JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
2007, i*First*, 1–21

Psychology Press
Taylor & Francis Group

Amnesic H.M.'s performance on the language competence test: Parallel deficits in memory and sentence production

Donald G. MacKay,[1] Lori E. James,[2] and Christopher B. Hadley[1]

[1]University of California, Los Angeles, CA, USA
[2]University of Colorado, Colorado Springs, CO, USA

To test conflicting hypotheses regarding amnesic H.M.'s language abilities, this study examined H.M.'s sentence production on the Language Competence Test (Wiig & Secord, 1988). The task for H.M. and 8 education-, age-, and IQ-matched controls was to describe pictures using a single grammatical sentence containing prespecified target words. The results indicated selective deficits in H.M.'s picture descriptions: H.M. produced fewer single grammatical sentences, included fewer target words, and described the pictures less completely and accurately than did the controls. However, H.M.'s deficits diminished with repeated processing of unfamiliar stimuli and disappeared for familiar stimuli—effects that help explain why other researchers have concluded that H.M.'s language production is intact. Besides resolving the conflicting hypotheses, present results replicated other well-controlled sentence production results and indicated that H.M.'s language and memory exhibit parallel deficits and sparing. Present results comport in detail with binding theory but pose problems for current systems theories of H.M.'s condition.

The behavior of a single individual has greatly influenced psychological theories since 1953. That individual is H.M., one of the most studied patients in the history of neuropsychology (Ogden & Corkin, 1991) and "a touchstone for research on amnesia and memory systems" (Manns, 2004, p. 411). In 1953, H.M. underwent bilaterally symmetric suction surgery that removed his amygdala and parts of his hippocampus and connected medial temporal lobe (MTL) structures, but virtually none of his neocortex (see Scoville & Milner, 1957).[1] Although largely successful in treating his epilepsy, H.M.'s 1953 operation had several tragic side effects. One is well established and uncontroversial in the field at large: H.M. has a *selective* memory deficit that impairs recall of new or never previously encountered information, but spares his ability to recall semantic information encountered frequently before and after his operation (see e.g., Gabrieli, Cohen & Corkin, 1988; James & MacKay, 2001).

The present article examines the more controversial hypothesis of Milner, Corkin, and Teuber (1968) that language-linked processes are intact

[1]Recent magnetic resonance imaging (MRI) data (Corkin, Amaral, González, Johnson, & Hyman, 1997) indicate a large cerebellar lesion due to H.M.'s use of dilantin for treating epilepsy since 1953, together with possible damage to lateral temporal neocortex that was not due to the original surgery. This possible but at most minimal damage may reflect either an age-linked effect or occurrence of transneuronal degeneration subsequent to H.M.'s 1953 surgical lesion.

The authors gratefully acknowledge support from the Samuel A. MacKay Memorial Research Fund and NIH Grant R01 AG 09755. We thank Dr. Suzanne Corkin for permission to study H.M., Kataryna Boese, Alicia Chang, Meghan Gould, Jenifer Taylor, and Diane Marian for general research assistance, Kethera Folger for running the preliminary rating study, and Pamela Crombie and Irwin Stein for providing control participants from the UCLA Cognition and Aging Laboratory and Dr. D. Burke for providing control participants from the Claremont Project on Memory and Aging. Finally, we thank an anonymous reviewer for suggesting "processing difficulty" as an alternate account of the present data (see the Discussion section).

Address correspondence to Donald G. MacKay, Psychology Dept., University of California Los Angeles, Los Angeles, CA 90095–1563, USA (E-mail: mackay@ucla.edu).

http://www.psypress.com/jcen DOI: 10.1080/13803390701380609

in H.M. relative to controls. If correct for all language-linked processes, this "intact-language hypothesis" carries important implications for two competing theoretical frameworks in widespread use in the cognitive and brain sciences (see e.g., MacKay, James, Taylor, & Marian, 2007).

Systems theories and the intact-language hypothesis

The intact-language hypothesis comports with current systems theories, where independent systems process language versus memory, and the hippocampus only subserves memory systems (see e.g., Schmolck, Stefanacci, & Squire, 2000). Under systems theories, a language system comprehends words and sentences and transmits the products of comprehension to an entirely separate system for long-term memory storage. A retrieval system later retrieves stored memories for transmission to the language production system, which enables verbal expression of the retrieved memory. If conclusive, evidence for the intact-language hypothesis therefore supports systems theory by dissociating the storage system (damaged in H.M.) from the retrieval, comprehension, and production systems for producing words and sentences (by hypothesis undamaged in H.M.).

Over the past thirty years, only two nonanecdotal reports have concluded in favor of systems theory and the intact-language hypothesis: Kensinger, Ullman, and Corkin (2001) concluded that H.M.'s language processing was unimpaired at grammatical levels in 1999–2000; and Skotko, Andrews, and Einstein (2005) concluded that H.M.'s spoken sentence production was "remarkably competent . . . . for his age and educational cohort" (p. 406–407).

However, evidence against the intact-language hypothesis is strong and extensive (and currently ignored in all major secondary sources that discuss H.M.). Evidence gathered from 1967–1999 in 15 experiments indicates selective deficits in H.M.'s language comprehension: impaired comprehension of new, noncliché, or never previously encountered phrases and sentences, but spared comprehension of phrases encountered frequently before and after his operation (see MacKay, Burke, & Stewart, 1998a; MacKay, Stewart, & Burke, 1998b; MacKay et al., 2007). Several other studies have provided corroborating evidence for deficits in H.M.'s language comprehension (see Corkin, 1984; Lackner, 1974; Schmolck et al., 2000) and for language deficits in other patients with hippocampal-MTL damage (Zaidel, Zaidel, Oxbury, & Oxbury, 1995).

Evidence gathered from 1970–1999 in 11 experiments also indicates that H.M. exhibits deficits relative to controls in spoken production of novel phrases and sentences, but spared production of phrases encountered frequently before and after his operation (MacKay et al., 1998a; MacKay & James, 2001, 2002; MacKay et al., 2007; and MacKay et al., 1998b). H.M.'s sentence production performance on the "reporter's" test of language production also indicates deficits relative to age-linked norms (see Corkin, 1984).

The intact-language hypothesis, binding theory, and H.M.'s sentence production

From a theoretical perspective, three aspects of the just-reviewed evidence are noteworthy. First, H.M.'s deficits in language comprehension and production are difficult to explain under the intact-language hypothesis and current versions of systems theory. Second, H.M.'s language deficits are selective and precisely parallel the nature of his memory deficits noted earlier: impaired processing of unfamiliar information but spared processing of information familiar to H.M. before his lesion. Third, binding theory predicted the parallel deficits and sparing in H.M.'s language and other aspects of cognition observed in MacKay et al. (1998a), MacKay and James (2001, 2002), MacKay et al. (2007), and MacKay et al. (1998b).

As applied to H.M.'s comprehension, production, and memory for phrases and sentences (see e.g., MacKay et al., 2007), binding theory predicts impaired processing for the initial encounter with unfamiliar information, including noncliché propositions and never previously encountered phrases. However, binding theory predicts spared processing of high frequency (HF) information familiar to H.M. before his lesion and spared "engrainment" processes that facilitate performance with unfamiliar information following massive repetition over many trials, encounters, or covert rehearsals. Results involving HF information familiar to H.M. before his lesion and unfamiliar information that is massively repeated are therefore consistent with both binding theory and the intact-language hypothesis. We return to this point in the concluding section of the present paper.

This study attempts to resolve the conflicting conclusions of Kensinger et al. (2001) and Skotko et al. (2005) versus MacKay et al. (1998a), MacKay and James (2001, 2002), MacKay et al. (2007), and MacKay et al. (1998b). Replication, extension, and cross-checking are the standard scientific procedures for resolving conflicting conclusions, especially when they carry important theoretical implications. The present study therefore reexamined H.M.'s sentence production and its implications for systems theory and binding theory in a standard sentence production paradigm never previously examined with H.M.

Under binding theory, H.M.'s hippocampal-MTL damage has impaired the subcortical binding mechanisms necessary to rapidly form new connections for creating never previously activated nodes or internal representations in the cortex, but has not impaired the mechanisms for activating already existing nodes (see MacKay et al., 2007, and MacKay & James, 2001, 2002, for detailed theoretical mechanisms). As a consequence, H.M. can produce frequently used propositions—for example, "that's wrong" and "I'll have some"—and frequently used phrases such as "to be" and "the same way" with proper word order because his syntax-based activation mechanisms for activating these units are intact and have been used frequently since his lesion. Similarly, H.M. can activate the already existing cortical representations for familiar words and morphemes without jumbling them together into morphological salads in the manner of agrammatic aphasics (Jackendoff, 2003, p. 264). However, H.M. lacks already formed cortical representations for producing propositions and phrases not used repeatedly before and after his lesion, an essential aspect of describing the unfamiliar word–picture stimuli in the standardized test for the present experiment. For example, one of our stimuli required use of the words *fresh, nor*, and *here* to describe two people conversing in a bakery displaying pies and bread. For this unfamiliar word–picture stimulus, the sentence, "The pie here looks neither fresh nor good" represents a model description. Under binding theory, producing such a sentence requires the formation of new connections to a node for representing this proposition and its component phrases: "neither fresh nor good," "the pie here," and "looks neither fresh nor good." Because H.M. lacks an already-formed node for representing this proposition and its connected phrase nodes, producing this sentence would be difficult for H.M. For example, only following massive repetition could H.M. bind or connect the words *fresh, good, nor*, and *neither* to a new phrase node for representing and producing the adjectival phrase "neither fresh nor good," an essential component for producing the sentence, "The pie here looks neither fresh nor good."

Our task was identical to Subtest 3 of the Test of Language Competence (TLC; Wiig & Secord, 1988): to describe pictures using prespecified lexical constraints (use of prespecified target words with clear relations to the picture), syntactic constraints (production of a single grammatical sentence), and referential constraints (accurate picture description). Our participants were H.M. and 8 memory-normal controls. The word–picture stimuli differed from the TLC stimuli in only two minor ways. First, we ensured that all of our target words were HF and entered English before 1953. This eliminated lexical unfamiliarity as a factor in our data because H.M. reads, understands, and produces without deficit isolated HF words familiar to him before his lesion (see James & MacKay, 2001; MacKay & James, 2001, 2002). Second, we only used TLC pictures that depicted objects, activities, and environments familiar to H.M. since young adulthood. This eliminated unfamiliarity with the pictured objects and actions as a factor in our results.

Our study had two goals. Goal 1 was to cross-check the Skotko et al. (2005) support for the intact-language hypothesis and resolve the conflicting claims related to H.M.'s language production. Goal 2 was to test three predictions derived from binding theory: that H.M.'s language production will be impaired relative to controls for unfamiliar stimuli and responses; that H.M.'s performance will be unimpaired for stimuli and responses familiar since early childhood; and that H.M.'s performance will improve as a function of repetition.

We manipulated three factors to test these predictions. One was *linguistic–situational familiarity*: For one set of word–picture stimuli, the situations and target words together depicted novel or unfamiliar scenarios with no links to familiar or cliché expressions that could accurately characterize these scenarios (see the example discussed earlier). For these stimuli, new binding or connection formation was in principle necessary to internally represent the novel relations between objects or events in the pictured situation and to create sentences using prespecified words to characterize this internal representation under binding theory. Because H.M.'s lesion has impaired his ability to form new connections to represent new situational relations and new phrase- and sentence-level verbal concepts, binding theory therefore predicted that H.M.'s descriptions of unfamiliar scenarios will incorporate fewer target words and will be ungrammatical more often than will controls' descriptions.

For a second set of word–picture stimuli, the situations and target words together were familiar: The situations depicted scenarios familiar since early childhood, and links between the target words and cliché phrases could accurately characterize the pictures. To illustrate one such familiar stimulus, a picture with the target words *before, first*, and *across* depicts a small boy, age about four years old, holding his father's hand on a sidewalk at the corner of an intersection. The traffic light before them reads, "Don't walk," and the boy is gazing up at his father and listening attentively to what his father is saying. This word–picture stimulus has high situational–linguistic familiarity because the scenario is familiar since childhood, and the target words *before, first*, and *across* participate in familiar clichés that could accurately characterize the picture—for example, "Before you walk across the street, you first have to wait for the light to turn green." Under binding theory, no new binding or connection formation is required to internally represent such a familiar scenario involving a parent and child at a crosswalk. Nor is new connection formation required to represent and produce clichés about what to do *first before* proceeding *across* a street at a red light (target words in italics).

To illustrate another familiar word–picture stimulus, the target words *sit, painted*, and *because* headed a picture of a boy gesturing toward a "wet paint" sign on a bench and saying something to a woman who is about to sit on the bench. This word–picture stimulus has high situational–linguistic familiarity because the target words *sit, painted*, and *because* participate in familiar clichés that could accurately characterize this familiar scenario. Example clichés are "Don't sit there," "just painted," and "it's just been painted." Under binding theory, no further binding or connection formation is required to add the target word *because* to configure these familiar clichés into a sentence such as "Don't *sit* there *because* it's just been *painted*" (target words in italics). Under binding theory H.M. will therefore accurately describe our familiar stimuli without deficit because he has no difficulty activating familiar words and clichés under binding theory.

Factor 2 was the level of lexical constraint on the picture descriptions: two versus three target words per picture. The main goal of this manipulation was to determine whether load (the number of target words occupying working memory) represents a primary determinant of H.M.'s performance relative to controls on the TLC.

Factor 3 involved repeated processing of input-side information and output-side information. To manipulate input-side repetition, H.M. described each picture twice on separate days, with two target words accompanying one presentation of the picture and the same two words (plus one additional word) accompanying the other. Because the controls described each picture and its target words only once (with number of target words per picture counterbalanced across participants), binding theory predicted that input-side repetition will improve H.M.'s performance relative to controls due to the engrainment learning that occurs with repeated processing of the identical target words and pictures.

To manipulate output-side repetition, the experimenter called for repeated attempts to achieve perfect target word performance. When H.M.'s initial and subsequent descriptions omitted some of the target words, the experimenter repeatedly drew H.M.'s attention to the target words and asked him to try again (up to seven times). Due to the engrainment learning that occurs with repeated output-side processing, binding theory predicted improvement in H.M.'s performance relative to controls as a function of output-side repetition.

METHOD

Participants

Table 1 shows the age, highest educational degree, and the Verbal and Performance IQ scores on the Wechsler Bellevue, Form 1 (W-B I) for the participants: H.M. and 8 memory-normal controls.

H.M.

We tested H.M. in 1999 when he was 72 years old and when his mean IQ score on the Verbal and Performance subtests of the W-B I was 112 (see Table 1).

Controls

We selected the 8 controls from more than 750 older adults in the participant pools of the UCLA Cognition and Aging Laboratory and the Claremont Project on Memory and Aging. Our criterion for selecting controls was to match H.M. as closely as possible on four factors: highest educational degree, mean age at time of test, mean IQ score on the Verbal and Performance subtests of the W-B I, and native language (all spoke English as children). We tested the controls from 1999–2003 when their mean age was 70, and their mean IQ score on the W-B I Verbal and Performance subtests was 113 (see Table 1). All controls reported an absence of neurological problems and participated for $10/hr.

Materials

The materials consisted of 22 word–picture stimuli: 20 experimental stimuli and 2 practice stimuli. The pictures were from the TLC (Subtest 3) but we omitted 8 of the 30 TLC pictures that depicted

TABLE 1
Age, verbal IQ, performance IQ, and highest educational degree for H.M. and control participants

Participants	*Age*[b]	*Verbal IQ (W-B I)*	*Performance IQ (W-B I)*	*Highest educational degree*
H.M.	72.5	107	117	High school
Control 5	70	117	130	High school
Control 21	77	115	104	High school
Control 31	77	129	113	High school
Control 51	71	99	101	High school not completed
Control 52	65	115	120	High school
Control 58	67	114	108	High school
Control 59	70	116	114	High school
Control 61	66	100	115	High school
Mean for controls[a]	70.38 (4.60)	113.13 (9.67)	113.13 (9.20)	High School

[a]Standard deviations in parentheses. [b]In years. W-B I = Wechsler Bellevue, Form 1.

concepts that would have been unfamiliar to H.M. prior to his operation—for example, a depiction of the inside of a music store with a record shelf labeled “Rock,” a post-1953 concept. We also substituted new words for TLC target words that entered English dictionaries after 1953. Finally, we visually modified the TLC stimuli to ensure against sensory-level misperception of the words and pictures: We increased the font size of the target words that headed each picture, and we xeroxed the new word–picture stimuli with enlargement onto separate sheets of paper.

All pictures featured two or more adults, teenagers, or children engaged in common activities—for example, students exiting a school and customers purchasing clothing. The target words for describing the pictures were either closed class words (e.g., *job, leg, hard*) or open class words (e.g., *or, and, some*) and were either HF or extremely HF. Table 2 shows the target words in the left-to-right order presented for each word–picture stimulus.

Experimental stimuli came in two versions labeled Level 1 and 2 in the TLC. The Level 1 version of a picture featured two target words for inclusion in the response, and the Level 2 version featured the same two words plus one additional word. The practice stimuli always contained two target words but otherwise resembled the experimental stimuli.

Preliminary rating study

To evaluate our stimuli, we presented the 20 experimental stimuli to 10 judges in a preliminary rating study. The judges were graduate and undergraduate students at the University of Colorado, Colorado Springs (2 male, 8 female; mean age 24.8 years).

The judges saw the word–picture stimuli in the same order as did participants in the main experiment and rated the stimuli on three dimensions. The first was relative frequency of scenario encounter: The judges rated how often in their lifetimes they had heard about or experienced the type of situation depicted in each picture using a 0–4 scale (0 = never before, 1 = once before, 2 = several times, 3 = once a year, 4 = more than once a year). Their mean ratings varied between 2.2 (several times) and 4.0 (more than once a year), indicating that H.M. is likely to have encountered the depicted situations from several times to more than 26 times prior to his age 26 lesion. This outcome suggests that H.M. and the controls were familiar with all 20 scenarios.

The second rating dimension focused on the relation between the pictures and individual target words: For each target word in turn, the judges used a 0–4 scale to rate the likelihood that they would use that word in describing the picture (0 = highly unlikely, 1 = somewhat unlikely, 2 = unlabeled, 3 = somewhat likely, 4 = highly likely).

The third rating dimension evaluated whether familiar or cliché expressions containing one or more of the target words could describe the depicted scenario: The judges rated the extent to which the target words for a picture brought to mind clichés or familiar expressions that could accurately characterize the picture. The scale was 0–3 (0 = no familiar expression including any of the words, 1 = one very familiar expression including one of the words, 2 = one very familiar expression for some but not all of the words, 3 = one or more very familiar expressions that included all of the words).

The mean ratings for clichés and likelihood of target word use served to categorize stimuli as familiar versus unfamiliar. We classified as familiar those stimuli with mean target word use ratings 3.5 or greater averaged across identical pictures in Sets 1 and 2 and mean cliché ratings 2.5 or greater

TABLE 2
The target words for the various stimuli with two versus three target words, with H.M.'s descriptions and trial number and a representative control description for comparison

H.M. trial number	*Mean target words per description (controls)*[a]	*Example control descriptions*	*Number of target words per description (H.M.)*	*I and F utterance of H.M.*
Practice stimuli				
		Set 1 Picture 1: Target words: *sad, moving*		Set 1 Picture 1: Target words: *sad, moving*
0	2 (0)	Control (I-F): Two people are looking at a moving truck and they're very sad because the neighbors are going away.	1	H.M. (I): Well . . . moving.
			1	H.M. (F): Mary's sad to see somebody move away.
		Set 2 Picture 1: Target words: *show, anyway*		Set 2 Picture 1: Target words: *show, anyway*
0	2 (0)	Control (I-F): Let's see the show anyway.	2	H.M. (I-F): And trying to buy tickets for the show anyway before they go in there.
Familiar experimental stimuli				
9		Set 1 Picture 10: Target words: *first, across*		Set 1 Picture 10: Target words: *first, across*
	1.75 (0.5)	Control (I-F): Let's see who can be first across the street.	1	H.M. (I-F): He wants to cross here . . . first.
19		Set 2 Picture 10: Target words: *before, first, across*		Set 2 Picture 10: Target words: *before, first, across*
	2.75 (0.5)	Control (I-F): And the man is telling the little boy that he must look first before he crosses the street.	3	H.M. (I): Before at first you cross across.
			2	H.M. (F): Before you cross the street you have to look both ways first.
12		Set 2 Picture 3: Target words: *sit, painted*		Set 2 Picture 3: Target words: *sit, painted*
	2 (0)	Control (I-F): Ooh, don't sit down, it's still painted.	1	H.M. (I-F): And that man is trying to tell that woman not to sit there because it's wet paint. (Exp.: Good.) He can uh see the sign better than she could and she's ready to sit down there.
2		Set 1 Picture 3: Target words: *sit, painted, because*		Set 1 Picture 3: Target words: *sit, painted, because*
	3 (0)	Control (I-F): Oh . . . he told her she shouldn't sit on the bench because it was just painted.	3	H.M. (I-F): Oh, don't sit because it's just been painted.
Unfamiliar two-word experimental stimuli				
1		Set 1 Picture 2: Target words: *fall, leg*		Set 1 Picture 2: Target words: *fall, leg*
	2 (0)	Control (I-F): If I fall and break my leg that's going, not going to be good.	0	H.M. (I): Seeing how somebody's climbing that mountain, they are discussing it themselves cause (unintelligible) stuff he should take.
			1	H.M. (F): If they don't use legs like he does . . . and his hands, they could fall.
3		Set 1 Picture 4: Target words: *some, and*		Set 1 Picture 4: Target words: *some, and*

(Continued)

TABLE 2
(Continued)

H.M. trial number	*Mean target words per description (controls)*[a]	*Example control descriptions*	*Number of target words per description (H.M.)*	*I and F utterance of H.M.*
	2 (0)	Control (I-F): I would like some of that ice cream and some of the, that ice cream.	1	H.M. (I-F): I like some her . . . what she had.
5		Set 1 Picture 6: Target words: *but, job*		Set 1 Picture 6: Target words: *but, job*
	1.75 (0.5)	Control (I-F): He's done a good job but it's not easy.	0	H.M. (I): I . . . she wants the house painted the same as him and he wants to mow the lawn.
			2	H.M. (F): I want that job . . . and . . . but *she* says, *he* gotta do the other part first.
7		Set 1 Picture 8: Target words: *because, hard*		Set 1 Picture 8: Target words: *because, hara*
	1.75 (0.5)	Control (I-F): Because this will get you in good shape, it will be hard but you'll like it.	0	H.M. (I): I want to exercise like these two are.
			2	H.M. (F): Because it's too hard to do it that way (Exp.: Okay.) [unintelligible]).
14		Set 2 Picture 5: Target words: *pie, either*		Set 2 Picture 5: Target words: *pie, either*
	2 (0)	Control (I-F): You can either have pie or cake.	1	H.M. (I): Since they've got their coffee already he isn't—they just want their uh *pie* and the piece of this *pie* up *here* because the cake is down *here*.
			1	H.M. (F): Well this *pie* is—or the *pie here* was back *here* (Exp.: uh-huh) and uh coffee is in there because heat a solid and this is only boiled milk say milk there and *this* is not liquid but *only* ice.
16		Set 2 Picture 7: Target words: *crowded, drive*		Set 2 Picture 7: Target words: *crowded, drive*
	2 (0)	Control (I-F): Oh, the school bus is so crowded, I'll just drive.	1	H.M. (I): A driving wanna drive some place and this bus is stopped up there.
			1	H.M. (F): Is it crowded and it just pointed out this bus is up here and it's crowded school bus.
18		Set 2 Picture 9: Target words: *although, wrong*		Set 2 Picture 9: Target words: *although, wrong*
	2 (0)	Control (I-F): Well, I think I'll take that one although it looks wrong.	0	H.M. (I): Well she's choosing the soup here—for him.
			1	H.M. (F): Yes. Because it's wrong for her to be and he's dressed just as this that he's dressed and the same way.
20		Set 2 Picture 11: Target words: *fresh, nor*		Set 2 Picture 11: Target words: *fresh, nor*

(Continued)

TABLE 2
(Continued)

H.M. trial number	*Mean target words per description (controls)*[a]	*Example control descriptions*	*Number of target words per description (H.M.)*	*I and F utterance of H.M.*
	1.75 (0.5)	Control (I-F): That pie doesn't look fresh, nor does it look good.	0	H.M. (I): Well you—she wants one thing and he wants another thing and the fresh are not—are not. Doesn't say that,it says not.
			0	H.M. (F): Yeah. Or she could say this. This is in [unintelligible] over here and this is just little things [unintelligible] a little spice you could call eclairs and stuff like that it's over here.
Unfamiliar three-word experimental stimuli				
4		Set 1 Picture 5: Target words: *pie, either, have*		Set 1 Picture 5: Target words: *pie, either, have*
	2.75 (1)	Control (I-F): Uh, there are two people getting pie, but there's only one piece of blueberry pie left, and so, either one of them will have to have it.	3	H.M. (I): I want some of that pie either some pie and I'll have some.
			2	H.M. (F): Any pie that either she either had.
6		Set 1 Picture 7: Target words: *crowded, drive, if*		Set 1 Picture 7: Target words: *crowded, drive, if*
	2.5 (0.58)	Control (I-F): If I had to drive that bus, I wouldn't have it that crowded.	1	H.M. (I): Melanie tra . . . on that bus, the *scrawny* bus and have it drive it off . . . it, it drives it off.
			1	H.M. (F): Well he has to go the same way as her if [unintelligible] . . . she wants to go on the bus . . . and it's crowded . . . it's crowded. (Exp.: Okay.) Too crowded to get on the bus. (Exp.: Okay.) [unintelligible] . . . one way out, it's on common street.
8		Set 1 Picture 9: Target words: *actually, although, wrong*		Set 1 Picture 9: Target words: *actually, although, wrong*
	2.25 (0.96)	Control (I-F): Actually I like that suit although I know it is wrong for me.	0	H.M. (I): She's taking that suit and he wants to take it . . . and he's trying to sell.
			2	H.M. (F): Actually it's best for him. It's wrong for her. They have 'em the same way.
10		Set 1 Picture 11: Target words: *fresh, nor, here*		Set 1 Picture 11: Target words: *fresh, nor, here*
	3 (0)	Control (I-F): Mmm . . . let's see . . . they have . . . they have fresh food here . . . they don't have fresh food here nor do they have anything that's really good.	1	H.M. (I): Once has to be trash in yellow [unintelligible] . . . is not here.
11		Set 2 Picture 2: Target words: *fall, leg, and*		Set 2 Picture 2: Target words: *fall, leg, and*

(Continued)

TABLE 2
(Continued)

H.M. trial number	*Mean target words per description (controls)*[a]	*Example control descriptions*	*Number of target words per description (H.M.)*	*I and F utterance of H.M.*
			0	H.M. (F): Gary is . . . almos . . . almost . . . hasn't been cut the same way. And his (unintelligible) just what they are there.
	3 (0)	Control (I-F): We hope he doesn't fall and break a leg.	2	H.M. (I): He's climbing that and he can fall.
			1	H.M. (F): And he has to use his legs to cli– climb.
13		Set 2 Picture 4: Target words: *some, and, get*		Set 2 Picture 4: Target words: *some, and, get*
	3 (0)	Control (I-F): Some have to come to the counter and get the food.	1	H.M. (I): Well he's putting the price of it and price of thing what it is and she wants to [unintelligible] in there and he's waitin' to be waited on.
			2	H.M. (F): He gettin' some of this and isn't plain [unintelligible] what it is and he is just waiting to get waited on.
15		Set 2 Picture 6: Target words: *job, but, easy*		Set 2 Picture 6: Target words: *job, but, easy*
	2.75 (0.5)	Control (I-F): It won't be easy but we need to do a thorough job.	2	H.M. (I-F): It is easy to paint the place even though it's been just a job and easy on the job part.
17		Set 2 Picture 8: Target words: *because, hard, like*		Set 2 Picture 8: Target words: *because, hard, like*
	2.75 (0.5)	Control (I-F): Because this is so hard, I don't like to do it.	1	H.M. (I-F): 'Cause he's doin' that and this one liked to do it this way to sit down. (Exp.: Um-hum.) And this could be hard here and soft here.

Note. I indicates H.M.'s initial utterance, F indicates his final utterance, and I-F indicates that his I utterance was his F utterance (see text for details).
[a]Standard deviations in parentheses.

averaged across identical pictures in Sets 1 and 2. The mean rating for target word use was 3.65 (*SD* = .12) for familiar stimuli (*N* = 4), indicating high to somewhat high likelihood of using each of the target words in describing these pictures. The mean cliché rating was 2.55 (*SD* = .30) for the familiar stimuli, indicating that these word–picture stimuli brought to mind an average of one or more very familiar expressions or clichés that included all of the target words and accurately characterized the pictures. The familiar stimuli are labeled in Table 2 (see e.g., Set 2 Picture 3: target words *sit, painted*).

We classified as unfamiliar those stimuli with mean cliché ratings less than 2.5 averaged across the identical pictures in Sets 1 and 2 and mean ratings for target word use less than 3.5 averaged across the identical pictures in Sets 1 and 2. The mean target word use rating for the unfamiliar stimuli (*N* = 16) was 3.26 (*SD* = 0.34), indicating somewhat likely use of each target word in describing these pictures. The mean cliché rating for the unfamiliar stimuli was 1.99 (*SD* = 0.38), indicating that these stimuli brought to mind only 0–1 familiar expressions or clichés that included some but not all of the target words for characterizing the pictures. The 16 unfamiliar stimuli are labeled in Table 2 (see e.g., Set 1 Picture 2: target words *fall, leg*).

Procedure

The experimenter first presented the instructions orally and then visually on a continuously displayed card: Make up a single grammatical sentence about each picture using the words at the top of the picture. To present each stimulus, the experimenter turned over the stimulus page, saying "OK" as soon as the participant could see the picture. Mean serial position in the series of pictures was equivalent for the familiar and unfamiliar stimuli, and the pictures were presented in the same order for H.M. and the controls.

The experimenter repeated the basic instructions orally to H.M. before presenting each stimulus and also during many of the trials. When H.M. failed to include one or more target words in his initial response, the experimenter asked H.M. to read the target words aloud or otherwise drew the words to H.M.'s attention and asked him to try again (up to seven times). Besides allowing engrainment learning (see the Introduction), this procedure served to ensure that H.M. was following the instructions, maintained continuous focus on the task, and received multiple opportunities on each trial to successfully complete the task.

H.M. saw the same picture twice, at Level 1 (two target words) on one day and at Level 2 (three target words) on another day, with a random half of the pictures occurring at Level 1 on Day 1 and at Level 2 on Day 2, and vice versa for the remaining pictures. We labeled stimuli for Day 1 as Set 1, and stimuli for Day 2 were labeled Set 2 (see Table 2).

Unlike H.M., the controls received no reminders of the instructions and only a single opportunity per stimulus to complete the task. Also unlike H.M., the controls saw either stimulus Set 1 or Set 2 on a single day, so that each picture was presented only once, with stimulus level counterbalanced across control participants. This counterbalancing procedure helped to avoid both memory-linked advantages and ceiling-level performance for the controls, outcomes that would render comparisons with H.M. problematic.

Responses were tape-recorded for subsequent transcription. Several primary listeners provided initial word-for-word transcriptions of the tapes for H.M. and the controls, using the label "???" when confidence in their transcription of a particular word was less than 100% and the label "unintelligible" when they could not decipher one or more words after repeated replay. A secondary listener then listened to segments of the tapes containing words labeled "???" or "unintelligible" and either agreed with the initial labels or asked a different primary listener to provide a third transcription. In the very few cases (N = 2) where all three listeners disagreed, the conservative label "unintelligible" was adopted in the final transcript (see Table 2).[2]

RESULTS

For all analyses, we defined a description or utterance as a string of words bounded by trial onset, trial offset, or a substantive comment from the experimenter—for example, a request to try again. Word strings containing unprompted self-corrections and nonsubstantive experimenter interjections—for example, "Mmm," "Uhuh," "Okay," or "Mmm-hm," were not counted as separate utterances. Table 2 shows for the practice stimuli and each familiar versus unfamiliar stimulus the full initial and final descriptions or utterances of H.M. together with the initial utterance of a typical control (because controls produced only one description, their initial and final descriptions were always identical). The typical control varied from trial to trial under our typicality criterion—that is, that control participant whose description included target words closest in number to the mean for all of the controls.

The initial utterances in Table 2 were the focus of our main analyses and illustrate our results in qualitative form. Note in Table 2 that with experimenter prompts, H.M. incorporated all of the target words for the Set 1 and Set 2 practice trials (albeit not always within a single sentence), indicating that H.M. was both motivated and able to follow the instructions prior to the experiment proper. Note also that H.M. performed well in describing the familiar stimuli in Table 2. For example, H.M. produced a single grammatical sentence containing all three target words (*sit, because*, and *painted*) to describe the familiar situation of what to say when someone is about to sit on fresh paint: "Oh, don't sit because it's just been painted." However, note that H.M. performed much worse in describing the unfamiliar stimuli, producing initial utterances that included fewer target words than the controls (see the means in Table 2). Note also that H.M.'s utterances were often ungrammatical, incoherent, and difficult to understand, as Initial Utterances 1–3 illustrate (see Table 2 for H.M.'s complete descriptions).

[2]All instances of the label "unintelligible" (N = 25) occurred for H.M., and two factors made H.M.'s speech difficult to decipher: incoherent and difficult to understand syntax (discussed shortly) and poor articulatory quality (undoubtedly related to H.M.'s cerebellar lesion noted earlier).

1. H.M.: “I like some her . . . what she had” (ungrammatical, incoherent, and difficult to understand utterance).
2. H.M.: “A driving wanna drive some place and this bus is stopped up there” (ungrammatical, incoherent, and difficult to understand utterance).
3. H.M.: “Well he’s putting the price of it and price of thing what it is” (ungrammatical, incoherent, and difficult to understand utterance).

Overall performance on the TLC

We examined overall performance on the TLC via three independent measures: target word performance, syntactic constraint performance, and referential constraint performance.

Target word performance

Table 2 shows the number of target words that H.M. incorporated in his description of each stimulus, along with the mean and standard deviations (*SD*s) for controls. Only target words with the precise form presented counted as correct, and our measure of overall target word performance was the mean number of target words that participants included across all 20 stimuli. H.M included a mean of 1.13 target words versus 2.30 for controls (SD = 0.43), a difference of 3.63 standard deviations. Under the standard convention that differences of 2.0 standard deviations or more reflect deficits, H.M. had a deficit relative to controls in target word performance. A by-stimulus analysis reinforced this conclusion: Using a sign test with stimuli as unit of analysis and target words per description as dependent variable, H.M. incorporated reliably fewer target words than the mean for controls across the 20 word–picture stimuli, $p < .01$.

Syntactic constraint performance

Our index of syntactic constraint performance was how often participants described a stimulus using a single grammatical sentence. We did not count as ungrammatical instances of stuttering (immediate repetition of word-initial speech sounds and words), corrected word substitution, and phonological informality, e.g., *gotta* instead of *got to*, and *gettin’* instead of *getting*. However, we counted as ungrammatical utterances that were incomplete or lacked a subject, a main verb, or an object in the case of transitive verbs and prepositions, and that contained malapropisms, errors in pronoun use, nonsequiturs, run-on clauses, free-associative phrases, and conceptually redundant repetitions and major violations of selection restrictions or agreement rules. Examples 4–8 provide illustrative utterances containing these types of ungrammaticality in various combinations.

4. H.M.: “Because it’s wrong for her to be and he’s dressed just as this that he’s dressed and the same way . . .” (incomplete, run-on, and ungrammatical sentence containing conceptually redundant repetition).
5. H.M.: “I want some of that pie either some pie and I’ll have some” (run-on, incomplete, and ungrammatical sentence).
6. H.M.: “. . . and this is just little things (unintelligible) a little spice you could call eclairs and stuff like that it’s over here” (ungrammatical sentence containing two word substitutions, a malapropism, and a major violation of verb–subject agreement).
7. H.M.: “Melanie tra . . . on that bus, the scrawny bus and have it drive it off . . . it, it drives it off” (incomplete and ungrammatical sentence containing conceptually redundant repetition and an error in pronoun use).
8. H.M.: “I . . . she wants the house painted the same as him and he wants to mow the lawn” (ungrammatical sentence with violation of an agreement rule).

H.M. produced single grammatical sentences in describing 20% of the stimuli, versus a mean of 70% for the controls (SD = 21), a difference of 2.38 standard deviations. H.M. therefore had a deficit relative to controls in syntactic constraint performance. A by-stimulus analysis reinforced this conclusion: Using a sign test with stimuli as unit of analysis, H.M.’s descriptions consisted of a single grammatical sentence reliably less often than the mean for controls across the 20 word–picture stimuli, $p < .01$. A by-stimulus analysis also indicated that H.M. produced reliably more ungrammatical descriptions (whether single sentences or not) than the mean for controls, $p < .01$, using a sign test with stimuli as unit of analysis.

Picture description accuracy

Our index of picture description accuracy was how often participants produced uncorrected referential errors in describing the pictures, regardless of whether the descriptions were complete, grammatical, relevant, coherent, or readily comprehended. Utterances 9–12 illustrate five such uncorrected referential errors (with explanations; see Table 2 for the target words):

9. H.M. (I, Set 2 Picture 5): "Since they've got their coffee already he isn't—they just want their uh pie and the piece of this pie up here because the cake is down here" (two uncorrected referential errors).

The picture for Utterance 9 shows a man and a woman at a cafeteria counter, with an apparently empty cup on the man's tray, and a piece of pie and a cup containing a (black) liquid on the woman's tray. The man is saying something to a female clerk behind the counter. Utterance 9 therefore contains two referential errors. One is that *they* don't already have *their* coffee because only the woman's cup contains coffee, and H.M. later describes the man's apparently empty cup as containing milk or ice (see Table 2). H.M.'s second referential error is that *they* don't want *their* pie: Although the man may be requesting pie, the woman is not requesting anything, and her tray already contains a piece of pie.

10. H.M. (I, Set 1, Picture 11): "Once has to be trash in yellow (unintelligible) . . . is not here" (uncorrected referential error).

Utterance 10 is inaccurate because the picture shows a clerk and two women in a bakery, but no sign of trash.

11. H.M. (I, Set 2, Picture 9): "Well she's choosing the soup here—for him" (uncorrected referential error, possibly phonological in origin).

Utterance 11 is inaccurate because the picture shows a man and a woman in a clothing store that contains suits but no soup.

12. H.M. (I, Set 2, Picture 4): "Well he's putting the price of it and price of thing what it is and she wants to in there and he's waitin' to be waited on" (uncorrected referential error).

Utterance 12 is inaccurate because the picture shows an ice cream parlor with two female customers and a male clerk or waiter who is waiting on the customers but is not "waitin' to be waited on."

H.M. produced nine uncorrected referential errors in describing the 20 pictures versus 0 for the controls (SD = 0), a difference in excess of 6 standard deviations.[3] H.M.'s initial utterances therefore exhibited a reliable deficit relative to the controls in referential constraint performance.

Conceptual completeness of the descriptions

To evaluate how completely the participants described the central concepts in the pictured scenario, the 10 judges in the preliminary rating study used a 0–3 scale (3 = as complete as possible for a single sentence, 2 = reasonably complete, 1 = somewhat incomplete, and 0 = extremely incomplete) to rate the 40 initial utterances in Table 2. Under each word–picture stimulus appeared (without labels as to speaker identity) H.M.'s description and the description of the typical control (see Table 2). Order of the two utterances was counterbalanced across judges, and order of the 20 word–picture stimuli was the same as that in the main experiment. Instructions informed the judges that the speakers were attempting to accurately describe the picture using a single grammatical sentence. However, the judges were to ignore the accuracy, grammaticality, coherence, and comprehensibility of the descriptions in forming their ratings of conceptual completeness.

The mean completeness rating for control descriptions was 2.43 (SD = 0.44) versus 1.09 for H.M., a deficit of 3.04 standard deviations relative to the controls. This difference was also reliable in a by-stimulus analysis using a sign test with stimuli as unit of analysis and the mean completeness rating for each description as dependent variable: The judges rated H.M.'s descriptions as less complete than the typical control descriptions for all 20 word–picture stimuli, $p < .01$.

Effects of situational–linguistic familiarity on H.M.'s target word inclusion

Table 3 shows mean target word inclusion in the initial descriptions of H.M. and the controls for the familiar versus unfamiliar stimuli. For the familiar stimuli, H.M incorporated a mean of 2.00 target words versus 2.38 for controls (SD = 0.60), a nonreliable difference of 0.63 standard deviations. For the unfamiliar stimuli, H.M. incorporated a mean of 1.75 target words versus 2.33 for controls (SD = 0.20), a reliable difference of 2.90 standard deviations. H.M. therefore had a deficit in target word performance relative to controls for unfamiliar but not familiar stimuli.

H.M.'s syntactic constraint performance relative to controls was also better for familiar than unfamiliar stimuli. In describing familiar stimuli, H.M. produced single grammatical sentences in describing 81% of the stimuli, versus a mean of 75% for the

[3]By standard convention, the magnitude of infinitely large deficits (as can occur when the controls perform a task with SD = 0) is indicated conservatively as "in excess of 6 SDs."

TABLE 3
Mean target word inclusion for H.M. and the controls as a function of stimulus familiarity and lexical constraint

	Familiar stimuli		*Unfamiliar stimuli*	
Lexical constraint	H.M.[a]	Controls[b]	H.M.[a]	Controls[b]
Two target words	1.00 [4.78]	1.88 (0.18)	0.38 [11.8]	1.91 (0.13)
Three target words	3.00 [0.67]	2.88 (0.18)	1.38 [4.69]	2.75 (0.26)
Overall	2.00 [0.63]	2.38 (0.60)	1.75 [2.90]	2.33 (0.20)

[a]H.M.'s deficits relative to controls (in *SD*s) appear in square brackets. [b]Standard deviations in parentheses.

controls (SD = 13), a nonreliable difference of 0.46 standard deviations. In describing unfamiliar stimuli, H.M. produced single grammatical sentences in describing 6% of the stimuli, versus a mean of 67% for controls (SD = 22), a reliable difference of 2.77 standard deviations. H.M. thus had a deficit in syntactic constraint performance relative to the controls for unfamiliar but not familiar stimuli.

Finally, H.M. produced eight uncorrected referential errors in describing unfamiliar stimuli and no referential errors in describing familiar stimuli. H.M.'s deficit in picture description accuracy therefore applied to unfamiliar but not familiar stimuli.

Effects of lexical constraint on H.M.'s target word inclusion

Table 4 shows mean target word inclusion for H.M. and the controls as a function of lexical constraint: pictures with two versus three target words. For two-word pictures, H.M. incorporated a mean of 0.50 target words, a deficit of 4.86 standard deviations relative to the controls (M = 1.90 target words; SD = 0.29). For three-word pictures, H.M. incorporated a mean of 1.70 target words, an unreliable deficit of 1.08 standard deviations relative to the controls (M = 2.78 target words, SD = 0.57). H.M.'s deficits were therefore 3.78 standard deviations smaller for three-word pictures than for two-word pictures. This outcome indicates that increasing the load or number of target words assumed to occupy working memory reduced rather than increased H.M.'s deficits in this task.

TABLE 4
Mean and percentage target word inclusion in descriptions of H.M. and the controls for pictures with two versus three target words

	Target words[a]			
	Two		*Three*	
Participants	*M*	%	*M*	%
Controls	1.90 (0.29)	95	2.78 (0.57)	93
H.M.	0.50	25	1.70	57
H.M.'s deficits[a]	4.86	—	1.08	—

Note. Standard deviations in parentheses.
[a]H.M.'s deficits relative to controls (in *SD*s).

An analysis of the percentage of target words incorporated in descriptions of two- versus three-word pictures reinforced this conclusion. H.M. incorporated 25% of the target words for two-word pictures versus 57% for three-word pictures, whereas controls exhibited a nonsignificant trend in the opposite direction: Controls incorporated 95% of the target words for two-word pictures versus 93% for three-word pictures (see Table 4). Thus, controls incorporated 2% fewer target words for three-word than two-word pictures, whereas H.M. incorporated 32% more target words for three-word than two-word pictures, confirming that H.M.'s target word deficit in this task decreased rather than increased with load or the number of target words assumed to occupy working memory.

Interestingly, H.M.'s superior three-word performance applied especially to familiar stimuli. For familiar stimuli, H.M. included a mean of 1.0 target words for two-word stimuli versus 3.0 target words for three-word stimuli (a large, 2.0-word difference; see Table 3). However, for unfamiliar stimuli, H.M. included a mean of 0.38 target words for two-word stimuli versus 1.56 target words for three-word stimuli (a smaller difference; see Table 3).

Effects of input-side repetition on H.M.'s target word inclusion

Table 5 shows H.M.'s mean and percentage of target word inclusion and his deficits relative to the controls for the first versus second presentation of the pictures. Overall H.M. incorporated a mean of 0.75 target words in describing the first presentation of a picture, versus 1.3 target words for the second presentation, a 21.2% improvement due to picture repetition. However, this improvement was greater for three- than for two-word pictures (see

TABLE 5
Mean and percentage of target word inclusion for H.M. as a function of picture repetition and lexical constraint

	Picture presentation				
	First		*Second*		
Pictures	*M*	*%*	*M*	*Target word inclusion*	*Improvement*[a]
Two-word	0.25 [5.72]	12.50	0.60 [4.51]	30.00	17.50
Three-word	1.25 [2.50]	41.70	2.0 [1.19]	66.70	25.00
Total	0.75 [3.60]	27.10	1.30 [2.33]	48.30	21.20

Note. Picture repetition: first versus second picture presentation. H.M.'s deficits relative to controls (in *SD*s) appear in square brackets.
[a]Improvement due to picture repetition, in percentages.

Table 5): H.M.'s improvement due to picture repetition was 17.5% for two-word pictures versus 25.0% for three-word pictures, and for three-word pictures, H.M.'s deficit relative to controls was reliable for the first (2.50 *SD*s) but not the second picture presentation (1.19 *SD*s). A by-stimulus analysis confirmed that H.M. incorporated fewer target words for the first than for the second presentation of three-word pictures, sign test $p < .05$.

Of course, picture repetition was almost certainly not the sole basis for this effect. Two covarying factors almost certainly contributed as well: target word repetition (because the first and second presentations of a picture contained two identical target words), and experimenter-induced repetition within each trial (discussed next).

Effects of experimenter-induced repetition on H.M.'s performance

Target word performance

Experimenter-induced repetition occurred on 16 trials (80%) where H.M. failed to include all of the target words in his initial response, and the experimenter asked him to try again. To evaluate the effectiveness of this manipulation, we calculated how many target word *tokens* H.M. produced per trial (across up to seven repeated attempts to include all of the target words in a single utterance). Mean number of target word tokens was 3.65 (range 1–16) for H.M. versus 1.96 (range 1.75–3) for the (unprompted) controls. Thus, even though controls included more target words than did H.M. per utterance, H.M. produced in his overall output almost twice as many target word tokens as did the controls. This outcome indicates that experimenter prompts led to extensive repetition and reprocessing of the target words, consistent with the occurrence of engrainment learning under binding theory. The fact that H.M. produced up to 16 target word tokens in a single trial also suggests that H.M. was strongly motivated to succeed in the present task.

To evaluate the effects of experimenter-induced repetition on target word inclusion *within a single utterance*, we compared H.M.'s initial utterance with his final utterance on the 16 trials in which the experimenter asked H.M. to try again (see Table 2). H.M. incorporated fewer target words in his initial utterances ($M = 1.10$) than in his final utterances ($M = 1.38$), with a considerably smaller deficit relative to controls for H.M.'s final (2.14 *SD*s) than initial (3.63 *SD*s) utterances. In combination with input-side repetition, experimenter-induced repetition therefore facilitated H.M.'s performance. Note, however, that H.M. nonetheless exhibited a reliable deficit of 2.14 standard deviations relative to controls in his final picture descriptions.

Referential constraint performance

Because experimenter requests to try again focused on target word inclusion rather than accuracy of the picture descriptions, it is perhaps not surprising that the number of referential errors in H.M.'s initial and final utterances did not differ. Interestingly, however, H.M.'s initial and final utterances often contained different referential errors. Final Utterances 13–14 illustrate new referential errors not seen in H.M.'s initial utterances discussed earlier (see Table 2 for the target words).

13. H.M. (Set 2 Picture 5): "Well this pie is—or the pie here was back here—(Exp.: Uh-huh) and uh coffee is in there because heat a solid and this is only boiled milk say milk there and this is not liquid but only ice" (uncorrected referential error).

Final Utterance 13 describes the cafeteria picture (discussed earlier for Initial Utterance 9) with a new referential error: The picture does not contain ice.

14. H.M. (Set 1 Picture 11): “Gary is . . . almos . . . almost . . . hasn’t been cut the same way” (uncorrected referential error).

Final Utterance 14 describes the clerk in the bakery shop picture (discussed earlier for Initial Utterance 10) with a new referential error: There are pies, loaves of bread, and two women customers in the picture, but no cuts of any kind.

In summary, H.M.’s overall performance was impaired relative to controls for three aspects of the TLC: target word inclusion (H.M. incorporated fewer target words than did the controls), syntactic constraint performance (H.M. produced fewer single grammatical sentences in his descriptions than did the controls), and referential constraint performance (picture descriptions were inaccurate more often for H.M. than for the controls).

However, H.M. only exhibited deficits on the TLC for unfamiliar stimuli that depicted novel or noncliché situations and contained target words with no links to familiar or cliché expressions that could accurately characterize the pictures. H.M.’s target word performance, syntactic constraint performance, and referential constraint performance exhibited no deficits for target words linked in H.M.’s semantic memory with cliché phrases and pictures depicting situations familiar to H.M. since early childhood.

Finally, two types of repetition facilitated H.M.’s performance on the TLC: picture repetition (especially when three target words accompanied the second picture presentation) and experimenter-induced repetition (relative to controls, H.M.’s target word deficit was smaller for his initial than for his final utterances following experimenter prompts to try again).

Present results therefore supported all three binding theory predictions discussed in the Introduction: (a) H.M. was impaired relative to controls when describing unfamiliar situations using target words with no links to cliché expressions; (b) H.M. was unimpaired relative to controls when describing familiar situations using target words with links to cliché expressions familiar to H.M. since early childhood; (c) Both input-side repetition and self-produced repetition of target words facilitated H.M.’s target word performance.

DISCUSSION

This section first summarizes H.M.’s selective deficits and sparing in language production on the TLC, together with related evidence indicating parallel deficits and sparing in H.M.’s language comprehension, sentence-reading, visual cognition, and memory. We next discuss the implications of these parallels for the major theories outlined in the Introduction: systems theory versus binding theory. Finally, we return to the initial goal of the present study: to cross-check the Skotko et al. (2005) support for the intact-language hypothesis and to resolve the conflicting conclusions regarding H.M.’s language production.

H.M.’s language production deficits

H.M.’s descriptions of unfamiliar word–picture stimuli in the TLC incorporated fewer target words than the mean for controls, a deficit of 2.90 standard deviations reminiscent of H.M.’s deficits on tests of semantic and symbolic verbal fluency (Corkin, 1984); H.M. produced fewer single grammatical sentences in his descriptions than the mean for controls, a difference of 2.77 standard deviations, and reliably more ungrammatical descriptions than the mean for controls (whether single sentences or not); judges rated H.M.’s descriptions as less conceptually complete than the descriptions of typical control participants, a difference of 3.04 standard deviations that was also reliable by sign test with stimuli as unit of analysis; Finally, H.M. produced more inaccurate descriptions than the mean for controls, a difference of at least 6.0 standard deviations that is reminiscent of H.M.’s deficits on the reporter’s test of language production (see Corkin).

The selective nature of H.M.’s language production deficits

Just as H.M.’s lesion has impaired his memory for novel or never previously encountered semantic information (Gabrieli et al., 1988) while sparing semantic information encountered frequently before and after his operation (see e.g., James & MacKay, 2001), H.M.’s language production deficits were selective rather than across the board in the present data. On the TLC, H.M. exhibited no deficits for familiar stimuli that depicted situations commonly encountered since early childhood and featured target words linked in H.M.’s semantic memory with familiar or cliché phrases: For familiar stimuli, H.M.’s descriptions were not reliably less accurate and did not include reliably fewer target words than the mean for controls. Nor did H.M. generate fewer single grammatical sentences than did controls for familiar stimuli.

Also unlike controls, H.M. exhibited a “reverse load” effect, with a reliably smaller target word deficit for pictures with three than two target words, especially for familiar stimuli. Moreover, just as H.M. improved with practice in the mirror-tracing task (see Milner, 1965), both input-side repetition and experimenter-induced repetition of target words facilitated H.M.’s target word performance on the TLC. Finally, H.M. had no difficulty producing HF words and cliché phrases per se.

H.M.’s pattern of selective deficits and sparing in language production carries theoretical significance because the same selective pattern reappears in many other cognitive domains—for example, sentence reading, language comprehension, and visual cognition. When reading sentences aloud (see MacKay & James, 2001), H.M. produces unusual prosodic pauses within unfamiliar but not familiar phrases in sentences and at major syntactic boundaries unmarked by commas. However, H.M. pauses normally at major syntactic boundaries marked by commas, an orthographic signal for pausing that children learn during grade school. These selective sentence-reading deficits indicate that H.M. only has difficulty with comprehending and producing novel aspects of sentences, just as he only exhibits deficits in describing novel situations with noncliché sentences (see Table 3; also MacKay et al., 1998a; MacKay & James, 2001). Visual cognition completes the nonmemory parallels: H.M. exhibits selective deficits in detecting unfamiliar but not familiar visual figures hidden in concealing arrays (MacKay & James, 2000).

Moreover, H.M.’s performance on memory tasks exhibits these same parallels. As noted in the Introduction, H.M. exhibits memory deficits for novel or never previously encountered information, exactly the information that must be retrieved in explicit, episodic, and declarative tests and in implicit memory tests for words that H.M. encountered for the first time after his 1953 lesion (Gabrieli et al., 1988). However, H.M. exhibits sparing for familiar and repeated information, exactly the information retrieved in repetition priming, eyeblink conditioning, and motor skills tasks, in tests of implicit memory involving preoperatively familiar words and in tasks involving frequently repeated semantic information (e.g., Gabrieli et al.; Keane, Gabrieli, & Corkin, 1987; Keane, Gabrieli, Mapstone, & Johnson, 1995; MacKay et al, 1998a; O’Kane, Kensinger, & Corkin, 2004; Skotko et al., 2004; and Spiers, Maguire, & Burgess, 2001). Thus, H.M. exhibits implicit memory deficits for unfamiliar words that lack preformed internal representations but not for preoperatively familiar words with intact internal representations. However, H.M. always exhibits deficits in episodic memory tasks because episodic encoding always requires the formation of new connections to represent the context of occurrence of unique events or episodes—for example, the fact that a particular word in a list occurred at a particular time or place.

In short, H.M.’s memory deficits parallel his pattern of deficits and sparing in language production, language comprehension, and visual cognition: spared activation of familiar or already formed representations but impaired binding or connection formation processes for creating never previously encountered representations. Any theory of H.M.’s condition must therefore explain these parallels in H.M.’s memory, visual cognition, and language, including sentence-level comprehension, production, and reading aloud.

Implications for systems theory

Under traditional systems theory accounts of H.M.’s performance (see e.g., Cohen & Eichenbaum, 1993; LeDoux, 1996), H.M.’s lesion has by chance damaged certain memory storage systems (involving, e.g., explicit, declarative, or consciously retrieved memories) but not others (e.g., memory systems for eyeblink conditioning, implicit recall, and memories involving procedural skills). Likewise by chance, H.M.’s lesion has completely spared systems for retrieving, comprehending, and producing words and sentences.

To address the present results, this systems theory must now add many additional assumptions. Systems theory must assume that H.M.’s lesion has damaged systems for representing and describing pictures of noncliché situations using target words with no links to familiar or cliché expressions that could accurately characterize the pictures. Systems theory must likewise assume that H.M.’s lesion has spared systems for spoken production of HF words and cliché phrases, together with systems for representing and describing pictures depicting situations familiar to H.M. since early childhood using target words with links to familiar expressions that could accurately characterize the pictures.

To address other data on H.M.’s cognitive abilities, systems theory must also assume that H.M.’s lesion has impaired a visual cognition system for detecting unfamiliar hidden figures (see MacKay & James, 2000), a language production system for generating prosody when reading unfamiliar phrases and novel sentences without commas (see MacKay & James, 2001), a word comprehension system for comprehending isolated HF words (see James & MacKay, 2001), and language systems for comprehending metaphors and lexically ambiguous words in sentences, for assigning thematic roles to the major constituents of sentences, and for detecting semantic and syntactic "errors" in anomalous sentences (see MacKay et al., 2007). Systems theory must likewise assume that H.M.'s lesion has spared a visual cognition system for detecting familiar hidden figures (see MacKay & James, 2000), a language production system for generating prosody when reading familiar phrases and novel sentences with commas (see MacKay & James, 2001), and language comprehension systems for understanding isolated HF words (see James & MacKay, 2001) and lexically ambiguous phrases presented in isolation (see MacKay et al., 2007).

In summary, systems theory can explain H.M.'s performance by assuming that his relatively circumscribed lesion has accidentally spared or damaged dozens of separate systems, many with no independently motivated raison d'etre (see MacKay et al., 2007). However, even this highly unparsimonious proliferation of systems fails to explain detailed aspects of present results. Why was H.M.'s target word deficit smaller for pictures with three than two target words, especially when the word–picture stimuli were familiar rather than unfamiliar? Why did picture repetition facilitate H.M.'s performance in present results, especially when three target words accompanied the second picture presentation? Systems theory provides answers to none of these questions (see MacKay et al., 2007, for other, more general shortcomings of current systems theories).

Nonetheless, present results do not challenge the concept of functional or neural systems per se: Systems theory can explain any results post hoc and remains a viable theory if one ignores issues of parsimony, significant empirical details, and the primary function of a theory—namely, to guide research, enable detailed predictions, and generate "new ideas and new discoveries" (Higgins, 2004, p. 138). On the other hand, present results indicate a clear need for principled criteria whereby systems theory can define and establish systems and the dividing lines between them (see MacKay et al., 2007, for illustrations of the type of "dividing line" criteria that seem needed; see also Barnard & Dalgleish, 2005).

Implications for binding theory

The parallels between H.M.'s deficits and sparing in language, memory, and visual cognition are nonaccidental and are simply explained under binding theory (see MacKay et al., 2007, for theoretical details): H.M.'s 1953 lesion has impaired the binding processes for representing new information in all three areas but has spared the activation processes that enable retrieval, repetition and use of familiar or preoperatively encountered information in language, visual cognition, and memory. Also spared is engrainment learning, the primitive and relatively inefficient process whereby repetition or rehearsal facilitates performance, even for unfamiliar information.

Three predictions derived from these general binding theory principles were supported in the present results: (a) H.M. was impaired relative to controls when describing unfamiliar situations via noncliché or novel expressions; (b) H.M. was unimpaired relative to controls when describing familiar situations using cliché expressions familiar to H.M. since early childhood; and (c) stimulus repetition and repeated attempts to try again facilitated H.M.'s target word performance.

Binding theory also readily explains more detailed results in the present study. One was H.M.'s smaller target word deficit for pictures with three than two interrelated target words, especially when the word–picture stimuli were familiar rather than unfamiliar. By hypothesis, H.M.'s binding deficit hindered the process of forming new connections required to represent a coherent sentence plans for describing both familiar and unfamiliar stimuli, and he was forced to activate and concatenate familiar phrases with preformed links in semantic memory to the target words and the situation depicted. However, concatenation of familiar phrases unguided by a coherent sentence plan often resulted in ungrammatical utterances (see Examples 1–8) that lacked essential ingredients for a sentence and contained free-associative phrases, nonsequiturs, run-on clauses, and violations of major selection restrictions or agreement rules.

Consider now H.M.'s superior performance for familiar than unfamiliar stimuli. By definition, target words for familiar but not unfamiliar stimuli had strong links to familiar or cliché expressions that could accurately characterize the pictures, so that H.M. could concatenate more clichés containing more target words to accurately describe familiar than unfamiliar stimuli. Moreover, increasing the number of target words enhanced this effect of stimulus familiarity: Because the target words applied to the same picture and therefore triggered interrelated clichés for familiar stimuli, increasing the number of target words served to constrain the coherence or grammaticality of H.M.'s descriptions of familiar stimuli, in addition to increasing target word inclusion.

To illustrate the relation between familiarity, number of target words, and coherence or grammaticality under binding theory, consider H.M.'s responses to the two- versus three-word versions of a familiar stimulus: someone saying something to someone about to sit on fresh paint. Note that for the two-word version (target words, *sit, because*), H.M.'s description ("and that man is trying to tell that woman not to sit there because it's wet paint . . . he can uh see the sign better than she could and she's ready to sit down there"; see Table 2) contains only one target word (*sit*), whereas for the three-word version (target words, *sit, because, painted*), H.M.'s description ("Oh, don't sit because it's just been painted"; see Table 2) contains all three target words. By hypothesis, adding the third target word (*painted*) helped constrain H.M.'s description by triggering several interrelated phrases with strong links in long-term semantic memory to each other, to the target words, and to the familiar situation depicted—for example, "Just painted" and "Don't sit." By conjoining these interrelated and familiar phrases, H.M.'s "Oh, don't sit because it's just been painted" not only included all three target words, but was also both coherent and grammatical.

However, for *unfamiliar* stimuli, adding a third target word did not facilitate H.M.'s descriptions because the unfamiliar stimuli depicted novel situations with target words with no links in H.M.'s semantic memory to familiar phrases that could help constrain his descriptions. As a result, H.M.'s descriptions included fewer target words and were less coherent or grammatical for unfamiliar than familiar stimuli, regardless of whether two or three target words headed the unfamiliar stimuli.

By contrast, adding a third target word helped the controls for neither familiar nor unfamiliar stimuli because the controls were describing both the familiar and the unfamiliar stimuli afresh by forming novel, situation-appropriate sentence plans rather than by retrieving and concatenating familiar clichés from long term semantic memory.

Another detailed result readily explained under binding theory is the fact that picture repetition facilitated H.M.'s target word performance, especially when three target words accompanied the second picture presentation. Under binding theory, the main effect reflects engrainment learning due to picture repetition and the extensive experimenter-induced repetition of the two overlapping target words from the first picture presentation. However, adding a third target word helped constrain H.M.'s descriptions above and beyond the effects of engrainment learning, so that facilitation due to repetition increased more when the second presentation of a picture contained three than when it contained two target words. When only two target words accompanied the second picture presentation, the reduced (two-word) constraint offset the facilitative effects of picture repetition, resulting in reduced improvement relative to H.M.'s description for the first picture presentation with three target words.

Consider now an alternate account of present data. Under this account, H.M. exposed himself to language (e.g., via reading) less than did the controls, perhaps because of his memory deficit following surgery. With less practice using language, H.M. therefore exhibited deficits that were indirectly rather than directly related to whatever caused his amnesia (unlike in the binding theory account). H.M.'s more limited exposure to language then introduced a nonlinear effect, affecting more difficult tasks (e.g., describing unfamiliar stimuli) more strongly than easy tasks (e.g., describing familiar stimuli).

This "processing difficulty" account faces several problems. One is that processing difficulty does not accurately characterize the nature of H.M.'s language deficits. For example, H.M. exhibits impaired comprehension for a wide variety of syntactic structures, including those that memory-normal participants find easy to comprehend and recall (see MacKay et al., 2007). Unlike controls, H.M. also exhibited a "reverse load" effect in the present data. A second problem is that the concepts "difficult" and "easy" lack noncircular definition in this processing difficulty account. If unfamiliar stimuli were in general more difficult to describe than familiar stimuli in the present study, then this should have been apparent in the control data. However, this was not the case: Unlike H.M., the controls were less successful overall when describing familiar than unfamiliar stimuli, albeit not significantly so (see Table 3). A third problem concerns parsimony: Reduced exposure to language cannot parsimoniously explain why H.M. exhibits parallel deficits and sparing in language, memory, and visual cognition. Finally, longitudinal data have conclusively ruled out reduced exposure as a general account of H.M.'s language deficits (see James & MacKay, 2001; and MacKay & James, 2002).

Conflicting conclusions regarding H.M.'s sentence production: A resolution

Present results did not support the intact-language hypothesis of Kensinger et al. (2001) and Skotko et al. (2005): H.M. exhibited deficits in target inclusion, syntactic constraint performance, and picture description accuracy and completeness relative to controls on the TLC. Moreover, H.M. produced ungrammatical utterances involving incomplete sentences, violations of major selection restrictions or agreement rules, malapropisms, pronoun reference errors, nonsequiturs, run-on clauses, free-associative phrases, and conceptually redundant repetitions more often than did controls.

How then can we resolve the conflicting conclusions regarding binding theory and the intact-language hypothesis in the present study versus Kensinger et al. (2001) and Skotko et al. (2005)? Present results do not fit the resolution proposed in Kensinger et al.: that factors contributing to memory load explain H.M.'s language deficits. H.M.'s target word deficit on the TLC decreased rather than increased with load or number of target words assumed to occupy working memory. Moreover, the reverse load effect in the present data was especially strong for familiar than for unfamiliar stimuli, even though a larger load might be expected for unfamiliar stimuli (see MacKay et al., 2007, for a review of other results that contradict memory load hypotheses).

We believe that resolution lies in the detailed procedures, stimuli, and responses in Kensinger et al. (2001) and Skotko et al. (2005) versus the present study and MacKay et al. (1998a), MacKay and James (2001, 2002), MacKay et al. (2007), and MacKay et al. (1998b). As noted earlier, results involving unfamiliar information that is massively repeated or HF information familiar to H.M. before his lesion are consistent with both binding theory and the intact-language hypothesis. Because critical stimuli and responses in Kensinger et al. were massively repeated and did involve HF information familiar since childhood (see MacKay, 2006), the Kensinger et al. results are in fact consistent with both binding theory and the intact-language hypothesis. We therefore focus here on the stimuli, responses, and procedures in Skotko et al., the only other nonanecdotal study that concluded in favor of the intact-language hypothesis and against binding theory.

The Skotko et al. evidence for the intact-language hypothesis

To test the intact-language hypothesis, Skotko et al. compared the utterances of H.M. and same-age controls in three independent interview studies. An interview study with H.M. examined 10 random samples from 5–6 hours of spontaneous four-way conversations between H.M. and three simultaneously present interviewers, and two interview studies with normal controls examined the one-way responses of older adults to prepared questions of a single interviewer. These control studies were not originally designed for comparison with H.M. and were conducted by Kemper, Kynette, Rash, O'Brien, and Sprott (1989) and Kemper and Sumner (2001): henceforth Kemper et al. (1989–2001) because both studies involved similar stimuli and procedures.

Validity of the Skotko et al. (2005) comparison of H.M. versus the controls therefore depends on methodological issues—for example, were the Skotko et al. versus Kemper et al. (1989–2001) procedures comparable?—and data analysis issues—for example, were appropriate statistical procedures adopted? We discuss these two types of issue in turn.

Methodological issues

Stimulus comparability. The controls in Kemper et al. (1989–2001) answered questions that differed fundamentally in form and in content from the questions that H.M. answered in Skotko et al. (2005), making it inappropriate to directly compare the resulting data sets for evaluating deficits. The single interviewer in Kemper et al. asked the controls a small number of fixed and nonoverlapping questions in identical order for each participant. The questions were open ended (none had a "correct" answer), and the interviewer never provided prompts as to the desired or expected answer. Finally, the control questions focused on two general topics: background information (employment history, current health, and activities) and an interesting experience or admired/prominent person that influenced the participant's life.

In contrast, three interviewers in Skotko et al. (2005) asked H.M. a large number of questions in unplanned order and on a wide variety of topics. Many were follow-up questions identical to or overlapping in content with prior questions. Moreover, many of the questions had a "correct" answer—for example, "Was that astronaut a man or a woman?", and the interviewers often provided prompts as to the correct answer. Finally, the questions differed in topic, content, and scope for H.M. in Skotko et al. versus the controls in Kemper et al. (1989–2001): Unlike the controls, H.M. answered questions such as: "Would you like to tell us anything?" "Did you know you're famous?" "Did you ever cut the grass?" "Do you like to listen to music?" "Do you remember Marilyn Monroe?" "So you love puzzles?" "We're not boring you, are we?"

Response comparability. H.M. and the controls did not produce outputs with comparable form and content, again making direct comparison difficult for evaluating deficits. The controls provided lengthy paragraphs in response to their open-ended questions—for example, 207 words in an illustrative example published in Kemper et al. (1989–2001). By contrast, H.M. provided short, often single-word responses, especially to yes–no questions such as, "Did you know you're amazing?"

Procedural comparability. Unlike Kemper et al. (1989–2001), Skotko et al. (2005) adopted procedures that functioned to discourage the production of novel sentences. The Skotko et al. interviewers often repeated the same question, sometimes several times on the same day, and on occasion introduced topics that they knew were highly familiar to H.M. For example, an interviewer familiar with H.M.'s tendency to repeat the cliché "live and learn" asked H.M., "Sometimes you just live and learn, right?" Procedures for the controls in Kemper et al. were quite different: The interviewer in Kemper et al. never asked the same question twice and never used background knowledge about a particular participant in formulating their questions. These contrasting procedures for H.M. versus the controls make it difficult to conclude that H.M.'s language production is intact relative to that of controls. Moreover, results using the Skotko et al. procedures are consistent with binding theory and the present results, because H.M. benefits when stimuli and responses are repeated and he processes familiar information without deficits.

Data analysis issues

Skotko et al. (2005) compared the interview responses of H.M. versus the controls on four dimensions: the type-token ratio (TTR, the number of different words used divided by the total number of words in a sample), the mean length of utterances (MLU), the mean number of clauses per utterance (MCU), and the number of left-branching clauses (LBC, in percentages). Results of these analyses indicated higher TTR scores for H.M. than for the controls in Kemper et al. (1989–2001) "at a level that was statistically significant" according to Skotko et al. (p. 403), but lower MLU, MCU, and LBC scores, albeit not reliably lower under the standard convention that only differences larger than 2.0 standard deviations count as significant (see e.g., MacKay & James, 2001). However, the TTR difference was nonsignificant, contrary to the claim for statistical significance in Skotko et al. Mean TTR for H.M. and the controls differed by only 0.69 standard deviations (see Skotko et al., Table 2, p. 403), well below the difference of 2.0 standard deviations required for significance.[4] The Skotko et al. data therefore yielded no reliable effects and do not support the intact-language hypothesis because null results are interpretively problematic in the absence of power analyses (see Cohen, 2003).

In summary, methodological and data analysis problems make it impossible for the Skotko et. al (2005) data to clearly support the intact-language hypothesis as opposed to binding theory. Moreover, as in Kensinger et al. (2001), stimuli and responses in Skotko et al. often involved repetition of HF information, so that even if Skotko et al. had demonstrated valid nondeficits under those conditions, their results would be consistent with binding theory. In short, all available evidence on H.M."s sentence production is consistent with the selective deficits and sparing predicted under binding theory and demonstrated in MacKay et al. (1998a), MacKay and James (2001, 2002), MacKay et al. (2007), MacKay et al. (1998b), and the present study.

Original manuscript received 12 December 2006
Revised manuscript accepted 29 March 2007
First published online day month year

REFERENCES

Barnard, P. J., & Dalgleish, T. (2005). Psychological-level systems theory—the missing link in bridging emotion theory and neurobiology through dynamic systems modeling: Commentary on Lewis. *Behavioral and Brain Sciences, 28*, 196–197.

Cohen, J. (2003). A power primer. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (3rd ed., pp. 427–436). Washington, DC: American Psychological Association.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: MIT Press.

Corkin, S. (1984). Lasting consequences of bilateral medial temporal lobectomy: Clinical course and experimental findings in H.M. *Seminars in Neurology, 4*, 249–259.

Corkin, S., Amaral, D. G., González, R. G., Johnson, K. A., & Hyman, B. T. (1997). H.M.'s medial temporal lobe lesion: Findings from MRI. *Journal of Neuroscience, 17*, 3964–3979.

[4]In fact, significantly superior TTR performance for H.M. relative to controls would be a curious result because H.M. scores low on working-memory measures whereas TTR and working-memory measures correlate positively for normal young and older adults (see Kemper & Sumner, 2001). Also problematic, Skotko et al. (p. 406) attribute H.M.'s low MLU, MCU, and LBC scores to his "inferior education," a claim that suggests a flaw in their matching procedures. Moreover, contrary to Skotko et al., inferior education cannot in principle simultaneously explain both superior TTR scores for H.M. than for controls and (nonsignificantly) inferior MLU, MCU, and LBC scores for H.M. than for controls.

Gabrieli, J. D. E., Cohen, N. J., & Corkin, S. (1988). The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. *Brain & Cognition, 7*, 157–177.

Higgins, E. T. (2004). Making a theory useful: Lessons handed down. *Personality and Social Psychology Review, 8*, 138–146.

Jackendoff, R. (2003). *Foundations of grammar: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

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

Keane, M. M., Gabrieli, J. D. E., & Corkin, S. (1987). Multiple relations between fact-learning and priming in global amnesia. *Society for Neuroscience Abstracts, 13*, 1454.

Keane, M. M., Gabrieli, J. D., Mapstone, H. C., & Johnson, K. A. (1995). Double dissociation of memory capacities after bilateral occipital-lobe or medial temporal lobe lesions. *Brain: A Journal of Neurology, 118*, 1129–1148.

Kemper, S., Kynette, D., Rash, S., O'Brien, K., & Sprott, R. (1989). Life-span changes at adults' language: Effects of memory and genre. *Applied Psycholinguistics, 10*, 49–66.

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging, 16*, 312–322.

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 H.M. *Hippocampus, 12*, 337–346.

Lackner, J. R. (1974). Observations on the speech processing capabilities of an amnesic patient: Several aspects of H.M.'s language function. *Neuropsychologia, 12*, 199–207.

LeDoux, J. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon & Schuster.

MacKay, D. G. (2006). Aging, memory and language in amnesic H.M. *Hippocampus, 16*, 491–494.

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

MacKay, D. G., & James, L. E. (2000, April). *Binding processes for visual cognition: A "hippocampal amnesic" (H.M.) exhibits selective deficits in detecting hidden figures and errors in visual scenes*. Poster presented to the Cognitive Neuroscience Society, San Francisco, CA.

MacKay, D. G., & James, L. E. (2001). The binding problem for syntax, semantics, and prosody: H.M.'s selective sentence-reading deficits under the theoretical-syndrome approach. *Language and Cognitive Processes, 16*, 419–460.

MacKay, D. G., & James, L. E. (2002). Aging, retrograde amnesia, and the binding problem for phonology and orthography: A longitudinal study of "hippocampal amnesic" H.M. *Aging, Neuropsychology and Cognition, 9*, 298–333.

MacKay, D. G., James, L. E., Taylor, J. K., & Marian, D. E. (2007). Amnesic H.M. exhibits parallel deficits and sparing in language and memory: Systems versus binding theory accounts. *Language and Cognitive Processes, 22*, 377–452.

MacKay, D. G., Stewart, R., & Burke, D. M. (1998b). H.M. revisited: Relations between language comprehension, memory, and the hippocampal system. *Journal of Cognitive Neuroscience, 10*, 377–394.

Manns, J. R. (2004). J.F.K., L.B.J., and H.M.: The famous memories of a famous amnesic. *Hippocampus, 14*, 411–412.

Milner, B. (1965). Visually-guided maze learning in man: Effects of bilateral hippocampal, bilateral frontal, and unilateral cerebral lesions. *Neuropsychologia, 3*, 317–338.

Milner, B., Corkin, S., & Teuber, H. L. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of H. M. *Neuropsychologia, 6*, 215–234.

Ogden, J. A., & Corkin, S. (1991). Memories of H.M. In W. C. Abraham, M. Corballis, & K. G. White (Eds.), *Memory mechanisms: A tribute to G. V. Goddard* (pp. 195–215). Hillsdale, NJ: Lawrence Erlbaum Associates.

O'Kane, G., Kensinger, E. A., & Corkin, S. (2004). Evidence for semantic learning in profound amnesia: An investigation with the patient H.M. *Hippocampus, 14*, 417–425.

Schmolck, H., Stefanacci, L., & Squire, L. R. (2000). Detection and explanation of sentence ambiguity are unaffected by hippocampal lesions but are impaired by larger temporal lobe lesions. *Hippocampus, 10*, 759–770.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*, 11–21.

Skotko, B. G., Andrews, E., & Einstein, G. (2005). Language and the medial temporal lobe: Evidence from H.M.'s spontaneous discourse. *Journal of Memory and Language, 53*, 397–415.

Skotko, B. G., Kensinger, E. A., Locascio, J. L., Einstein, G., Rubin, D. C., Tupler, L. A., et al. (2004). Puzzling thoughts for H.M.: Can new semantic information be anchored to old semantic memories? *Neuropsychology, 18*, 756–769.

Spiers, H. J., Maguire, E. A., & Burgess, N. (2001). Hippocampal amnesia. *Neurocase, 7*, 357–382.

Wiig, E. H., & Secord, W. (1988). *Test of language competence: Expanded edition*. New York: The Psychological Corporation, Harcourt, Brace, Jovanovich, Inc.

Zaidel, D. W., Zaidel, E., Oxbury, S. M., & Oxbury, J. M. (1995). The interpretation of sentence ambiguity in patients with unilateral focal brain surgery. *Brain and Language, 51*, 458–468.