H.M. Revisited: Relations between Language Comprehension, Memory, and the Hippocampal System

Donald G. MacKay
University of California, Los Angeles

Rachel Stewart and Deborah M. Burke
Pomona College

Abstract

Three studies tested the claim that H.M. exhibits a "pure memory deficit" that has left his ability to comprehend language unimpaired relative to memory-normal controls. In Study 1, H.M. and memory-normal controls of comparable intelligence, education, and age indicated whether sentences were ambiguous or unambiguous, and H.M. detected ambiguities significantly less often than controls. In Study 2, participants identified the two meanings of visually presented sentences that they knew were ambiguous, and relative to controls, H.M. rarely discovered the ambiguities without help and had difficulty understanding the first meanings, experimenter requests, and his own output. Study 3 replicated these results and showed that they were not due to brain damage per se or to cohort effects: Unlike H.M., a patient with bilateral frontal lobe damage detected the ambiguities as readily as young and same-cohort older controls. These results bear on two general classes of theories in use within a wide range of neurosciences and cognitive sciences: The data favor "distributed-memory theories" that ascribe H.M.'s deficit to semantic-level binding processes that are inherent to both language comprehension and memory, over "stages-of-processing theories," where H.M.'s defective storage processes have no effect on language comprehension.

INTRODUCTION

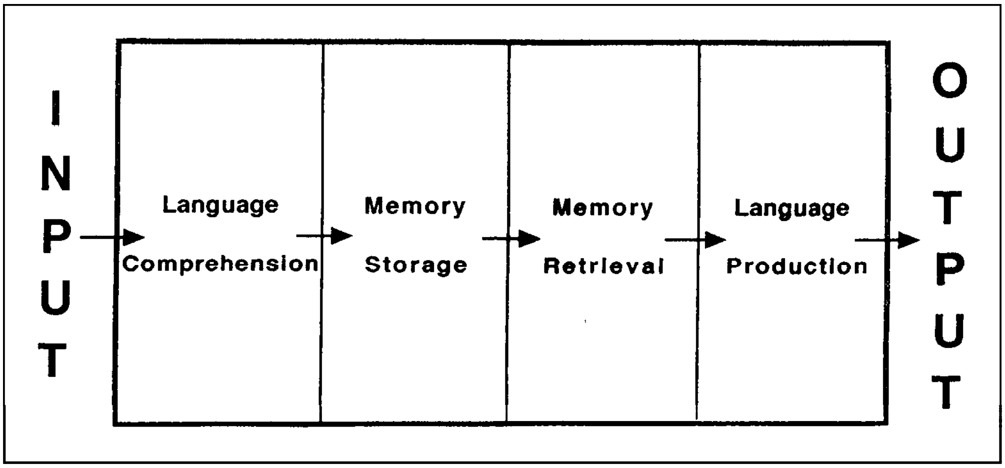
In 1953, a "frankly experimental operation" removed parts of the hippocampal system of a 27-year-old patient known as H.M. (Scoville & Milner, 1957). This bilateral operation partially alleviated H.M.'s debilitating and otherwise untreatable epileptic condition but had an unexpected and tragic side effect: a severe amnesia that has prevented recall of new information and has confined his subsequent life to the immediate present and the distant past preceding his operation. Because Scoville's operation is no longer performed, H.M. has provided especially valuable evidence regarding the bilateral functioning of the hippocampal system in human cognition, and he has become one of the most famous patients in the history of neuropsychology (Ogden & Corkin, 1991). His condition has also had a major impact on theories of memory and on neural and psychological theories in general.

The prevailing view is that H.M. suffers from a pure memory deficit that prevents acquisition or consolidation of new information for subsequent voluntary recall but completely spares other processes, including language comprehension. This view originated with Milner, Corkin, and Teuber (1968), who claimed that H.M.'s comprehension of language is "undisturbed" (see also Lackner, 1974). However, we had two reasons for reexamining H.M.'s language comprehension. One was its theoretical significance, discussed next. The other was that one of Lackner's findings suggested a possible comprehension deficit.

WHY H.M.'S LANGUAGE COMPREHENSION IS THEORETICALLY IMPORTANT

How important is the status of H.M.'s language comprehension from a theoretical point of view? At stake are two general and contrasting types of theories in use within a wide range of neurosciences and cognitive sciences. One type, known as stages-of-processing theories, has been around in various forms since Descartes (see MacKay, 1998, for historical overview). These theories postulate separate units and processes for language and memory and organize human information processing into at least four sequentially ordered stages or processing modules, labeled comprehension, storage, retrieval, and production in Figure 1. Thus, verbal inputs are first comprehended, with the products of compre-

Figure 1. A standard information processing flow chart with stages for comprehension, storage, retrieval, and production of verbal materials.



hension subsequently stored in memory. Later, for verbal recall, people first retrieve the stored memory in stages-of-processing theories and then express the memory in language during the final production stage. Thus, if the prevailing view is correct that H.M. exhibits a pure memory deficit with unimpaired language comprehension, H.M.'s condition strongly supports stages-of-processing theories by dissociating the storage stage (damaged) from the comprehension stage (undamaged) in Figure 1.

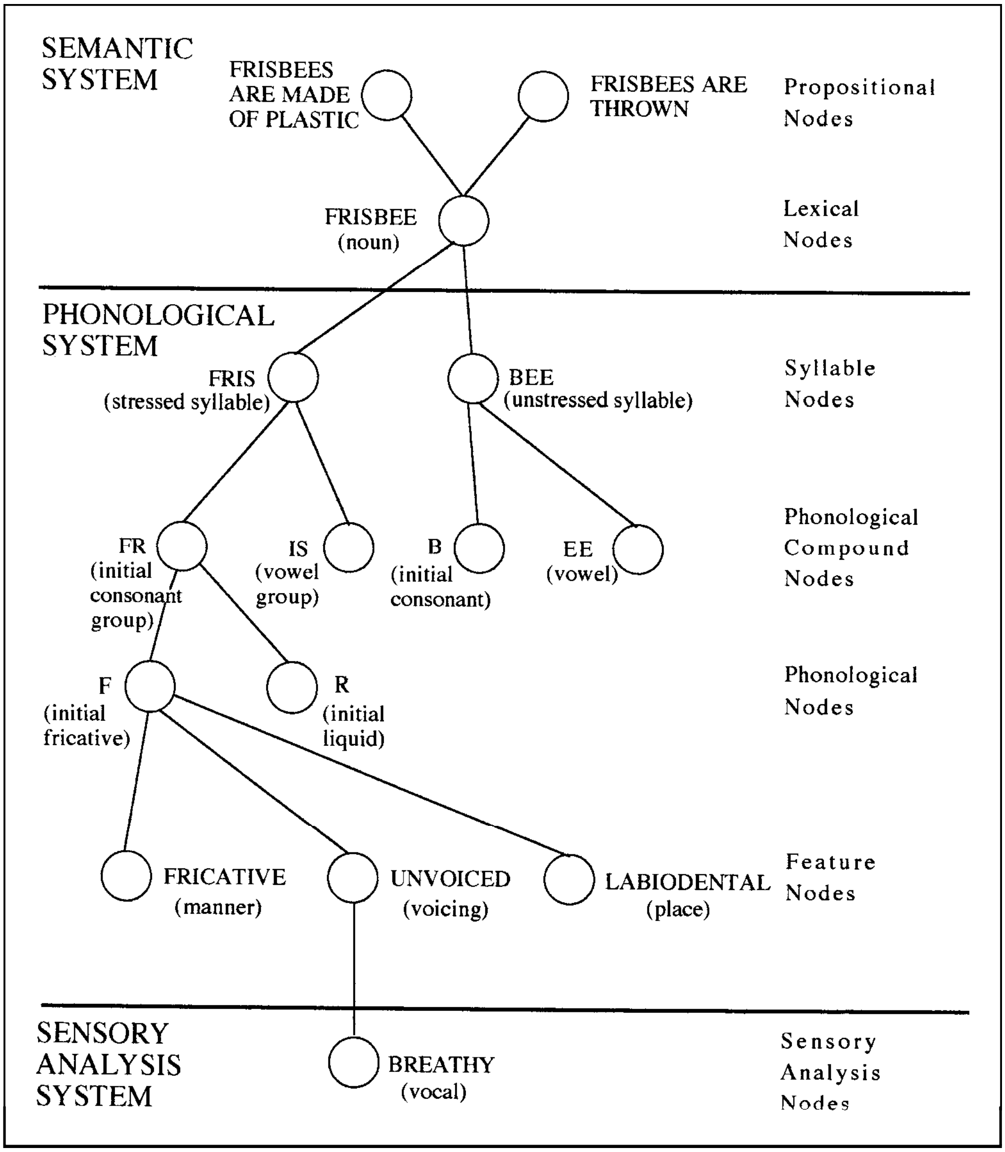
One attraction of the basic stages-of-processing framework in Figure 1 is that new theories are readily created by subdividing one or more stages or by adding new stages in parallel. In this manner, the past 15 years have seen a proliferation of memory stages or modules for episodic versus semantic memory, sensory and perceptual versus modality-independent memory, explicit versus implicit memory, procedural versus declarative memory, reference versus working memory, and short-term versus long-term versus very-long-term memory. However, like the original stages-of-processing framework, data from H.M. have provided central support for all of these modules (see MacKay, Burke, & Stewart, 1998).

A newer and more complex conception of human information processing views relations between language, memory, and the hippocampal system quite differently. Within this "distributed-memory" framework (e.g., Carpenter, & Grossberg, 1993; Grafman, & Weingartner, 1996; McClelland, McNaughton, & O'Reilly, 1995; MacKay & Miller, 1996), memory for verbal materials is not localized within discrete or isolable stages or modules (e.g., for short-term versus long-term memory) but depends fundamentally on the strength of connections between millions of neural units or nodes distributed throughout a vast interactive activation network. To illustrate this distributed-memory framework and its relation to language comprehension, we will focus on one from among many recently developed and equally viable distributed-memory accounts. Under this "Node Structure theory" (described in detail in MacKay, 1990), memory storage involving verbal materials results from forming and strengthening connections between nodes, processes that occur automatically many times a day during normal language comprehension. Nodes are organized hierarchically in the manner illustrated in Figure 2, a detailed representation of bottom-up connections for comprehending the word *frisbee* (adapted from Burke, MacKay, Worthley, & Wade, 1991), and the strength of connections between nodes determines one's ability to retrieve verbal memories via an activation process that is indistinguishable from what happens during normal language production (see MacKay, 1990).

Within this distributed-memory account, two fundamentally different aspects of language comprehension must be distinguished: activation of old or already established connections versus formation of new or never previously used connections (see MacKay & Burke, 1990). Many aspects of language use (e.g., comprehension of everyday words and phrases such as "bus stop") involve old or already established connections between nodes that have been formed during childhood and strengthened throughout a lifetime of use. However, sentences often communicate new or never previously expressed or encountered ideas, and representing a new idea in comprehension and memory requires formation of new cortical connections. Unlike automatic activation processes involving old or already established connections, forming new connections within the language cortex normally involves input from special mechanisms known as binding nodes located within the hippocampal system (see MacKay, 1990, for details).

Like stages-of-processing theories, distributed-memory theories have accumulated various sources of support that we review later. Here we simply spell out some distributed-memory predictions regarding H.M.'s comprehension abilities. By assumption, H.M.'s operation destroyed some (but perhaps not all) of the binding nodes required for normal language comprehension (see MacKay, 1990). Because intact hippocampal binding nodes are necessary to comprehend a full range of novel verbal concepts, H.M.'s lesion should impair his compre-

Figure 2. A sample of bottom-up connections for comprehending the word *frisbee* (from Burke et al., 1991).



hension of concepts that are new to him, but his comprehension of words, phrases, and propositions familiar to him before his operation should be intact, because reactivating already formed connections does not require input from hippocampal binding nodes.

IS H.M.'S LANGUAGE COMPREHENSION NORMAL? LACKNER (1974, EXPERIMENT 2)

Two factors motivated our focus on Lackner (1974) as the first place to look for comprehension deficits predicted under distributed-memory theories: Lackner's work is the only published study of H.M.'s language comprehension, and it preceded the decline in H.M.'s language abilities that Corkin (1984, p. 254) observed and attributed to late-onset deterioration in "general cognitive capacities" that is "characteristic of premature aging" and independent of H.M.'s particular lesion. Corkin's 1984 assessment therefore suggested that only data collected from H.M. before 1983 (e.g., Corkin, 1973; Lackner, 1974; MacKay, 1972; Marslen-Wilson, 1970) could definitively test distributed-memory accounts of the relation between language and memory.

Lackner's task required H.M. to discriminate between two types of auditorily presented sentences: ambiguous versus unambiguous. Lackner first made certain that H.M. knew what it meant for a sentence to have "more than one meaning" and then gave H.M. practice examples involving the five types of ambiguity that he would be hearing (deep structure, surface structure, particle-preposition, lexical, and phonetic ambiguities), explaining the basis for ambiguity after each. In the experiment proper, H.M. heard a mixed series of ambiguous and unambiguous sentences, and after each sentence, Lackner asked H.M. to indicate whether the sentence had one or two possible interpretations, and, if it had two, to indicate what the interpretations were. A prompt card in front of H.M. contained the instruction, "Does this sentence have one or two meanings?" After every 10 sentences, Lackner reinstructed H.M. and repeated earlier information concerning ways that sentences can be ambiguous. It is unlikely that forgetting of the instructions influenced the results because H.M. would often respond, "Nope, just one meaning," and frequently complained that repeating the instructions was unnecessary because he knew what he was supposed to be doing, pointing to the instruction card.

Turning to results, H.M. responded "two meanings" for 4 of the 20 lexical ambiguities (20% correct) and for 18 of the remaining 45 ambiguities (40% correct), for an overall hit rate of 33.8%. To evaluate these data, we compared H.M.'s hit rates with $p = 0.5$, the probability of hits based on tossing an unbiased coin on each trial and responding "two meanings" for "heads" and "one meaning" for "tails." Hit rates for "coin-toss responses" were significantly higher than for H.M. for all 65 ambiguities, $X^2 = 3.39$, and for the 20 lexical ambiguities, $X^2 = 3.60$, both p's < 0.05 using one-tail chi-square tests.[1] However, hit rates in the absence of false alarm rates are difficult to interpret, and Lackner did not provide data on H.M.'s responses to unambiguous sentences by which to assess false negatives. One might nevertheless suspect a comprehension deficit when ambiguity detection falls below coin-toss levels, and Lackner did note that H.M.'s "low frequency of ambiguity detection requires explanation." Nonetheless, Lackner's conclusions stressed that H.M.'s language comprehension is "essentially normal," that he "exhibits the same range of linguistic capabilities as normal subjects with certain quantitative differences," and that "in the absence of better norms concerning the performance of normal subjects of the same age, educational status, and intellectual level, it is uncertain how unusual H.M.'s performance should be considered."

In short, the status of H.M.'s language comprehension abilities in the only published paper on this topic remains ambiguous pending data from normal controls, and providing such control data was one of the goals of the present research. Studies 1 and 2 compared the language comprehension abilities of memory-normal controls and H.M. in Lackner (1974), and in Corkin (1973), an unpublished ambiguity detection study that adopted somewhat different procedures from Lackner's. Controls for Studies 1 and 2 were six native English speakers who were recruited through their places of employment in clerical or physical plant positions and were paid for participating. They matched H.M. as closely as possible in 1973 in age,[2] highest educational level, and verbal IQ (see Table 1) but had lower performance IQ than H.M. (see Table 1).

STUDY 1: H.M.'S LANGUAGE COMPREHENSION IN LACKNER (1974)

Study 1 replicated as closely as possible Lackner's ambiguity detection procedures with memory-normal controls for comparison with H.M. The stages-of-processing account predicted equivalent ambiguity detection by H.M. and controls whereas the distributed-memory account predicted a comprehension deficit (i.e., inferior ambiguity detection by H.M. than controls because MacKay's, 1990, distributed-memory theory requires many new connections to represent the two meanings of ambiguous sentences). Figures 3a and 3b illustrate the

Table 1. Participant characteristics: age at time of study, verbal IQ, performance IQ, and highest educational level for H.M. (Corkin, 1984), for controls 1 through 6 in Studies 1 and 2 (mean and individually), and for A.N. and same-cohort controls (7 through 12) in Study 3.

Participants	*Verbal IQ*	*Performance IQ*	*Highest degree*	*Age at time of study*
H.M.	107	126	High school	47 (in 1973) 40 (in 1966)
Controls 1–6	M = 110	M = 111	High school	M = 48 (in 1996)
Control 1	104	96	High school	50
Control 2	116	114	High school	47
Control 3	112	120	High school	46
Control 4	104	111	High school	49
Control 5	114	101	High school	49
Control 6	112	122	High school	47
Controls 7–12	NA[a]	NA[a]	High school	M = 71 (in 1995)
A.N.	127	122	High school	44 (in 1966)

[a] NA = not available.

Figure 3a. Critical connections (shown with broken lines) that differentiate the two meanings of the syntactically ambiguous sentence, *Hortense defended the man she loved with all her heart.*

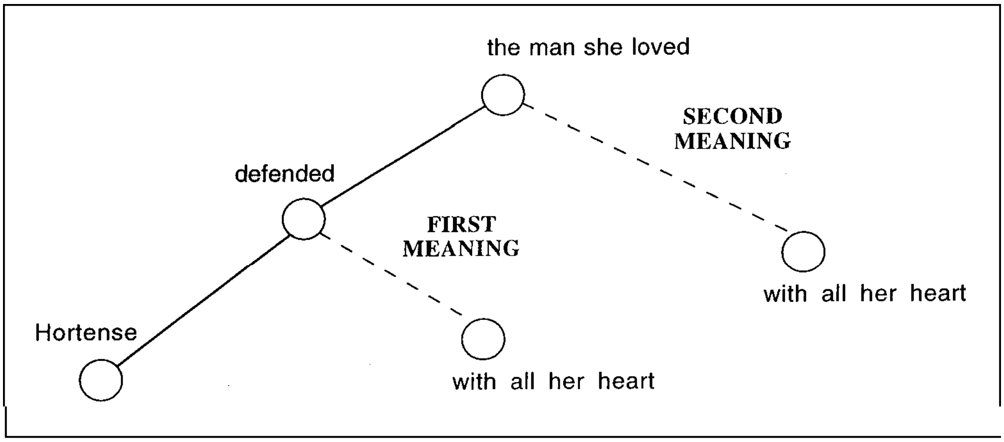
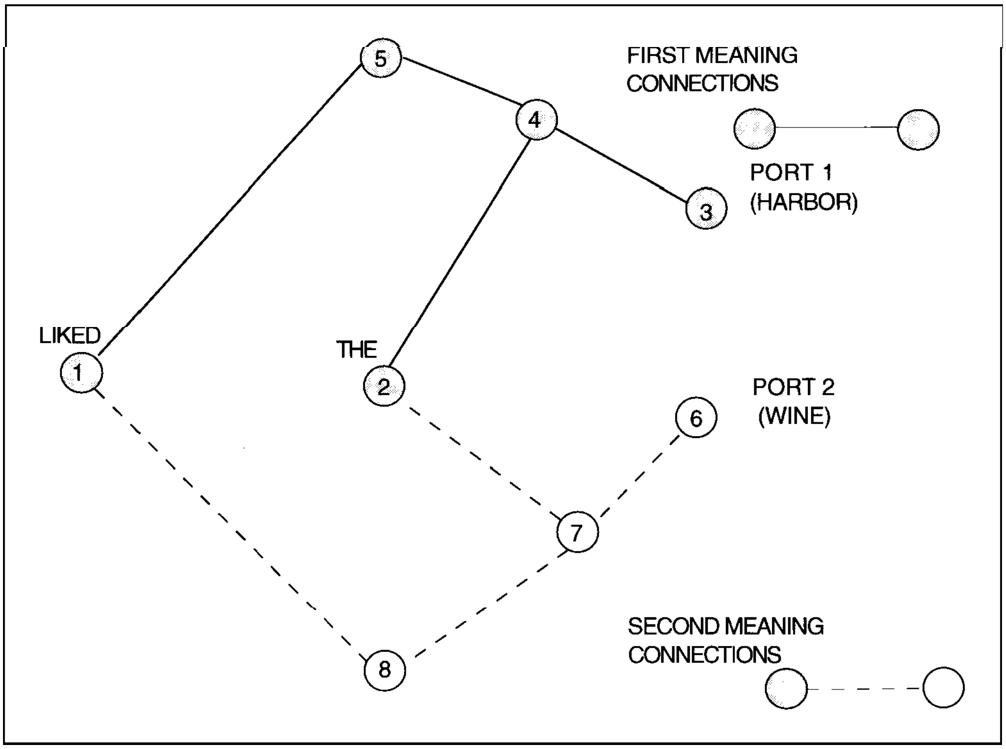


Figure 3b. Solid lines represent connections for comprehending the first meaning of the lexically ambiguous sentence, *The sailors liked the port in the evening* (by assumption, "harbor"), and broken lines represent new connections needed for comprehending the second meaning ("wine").



nature and number of such new connections for two sentences from Lackner: a syntactic ambiguity, "Hortense defended the man she loved with all her heart," and a lexical ambiguity, "The sailors liked the port in the evening." If the initially comprehended interpretation of the syntactic ambiguity is "defended with all her heart," the listener connects the node for "with all her heart" first to "defended" and then to "loved" for comprehending the second interpretation, "loved with all her heart" (see Figure 3a). Figure 3b shows the nodes for representing first the "harbor" and then the "wine" meaning of *port,* and as can be seen there, after forming the connections to represent the "harbor" meaning, at least three additional new connections must be formed to represent the "wine" meaning as a viable interpretation of this sentence. Distributed memory theories therefore predict that H.M. will experience difficulty in detecting lexical as well as syntactic ambiguities: Forming the new connections linking lexical nodes to their novel context in lexically ambiguous sentences will be difficult for H.M., even though the lexical nodes themselves were formed long before his operation.

Results and Discussion

Table 2 shows the number of correct responses to ambiguous sentences (in %) for controls and H.M. (from Lackner, 1974). H.M. responded "two meanings" for about 34% of the 65 ambiguities, whereas controls on average responded "two meanings" for about 81% of the ambiguities (see Table 2), a difference reliable at $p < 0.001$, $z = 8.74$. This difference indicates a deficit in

Table 2. Ambiguities detected and accurately described (in %) by H.M. and controls in Study 1, with SD in parentheses.

Level and type of detection	*H.M.*	*Controls (mean)*	*Control 1*	*Control 2*	*Control 3*
Ambiguities with both meanings detected	33.8	81.0 (3.51)	86.2	75.4	81.5
Ambiguities with both meanings detected and described	Not available	77.5 (25.38)	48.2	91.8	92.5
Unambiguous sentences identified as ambiguous	Not available	25.0 (7.22)	33.3	20.8	20.8

H.M.'s ability to detect the two meanings of ambiguous sentences, consistent with predictions of distributed-memory theories.

Although Lackner did not provide comparable data for H.M, Table 2 shows that controls could subsequently *describe* both meanings for only about 78% of the sentences that they accurately "detected" as ambiguous. This difference between detection versus description data indicates a potential role for guessing in the "yes-no" responses of both controls and H.M. Another problem with Lackner's "detection" paradigm concerns the "unambiguous" sentences. Controls responded "ambiguous" for about 25% of the unambiguous sentences (see Table 2), but some of their subsequent meaning descriptions indicated that these sentences inadvertently contained uncommon but genuine ambiguities: The ubiquity of ambiguity in natural language (see e.g., MacKay & Bever, 1967) makes absence of ambiguity impossible to guarantee when creating "unambiguous" sentences against which to measure guessing or false negative rates in Lackner's paradigm.

STUDY 2: H.M.'S LANGUAGE COMPREHENSION IN CORKIN (1973)

Study 2 replicated the ambiguity detection procedures of Corkin (1973) as closely as possible with memory-normal controls for comparison with H.M. In Corkin's task, H.M. knew that he would only receive ambiguous sentences and that he was to read each sentence, find the two meanings, and briefly describe them to the experimenter in a tape recorded session. Responses for the controls and H.M. in Corkin were scored from detailed verbatim transcripts that included errors, pauses, word repetitions, filled pauses ("ers" and "uhmms"), for experimenter and participants alike. Stages-of-processing accounts predicted equivalent ambiguity detection by H.M. and controls, whereas distributed-memory accounts predicted that without help, H.M. would discover both meanings in the ambiguous sentences less often than controls.

Results and Discussion

To illustrate typical response characteristics of H.M. and controls, Tables 3a and 3b provide the complete transcripts for two sentences chosen to represent different types of ambiguity, early versus late responses (numbers 6 and 30 in the sequence of 32 sentences), H.M.'s typical response length (193 words versus his overall mean, 188.7 words), and the typical coherence, grammaticality, and comprehensibility of H.M.'s responses (see MacKay et al., 1998). These transcripts also illustrate the nature of Corkin's interjections (see Tables 3a and 3b), which the present experimenter (R.S.) studied and attempted to introduce wherever appropriate for controls.

Table 4 shows ambiguity detection responses (in %) for H.M. and controls sorted into five categories: both meanings discovered without experimenter help; both meanings discovered, one with experimenter help; both meanings discovered, both with experimenter help; only one meaning discovered, but with no help; and only one meaning discovered and with experimenter help. Controls discovered both meanings without help more often than H.M. (76.8 versus 37.5% of the trials), a difference reliable at $p < 0.03$, $z = \quad 1.85$.[3] These findings reinforce the conclusion of Study 1 that H.M. exhibits a comprehension deficit, as predicted under distributed-memory accounts.

Consistent with this conclusion, H.M. exhibited an unusual strategy for detecting ambiguity: He often free associated to one word or phrase after another in the stimulus sentences until Corkin gave him the two meanings or announced success (see, e.g., SC17 in Table 3a and SC11 in Table 3b). Consider HM5 in Table 3b, for example: "Any man who is trying, they, uh, people, they say they are pleasing salesmen, well, that pleasing salesman. And uh.." Here H.M.'s *people, they, any man,* and *salesman* seem to be free associations to the word *salesmen* in *I just don't feel like pleasing salesmen* and bear no relation to the *I don't want to please* meaning that Corkin had asked him to describe (see HM5 in Table 3b). Excluding multiple free associations to the same word in the same sentence, free associations were significantly more common for H.M. than for controls (see

Table 3a. Complete transcripts for H.M. and controls responding to the ambiguous sentence *We are confident that you can make it* in Study 2. Responses of H.M. and S.C. are numbered for ease of reference in the text.

	Control 4 Response
1:	Make it—it's a goal, to get somewhere, or to make it, to do something physical.
	Control 5 Response
2:	Well they're confident that you can either arrive or that you can make something physically, like bake a pie, or . . . okay.
	Control 6 Response
3:	We have faith that you're able to attain something, or . . . we're sure that you can . . . meet an appointment perhaps? I don't know, I guess that's kind of the same thing.
	H.M. Response (from Corkin, 1973)
SC1:	Good. There's the next one. Now you're going to read the sentence, but don't say anything until you know the two meanings then tell me what they are. Go.
HM1:	Well, that is like, um people are sure that a person can do something.
SC2:	Uhmmm.
HM2:	And also I thought of, that a person is sure himself that others are sure that he can do it.
SC3:	Now, don't go beyond the sentence. I don't want you to go beyond those words right there.
HM3:	Well, uh, well
SC4:	Umhmmm
HM4:	Well, that person is confident that he can make it, too.
SC5:	Yes, but the sentence doesn't say anything about that.
HM5:	Well, that he makes it, he can make it. Well, I say, you
SC6:	I want you to stick right to that sentence.
HM6:	You can make it, means a person can make it.
SC7:	What does that mean?
HM7:	He can do it and you're, he's confident he can do it, he's sure he can do it.
SC8:	Well, we don't know anything about him, how he feels because this sentence says 'we are confident that you'
HM8:	Well, we and they're sure that he can do it, whichever it is. They're sure that that person can do that and the person is sure that he can do it. Well, I say he
SC9:	Well, we don't know about that and we're only interested in what this sentence says. We don't really care, for our present purposes, what he thinks about it.
HM9:	Well, they're confident
SC10:	This sentence right here, those very words, can be interpreted in two different ways.
HM10:	Well, if they were to speak to that, are right, have it written down like that, but they're sure or confident that a person can do it.
SC11:	That's one meaning.
HM11:	That's one meaning.
SC12:	That he can do something, that he can make something.
HM12:	To make it. And
SC13:	That he can make a model airplane.
HM13:	Yeh.
SC14:	We are confident that you can make a model airplane. That's the meaning that you saw first. Is that right?
HM14:	The people that have given that person the job, whatever it is, to do
SC15:	Uhmm, mmhumm, and he can accomplish that task.
HM15:	and
SC16:	Now, there is another meaning, a different meaning.
HM16:	Well, they are both, they do a different task, like I said before do a different thing, do a task but they can also, a person can also reach some place.
SC17:	That's right, that's the other meaning.

Table 3b. Complete transcripts for H.M. and controls responding to the ambiguous sentence, *I just don't feel like pleasing salesmen* in Study 2. Responses of H.M. and S.C. are numbered for ease of reference in the text.

	Control 4 Response
4:	I just don't feel like pleasing salesmen. I just don't feel like um making their day, making them happy, giving them a sale.
RS:	Okay
4:	And the other one's just pretty absurd. I don't, I just don't feel like pleasing salesmen—I don't feel like one of them.
	Control 5 Response
5:	I just don't feel like pleasing salesmen. Okay, you don't want to make salesmen happy or you just don't want to see pleasing salesmen. I just don't feel like pleasing salesmen. They are pleasing salesmen around you or you don't want to make the salesmen that are there happy.
	Control 6 Response
6:	I don't feel like pleasing salesmen. Uh, you don't feel like being nice to them. Uh, you don't feel like pleasing salesmen—you don't feel like talking to salesman that is sugary-sweet and trying to please you.
	H.M. Response (from Corkin, 1973)
HM1:	'I just don't feel like pleasing salesmen.' Well, I think of one thing, the person doesn't like salesmen that are pleasing to him. Uh, and that personally he doesn't like them and and personally he doesn't like them and then I think of a phrase that he would say himself, he doesn't, uh, pleasing, as conglamo, of all of pleasing salesmen.
SC1:	Uhmmm. That's one meaning.
HM2:	You say that's one, there's two meanings to it.
SC2:	Why, what's the second one?
HM3:	Because the second one I think of is, uh, salesmen that are pleasing, they are pleasing to, he doesn't like them.
SC3:	OK, that's the same meaning.
HM4:	No, it isn't.
SC4:	What's the other one then?
HM5:	Well, he doesn't like to see them around. Any man who is trying, they, uh, people, they say they are pleasing salesmen, well, that pleasing salesman. And uh..
SC5:	Read the sentence.
HM6:	'I don't like pleasing salesmen.'
SC6:	No, read it again.
HM7:	'I just don't like pleasing salesmen.'
SC7:	You're leaving out a word.
HM8:	'I just don't feel like pleasing' yep.
SC8:	Read it again, then.
HM9:	'I just don't feel like pleasing salesmen.'
SC9:	Now think about that for a second before you say anything.
HM10:	I think of, ah, that person, I, doesn't want to please the salesmen.
SC10:	Right.
HM11:	And that person, I, doesn't feel like a pleasing salesman.
SC11:	Very good. Perfect.

Table 5), $p < 0.001$ using a sign test with sentences as unit of analysis. In response to H.M.'s free associations, Corkin repeatedly asked H.M. to "make sure you have two independent meanings" or "two clear pictures of different meanings" before responding (see, e.g., SC1 in Table 3a). In addition, Corkin provided "stop signals" to indicate that H.M. had finally expressed both meanings, on one trial asking H.M. to "Say no more" when he seemed determined to continue beyond the second meaning.

Also consistent with a free association strategy were H.M.'s unusual reactions to the discovery of ambiguity. When normal individuals detect the second meaning in this task, their reactions resemble the "aha" experience in studies of problem solving: They spontaneously generate "recognition markers" such as "Okay," "Yes," "Alright," "Oh," or "Yep," often accompanied by emotional

Table 4. Detection of ambiguity (in %) as a function of experimenter help for H.M. and controls in Study 2, with SD in parentheses.

Level of detection	*H.M.*	*Controls (mean)*	*Control 4*	*Control 5*	*Control 6*
Both meanings, no help	37.5	76.8 (18.9)	59.4	74.2	96.9
Both meanings, one with help	25.0	10.6 (9.2)	15.6	16.1	0.0
Both meanings, both with help	9.4	3.2 (5.6)	0.0	9.7	0.0
One meaning, no help	15.6	9.4 (13.6)	25.0	0.0	3.1
One meaning, with help	12.5	0.0 (0.9)	0.0	0.0	0.0

reactions such as laughter, and when asked, they describe a mild sense of elation and certainty or confidence in their discovery or "breakthrough." However, such "aha" reactions were notably absent from H.M.'s transcript. When H.M. discovered ambiguities on his own, his reactions to congratulations from Corkin were subdued and uncertain. For example, H.M.'s response to Corkin's, "Right, very good, superb," was, "I wonder," as if his success had been accidental. Moreover, H.M. seemed unable to understand Corkin's descriptions of ambiguities that he failed to detect on his own. For example, consider Corkin's detailed explanation of the meanings of *position* in *The marine captain liked his new position:* "All right, I'm going to tell you what the other meaning is. One meaning, that one that you have, means that he likes his new job. The other meaning is that he likes his new physical position. In other words, he may have been standing up on watch for a great number of hours and then he gets to sit down and he likes that new position of being able to sit down. The position of his body." When asked, H.M. agreed that Corkin's just-described meanings differed but continued to use the word *position* ambiguously, as if he hadn't really understood Corkin's account.

Table 5. Six indicators of comprehension problems involving single meanings in Study 2, with total frequency for H.M. and mean frequency for controls

Comprehension problem categories	*H.M. (total)*	*Controls (mean)*
Free associations	62	0.33
Impossible interpretations	37	0.33
Unusual pronoun use	14	0.67
Failure to follow experimenter requests	22	0.33
Misreadings	10	0.33
Self-miscomprehensions	25	0.0

H.M.'s data were also unusual in other respects. Normal adults in MacKay and Bever (1967) took reliably less time to detect lexical ambiguities with conceptually related meanings (e.g., *position*; abstract versus concrete) than those with conceptually unrelated meanings (e.g., *bark*; of a dog versus a tree), whereas H.M. showed the opposite direction of difficulty for these same sentences. Also, the relation between trial number and the time to begin description for normal adults in MacKay and Bever was negative, strong and reliable, suggesting a practice effect, whereas the correlation between trial number and a measure of H.M.'s difficulties in ambiguity detection and description was positive, weak ($r = 0.13$) and unreliable ($p > 0.50$).

COMPREHENSION PROBLEMS INFERRED FROM PRODUCTION

Other aspects of H.M.'s responses, discussed next, suggested that his comprehension deficit is general in nature and not confined to detecting ambiguity.

Problems with the First Meaning

How often H.M. required help in finding the first meaning of the sentences can be inferred from two categories in Table 4: one meaning discovered, with help (12.5% of the trials), and both meanings discovered, both with help (9.4% of the trials). Combining these categories, H.M. required help in finding the first meaning on 22% of the trials, which was over 11 standard deviations more often than for the controls (3% of the trials). This finding suggests that H.M.'s comprehension deficit applies not just to ambiguity detection but to language comprehension in general because virtually all sentences are potentially ambiguous (see MacKay & Bever, 1967).

Inability to Repeat Single Meanings

On several trials, Corkin explained the two meanings of an ambiguity and then asked H.M. to give her the two meanings that she had just given him, and he was unable to do so, indicating a comprehension or memory problem associated with dual meanings. However, H.M. was also unable to comply on trials where Corkin asked him to reiterate a single meaning that she had just explained to him, strongly suggesting an inability to understand single meanings.

Free Associations to His Own Output

Besides free associating to words and phrases in the stimulus sentences, H.M. also free associated to his own output in a way that reflects on his comprehension abilities. For example, in explaining the ambiguity in *The marine captain liked his new position,* H.M. said, "because he was above them and of all, most of all . . ." Here H.M.'s "most of all" bears no relation to his topic (why the captain liked his new position) and is clearly a free association to his immediately prior "of all," as if H.M. lacked a clear understanding of what he himself was saying. Such free associations extend H.M.'s comprehension deficit to single meanings in his own just-produced output.

Impossible Interpretations

H.M. often interpreted sentences in distorted, ungrammatical, or impossible ways. For example, H.M. interpreted *We are confident that you can make it,* as, "a person is sure himself that others are sure that he can do it" (see HM2 in Table 3a), a grammatically impossible interpretation of this sentence. Such impossible interpretations were more common for H.M. than for controls (see Table 5), a difference significant at $p < 0.001$ using a sign test with stimulus sentences as the unit of analysis. H.M.'s impossible interpretations again suggest that his deficit applies to comprehending single meanings as well as to detecting ambiguity.

Unusual Pronoun Use

H.M. often used pronouns that had no clear referent or were otherwise inappropriate in describing a particular sentence. For example, when H.M. said, "because he was above them and of all, most of all . . . " in explaining why the marine captain liked his new position, the referent for the pronoun *them* is unclear and so is the referent for *of all* (see HM2, 4, 8, and 9 in Table 3b for other examples). Excluding multiple uses of the same pronoun in the same response, inappropriate pronoun use was over 33 standard deviations more common for H.M. than for controls (see Table 5). Counting each misused pronoun would have amplified this difference. For example, H.M.'s response to *We are confident that you can make it* (see Table 3a) included 20 uses of the human pronouns *he* and *they* and only two uses of the human pronouns that actually appear in this sentence, namely, *we* and *you.* Further, H.M.'s *"you're confident"* misuses the occurring pronoun *you* in describing the sentence, *We are confident that you can make it.* Moreover, H.M. continued to misuse pronouns even when Corkin repeatedly pointed out the problem to him (see, e.g., SC5, 8, and 9 in Table 3a). These findings suggest that H.M. understood neither the stimulus pronouns nor Corkin's problem with his pronoun use nor his own use of pronouns, again implicating a comprehension-production deficit for single as well as multiple meanings (see also MacKay et al., 1998).

Conceptual Conflations

Conceptual conflations are instances where H.M. seemed to blend two distinct concepts, further suggesting a problem in comprehending single meanings. An example is H.M.'s "he liked the new position on a boat that he was in charge of, the size and kind it was" in describing the "job" interpretation of *position* in *The marine captain liked his new position.* Here, the syntax of H.M.'s response equates the concepts "position" and "a boat, the size and kind it was," as if H.M. was unable to distinguish these concepts. Later in his description of the same "job" meaning, H.M. generated another conceptual conflation when he said, "he liked the new position because of being, being a passenger line," as if H.M. was unable to distinguish the marine captain's "position" from "a passenger line." Controls made no such errors, but in general, interjudge reliability was low for conceptual conflations, and unambiguous cases were too few in number to enable statistical analysis.

Failure to Follow Experimenter Requests

H.M. failed to follow experimenter requests over 40 standard deviations more frequently than controls (see Table 5). These requests that H.M. either failed to understand or ignored or both fell into three categories. One common request was to give the second meaning of an ambiguity rather than to repeat the same meaning twice. However, H.M. either did not understand these requests or did not realize that he was re-expressing the same meaning because he immediately re-expressed the same meaning again (see, e.g., HM3 in Table 3b). Despite Corkin's frequent admonitions (see, e.g., SC3 in Table 3b) and despite being specifically asked not to, H.M. repeated the same meaning more often than controls (see Table 5), a difference reliable at $p < 0.01$ using a sign test with stimuli as unit of analysis.[4]

Calls for clarification were another common experimenter request. H.M. frequently used key ambiguous words from the stimulus sentences without disambiguation, whereupon Corkin either indicated that she did not understand H.M.'s use of that particular word or asked H.M. to use a different word in explaining the two meanings. However, H.M. seemed not to understand these requests or the nature of Corkin's problem in understanding him because he typically continued to use the same ambiguous word, again ambiguously (see, e.g., HM3 in Table 3b).

A third category concerned Corkin's repeated requests to read a stimulus sentence aloud before beginning his response (see, e.g., SC1 in Table 3a): H.M. failed to comply with 71% of these requests, as if reading a short sentence aloud was difficult for him.

Misreadings

After H.M. repeatedly tried and failed to describe the two meanings of a sentence, Corkin generally insisted that he read the sentence aloud, whereupon H.M. misread it, often several times (see, e.g., HM6–7 in Table 3b), and he had to read one sentence six times before getting it right. Misreadings were more than 16 standard deviations more common for H.M. than the mean for controls (see Table 5), even though H.M. actually read fewer sentences aloud than did controls. Interestingly, H.M.'s single-word misreadings (e.g., *Professor Smith* misread as *Mr. Smith*) were usually semantically similar to the original, but his other misreadings usually altered the original sentence in syntax and eliminated one of the ambiguous meanings (e.g., "John is the one to help today" misread as "John is the one that helped today," and "On top of everything there was a tarpaulin" misread as "On top of everything was a tarpaulin.") H.M.'s misreadings therefore suggest a problem in registering sentential meaning and syntax but not in perceiving orthography or phonology per se.

Self-Miscomprehensions

H.M. generated one meaning for a sentence on 88% of the trials, but for the second meaning, he often reiterated the first meaning, sometimes immediately and with only minor rewording, and insisted that his restatement differed from his first interpretation (see, e.g., HM2 in Table 3b). In a conservative analysis that only scored instances where participants misdescribed a restated meaning as denoting two different meanings,[5] self-miscomprehensions were more common for H.M. than for controls (see Table 5), a difference reliable at $p < 0.001$ using a sign test with stimulus sentences as the unit of analysis. This finding suggests that H.M. often failed to clearly grasp his first interpretation or the fact that it was identical to his second.

To summarize, nine aspects of H.M.'s responses suggested that H.M. has a comprehension deficit that is general in nature rather than specific to ambiguity detection. Corroborating this conclusion, many of these same response characteristics (free associations, unusual pronoun use, conceptual conflations, miscomprehensions, failures to follow experimenter requests, and self-miscomprehensions) also appeared in Marslen-Wilson's (1970) transcript of H.M.'s language production as he described pictures and conversed about early childhood experiences (see MacKay et al., 1998).

STUDY 3: H.M.'S LANGUAGE COMPREHENSION IN MACKAY (1972)

Study 3 describes ambiguity detection tests conducted in 1965-1966 involving 20 Harvard undergraduates, H.M., and A.N., a patient with bilateral damage to the frontal lobes who matched H.M. in age, education, and performance IQ (see Table 1). Procedures otherwise resembled Study 2 except for the dependent measures: time to begin describing both meanings of the ambiguous sentences and how many descriptions were attempted in the 90-sec time limit. These measures circumvent a "memory interpretation" applicable to Study 2, namely, that expressing the first perceived meaning caused forgetting of the second, already perceived meaning of the ambiguity: Time to begin expressing the first perceived meaning in Study 3 cannot reflect forgetting of the second meaning, perceived or not. Comparing H.M. and A.N. also tested whether H.M.'s problems in detecting ambiguities are specific to his particular lesion or result from brain damage per se. Despite obvious differences in age, IQ, and educational status, comparing A.N. with the Harvard undergraduates helped to highlight A.N.'s competence in detecting ambiguities. To test whether H.M.'s comprehension deficits reflect a cohort effect, Study 3 also compared ambiguity detection of H.M. and Controls 7 through 12 in Table 1, memory-normal older adults who were born about the same year as H.M.

Results and Discussion

There was no overlap in distribution of means between H.M. and the other three classes of participants in either mean time to begin ambiguity description or the number of descriptions attempted (see Table 6). A.N. discovered and accurately described both meanings for all 32 ambiguities within 4.5 sec, that is, at least as fast as the Harvard undergraduates (4.7 sec) and much faster than H.M. (49.2 sec), who neither indicated detection nor attempted to describe the two meanings within the 90-sec time limit for 14 of the 32 ambiguities (44%). Moreover, unlike H.M.'s ambiguity descriptions, A.N.'s descriptions were invariably clear and never left any doubt as to which meaning he had discovered first or whether he had discovered both meanings. These findings suggest that H.M.'s deficits in comprehending ambiguities are attributable to his particular lesion, and

Table 6. Time to begin ambiguity description and number of descriptions attempted in 90 sec by H.M., a frontal patient (A.N.), Harvard undergraduates in MacKay (1972), and same-cohort older adults run in 1995.

MacKay (1966) response measures	*H.M. (bilateral damage to amygdala and hippocampal system)*	*A.N. (bilateral frontal lobe damage)*	*Harvard undergraduates*	*Same-cohort older adults*
Time to begin ambiguity description	49.2 sec	4.5 sec	4.7 sec	19.7 sec
Number of descriptions attempted in 90 sec	18 (56%)	32 (100%)	32 (100%)	32 (100%)

not to brain damage per se. Indeed, the similarity of data for A.N. and the undergraduates suggests that damage to frontal lobe structures may have little effect on ambiguity detection, contrary to Cohen and Eichenbaum's (1993, p. 196) suggestion that frontal but *not* hippocampal structures organize abstract semantic processes.

Finally, despite being 71 rather than 40 years old when tested in 1995, the same-cohort older adults responded within the 90-sec time limit for 100% of the ambiguities and began to describe the meanings in 19.7 sec, that is, much faster than H.M. (49.2 sec). Like the age-matched controls in Study 1, these same-cohort older adults accurately *described* both meanings for 90.6% of the sentences, and their meaning descriptions were always comprehensible as to which meaning(s) they had discovered (unlike H.M.), replicating results of Studies 1 and 2 and ruling out cohort accounts of H.M.'s deficit.

GENERAL DISCUSSION

Does H.M. Have a Language Comprehension Deficit?

More than 10 sources of data indicate that H.M. had a comprehension deficit in 1966–1973: His ambiguity detection in Study 1 was below coin-toss levels of accuracy (50%) and below the least proficient normal control. His ambiguity detection also lagged behind normal controls in Study 2 and the same-cohort older controls in Study 3. As an indirect reflection of this comprehension deficit, H.M. required significantly more help than controls in discovering the initial meanings of sentences in Study 2, and relative to controls, generated significantly more impossible interpretations, free associations, misreadings, and unusual pronominal references. H.M.'s failures to follow Corkin's requests (e.g., to read sentences aloud) is likewise consistent with a comprehension deficit. So are H.M.'s self-miscomprehensions, cases where H.M. insisted that repeating an interpretation constituted a different interpretation. None of these results are readily explained in terms of a "pure memory" deficit as defined within stages-of-processing theories. For example, the main results of Studies 2 and 3 are difficult to explain in terms of an inability to remember the stimulus sentences: The sentences were short and syntactically simple, H.M. had unlimited opportunity to read and reread them, and Corkin repeatedly drew H.M.'s attention to the words in front of him (see, e.g., SC5 in Table 3a). Subsidiary results such as H.M.'s misreadings, impossible interpretations, and free associations to his own immediately prior output are also difficult to ascribe to storage rather than comprehension processes.

Several additional observations are of interest in this context: After completing the present study, we learned that the 1984 decline in H.M.'s language abilities almost certainly reflected a change in H.M.'s 1984 IQ test rather than a genuine change in his language abilities (Corkin, personal communication, January 1997). This means that H.M.'s deficits on the reporter's test of language production, on the token test of language comprehension, and on tests of semantic and symbolic verbal fluency (Corkin, 1984) may be part of the same pattern of deficits seen in his 1966–1973 performance examined here. In short, all of H.M.'s language deficits observed to date may stem from his 1953 lesion rather than some late-onset factor such as premature aging.

Another observation of interest is that ambiguity detection was good but not perfect for age-matched controls in Study 1 and 2 and for the 71-year-olds in Study 3. Moreover, ambiguity detection required over 4.5 sec for A.N. and young normals in Study 3, an extremely long time in the realm of sentence comprehension. Why is ambiguity detection so difficult, especially for older adults, and even more so for H.M.? Ambiguity detection always requires the formation of two sets of new connections from the same set of nodes (see Figures 3a and 3b), and young adults under extreme time pressure are known to experience difficulty with this double connection process (see, e.g., MacKay & Miller, 1996). Moreover, this double connection process causes greater difficulty for older than young adults (see MacKay, Miller, & Schuster, 1994) because older adults suffer from a general deficit in the ability to form new connections that is not specific to hippocampal functioning (for recent reviews, see MacKay & Abrams, 1996; and Burke & MacKay, 1997). It thus makes sense that a 47 year old with an additional and much more serious connection formation deficit attributable to his bilateral hippocampal lesion would experience greater difficulty with this double connection process than young adults, age-matched controls, and 71-year-old adults.

Were H.M.'s Comprehension Deficits Motivational?

In view of established relations between limbic structures and motivation, initiative, and affect (see, e.g., Hebben, Corkin, Eichenbaum, & Shedlack, 1985), it is important to determine whether lack of motivation underlies H.M.'s comprehension difficulties in the present work. Perhaps H.M. was simply not interested in detecting the two meanings of ambiguous sentences. Contrary to this motivational hypothesis, however, H.M. often expressed frustration with his inability to find the two meanings of sentences in Corkin's task, and he always remained focused on the general task or goal, with no tendency to intrude entirely new and unrelated topics of discussion due to boredom or waning interest (see, e.g., Tables 3a and 3b). Also, H.M. generated over 4 times as many words per response as did controls in Corkin's task (see MacKay et al., 1998), suggesting that H.M. was at least as motivated as controls.

H.M.'s "Sense of Humor"

Present results call for detailed scrutiny of Milner et al.'s (1968) evidence that in everyday life, H.M. "gets the point of jokes, including those turning on semantic ambiguity." Consider two examples of H.M.'s "New England" sense of humor discussed in Hilts (1995, pp. 115–116). In one example, Corkin calls H.M. "the puzzle king," and his response, "Yes, I'm puzzling," is interpreted as a verbal quip. However, present findings suggest some simpler possibilities, namely, that H.M. misheard "puzzle king" as "puzzling" or that he misinterpreted the novel noun phrase *puzzle king* to mean "puzzling." The second example of H.M.'s New England humor occurred after a lengthy car trip; Teuber says, "That was a long drive . . . Are you stiff, Henry?" and H.M. replies, "Nope, I haven't had a drop," as if he comprehended but chose to ignore the obviously intended "psycho-physiological stiffness" meaning of *stiff* and created a quip based on a rare interpretation of *stiff* that resembles "inebriated." Again, however, present findings suggest some simpler possibilities. One is that H.M. did not link the long car trip with psycho-physiological stiffness, as required for comprehending this intended meaning of *stiff,* and his "haven't had a drop" was a free association to the common noun meaning of *stiff* (i.e., "a drunk"). Consistent with this comprehension deficit hypothesis, but not with the creative quip hypothesis, Hebben et al. (1985) showed that H.M. finds his own psycho-physiological states remarkably difficult to detect and evaluate. Also difficult to explain under the creative quip hypothesis, H.M.'s ambiguity detection fell below coin-toss levels of accuracy in Study 1, fell well short of even the least skilled normal controls in Studies 1 through 3, and was so slow in Study 3 as to render the concept "quip" inapt.

Are Memory and Language Autonomous?

H.M.'s comprehension deficit contradicts the separation of comprehension, storage, and retrieval postulated in stages-of-processing theories. What then of earlier dissociations between memory disorders and comprehension disorders? Wernicke (1874, p. 36–46) and others argued that comprehension disorders (Wernicke's aphasia) are independent of memory disorders (anterograde amnesia), and vice versa, as if comprehension represents a separate stage involving units and processes that are dissociable from memory storage and retrieval. However, more recent techniques and data have called these earlier stage dissociations into question (see, e.g., Kempler, Curtiss, Metter, Jackson, & Hanson, 1991; and MacKay, 1998, for a general review). For example, Tyler (1988) used highly sensitive implicit measures of comprehension to directly address the separability of memory and comprehension as distinct, sequentially ordered stages. Tyler first replicated the observation that, unlike normal controls, Wernicke's aphasics are unable to respond accurately to after-the-fact questions about a presented sentence, such as, "Is that sentence grammatical?" Tyler's next step was to present the aphasics with similar sentences in a priming task where effects of ungrammaticality could be determined implicitly *at the time of processing.* In this on-line task, the Wernicke's aphasics responded like normal controls, contrary to the view that their "pure comprehension deficit" is separate from "memory" assessed after the fact (see also MacKay & Abrams, 1996, who review a wide range of data that further undermine Wernicke's arguments for sharply separating comprehension and memory).

To summarize, on the basis of case H.M., memory storage and language comprehension cannot be considered autonomous, sequentially ordered processing stages. H.M.'s semantic-level *production deficits* (see MacKay et al., 1998; also Table 5) raise parallel questions regarding the autonomy of memory storage and language production and add an even more difficult question for stages-of-processing theories: How can H.M.'s relatively restricted lesion simultaneously cause related deficits in both language comprehension and language production, supposedly autonomous stages that occupy opposite ends of the stages-of-processing spectrum (see Figure 1)?

Can Response Interference Explain H.M.'s Language Comprehension Deficit?

Because amnesics commonly exhibit increased sensitivity to interference in encoding and retrieval (see, e.g., Shapiro & Olton, 1994), we developed an interference hypothesis for explaining aspects of Study 2 data. Under this hypothesis, H.M. perceived and weakly represented both meanings of the ambiguous sentences via extremely fragile new connections, and the process of describing the first meaning interfered with his ability to retrieve and describe the second meaning. Describing the first meaning automatically strengthened the connections to nodes for that meaning (see Figures 3a and 3b), thereby reducing the "relative competitiveness" of nodes for the second meaning (see MacKay, 1990). As a result, the first meaning was activated repeatedly (explaining H.M.'s meaning repetitions), and the second meaning became increasingly difficult to activate. However, results of Studies 1 and 3 are problematic for this response interference hypothesis: Effects of describing one meaning on perceiving the other are irrelevant to the yes-no responses in Study 1 and so is the time to begin describing meaning one in Study 3.

How General Is the Link between Comprehension and the Hippocampal System?

H.M.'s comprehension deficit raises two questions: Do other patients with medial temporal lobe damage exhibit similar difficulties in comprehending ambiguous sentences? And can brain mechanisms underlying H.M.'s comprehension deficit be localized more precisely? Recent data provide a clear answer to both questions. Zaidel, Zaidel, Oxbury, and Oxbury (1995) presented an ambiguity comprehension task resembling Corkin's to 33 patients with unilateral left- or right-sided surgical lesions to the amygdala and anterior hippocampus. Patients with left-side lesions discovered both interpretations of the ambiguities less often than patients with right-side lesions, who did not differ from age-matched norms. These data support the hypothesis of Milner (1975), O'Keefe and Nadel (1978), and others that the left hippocampal system in humans specializes in language, whereas the right hippocampal system specializes in other functions. Also, even though patients with left-side lesions in Zaidel et al. were relatively young when tested (mean = 24.5 years), had epilepsy for a relatively brief time, and were tested relatively soon after surgery (mean = 4.4 years), they exhibited the same comprehension deficit as H.M. This suggests that H.M.'s left hippocampal system damage, rather than his cerebellar damage or residual epilepsy (see "Methods" section), is responsible for his language comprehension deficit. However, it remains to determine whether patients with unilateral left medial temporal lobe damage exhibit *production* deficits resembling H.M.'s (see Study 2; also MacKay et al., 1998) and comprehension deficits of the same nature, generality, and severity as H.M.'s.

How Do Language Comprehension and Memory Overlap?

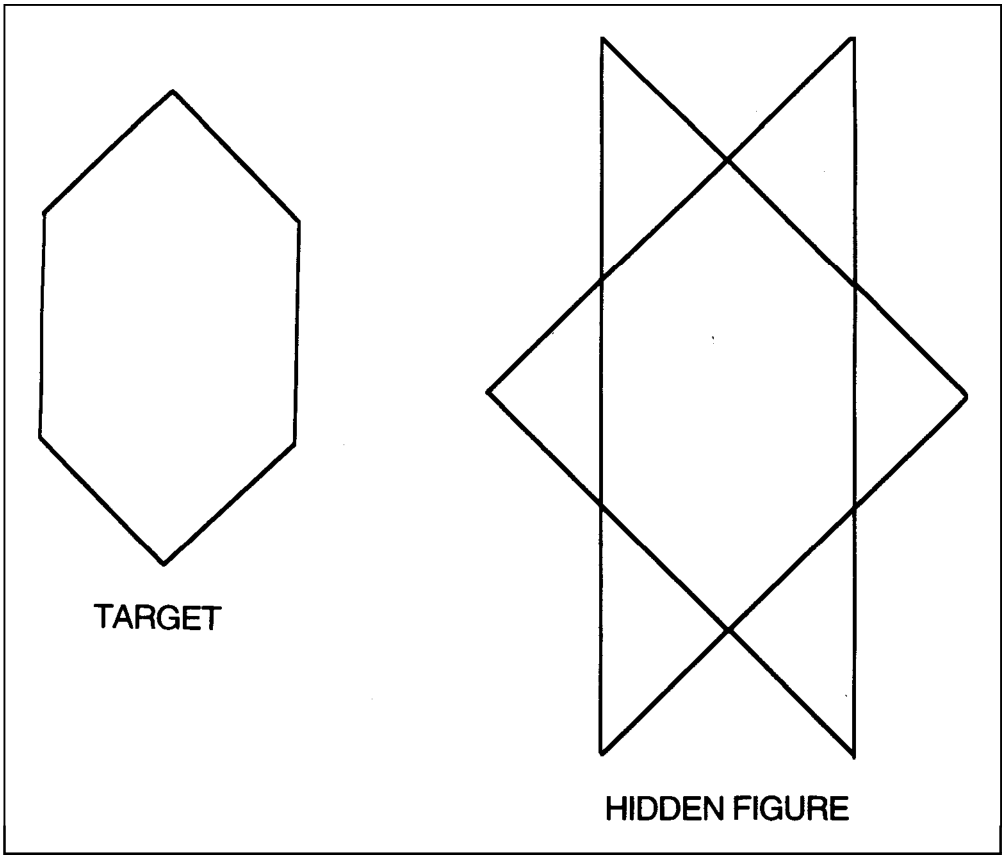
Why should someone with a binding deficit, or difficulty in forming new connections, experience problems in comprehending novel aspects of language? The answer is that forming new connections is integral to comprehending new ideas in nonformulaic language. However, not all language comprehension requires new connection formation: comprehending familiar words and phrases only requires activation of nodes with already formed connections. Activation in turn increases the strength of these already formed connections, a process known as engrainment learning that does not require hippocampal input (MacKay, 1990).

All of these processes involving existing nodes and connections are intact in case H.M., which explains why he readily comprehends familiar words and phrases and is unimpaired on indirect tests of memory involving these words. For example, H.M. will recognize the word *magic* presented in noise more readily if he has recently processed the word *magic* in an incidental context earlier in the experiment (see Keane, Gabrieli, & Corkin, 1987). This implicit memory effect is known as repetition priming and is a direct consequence of engrainment learning involving well-established cortical connections that were formed in early childhood for perceiving and comprehending the word *magic.* However, H.M. is unable to voluntarily recall having seen the word *magic* earlier (see Keane et al., 1987) because this explicit memory task requires the formation of new connections that link *magic* to its earlier context of use, not unlike the new connections required for comprehending novel sentences composed of familiar words (see Grafman & Weingartner, 1996, for a similar view).

Does H.M. Exhibit Similar Binding Deficits in Visual Perception?

H.M.'s visual perception seems to mirror his binding deficits in language comprehension. Consider H.M.'s deficit in discovering "hidden figures" such as the one illustrated in Figure 4 (see Milner et al., 1968). This hidden figure deficit is theoretically important because H.M. exhibits "superior performance on many perceptual tasks" (Milner et al., 1968) and because it contradicts the claim that H.M. has entirely intact perceptual representations (Schacter, 1990) and entirely intact perceptual skills or procedural memory (Squire, 1987, pp. 152–169; and Cohen & Eichenbaum, 1993, pp. 49–219). Milner et al. (1968) suggested that H.M.'s "need to keep one figure in mind whilst analyzing another more complex figure, might . . . overtax his short-term memory system," but this account of H.M.'s hidden figure deficit conflicts with two observations: Both the target and hidden figure are continuously available to perception in this task, and

Figure 4. The hidden figure task: H.M. must discover and trace out the target figure in the simultaneously presented hidden figure display.



H.M.'s immediate memory is intact for both verbal and nonverbal materials (Wickelgren, 1968). A more plausible explanation is that H.M.'s binding deficit impairs the connection formation processes necessary to represent the hidden figure and to inhibit the embedded and overlapping distractor lines and figures in this task. In general, however, H.M.'s perceptual abilities are intact because most perceptual tasks do not require new connection formation but only H.M.'s unimpaired ability to activate already formed nodes for recognizing familiar forms and objects. In short, H.M.'s data are consistent with the hypothesis that forming new connections is central to the "binding problem" in both language and visual perception.

To conclude, several unique aspects of H.M.'s language deficits must be emphasized. One is that the hippocampal system contributes to language comprehension and production by helping to form new cortical connections, but it *does not* activate already formed nodes for expressing these cognitive skills (see MacKay, 1990). Another important point is that existence of a comprehension deficit *does not* imply that H.M.'s memory deficits are solely due to his comprehension deficit, a stages-of-processing argument. Rather, our data support a distributed-memory argument that there is no dividing line between language and memory for verbal materials and that H.M.'s binding deficits are simultaneously responsible for both his language deficits and his memory deficits.

A final caveat concerns the relation between H.M.'s lesion and the binding nodes for language. According to MacKay (1990), thousands of binding nodes are required for normal comprehension, production, and acquisition of language (at all ages), and different sorts of binding nodes are specialized for conjoining different classes of never previously linked units. For example, when a normal person initially comprehends, produces, or learns the phrase *binding nodes,* a binding node is engaged for conjoining participles (e.g., *binding*) and nouns (e.g., *nodes*) to form a noun phrase (here, *binding nodes*). If the entire left hippocampus in humans is devoted to language, some of the language-specific binding nodes in H.M.'s posterior hippocampus have probably been spared. Nonetheless, we cannot be certain how many or which language-specific binding nodes have been destroyed or spared in case H.M.

Why Were H.M.'s Comprehension Deficits so Difficult to Detect?

The number and magnitude of differences between H.M. and memory-normal controls in the present data raise one final question: Why was H.M.'s language comprehension deficit overlooked in so many previous studies?

Some earlier observations may have allowed H.M. to compensate for his language comprehension deficits by use of situational or nonlanguage cues, cues that are not available in a specifically semantic task such as detecting and describing the meanings of ambiguous sentences. In addition, the widely held stages-of-processing framework may have encouraged "pure memory" interpretations of behaviors that have other possible explanations. An example is the observation that H.M. never wrote notes to himself as reminders, even though he repeatedly complained of his difficulties in remembering (see Hilts, 1995). Under the stages-of-processing framework such notes were not worth creating because H.M. would forget to use them, but present results suggest that H.M. may have had trouble reading and comprehending such notes even if he or others had created them. Finally, previous researchers may have considered H.M.'s comprehension deficits relatively trivial given his normal verbal IQ and not worth mentioning relative to the problems of aphasics with more conspicuous language deficits or relative to H.M.'s other debilitating problems (see Milner et al., 1968). Deficits that would suggest impairment in an otherwise normal individual may thus have been overlooked, a process that the generally accepted stages-theory of language may have facilitated. Then, too, the converse role of new theories in highlighting overlooked facts as facts deserves equal emphasis.

METHODS: STUDIES 1, 2, AND 3

Participants, Stimuli, and Procedures: Study 1

Participants were H.M. and Controls 1, 2, and 3 in Table 1. Based on MRI data (Corkin, Amaral, González, Johnson, & Hyman, 1997), H.M.'s bilaterally symmetric surgical lesion included virtually all of the amygdaloid complex and the anterior rostrocaudal extent of the intraventricular portion of the hippocampal formation. Spared were the parahippocampal cortex, temporal stem, collateral sulcus, including portions of the ventral perirhinal cortex, and the caudal 2 cm of the hippocampal body, although the functional status of this spared 2 cm is currently unknown. The suborbital route of the surgery also spared virtually the entire neocortex, including all neocortex correlated with language comprehension (see Hart & Gordon, 1990). However, MRI also revealed a major cerebellar lesion attributable to H.M.'s large and long-term doses of dilantin and other drugs for treating epilepsy (Corkin et al., 1997).

The materials were 65 ambiguous sentences from Lackner (1974, Appendix 2a and b) plus 24 unambiguous sentences of similar style and length created in our lab to replicate Lackner's procedures. Controls followed Lackner's procedures except for the following steps, which were unnecessary for memory-normal controls: a large number of practice examples, frequent reinstruction and prominent positioning of written instructions to ensure recall of the task, and repeated tutorials on different types of ambiguity. After three practice sentences with different types of ambiguity, controls heard the 89 sentences read in a neutral or monotone voice, one at a time in the same order as H.M. Participants listened to each sentence and indicated as soon as possible whether it had "one or more than one possible meaning." Participants who responded "more than one" or "two" briefly described both meanings in any order.

Participants, Stimuli, and Procedures: Study 2

Materials used by MacKay (1972) and Corkin (1973) consisted of 32 ambiguous sentences developed by MacKay and Bever (1967) to have approximately equal "bias" or likelihood of eliciting either meaning. All were short (mean length, 7.7 words; range, 7 to 9 words), and syntactically simple (i.e., simple active declaratives, with no subordinate or embedded clauses). Controls 4 through 6 in Table 1 saw the sentences one at a time typed in large font on sheets of paper that remained in continuous view until the next sentence. Presentation order was the same as in Corkin: lexically ambiguous sentences (e.g., "The office of the president is vacant"), followed by surface structure ambiguities (e.g., "They talked about the problem with the mathematician"), followed by deep structure ambiguities (e.g., "The mayor asked the police to stop drinking"); (see MacKay & Bever, 1967, for all 32 sentences). Three practice sentences illustrated the types of ambiguity (fewer than Corkin gave H.M.), with instructions asking controls to find both meanings of each sentence and briefly describe them in any order.

Participants, Stimuli, and Procedures: Study 3

Participants were H.M.; A.N., a patient with a large, bilateral injury to the frontal lobes; 20 Harvard undergraduates whose data were reported in MacKay and Bever (1966); and Controls 7 through 12, healthy older adults (mean age 71 in 1995) who were recruited via advertisements and were paid $10 plus travel costs. Their highest education level (high school) and date of birth (mean, 1924; range, 1919–1931) resembled H.M.'s (see Table 1), but their forward digit span was 7, slightly higher than H.M.'s (6). A.N. (age 44 in 1966) had a 1964 verbal IQ (Wechsler-Bellevue II) of 127, performance IQ of 122, and high school as highest degree (see Table 1).

MacKay and Bever (1967) describe in detail the materials and procedures used in Study 3 for H.M., A.N., and the undergraduates: The 32 ambiguous sentences were typed on 3 × 5-inch index cards that were thoroughly shuffled for each participant, who found the two meanings of each sentence as quickly as possible and described them in the order perceived. Time from sentence presentation to when participants said "Yes," indicating ambiguity detection, was determined by stop watch. If participants failed to respond within 90 sec, they were told the two meanings and the next trial was begun. For H.M., the instructions were summarized on a card taped to the table in front of him and periodically repeated by the experimenter. Unlike Corkin (1973), the experimenter never interjected during H.M.'s responses, making it difficult to determine whether H.M. had in fact discovered either or both meanings before beginning his response (see Studies 1 and 2). For the six same-cohort controls, the 32 ambiguous sentences were modified to equate the bias of the two meanings of the ambiguities for young and older adults tested in 1995. Specifically, three groups of 15 to 26 young and older adults estimated the relative likelihood of the two meanings of each sentence (in %) in successive rating sessions. After the initial rating session, the sentences were modified to equate bias, and rated again in session 2, and by session 3, bias was approximately equal for young and older participants. In the experiment proper, the six older adults saw the sentences presented via Macintosh II using Psyscope software. To indicate ambiguity detection, they pressed a key that stopped the computer clock. Then, with the sentence still visible on the screen, they reported the two meanings in the order detected.

Acknowledgments

The authors thank Dr. Suzanne Corkin for generously providing her unpublished transcript of H.M.'s ambiguity detection responses, for furnishing background information on A.N., and for helpful comments on a previous version of this paper. We thank Dr. William Marslen-Wilson for generously providing his unpublished transcript of H.M.'s conversational speech; Drs. Barbara Knowlton, Yvonne Lippa, and Michelle Miller for helpful comments on an early draft of this paper; and Dr. Endel Tulving for his important observations regarding the correct but frequently misunderstood way to represent chance levels of responding in two-alternative forced choice tasks (see Note 2). We thank Dr. Lise Abrams, Dr. Lori James, and Elizabeth Bloombaum for technical and research assistance. Don MacKay thanks Drs. Hans-Lucas Teuber and Brenda Milner for permission to study A.N. and H.M. in 1966 and gratefully acknowledges support from National Institute on Aging grant R01 AG 09755 and from the Max-Planck-Institute for Psychological Research in Munich, Germany. Deborah Burke gratefully acknowledges support from National Institute on Aging grant R01 AG 08835.

Reprint requests should be sent to Donald G. MacKay, Psychology Dept., University of California, Los Angeles, Los Angeles, CA 90095-1563.

Notes

1. Note that we do not refer to 50% as a chance level of responding in Lackner's two-alternative forced choice decision task. The only logically correct definition of chance level responding is a 0% difference between hit rates and false alarm rates: chance levels for hit rates do not exist in the absence of false alarm rates. This relieves us of the difficult task of developing a substantive explanation of how H.M. or anyone else who is not deliberately falsifying his or her responses can perform significantly below chance: H.M.'s below-coin-toss-level of hits can be dismissed as a possible bias in favor of the response "unambiguous" in Lackner (1974).
2. We chose 1973 as the closest date to the Lackner (1974) and Corkin (1973) studies. However, H.M.'s closest IQ test was conducted in 1977 when H.M. was 51 rather than 47 years old. Consequently, H.M.'s "true" verbal IQ in 1973 may have been slightly higher than the 1977 estimate in Table 1 (see Corkin, 1984). Also, because H.M.'s 1977 IQ test (the Wechsler-Bellevue, Form 1) is currently out of print, we tested controls on the (1983) Wechsler Adult Intelligence Scale-R (WAIS-R).
3. Correlations between how often controls in Studies 1 and 2 detected both meanings of the ambiguities without help and their age (range, 46 to 50 years), verbal IQ (range, 104 to 112), and performance IQ (range, 96 to 122) were small and nonsignificant ($r = 0.07, p > 0.45$ for age; $r = 0.23, p > 0.33$ for verbal IQ; and $r = 0.23, p > 0.33$ for performance IQ) and for Full Scale IQ ($r = 0.23, p > 0.33$), which was virtually identical to the same correlation for the 33 participants in Zaidel et al., 1995 ($r = 0.26, p > 0.1$). These weak and unreliable correlations make it difficult to attribute present ambiguity detection results to minor differences in age or IQ between H.M. and controls (see Table 1).
4. For further data on H.M.'s repetition tendencies in language production and memory tasks, see MacKay, Burke, and Stewart (1998), who argue that H.M.'s repetitions reflect an attempt to compensate for his binding deficit rather than a general tendency to perseverate.
5. This criterion served to exclude cases such as HM1 in Table 3b, where H.M.'s "personally he doesn't like them and and personally he doesn't like them" could be interpreted as sentence-level stuttering.

REFERENCES

Burke, D. M., & MacKay, D. G. (1997). Memory, language, and aging. *Philosophical Transactions of the Royal Society, London, Biological Sciences, 352,* 1845–1856.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30,* 542–579.

Carpenter, G., & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in the Neurosciences, 16,* 131–137.

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

Corkin, S. (1973). H. M.'s detection and description of ambiguous meanings. Unpublished transcript, Dept. of Brain and Cognitive Sciences, M.I.T., Cambridge, MA.

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.

Grafman, J., & Weingartner, H. A. (1996). A combinatorial-binding and strength (CBS) model of memory: Is it a better framework for amnesia? In D. Herrmann, M. Johnson, C. McEvoy, C. Hertzog, & P. Hertel (Eds.), *Basic and applied memory: Theory in context* (pp. 259–276). Hillsdale, N. J.: Lawrence Erlbaum.

Hart, J., & Gordon, B. (1990). Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology, 27,* 226–231.

Hebben, N., Corkin, S., Eichenbaum, H., & Shedlack, K.

(1985). Diminished ability to interpret and report internal states after bilateral medial temporal resection: Case H. M. *Behavioral Neuroscience, 99,* 1031–1039.

Hilts, P. J. (1995). *Memory's ghost: The strange tale of Mr. M. and the nature of memory.* New York: Simon & Schuster.

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.

Kempler, D., Curtiss, S., Metter, E. J., Jackson, C. A., & Hanson, W. R. (1991). Grammatical comprehension, aphasic syndromes, and neuroimaging. *Journal of Neurolinguistics, 6,* 301–318.

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.

MacKay, D. G. (1972). *The search for ambiguity by an amnesic patient: Implications for the theory of comprehension, memory and attention.* Unpublished manuscript, Psychology Dept., University of California, Los Angeles.

MacKay, D. G. (1990). Perception, action, and awareness: A three-body problem. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 269–303). Berlin: Springer-Verlag.

MacKay, D. G. (1998). Stage theories refuted. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science.* Oxford: Blackwell (in press).

MacKay, D. G., & Abrams, L. (1996). Language, memory, and aging: Distributed deficits and the structure of new-versus-old connections. In J. E. Birren & W. K. Schaie (Eds.), *Handbook of the psychology of aging,* 4th Ed. (pp. 251–265). San Diego: Academic Press.

MacKay, D. G., & Bever, T. G. (1967). In search of ambiguity. *Perception & Psychophysics, 2,* 193–200.

MacKay, D. G., & Burke, D. M. (1990). Cognition and aging: A theory of new learning and the use of old connections. In T. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 281–300). Amsterdam: North-Holland.

MacKay, D. G., Burke, D. M., & Stewart, R. (1998). 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., & Miller, M. D. (1996). Can cognitive aging contribute to fundamental psychological theory? Repetition deafness as a test case. *Aging, Neuropsychology, and Cognition, 3,* 169–186.

MacKay, D. G., Miller, M. D., & Schuster, S. P. (1994). Repetition blindness and aging: Evidence for a binding deficit involving a single, theoretically-specified connection. *Psychology and Aging, 5,* 52–65.

Marslen-Wilson, W. (1970). Biographical interviews with H. M., Unpublished transcript, Applied Psychology Unit, Cambridge, England.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why are there complementary learning systems in the hippocampus and neocortex: Insights from the success and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457.

Milner, B. (1975). Psychological aspects of focal epilepsy and its neurological management. In D. P. Purpura, J. K. Penny, & R. D. Walter (Eds.), *Advances in neurology, vol. 8* (pp. 299–320). New York: Raven.

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: Erlbaum.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map.* Oxford: Clarendon.

Schachter, D. L. (1990). Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate. In A. Diamond (Ed.), *Development and neural basis of higher cognitive functions* (pp. 543–571). New York: New York Academy of Sciences.

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

Shapiro, M. L., & Olton, D. S. (1994). Hippocampal function and interference. In D. L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 87–117). Cambridge, MA: MIT Press.

Squire, L. R. (1987). *Memory and brain.* New York: Oxford University Press.

Tyler, L. K. (1988). Spoken language comprehension in a fluent aphasic. *Journal of Cognitive Neuropsychology, 5,* 375–400.

Wernicke, C. (1874). *Der aphasische symptomkomplex.* Breslau: Cohn & Weigert.

Wickelgren, W. A. (1968). Sparing of short-term memory in an amnesic patient: Implications for strength theory of memory. *Neuropsychologia, 6,* 235–244.

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.