -3.59$). Lastly, significant effects of the fixed AR(1) term ($\beta = 9.17$, $z = 62.96$) and the time (trend) effect ($\beta = -0.17$, $z = -28.28$) are due to dependency in fixations from time-point to time-point, and a decrease in fixations to the competitor over time, respectively.

To calculate the sample size that would be needed to detect an effect as large or larger than the non-significant Condition by Group interaction ($\beta = 0.043$), we conducted a simulation-based power analysis at alpha=0.05 using the simr package (Green & MacLeod, 2016) in R (R Core Team, 2016) based on the model presented in Table 5. This analysis revealed that even if we were to double our sample size to 20, this would only result in an estimated power of 0.125 (95% CI: 0.083, 0.179) to detect an effect of that size. Given the difficulty in recruiting this participant group, pursuing an effect size that small is impractical with sample sizes that are feasible given this population. A separate power analysis for Experiment 1 estimated power to replicate the Condition effect in the Amnesia group (dummy coded as the reference level); that analysis estimated that with $N=20$, power to replicate the Condition effect approached 1 (95% CI: 0.982, 1).

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-5.930	0.059	-100.544	<.0002
AR(1)	9.169	0.146	62.963	<.0001

Time	-0.170	0.006	-28.277	<.0000
Group (Amnesia = 0, Comparison = 1)	-0.145	0.068	-2.142	0.032
Condition 1 (competition = -0.66, unrelated with competitor = 0.33, unrelated without competitor = 0.33)	-0.272	0.044	-6.242	<.0001
Condition 2 (competition = 0, unrelated with competitor = -0.5, unrelated without competitor = 0.5)	-0.077	0.052	-1.459	0.145
Condition 1*Group	0.043	0.077	0.564	0.573
Condition 2*Group	-0.018	0.092	-0.195	0.845
Random Effects	Variance	Std.Dev.	Corr	
Trial (intercept)	0.000	0.004		
Item (intercept)	0.012	0.109		
Participant (intercept)	0.016	0.128		
<i>participant*AR(1) (slope)</i>	0.197	0.444	-0.81	

Table 5. Experiment 1. Results of dynamic GLMM for participants with Amnesia ($N=5$) and healthy comparison participants ($N=5$), 618 trials, 51 items and 510,600 observations. Bolded values indicate significance at p -value $<.05$.

Finally, a supplemental analysis tested for practice effects in the Amnesia group, as they completed the task twice. While there were fewer overall fixations to the competitor the second time the patients with Amnesia did the task ($\beta = -0.14$, $z = -3.18$), this did not interact with the condition effect ($\beta = 0.08$, $z = 0.96$), indicating that the preserved competition effects in the Amnesia group were not a result of practice in the task.

Interim Discussion

The results of Experiment 1 demonstrate a robust semantic competition effect as participants were interpreting individual spoken words. This effect was present in all 3 groups of participants that were tested. The magnitude of the semantic competition effect was not significantly attenuated in amnesia. These findings suggest that during interpretation of individual spoken words, that activation of semantically related

competitors remains relatively intact, even in the face of severe declarative memory impairment. In Experiment 2 we test for semantic activation in sentence contexts.

Experiment 2

Experiment 2 examines lexico-semantic mapping processes during interpretation of sentences that do or do not lead to a strong expectation for an upcoming word such as “*She will hunt the deer*”. Whereas Experiment 1 investigated lexico-semantic mapping as participants localized a referent as it was named, in contrast, Experiment 2, we examine lexico-semantic mapping between the verb and candidate referents.

Prior findings that real-time processing of sentences in discourse contexts is impaired following hippocampal damage (Kurczek et al., 2013; Covington et al., 2020) raise the possibility that the processing of even simple sentences as they unfold in time will be impaired in amnesia. Hippocampus clearly plays a role in relational binding (Eichenbaum & Cohen, 2001; Hannula, Tranel, & Cohen, 2006; Konkel, Warren, Duff, Tranel, & Cohen, 2008), even over short time-scales (Barense, Gaffan, & Graham, 2007; Hannula & Ranganath, 2008; Hannula et al., 2006). Thus, a clear prediction is that damage to hippocampus would confer deficits in the ability to process the meaning of a sentence as it builds over time, linking meaning of words together, and to pictorial representations in the corresponding visual world. If hippocampal damage impairs lexico-semantic mapping processes that occur as listeners construct the meaning of an unfolding sentence, the ability to use semantic information contained in the initial part of the sentence to guide expectations for upcoming referents should be attenuated in patients with bilateral hippocampal damage and dense amnesia. Alternatively, if

patients with hippocampal damage are successful, this would suggest that, at least for linking simple sentence meanings to simple contexts, intact hippocampal functioning is not required.

Method

Participants

A sample of 16 healthy young adults from the student community at the University of Illinois Urbana-Champaign participated in this study in exchange for \$8 / hour payment or partial course credit. As in Experiment 1, the young adult sample was tested in the first author's laboratory at the University of Illinois, and allowed us to vet the materials.

The focus of the research was on the same participants with bilateral hippocampal damage ($N=5$) that participated in Experiment 1. These patients were matched to 10 healthy comparison participants; the additional healthy comparison participants provided us with a better estimate of the effect size in the healthy population in this age range. The patients and comparison participants were tested either in the last author's lab at the University of Iowa, or at a location convenient to the patient (e.g., in a private conference room at a hotel near their home). The patients and healthy comparison participants were tested between the years 2014 and 2016. As in Experiment 1, the patients and healthy comparison participants were compensated \$15 per hour of participation.

Procedure

Participants were tested individually. After signing a consent form, the participant was seated at a computer with a desktop-mounted Eyelink 1000 eye-tracker.

Participants were instructed that on each trial, they would hear a sentence describing one of four pictures on the screen. When they heard the picture name, they were asked to click on it as quickly as they can. All participants completed the task successfully.

The experiment was composed of a total of 232 trials. Young adults completed a single set of 232 trials in less than an hour. Patients and their matched comparison participants each completed 3-4 sessions depending on availability. Breaks were provided between sessions and as needed within a session. In the event of scheduling constraints, sessions were completed on different days. Each trial began with a drift-check of the eye-tracker. If the drift-check failed, the eye-tracker was re-calibrated. Next, 4 pictures appeared on the screen, followed by a 1-second delay, and auditory presentation (over the computer speakers) of the sentence, e.g., “*She will hunt the deer.*” Once the computer recorded the participant’s click response, there was another 1-second delay before the drift-check for the next trial.

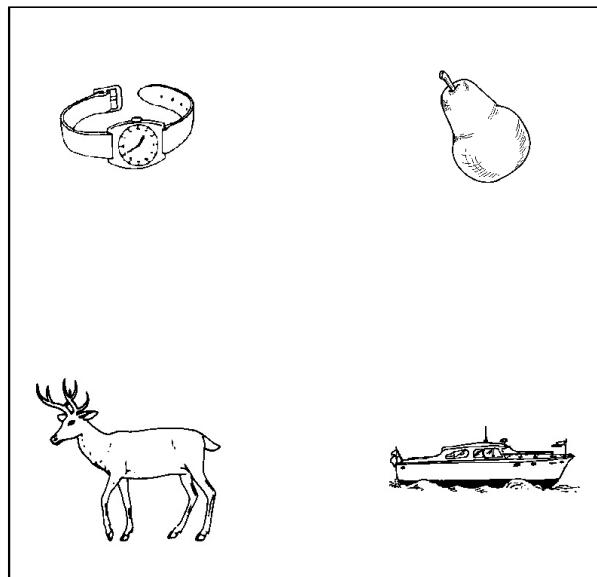


Figure 3. Experiment 2: Example experimental display.

Materials

Each trial presented 4 unrelated pictures on-screen (Figure 3), and played a sentence that referenced one of the pictures, e.g., “*She will hunt the deer.*” The critical manipulation was whether the verb in the sentence strongly predicted the direct object (e.g., *hunt*, restrictive trials), or did not strongly predict any of the four pictures in the scene (e.g., *paint*, non-restrictive trials).

The materials were adapted from Nozari, et al. (2016). We created 29 sets of 4 easy-to-name pictures. Each set was designed to generate 8 trials, where each picture was the target twice, once with a restrictive verb and once with a non-restrictive verb, for a total of $29 \times 8 = 232$ total trials. The materials were selected on the basis of two norming studies in which a separate group of participants were given the preamble, e.g., “*She will tune the...*”, and were asked to select the picture representing the most likely continuation out of 4 possible choices (see Appendix for details on the norming studies and the complete list of materials).

For Restrictive trials (102 total), the target was selected as the best continuation over on average 98.3% (range: 96% to 100%). For Non-Restrictive trials (75 total), the target was selected as the best continuation over on average 28.3% (range: 57% to 0%). An additional constraint on the Non-Restrictive trials was that none of the three non-target items was selected as the best continuation more than 60% of the time in the norming studies. Identifying stimuli that met these constraints was challenging, thus there were fewer Non-Restrictive trials. The remaining trials (55 total) were of the same format (verbs designed to be Restrictive vs. Non-Restrictive) but were not included in

the planned analysis because they did not meet these criteria based on the norming data.

Altogether participants heard 232 sentences, each of which ended in one of 31 different target words (targets were repeated 6-8 times across the entire experiment). Each target was paired with one of 31 Restrictive verbs, and one of 31 Non-Restrictive verbs. The same 31 pictures served as non-targets on other trials, such that participants saw each of the 31 pictures 24-32 times across the entire study. The pictures were the same black and white line drawings as used in Nozari et al., (2016), which were taken either from Snodgrass and Vanderwart (1980) or from the IPNP corpus (Szekely et al., 2004).

The auditory stimuli were recorded by the first author. For each target within a set, the same recording of the preamble “*She will*” was spliced onto the restrictive (e.g., ... *hunt the deer*), and the non-restrictive (e.g., ... *paint the deer*) sentence versions. The average delay between verb onset and noun onset was 948ms (SD=131ms, max=1314, min=619). Note that to preserve the naturalness of the stimuli as much as possible, the verb-*the*-noun sequence was not spliced, thus some co-articulatory information was likely present leading into the noun.

Predictions

The aim of Experiment 2 was to examine the impact of bilateral hippocampal damage and declarative memory impairment on the use of semantic information to guide expectations for how a sentence will unfold in time. Findings from the healthy young adult sample are expected to replicate findings from related paradigms in the literature (e.g., Altmann & Kamide, 1999), such that upon hearing the verb, listeners will

make more fixations to the object corresponding to the upcoming noun (e.g., *deer*) when the verb more clearly predicted the noun (restrictive trials) than when it did not (non-restrictive trials).

While the results of Experiment 1 did not find any evidence that bilateral hippocampal damage impaired semantic activation during processing of individual words, this preservation of function may be due to the fact that the bottom-up input provided by the target words in Experiment 1 was sufficiently strong to activate semantic associates of that word. In contrast, the sentence contexts tested in Experiment 2 require the listener to interpret the verb in the sentence context and then guide the eyes to a to-be-named referent. This more integrative and contextually constrained prediction process may make greater demands on the hippocampal-dependent declarative memory system (Bonhage et al., 2015; Jafarpour et al., 2017). Indeed, in our prior work, linking words across sentences (e.g. to interpret an ambiguous pronoun using the prior discourse context) was profoundly impaired in amnesia (Kurczek, et al. 2013; Rubin, et al. 2011), thus the need to interpret the verb and noun in context here may be similarly hippocampal demanding. If so, individuals with hippocampal amnesia should show an attenuated condition effect (i.e., a smaller difference between restrictive and non-restrictive trials).

Alternatively, processing of simple sentences and activation of semantically related concepts as sentences unfold in time may be a hippocampal-independent process for previously learned associations (e.g. Gluck & Meyers, 1993). If so, amnesic patients should show a condition effect of the same magnitude as their healthy comparisons. Patients' successful use of information carried on the verb to activate

information that is closely semantically related (e.g., fly → looks to kite) would further circumscribe the role of hippocampus in language processing, potentially pointing towards a more prominent role in more distant relations or in generation of relations.

Analysis and Results

Accuracy in clicking on the picture corresponding to the target word was >90% for all participant groups (Table 7)².

Participant Group	Accuracy	N	Trials in Analysis
Young adults	99.2% (SD=1.7%)	16	2,832
Amnesia patients	94.2% (SD=9.4%)	5	3,540
Amnesia comparisons	99.8% (SD=0.3%)	10	6,726

Table 7. Experiment 2: Accuracy by participant group.

Eye-gaze analyses focus on the saccades and fixations made as participants interpreted the unfolding sentence starting at verb onset. The dependent measure for the analysis is the fixations made to the *target* object. The time-course of fixations to the target object following verb onset for the three groups is presented in Figure 4.

² Average click accuracy in Experiment 2 was slightly lower in the Amnesia group, due to more click errors in the two HSE patients, 1951 (94% correct) and 2308 (78% correct); accuracy was high in the anoxic patients, 1846 (100%), 2363 (99%), 2563 (99.7%). The HSE patients have little computer experience and sometimes have difficulty manipulating the mouse (87% of 2308's errors were due to clicking the bottom right corner of the screen). We report the click data for completeness; our planned analyses focus on the on-line measures of language understanding, rather than the clicks.

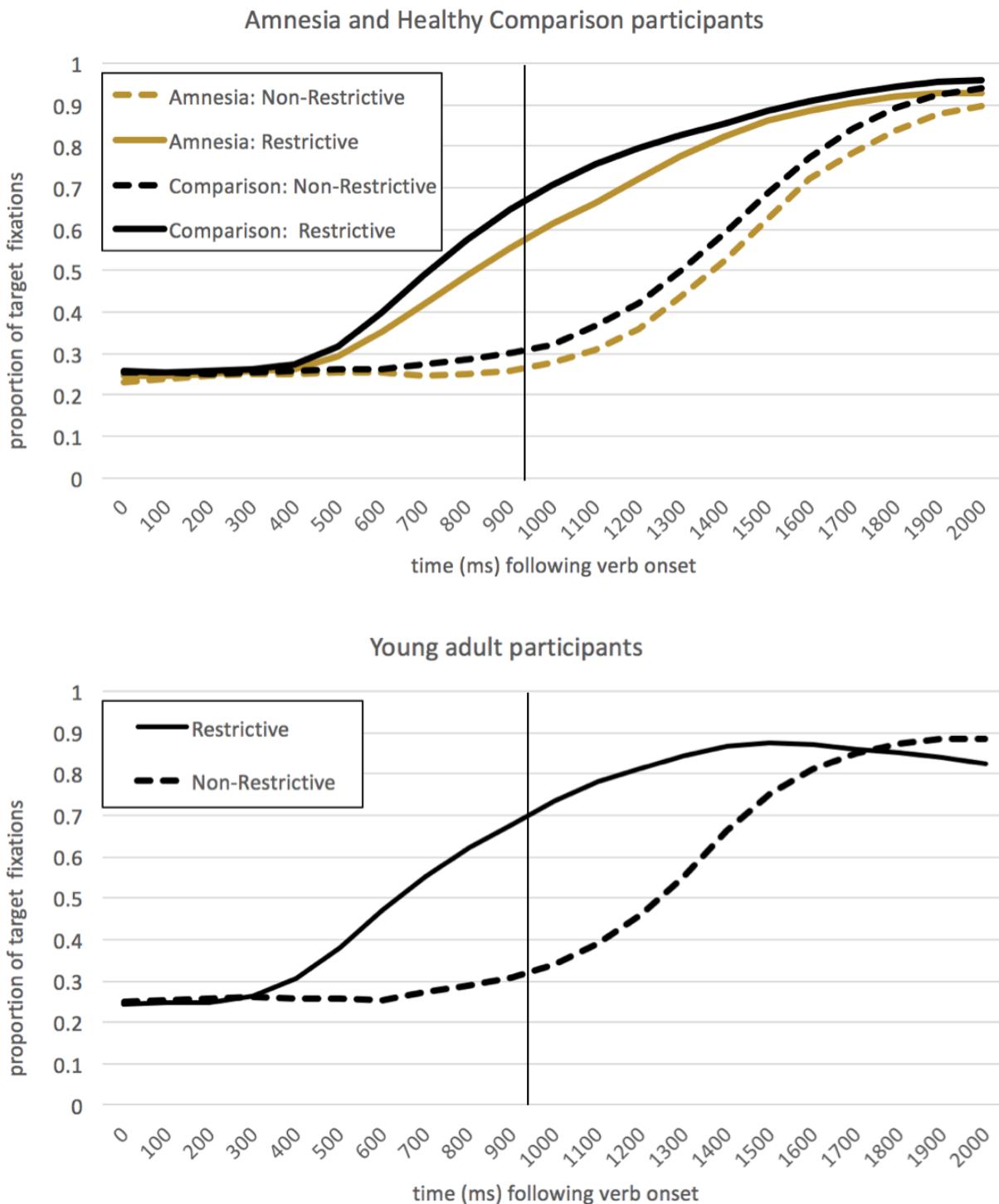


Figure 4. Experiment 2: Time-course of fixations to the target object following the onset of the verb in sentences like “She will paint/ hunt the deer”, separately for sentences with restrictive (e.g. hunt) and non-restrictive (e.g. paint) verbs. Top panel: young adult participants. Bottom panel: Participants with amnesia and matched healthy comparison participants. Vertical line indicates average noun onset following the verb (948ms).

The time-course data were analyzed using a binary measure of fixations to the target in a series of 10ms time-bins from 180-1700ms following the onset of the verb. Experiment 2 used a longer window of analysis as the spoken stimulus is a full sentence, rather than just a single word as in Experiment 1. The average time between the onset of the verb and the offset of the target noun was 1,435ms, thus given a 200ms eye movement delay, this time-window is expected to capture processing of both the verb and noun together. Critically note that the noun is the same across the restrictive and non-restrictive conditions—only the verb varied. As in Experiment 1, gaze was modeled using a *dynamic GLMM* (Cho et al., 2018), with a 20ms baseline (180-200ms) in order to define the beginning of the AR(1) process. As before, the model was fit using the *glmer* function in R (Bates et al., 2018; R Core Team, 2016).

Fixed effects included the AR(1), and a trend (time) effect capturing the tendency to fixate the target *more* over time as the sentence unfolded. Orthogonal contrasts compared the restrictive (= .5) to the non-restrictive condition (= -.5). Time following verb onset was included to model trend effects in the data; time was scaled (ms/100) and centered. Analyses are presented below for (1) young adults, and (2) amnesia patients and their matched comparisons. As in Experiment 1, for each analysis, models were specified by first exploring trend and AR(1) effects, and then identifying a set of random effects. The final selected models are presented below.

Young Adults

Consistent with many findings in the literature, young adult participants made more fixations to the target in the Restrictive compared to the Non-Restrictive condition

($\beta = .63$, $z = 10.33$). Significant effects of time ($\beta = .134$, $z = 33.72$) and the AR(1) ($\beta = 9.39$, $z = 80.31$) reflect an increase in target fixations over time within the trial, and dependency between adjacent time-points, respectively.

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-4.109	0.104	-39.420	<.0001
AR(1)	9.391	0.117	80.310	<.0001
Time	0.134	0.004	33.720	<.0001
Condition (restrictive = 0.5; non-restrictive = -0.5)	0.630	0.061	10.330	<.0001
Random Effects	Variance	Std.Dev.	Corr	
Trial (intercept)	0.008	0.091		
Item (intercept)	0.034	0.185		
Subject (intercept)	0.157	0.396		
AR(1) slope	0.196	0.443	-0.52	

Table 8. Experiment 2: Results of dynamic GLMM for young adults ($N=16$), 177 trials, 59 items (verbs) and 427,632 observations. Bolded values indicate significance at p -value $<.05$.

Amnesia and Healthy Comparison Participants

The analysis of fixations for participants with amnesia and the matched healthy comparison participants included participant Group as a factor. Group was dummy coded with participants with amnesia as the reference level, allowing the fixed condition effect to be interpreted as the effect in the patient group; the condition by group interaction tests whether the condition effect was larger in the healthy comparison group. In addition, recall that patients and matched comparison participants repeated the trials across multiple test blocks, thus test block was included as an additional factor.

Significant effects of time ($\beta = .153$, $z = 28.56$) and the fixed AR(1) ($\beta = 9.80$, $z = 99.39$) reflected increasing fixations throughout time within each trial, and serial

dependency between adjacent time-points, respectively. A positive interaction between time and testing block ($\beta = .004$, $z = 2.16$) indicated that the temporal effects within trials grew in magnitude across blocks, consistent with practice effects. Lastly the AR(1) decreased across blocks ($\beta = -0.072$, $z = -3.80$), possibly indicating that participants were more willing to look around the screen more with more practice at the task.

Participants with amnesia showed a robust condition effect, with more fixations to the target in the Restrictive compared to the Non-Restrictive condition ($\beta = .636$, $z = 6.59$). The interaction between Condition and Group was not significant, indicating that the magnitude of the condition effect was not significantly smaller in the amnesia group ($\beta = .07$, $z = .71$), though we acknowledge that the study may be underpowered to test for a small interaction effect. That said, a supplemental analysis in which the healthy Comparison participants were set as the reference level revealed a comparably sized condition effect for Comparison participants ($\beta = .700$, $z = 8.48$). An additional supplemental analysis examined fixations made following verb onset, but prior to the onset of the noun. This analysis, which is reported in the Appendix, revealed the same pattern of effects as reported in the primary analysis, with a significant effect of Condition that did not interact with Group.

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-4.485	0.142	-31.647	<.0001
Condition (restrictive = .5, non-restrictive = -.5)	0.636	0.096	6.592	<.0001
Group (Amnesia = 0, Healthy Comparison = 1)	0.118	0.148	0.798	0.425
Test Block	0.031	0.029	1.064	0.287
AR(1)	9.803	0.099	99.392	<.0001

time	0.153	0.005	28.561	<.0001
Condition*Group	0.066	0.093	0.709	0.478
Condition*Test Block	-0.001	0.028	-0.041	0.967
Group*Test Block	0.042	0.035	1.215	0.224
Test Block*AR(1)	-0.072	0.019	-3.798	<.001
Test Block*time	0.004	0.002	2.157	0.031
Condition*Group*Test Block	0.002	0.034	0.064	0.949
Random Effects	Variance	Std.Dev.	Corr	
Trial (intercept)	0.018	0.133		
Test Block (slope)	0.001	0.037	-0.510	
Item (intercept)	0.048	0.219		
Participant (intercept)	0.131	0.362		
Test Block (slope)	0.003	0.055	-0.600	
AR(1) (slope)	0.106	0.326	-0.710	0.020

Table 10. Experiment 2: Results of dynamic GLMM for patients with amnesia ($N=5$) and healthy comparison participants ($N=10$), 177 trials, 59 items (verbs) and 1,550,166 observations. Bolded values indicate significance at p -value $<.05$.

General Discussion

In two experiments, we find that patients with bilateral hippocampal damage and severe declarative memory impairment successfully activate semantic knowledge from spoken words and sentences with a time-course similar to healthy matched comparison participants. How can we reconcile this evidence of sparing in activating semantic knowledge from spoken language, given previous findings for a role for hippocampus in generating semantic associations (Butler et al., 2009), in semantic interference during naming (deZubicaray et al., 2014), in semantic feature generation (Klooster & Duff, 2015), and in prediction (Bonhage et al., 2015; Chen et al., 2013; Henson & Gagnepain, 2010)?

One explanation is that hippocampus contributes to semantic processing through the lifelong tuning of semantic knowledge, and thus maintenance of the integrity of the semantic network. This proposal fits with work showing hippocampal contributions to relational processing of information across time or space (Hannula et al., 2006) and the updating of previously acquired information through reconsolidation (McKenzie & Eichenbaum, 2011). A role in life-long tuning would predict that use and processing of remote semantic connections will fail in amnesia due to impairments in hippocampal-mediated maintenance of network connections. Such a mechanism would predict graded effects, with close and frequent semantic relations being intact, and distant relations impaired. On this view, then, the close and frequent semantic relations tested in the present study had not yet degraded in amnesia. In contrast, generation tasks may place more demands on remote or infrequently used connections within the semantic network, thereby revealing substantial impairment (e.g., Klooster & Duff, 2015; Cutler et al., 2019; Hilverman & Duff, in revision).

Along these lines, we note that the patients with amnesia scored normally on standard neuropsychological assessments of semantic memory. Their ability to name common objects or provide a simple definition for common words appears intact. Their surface-level knowledge of common words and concepts appears normal. On more sensitive measures of semantic richness and depth of knowledge, however, the patients' knowledge is severely impoverished (Klooster & Duff, 2015). Similarly, in the current studies, patients' surface level of knowledge appears intact. Patients and comparison participants were sensitive to strong semantic associates of target words.

Whether patients would be as sensitive to more remote associates remains an open question.

The critical words we tested mapped onto simple images in the immediate environment (e.g., *banana*, *lightbulb*), and the semantic relationships we tested were all simple, close relations (e.g., *lock – key*; *fly – kite*). In prior work, language processing impairments have been observed in patients with hippocampal amnesia when meaning builds across words in a phrase or discourse (Kurczek et al., 2013; Rubin et al., 2011; Covington et al., 2020; Covington & Duff, 2017). Thus, language comprehension processes may be more likely to place demands on hippocampus when tasks go beyond those that involve basic one-to-one mappings in the immediate world, or that involve linking linguistic components across time or sentences (e.g., Kurczek et al. 2013; Covington et al., 2020; Rubin, et al., 2011; also see Nieuwland & Martin, 2017). An unknown is whether patients with hippocampal damage would exhibit problems in semantic processing more generally and in the absence of linguistic input, such as identifying semantic relations among objects when no lexical input is provided.

Another key feature of the present research is that it involved *comprehension* of simple spoken language, rather than *production* of language. Recall that Hilverman and Duff (in revision) reported deficits in single word production in patients with amnesia, in a task where the to-be-named items varied on dimensions including frequency, imageability, abstractness, and that Hamamé et al. (2014) found that naming latency was related to hippocampal activity. A key difference between the present results – which offer no evidence of hippocampal necessity, and these findings is that both Hilverman and Duff (in revision), and Hamamé et al. (2014) tested participants in a task

where patients had to *generate* a lexical item. Similarly, Solomon et al. (2019), Jafarpour, et al. (2017), and Piai et al. (2016), all of whom report hippocampal involvement in language tasks, similarly employed tasks with a production component.

In healthy adults, the generation and production of information has well known memorial benefits as compared to comprehension or reading of the same information (Jacoby, 1978; Slamecka & Graf, 1978; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010). These generation benefits extend to conversational language use (McKinley, Brown-Schmidt, Benjamin, 2017), picture naming (Zormpa, Brehm, Hoedemaker, & Meyer, 2019), and even commenting on social media posts (Zimmerman & Brown-Schmidt, 2020), with better memory when persons talk about, describe, and otherwise engage with material as compared to when they view or hear descriptions of it. By contrast, prior work examining recognition memory in patients with hippocampal amnesia finds that the generation benefit is absent (Verfaellie & Treadwell, 1993; Turriziani, et al 2008), pointing to hippocampal involvement in this process. The present work shows that during spoken language understanding, patients' spontaneous eye-gaze was drawn to semantically related images in the immediate visual context, similar to healthy matched comparison participants. By contrast, in a task that required generation or simulation of even simple and close semantic associates, we may expect to see evidence of a pronounced deficit in amnesia. The active processes involved with, e.g. generating a word or a picture name may be more hippocampal demanding than more passive processes involved with understanding what another person has to say. An open question, then, is whether preserved use of semantic association in linguistic prediction, would be linked to sparing in the *production* (generation) of similarly close

semantic associates. If so, we may also expect to find that language production tasks that produce in-the-moment generation deficits in amnesia will be linked to a lack of a subsequent generation benefit in memory.

Our findings are potentially consistent with Gluck and Myers (1993) view of hippocampus as a predictive autoencoder that, over time, supports the development of cortical representations that are linear combinations of those developed in the hippocampus. If we construe language comprehension as mapping an input (acoustic signal) to an output (semantics), it is reasonable to assume that hippocampus may have mediated the formation of the initial representations that accomplished such mappings. In the face of hippocampal damage, these old representations can remain intact and functional. If true, this would mean that the ability to map sound to meaning remains intact following hippocampal damage, as it draws on these representations which are now hippocampal-independent. On this view, we would expect to find deficits in amnesia when the language processing task required calculating new predictive relationships, or the reversal of existing predictive relationships (see Shohamy, et al. 2009). For example, language tasks that involve learning mappings between unfamiliar talkers and their preferences (Creel & Tumlin, 2011), or that involve tracking individual referents that change in form over time (Altmann, 2017) may place greater demands on hippocampus.

Conclusion

The results of two visual-world eye-tracking studies find no evidence of impairment in the ability of patients with bilateral hippocampal damage and severe

declarative memory impairment to activate semantic knowledge from spoken words and sentences, and shape predictions about upcoming referents. These findings offer considerable utility in characterizing the nature of hippocampal contributions to language use and processing and for understanding the neurobiology of various disorders where language is affected (aphasia, semantic dementia, traumatic brain injury, Alzheimer's disease) with and without overt or known hippocampal involvement. Our findings show that during spoken language comprehension, the ability to spontaneously generate lexico-semantic mappings is intact in the face of severe declarative memory impairment and bilateral hippocampal damage. This sparing is present in the context of known contributions of hippocampus to language use and processing more generally (for reviews, see Covington & Duff, 2016; Duff & Brown-Schmidt, 2012). The sparing that was observed here offers a useful signpost in interpreting the broader range of language deficits seen in patients with hippocampal damage (e.g., Kurczek et al. 2013; Rubin et al., 2011; Covington et al., 2020; MacKay, et al., 1998; Kurczek & Duff, 2011), as well as pervasive evidence from direct hippocampal recordings of hippocampal involvement in the processing of semantic relations (Solomon et al., 2019; Jafarpour, et al., 2017; Piai et al., 2016). The fact that the patients here were successful at mapping words to semantically related images in the immediate environment provides grounding for future work examining how deep a semantic deficit might be, given evidence of impaired semantic feature generation, particularly for distant semantic relations (Klooster & Duff, 2015; Cutler et al., 2019).

Acknowledgments

This work was supported by NIDCD grant R01 DC011755 to M. C. D. and S. B. S. and by NSF grant SES 1851690 to S. J. C. and S. B. S.

References

- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, 38(4), 419-439.
- Altmann, G. T. (2017). Abstraction and generalization in statistical learning: implications for the relationship between semantic types and episodic tokens. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160060.
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of memory and language*, 57(4), 502-518.
- Barense, M. D., Gaffan, D., & Graham, K. S. (2007). The human medial temporal lobe processes online representations of complex objects. *Neuropsychologia*, 45(13), 2963-2974.
- Bates, D., Mächler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., et al. (2018). *Package “lme4”: Linear mixed-effects models using ‘eigen’ and s4*. Retrieved May 20th, 2019 from <https://cran.r-project.org/web/packages/lme4/lme4.pdf>.

- Bonhage, C., Mueller, J., Friederici, A., & Fiebach, C. (2015). Combined eye-tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*, 58, 33-47.
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS one*, 5(5), e10773.
- Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of standardized stimuli (BOSS) phase II: 930 new normative photos. *PLoS One*, 9(9), e106953.
- Brown-Schmidt, S. & Duff, M. C. (2016). Memory and common ground processes in language use. *Topics in Cognitive Science*, 8, 722-736.
- Buckner, R. (2010). The role of the hippocampus in prediction and imagination. *Annual Review of Psychology*, 61, 27-48.
- Butler, C. R., Brambati, S. M., Miller, B. L., & Gorno-Tempini, M. L. (2009). The neural correlates of verbal and non-verbal semantic processing deficits in neurodegenerative disease. *Cognitive and behavioral neurology: official journal of the Society for Behavioral and Cognitive Neurology*, 22(2), 73.
- Chen, J., Dastjerdi, M., Foster, B. L., LaRocque, K. F., Rauschecker, A. M., Parvizi, J., & Wagner, A. D. (2013). Human hippocampal increases in low-frequency power during associative prediction violations. *Neuropsychologia*, 51(12), 2344-2351.
- Cho, S. J., Brown-Schmidt, S., & Lee, W. Y. (2018). Autoregressive generalized linear mixed effect models with crossed random effects: an application to intensive binary time series eye-tracking data. *Psychometrika*, 83(3), 751-771.
- <https://doi.org/10.1007/s11336-018-9604-2>

- Covington, N., & Duff, M.C. (2017, June). Impaired Collocation Knowledge in Amnesia: Evidence for hippocampal contributions to statistical learning. Poster presentation at the Clinical Aphasiology Conference, Snowbird, Utah.
- Covington, N, Kurczek, J. Duff, M.C. Brown-Schmidt, S. (2020). The Effect of Repetition on Pronoun Resolution in Patients with Memory Impairment. *Journal of Clinical and Experimental Neuropsychology*, 42(2), 171-184.
- Covington, N. V., & Duff, M. C. (2016). Expanding the language network: Direct contributions from the hippocampus. *Trends in Cognitive Sciences*, 20(12), 869-870.
- Cutler, R., Duff, M., & Polyn, S. M. (2019). Searching for semantic knowledge: A vector space semantic analysis of the feature generation task. *Frontiers in human neuroscience*, 13, 341.
- Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. *Journal of Memory and Language*, 65(3), 264-285.
- Davachi, L., & DuBrow, S. (2015). How the hippocampus preserve orders: the role of prediction and context. *Trends in Cognitive Sciences*, 19 (2), 92-99.
- de Zubicaray, G., Johnson, K., Howard, D., & McMahon, K. (2014). A perfusion fMRI investigation of thematic and categorical context effects in the spoken production of object names. *Cortex*, 54, 135-149.
- Duff, M. C., & Brown-Schmidt, S. (2012). The hippocampus and the flexible use and processing of language. *Frontiers in human neuroscience*, 6, 69.
- Duff, M.C., & Brown-Schmidt, S. (2017). Hippocampal contributions to language use and processing. In Hannula, D. & Duff, M.C. (Eds.). *The hippocampus from cells*

to systems: Structure, connectivity, and functional contributions to memory and flexible cognition. pp. 503-536. Springer International Publishing.

Duff, M. C., Covington, N. V., Hilverman, C., & Cohen, N. J. (2020). Semantic memory and the hippocampus: Revisiting, reaffirming, and extending the reach of their critical relationship. *Frontiers in Human Neuroscience*, 13, 471.

Eichenbaum, H., & Cohen, N. J. (2001). From conditioning to conscious recollection: Memory systems of the brain. New York: Oxford University Press.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of memory and Language*, 41(4), 469-495.

Gabrieli JD, Cohen NJ, Corkin S (1988) The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. Special issue: single-case studies in Amnesia: theoretical advances. *Brain Cogn* 7:157–177.

Glenberg, A. M. (1997). What memory is for. *Behavioral and brain sciences*, 20(1), 1-19.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of memory and language*, 43(3), 379-401.

Gluck, M. A., & Myers, C. E. (1993). Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus*, 3(4), 491-516.

Gluck, M. A., Meeter, M., & Myers, C. E. (2003). Computational models of the hippocampal region: linking incremental learning and episodic memory. *Trends in cognitive sciences*, 7(6), 269-276.

- Gluck, M. A., Myers, C., & Meeter, M. (2005). Cortico-hippocampal interaction and adaptive stimulus representation: A neurocomputational theory of associative learning and memory. *Neural Networks*, 18(9), 1265-1279.
- Green, P. & MacLeod, C.J. (2016). "simr: an R package for power analysis of generalised linear mixed models by simulation." *Methods in Ecology and Evolution*, 7(4), 493–498. doi: [10.1111/2041-210X.12504](https://doi.org/10.1111/2041-210X.12504), <https://CRAN.R-project.org/package=simr>.
- Gruber, M., & Ranganath, C. (2019). How curiosity enhances hippocampal-dependent memory: The prediction, appraisal, curiosity, and exploration (PACE) framework. *Trends in Cognitive Sciences*, 23(12), 1014-1025.
- Hallett PE. (1986). Handbook of perception and human performance. In: Boff KR, Kaufman L, Thomas JP, editors. Eye movements. New York, NY: Wiley; pp. 10.1–10.112.
- Hamamé, C. M., Alario, F. X., Llorens, A., Liégeois-Chauvel, C., & Trébuchon-Da Fonseca, A. (2014). High frequency gamma activity in the left hippocampus predicts visual object naming performance. *Brain and language*, 135, 104-114.
- Hannula, D., Tranell, D., & Cohen, N. J. (2006). The long and the short of it: Relational memory impairments in amnesia, even at short lags. *The Journal of Neuroscience*, 26(32), 8352-8259.^[1]
- Hannula, D., & Raganath, C. (2008). Medial temporal lobe activity predicts successful relational memory binding. *Journal of Neuroscience*, 28, 116-124.^[2]
- Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, 20(11), 1315-1326.

Hilverman, C., & Duff., M.C. (in revision). Evidence of impaired naming in patients with hippocampal amnesia. *Hippocampus*.

Huetting, F., & Altmann, G. T. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, 96(1), B23-B32.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 17, 649–667.

Jafarpour, A., Piai, V., Lin, J. J., & Knight, R. T. (2017). Human hippocampal pre-activation predicts behavior. *Scientific reports*, 7(1), 5959.

James, L. E., & MacKay, D. G. (2001). HM, word knowledge, and aging: support for a new theory of long-term retrograde amnesia. *Psychological Science*, 12(6), 485-492.

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and language*, 49(1), 133-156.

Kensinger, E. A., Ullman, M. T., & Corkin, S. (2001). Bilateral medial temporal lobe damage does not affect lexical or grammatical processing: Evidence from amnesic patient HM. *Hippocampus*, 11(4), 347-360.

Klooster, N.B. & Duff, M.C. (2015). Remote semantic memory is impoverished in hippocampal amnesia. *Neuropsychologia* 79(A), 42 – 52.

Klooster, N. (2016). "The hippocampus and semantic memory beyond acquisition: a lesion study of hippocampal contributions to the maintenance, updating, and use

of remote semantic memory." PhD (Doctor of Philosophy) thesis, University of Iowa, 2016.

Klooster, N. B., Tranel, D., & Duff, M. C. (2020). The hippocampus and semantic memory over time. *Brain and language*, 201, 104711.

Konkel, A., Warren, D. E., Duff, M. C., Tranel, D., & Cohen, N. J. (2008). Hippocampal amnesia impairs all manner of relational memory. *Frontiers in human neuroscience*, 2, 15.

Kurczek, J., & Duff, M.C. (2011). Cohesion, coherence, and declarative memory: Discourse patterns of patients with hippocampal amnesia. *Aphasiology*, 25 (6-7), 700-712.

Kurczek, J., Brown-Schmidt, S., & Duff, M. (2013). Hippocampal contributions to language: Evidence of referential processing deficits in amnesia. *Journal of Experimental Psychology: General*, 142, 1346-1354. 

MacKay, D. G., Stewart, R., & Burke, D. M. (1998). HM revisited: Relations between language comprehension, memory, and the hippocampal system. *Journal of cognitive neuroscience*, 10(3), 377-394.

MacKay, D. G., Burke, D. M., & Stewart, R. (1998). HM's language production deficits: Implications for relations between memory, semantic binding, and the hippocampal system. *Journal of Memory and Language*, 38(1), 28-69.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 671–685.

- Maguire, E.,A. (2001). Neuroimaging studies of autobiographical event memory. *Phil Trans R Soc Lond B Biol Sci*, 356,1441–51.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102(3), 419.
- McKenzie, S., & Eichenbaum, H. (2011). Consolidation and reconsolidation: two lives of memories? *Neuron* 71 (2), 224–233.
- Manns JR, Hopkins RO, Squire LR (2003) Semantic memory and the human hippocampus. *Neuron* 38(1):127–133.
- McKinley, G. L., Brown-Schmidt, S., & Benjamin, A. S. (2017). Memory for conversation and the development of common ground. *Memory & cognition*, 45(8), 1281-1294.
- Nieuwland, M. S., & Martin, A. E. (2017). Neural oscillations and a nascent corticohippocampal theory of reference. *Journal of cognitive neuroscience*, 29(5), 896-910.
- Nozari, N., Mirman, D., & Thompson-Schill, S. L. (2016). The ventrolateral prefrontal cortex facilitates processing of sentential context to locate referents. *Brain and language*, 157, 1-13.
- Oppenheim, G.M., Tainturier, M.J., & Barr, P. (2015). Preserved cumulative semantic interference despite amnesia. *Front. Psychol. Conference Abstract: Academy of Aphasia 53rd Annual Meeting*. doi: 10.3389/conf.fpsyg.2015.65.00002
- O'Reilly, R. C., & Rudy, J. W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus*, 10(4), 389-397.

Piai, V., Anderson, K. L., Lin, J. J., Dewar, C., Parvizi, J., Dronkers, N. F., & Knight, R.

T. (2016). Direct brain recordings reveal hippocampal rhythm underpinnings of language processing. *Proceedings of the National Academy of Sciences*, 113(40), 11366-11371.

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved November 11, 2017, from <https://www.R-project.org/>

Rossoni, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33, 217–236. doi:10.1068/p5117

Rubin, R. D., Brown-Schmidt, S., Duff, M. C., Tranel, D., & Cohen, N. J. (2011). How do I remember that I know you know that I know? *Psychological Science*, 22, 1574-1582.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron*, 76(4), 677-694.

Schmolck, H., Kensinger, E. A., Corkin, S., & Squire, L. R. (2002). Semantic knowledge in patient HM and other patients with bilateral medial and lateral temporal lobe lesions. *Hippocampus*, 12(4), 520-533.

Shohamy, D., Myers, C. E., Hopkins, R. O., Sage, J., & Gluck, M. A. (2009). Distinct hippocampal and basal ganglia contributions to probabilistic learning and reversal. *Journal of Cognitive Neuroscience*, 21(9), 1820-1832.

- Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, 4, 592–604.^[1]
- Solomon, E. A., Lega, B. C., Sperling, M. R., & Kahana, M. J. (2019). Hippocampal theta codes for distances in semantic and temporal spaces. *bioRxiv*, 611681.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6 (2), 174.
- Stachenfeld, K., Botvinick, M., & Gershman, S. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, 20, 1643-1653.
- Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., ... Wicha, N. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, 51(2), 247–250.^[1]
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632-1634.
- Theves, S., Fernandez, G., & Doeller, C. F. (2019). The hippocampus encodes distances in multidimensional feature space. *Current Biology*, 29(7), 1226-1231.
- Trude, A. M., Duff, M., & Brown-Schmidt, S. (2014). Talker-specific learning in amnesia: Insight into mechanisms of adaptive speech perception. *Cortex* 54, 117-123.
- Turriziani, P., Serra, L., Fadda, L., Caltagirone, C., & Carlesimo, G. A. (2008). Recollection and familiarity in hippocampal amnesia. *Hippocampus*, 18(5), 469-480.

- Verfaellie, M., & Treadwell, J. R. (1993). Status of recognition memory in amnesia. *Neuropsychology*, 7(1), 5.
- Verfaellie, M., Bousquet, K., & Keane, M. M. (2014). Medial temporal and neocortical contributions to remote memory for semantic narratives: evidence from amnesia. *Neuropsychologia*, 61, 105-112.
- Wang, L., Hagoort, P., & Jensen, O. (2018). Language prediction is reflected by coupling between frontal gamma and posterior alpha oscillations. *Journal of cognitive neuroscience*, 30(3), 432-447.
- Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 1.
- Zormpa, E., Brehm, L. E., Hoedemaker, R. S., & Meyer, A. S. (2019). The production effect and the generation effect improve memory in picture naming. *Memory*, 27(3), 340-352.
- Zimmerman, J., & Brown-Schmidt, S. (2020). # foodie: Implications of interacting with social media for memory. *Cognitive Research: Principles and Implications*, 5, 1-16.

Appendix

Experiment 1 Stimulus Norming

Candidate pairs of semantically related words were initially selected from stimulus materials reported in the literature (Yee & Sedivy, 2006; Huettig & Altmann, 2005; Huettig, et al., 2006), and new semantically related pairs. Some pairs identified from the literature were Americanized (e.g. *teapot* instead of *kettle*). Using the LSA semantic similarity tool (<http://lsa.colorado.edu/>) we calculated the semantic similarity of the candidate pairs, using pairwise comparison of pairs of words, based on the “General reading up to 1st year of college”, 300 factors. We then paired each semantically related pair (e.g. Set 1: candle-lightbulb) with two words that were unrelated to each other (e.g. peacock-shoe), each of which came from a different semantically related-pair (e.g. Set 39: peacock-owl, and Set 32: jacket-shoe), see Table A1. Each of the 51 stimulus sets was designed to maximize the semantic relatedness of the critical pair of words (e.g. candle-lightbulb) while minimizing the semantic relationships among the other words. The average similarity across the 51 pairs of critical semantically related pairs was .361 (SD=.16), vs. .076 (SD=.06) for the relationship between the critical words and the unrelated and filler words. We also avoided initial phonological overlap between the words in the set (i.e. so that the words were not cohort-competitors, Allopenna, et al., 1998).

One question is whether the semantically related pairs were visually distinguishable, as items that are highly similar in shape can elicit competition effects (Dahan & Tanenhaus, 2005). To this end, we avoided words for which the associated pictures would be visually hard to distinguish on a computer screen (e.g. we avoided

pairs like *apple - peach*). Further, to evaluate the visual similarity of the semantic competitors to the image evoked by target words, we conducted a norming study with 12 participants (an additional two participants were run but not include in the analysis due to having done the study before, and falling asleep). Participants saw the target and competitor words, in a random order one at a time. The words were visually presented on the screen, e.g. “candle”, and participants were asked to form a mental image of the object the word referred to. A picture of one of the other words from that item set was then presented (e.g. lightbulb, peacock or shoe) and participants were asked to “Rate the picture’s shape according to how similar it is to the mental image you formed.”, on a scale of 1=not similar to 7=highly similar. The aim of this norming study was to ensure that the target word was not visually confusable with the other words in the set (note we did not test the filler items). The average similarity rating for target and competitor items (e.g. the similarity of a picture of “lightbulb” after imagining “candle”) was 2.21 (SD=.94), vs. 1.20 (SD=.29) for pairings of targets with unrelated items, and competitors with unrelated items. While the difference was significant ($p < .01$) the ratings were, on average, close to floor. We suspect one reason for their higher similarity ratings for the semantically related items is that none of the pairs tested in the norming study matched (i.e. we never tested a case where the participant imagined “candle” and then saw a picture of a candle; we also did not test items that were intended to be shape competitors). Thus while the semantic similarity may have driven these visual similarity scores slightly higher for the semantically related items, on the whole, as intended, the stimuli were unlikely to cause visual confusion.

Set	Source	LSA	Target	Competitor	other	other	other	other
1	YS06	0.32	candle	lightbulb	peacock	shoe	mailbox	bear
2	YS06	0.26	car	bicycle	jacket	piano	tiger	dog
3	H06	0.46	carrot	tomato	leg	racket	basketball	horn
4	YS06	0.72	cat	mouse	lock	bed	french fries	anchor
5	YS06	0.5	hammer	nail	cup	orange	sandwich	brick
6	YS06	0.43	lock	key	matches	broccoli	airplane	soap
7	YS06	0.17	matches	lighter	car	doll	fence	book
8	YS06	0.33	pie	ice cream	butterfly	wagon	ladybug	ghost
9	YS06	0.36	robe	slippers	pear	trumpet	fork	motorcycle
10	new	0.38	scissors	paper	grapes	tie	igloo	clown
11	YS06	0.17	tape	glue	mitten	saw	chair	football
12	YS06	0.57	telescope	binoculars	glove	kite	ashtray	ear
13	new	0.45	tree	ax	bat	spider	hanger	camera
14	YS06	0.4	wallet	purse	rattle	toaster	cow	helicopter
15	YS06	0.66	window	door	robe	lighter	monkey	fox
16	HA05	0.32	accordion	flute	train	pan	burger	eye
17	HA05	0.14	balloon	doll	tree	potato	hotdog	gun
18	YS06	0.42	bat	racket	slippers	donut	necklace	cannon
19	new	0.16	battery	robot	ostrich	paper	clip	moon
20	HA05	0.46	bee	spider	accordion	ice cream	foot	horse
21	HA05	0.52	butterfly	ant	mushroom	refrigerator	cane	scale
22	HA05	0.58	celery	potato	battery	clarinet	table	windmill
23	HA05	0.27	chicken	penguin	violin	ax	clock	ring
24	H06	0.46	corn	broccoli	teapot	eagle	lamp	hand
25	HA05	0.39	cup	knife	pants	alligator	bus	rocket
26	HA05	0.18	desk	bed	pig	knife	cheese	handcuff
27	H06	0.4	elephant	alligator	purse	bee	wine	cone
28	new	0.36	refrigerator	toaster	window	nail	briefcase	gorilla
29	HA05	0.27	glove	tie	chicken	door	whale	sheep
30	HA05	0.18	grapes	watermelon	rabbit	flute	tent	house
31	H06	0.68	guitar	clarinet	pie	ladder	truck	bell
32	HA05	0.32	jacket	shoe	elephant	hammer	cucumber	cellphone
33	H06	0.12	mitten	scarf	telescope	lightbulb	faucet	crab
34	YS06	0.14	muffin	donut	teepee	coat	frog	lollipop
35	HA05	0.2	mushroom	peanut	skirt	binoculars	flag	tire
36	HA05	0.2	nose	leg	drum	candle	egg	ribbon

37	HA05	0.49	ostrich	eagle	wallet	key	crayon	leaf
38	new	0.33	pants	boot	cat	watermelon	lion	magnifying glass
39	HA05	0.24	peacock	owl	celery	shirt	glasses	basket
40	HA05	0.18	pear	orange	harmonica	ant	mask	grasshopper
41	HA05	0.62	piano	trumpet	muffin	robot	lemon	compass
42	HA05	0.34	rabbit	pig	balloon	guitar	skis	caterpillar
43	H06	0.12	rattle	kite	mouse	peanut	onion	dice
44	HA05	0.28	saw	ladder	tape	bicycle	eraser	globe
45	HA05	0.67	shirt	hat	carrot	owl	chain	bread
46	HA05	0.43	skirt	coat	whistle	glue	pizza	baby
47	HA05	0.16	teapot	pan	buffalo	scarf	wave	crown
48	new	0.43	teepee	buffalo	corn	hat	wrench	snake
49	new	0.4	train	wagon	nose	boot	pencil	corkscrew
50	HA05	0.53	violin	drum	scissors	tomato	ruler	octopus
51	H06	0.24	whistle	harmonica	desk	penguin	raccoon	star

Table A1. Complete stimulus set for Experiment 1. Critical item pair source is listed in the “Source” column (YS06=Yee & Sedivy, 2006; HA05=Huettig & Altmann, 2005; H06=Huettig, et al., 2006). The semantic similarity of the critical target-competitor pair based on the LSA analysis is provided. Note that target and competitor words are separately paired with other item sets to serve as unrelated items (see first two columns of “other” items). The final two columns of “other” serve as targets on filler trials.

Experiment 2 Stimulus Norming

Items from Nozari, et al., (2016) were adapted so that each participant could be exposed to the same set of stimuli multiple times without generating expectations for the target item. A pair of norming studies were used to create 29 sets of items. Each item set had 4 pictures, where each picture was the target twice, once with a restrictive verb and once with a non-restrictive verb.

Participants in the norming studies were workers on Mechanical Turk. Two groups of 54 workers participated in exchange for \$1.50 or \$0.75 (group 1 and 2, respectively). They saw a series of 280 trials (group 1) or 102 trials (group 2) where

they viewed a scene with four pictures and were asked to pick the best completion to a sentence fragment, e.g. "She will tune the..." (Figure A1).

Figure A1. Example display for the norming studies for Experiment 2.

Responses to the questions were used to select two verbs that could be used to reference each item in the context, one verb which generated high rates of the target selection, and one which did not create an expectation for the target. The full list of stimuli, including the norming data is available *<link to raw data and stimuli to be included in final published version>*.

Experiment 2 Supplemental Analysis

This supplemental analysis was conducted to examine eye-gaze prior to the onset of the noun in the test sentences, e.g. prior to the onset of "deer" in "*She will hunt the deer*". Across all of our stimuli, the minimum time between verb onset and noun onset was 619ms (average 939ms, max 1156ms). In this supplemental analysis, then, we

analyze fixations beginning at verb onset +180ms to verb onset +800ms, which, taking into account a ~200ms delay to move the eyes should capture fixations made following the verb but prior to the noun. One caveat, however, is that coarticulatory information may be present in the stimuli prior to the noun.

The results of this analysis (Tables A2 and A3) were similar to the primary analyses presented in the paper. For young adults, an effect of condition ($z = 11.21$) was due to more target fixations in the Restrictive vs. the Non-restrictive condition. For patients with amnesia (coded as the reference level in this analysis), the condition effect was significant ($z = 4.60$), and the group by condition interaction was not significant ($z = 1.73$).

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-4.686	0.096	-48.66	<.0001
AR(1)	9.274	0.134	69.26	<.0001
Time	0.188	0.013	13.84	<.0001
Condition (restrictive = 0.5; non-restrictive = -0.5)	0.850	0.076	11.21	<.0001
Random Effects	Variance	Std.Dev.	Corr	
Trial (intercept)	1.05E-09	3.24E-05		
Item (intercept)	4.43E-02	2.11E-01		
Subject (intercept)	1.20E-01	3.46E-01		
AR(1) slope	2.39E-01	4.88E-01	-0.71	

Table A2. Experiment 2: Analysis of data prior to noun onset. Results of dynamic GLMM for young adults ($N=16$), 177 trials, 59 items (verbs) and 175,584 observations. Bolded values indicate significance at p -value $<.05$.

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	-5.072	0.137	-36.918	<.0001
Condition (restrictive = .5, non-restrictive = -.5)	0.591	0.129	4.604	<.0001
Group (Amnesia = 0, Healthy Comparison = 1)	0.055	0.119	0.459	0.646
Test Block	0.029	0.025	1.128	0.259
AR(1)	9.650	0.140	68.813	<.0001
time	0.189	0.019	10.056	<.0001
Condition*Group	0.250	0.145	1.727	0.084
Condition*Test Block	0.004	0.043	0.103	0.918
Group*Test Block	0.034	0.028	1.219	0.223
Test Block*AR(1)	-0.073	0.027	-2.687	0.007
Test Block*time	0.000	0.007	-0.001	0.999
Condition*Group*Test Block	-0.046	0.053	-0.882	0.378
Random Effects	Variance	Std.Dev.	Corr	
Trial (intercept)	0.004	0.059		
Test Block (slope)	0.001	0.035	-0.59	
Item (intercept)	0.032	0.179		
Participant (intercept)	0.150	0.388		
Test Block (slope)	0.000	0.020	-0.71	
AR(1) (slope)	0.213	0.462	-0.90	0.33

Table A3. Experiment 2: Analysis of data prior to noun onset. Results of dynamic GLMM for patients with amnesia ($N=5$) and healthy comparison participants ($N=10$), 177 trials, 59 items (verbs) and 636,492 observations. Bolded values indicate significance at p -value <.05.