LANGUAGE AND COGNITIVE PROCESSES, 2001, *16* (4), 419–460

The binding problem for syntax, semantics, and prosody: H.M.'s selective sentence-reading deficits under the theoretical-syndrome approach

Donald G. MacKay and Lori E. James
University of California, Los Angeles, USA

In this case study, a ''hippocampal amnesic'' (H.M.) and memory-normal controls of similar age, background, intelligence, and education read novel sentences aloud in tasks where fast and accurate reading either was or was not the primary goal. In four experiments, H.M. produced more misreadings than normal and cerebellar controls, usually without self-correction. H.M.'s misreadings typically reduced semantic and syntactic complexity and caused ungrammaticality by omitting short high-frequency function-words. H.M. also produced each word more slowly and paused longer than controls at three points: before beginning to produce a sentence, between words in unfamiliar phrases, and at major syntactic boundaries unmarked by commas. H.M.'s selective sentence-reading deficits were unrelated to word-specific factors, ambiguity, and sentence length, and were not attributable to his cerebellar damage, speed-accuracy trade-off, general slowing, general cognitive decline, left-to-right reading processes, or limitations in working-memory capacity. However, present results supported a ''theoretical-syndrome approach'' under which all of H.M.'s deficits (in reading sentences, in comprehending and producing spoken sentences, in reading isolated words and pseudo-words, in visual cognition, and in recall from episodic memory) form part of a general, theoretically coherent syndrome that generalises to other patients.

Requests for reprints should be addressed to Donald G. MacKay, Psychology Dept., University of California, Los Angeles, Los Angeles, CA 90095–1563, USA.

Don MacKay gratefully acknowledges support from National Institute on Aging grant R01 AG 09755. The authors thank Drs Suzanne Corkin and Richard Ivry for generously providing the venues and opportunities to test H.M. and the cerebellar patients (respectively). We also thank Angela Kessell, Christine Rho, and Dr Deborah Burke for helpful comments on earlier versions of this paper, Irwin Stein for recruiting participants, and Amy Flynt and Bryan Chae for research assistance.

http://www.tandf.co.uk/journals/pp/01690965.html DOI: 10.1080/01690960042000102

INTRODUCTION

This case study examined sentence-reading abilities of H.M., the famous ''hippocampal amnesic'' who has acquired virtually no new long term verbal information since 1953 when bilateral portions of his midbrain were surgically ablated (see Corkin, Amaral, González, Johnson, & Hyman, 1997; Milner, Corkin, & Teuber, 1968; Ogden & Corkin, 1991). The study addressed three general issues.

Issue I: Does H.M. pause normally in speech and reading?

Issue I originated with Milner et al.'s (1968) suggestion that H.M. may speak with abnormal prosody (pauses, pitch, stress, and intonation) but for methodological reasons (see Cooper & Paccia-Cooper, 1980, pp. 5–7), pauses were our main empirical measure of prosody. What determines whether and when speakers pause when producing sentences? One determinant is planning trouble or problems in formulating what one wants to say (see Cooper & Paccia-Cooper, 1980, pp. 59–65; Deese, 1984; Gee & Grosjean, 1983). Other determinants are syntax and ambiguity: Speakers normally group words into syntactic units by introducing *prosodic pauses* at syntactically appropriate points within sentences. Prosodic pauses carry important psychological functions (see Cutler, 1987; Levelt, 1989, pp. 363–412; Miller & MacKay, 1994; Wingfield & Butterfield, 1984), e.g., enabling listeners to rapidly encode the syntax and meaning of ambiguous sentences such as, ''Watson fed her dog biscuits''. If a prosodic pause follows *her*, as in, ''Watson fed her _____ dog biscuits'', listeners quickly link *dog* and *biscuits* to form the noun compound *dog biscuits*. Or if a prosodic pause follows *dog*, as in, ''Watson fed her dog_____ biscuits'', listeners quickly link *her* and *dog* to form a quite different proposition. However, equal-length pauses after *her* and *biscuits* slow comprehension and impair immediate recall (see MacKay & Miller, 1996).

Major constituent boundaries (see Cooper & Paccia-Cooper, 1980, pp. 54–65) represent important loci for prosodic pauses, and punctuation often signals these loci in written sentences, but not always. For example, the major constituent boundary between *dogs* and *got* in *The boys who were fed hot dogs got stomach aches*. lacks punctuation, but a period marks the major constituent boundary at sentence end.

Issue I therefore had two parts: Does H.M. produce abnormal-duration pauses at inappropriate points in reading sentences, analogous to planning-trouble pauses in spoken speech? And does H.M. produce normal-duration prosodic pauses at syntactically informative points, regardless of whether punctuation marks those points? Consistent with this focus on

supra-lexical prosodic/syntactic processes, sentences in present research only contain short, high-frequency (HF) words because H.M. often misreads isolated low frequency (LF) words (i.e., below 1 per million in Francis & Kucera, 1982; see MacKay & James, 1999).

Issue II: Does H.M. produce analogous errors in speech and reading?

Issue II began with some curious omission errors that H.M. produced when reading a small set of ambiguous sentences in Corkin (1973). For example, H.M. misread the sentence, *I just don't feel like pleasing salesmen*, as ''I don't like pleasing salesmen.'', distorting the meaning and syntax of this ambiguous sentence by omitting the words *just* and *feel*. MacKay, Burke, and Stewart (1998a) suggested that these reading errors resembled the large number of semantic-level omissions that H.M. produced when describing pictures, when answering autobiographical questions related to childhood experiences, and when explaining the two meanings of ambiguous sentences. For example, when describing the ''job'' interpretation of, *The marine captain liked his new position*, H.M. omitted the function-words ''its'' and ''on'' in an intended output resembling, ''He liked the new position because of its being on a passenger line'', rendering what he said incoherent and ungrammatical, i.e., ''He liked the new position because of being a passenger line''. Another striking reading-speaking resemblance was H.M.'s failure to detect and correct his speech and reading errors, unlike memory-normal controls (see MacKay et al., 1998a).

Issue II therefore had three parts: What is the nature of H.M.'s sentence-reading errors when examined systematically and in large numbers? Do all of H.M.'s reading errors resemble his errors in spoken speech? And are H.M.'s errors attributable to ambiguity, or does H.M. also make more errors than controls when reading unambiguous sentences?

Issue III: Does a coherent theoretical syndrome underlie all of H.M.'s deficits?

Issue III called for a theoretical comparison between H.M.'s episodic memory deficits (see the review in MacKay et al., 1998a), his visual cognition deficits (see MacKay, Stewart, & Burke, 1998b), his language production deficits, his sentence-reading deficits, and his deficits in reading isolated words and pseudo-words (MacKay & James, 1999). The question was whether all of H.M.'s deficits reflect a unified theoretical syndrome with a common cause or set of causes having direct relations to his ''hippocampal amnesic syndrome'' (Milner et al., 1968).

THEORETICAL PREDICTIONS REGARDING H.M.'S SENTENCE-READING

Because Node Structure Theory (NST; e.g., MacKay, 1990; MacKay et al., 1998a) bears on all three of our general issues and served to guide present research, we will examine NST's theoretical mechanisms and predictions in some detail, omitting extraneous aspects and the extensive data that originally motivated these mechanisms wherever possible (but see e.g., MacKay, 1987). The General Discussion section deals with contrasting theories and hypotheses with more restricted relevance to present research.

Under NST, H.M.'s brain damage has made it difficult for him to bind or form new connections between cortical units, and these "binding deficits" underlie both his episodic memory deficits and his language-related deficits. For example, H.M.'s incoherent and ungrammatical spoken output, "He liked the new position because of being a passenger line" in MacKay et al. (1998a), reflects a failure to form the new semantic-level connections required to link "its" and "on" to a coherent overall plan representing his intended utterance "He liked the new position because of its being on a passenger line".

NST predicts that H.M.'s binding deficits will cause similar omissions in sentence-reading, together with abnormal prosody and patterns of pausing. To illustrate these predictions, consider the sentence, *The boys who were fed hot dogs got stomach aches*. Under NST, a hierarchy of top-down connections between cortical units organised as in Figure 1 is required to accurately represent and produce the words in this sentence with the appropriate pattern of stress and prosodic pauses (see MacKay et al., 1998a).[1] What is important to note in Figure 1 is that some of the connections are new whereas others are old or already-formed before encountering this sentence. For example, adults have already-formed connections for representing the phrases, *the boys*, *were fed*, *hot dogs*, and *stomach aches*, which they have encountered many times during everyday life outside the laboratory. However, the proposition, *The boys who were fed hot dogs got stomach aches*, combines these familiar phrases in a novel or never previously encountered way, and many new bottom-up and top-down connections must be formed in order to read this novel combination

[1] As in Figure 1, we capitalise entire words, e.g., [CHICKEN], to indicate sentential stress. To distinguish phonological from orthographic units, we place phonological units between slashes, e.g., /-êr/, and orthographic units between brackets, e.g., [-er], and we capitalise syllables, e.g., [LAbel], to indicate main stress within a word. However, for phonological transcriptions of non-words, we follow current IPA rules (available at *www.arts.gla.ac.uk/IPA/fullchart.html*), with /'/ indicating that the subsequent syllable carries primary stress.

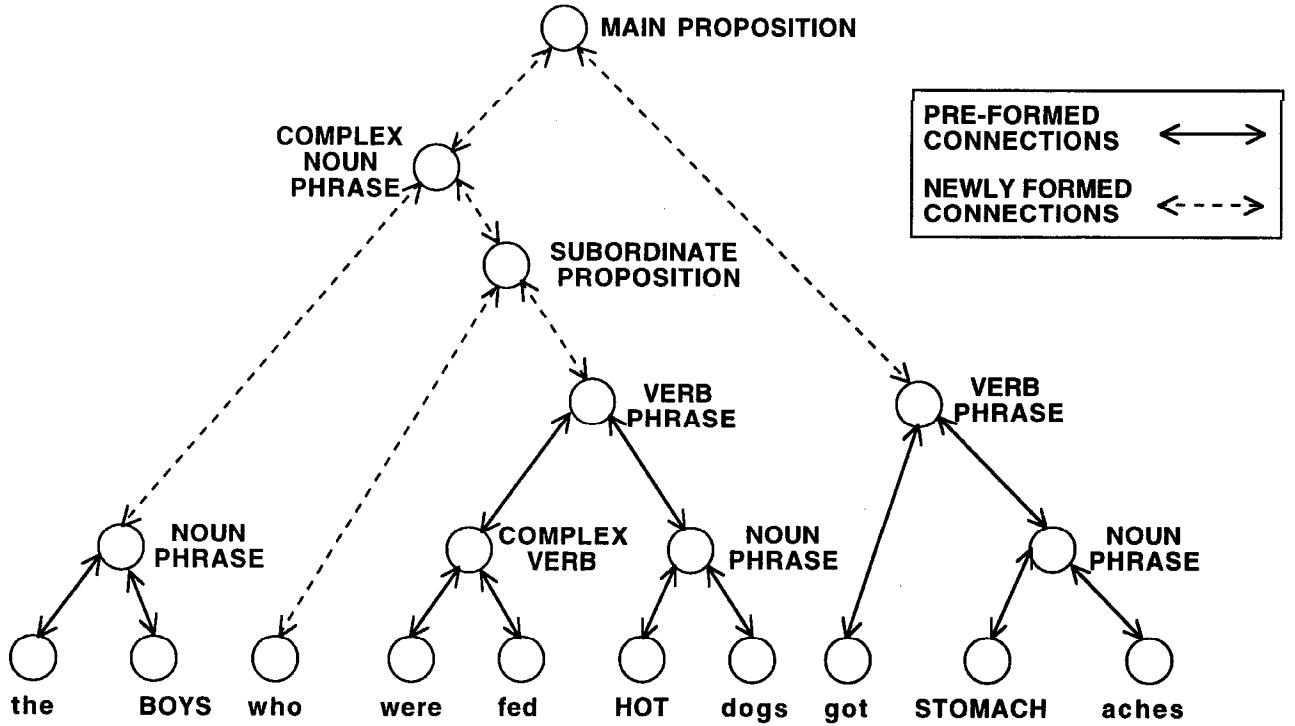

Figure 1. Two-way connections between hypothetical cortical units in NST's semantic system for comprehending and producing the proposition, *The boys who were fed hot dogs got stomach aches* (from sentence 9). Solid lines indicate connections formed (by hypothesis) during childhood. Broken lines indicate new connections formed at first encounter with the sentence. Binding nodes for speeding up the formation of these new cortical connections are not shown.

of phrases with prosodic pauses at syntactically appropriate points such as the major constituent boundary between *boys* and *who* in Figure 1 (see Cooper & Paccia-Cooper, 1980, pp. 54–65).

NST provides two ways to form new connections between cortical units: engrainment learning and binding. Engrainment learning is a primitive process that occurs when nodes are repeatedly activated during overt or internal production of, say, words in a phrase (see MacKay, 1990). For example, if *The boys who were fed hot dogs* is repeated a large number of times, engrainment learning will eventually "burn in" the new connections for representing this complex noun phrase. Repeating this noun phrase many times together with *got stomach aches* can then burn in the remaining new connections for the entire proposition in Figure 1.

Binding, the second means of forming new connections, is more efficient than engrainment learning, but also more complex. Binding requires engagement of a specific "binding node" from among many different types of binding nodes that are specialised for conjoining different classes of units.[2] Once engaged, binding nodes direct supplementary input to specific cortical units that enables formation of new bottom-up and top-down connections in a single pass or repetition (see MacKay, 1990, for details).

NST predictions regarding H.M.'s sentence-reading

NST predicts that H.M. will be unimpaired relative to memory-normal controls in reading HF words in familiar phrases such as *the boys*, *were fed*, *hot dogs*, and *stomach aches* for two reasons: because top-down connections for producing phrases learned before H.M.'s operation have been strengthened via engrainment learning over the course of his lifetime, and because activation processes involving connections formed before H.M.'s operation are intact (see e.g., MacKay & James, 1999). For

[2] Just as "hippocampal system damage" is a summary term for H.M.'s surgical midbrain lesion, "semantic binding nodes" is a summary label for the binding mechanisms required to read and produce the syntax, prosody, and semantics of sentences. However, two other general classes of binding node are relevant to reading unfamiliar and (invented) pseudo-words: orthographic and phonological binding nodes (see MacKay & James, 1999). Note that binding nodes are a theoretical construct with a hypothetical locus that almost certainly includes subcortical structures such as the hippocampus, and in the case of semantic binding nodes may include only *left* hippocampus and connected structures (see e.g., Milner, 1975; O'Keefe & Nadel, 1978). However, these neuroanatomical issues are not central to NST. If future evidence indicates, say, that some binding nodes are located in additional structures, or in *right* rather than *left* hippocampal systems in individuals with right hemisphere language dominance, this would not affect NST (see MacKay & James, 1999).

example, H.M. should read the familiar phrase *stomach aches* like normal controls: without error, with a normal-duration pause between *stomach* and *aches*, with greater stress on *stomach* than *aches*, and with falling prosodic pitch on *aches*. However, H.M. should produce HF words in unfamiliar phrases such as *cotton farmers* word-by-word, i.e., as isolated units lacking normal supra-lexical prosody (pauses, pitch contours, and lexical stress). The reason is that H.M.'s lesion has destroyed some (but perhaps not all) of the binding nodes required to efficiently establish the new connections for integrating familiar words into novel phrases with syntactically appropriate prosody. For example, H.M. will either fail to form the connections labelled "new" in Figure 1, or will take much longer to form them than memory-normal controls, resulting in abnormally long "planning-trouble" pauses because connection-formation via engrainment learning is so slow. H.M.'s slowness or failure to form new connections will likewise hamper selective assignment of prosodic stress to appropriate words in a sentence (e.g., the word *boys* in the sentence in Figure 1), and should often result in omission errors. These omission errors should tend to simplify the syntactic and semantic structure of the sentence, and should often involve words that *only* link to the structure of a sentence via new connections. An example is the extremely HF connector *who* in *The boys who were fed hot dogs got stomach aches*: Only a new connection links *who* to other units in this sentence (see Figure 1). However, H.M. should be able to read familiar phrases such as *hot dogs* without errors because connections for reading such phrases were formed before his operation and used throughout his lifetime.

Slowness in forming new connections should increase H.M.'s sentence-reading time overall, but NST predicts especially large deficits in H.M.'s planning times (the time to begin to read a sentence) relative to his deficits in production time (the acoustic duration of each HF word). The reason is that forming new connections is the primary determinant of planning times, whereas activating old or already formed connections is the primary determinant of production times for HF words (see MacKay & James, 1999, for details). NST also predicts that H.M. will pause for abnormally long times at unusual points in the sentences, an indication of planning trouble or problems in forming the new connections to represent what he wants to say. H.M. will also fail to produce normal-duration prosodic pauses at syntactically informative points because his binding deficits make it difficult for him to represent the syntax of novel sentences. However, H.M. should introduce normal-duration prosodic pauses at syntactic junctures marked by punctuation e.g., commas: Because H.M. has responded appropriately to commas since early childhood, commas will trigger pause responses via activation of old connections and independently of H.M.'s binding deficits.

EXPERIMENT 1: AN EXPLICIT TEST OF H.M.'S SENTENCE-READING ABILITIES

Experiment 1 compared abilities of H.M. and memory-normal controls to read novel sentences containing HF words. Our sentences (shown in the Appendix) varied greatly in length, syntactic-complexity, and semantic-coherence for three reasons. First, we wanted H.M. and controls to read some (presumably short, syntactically simple, semantically coherent) sentences *without errors* to eliminate effect of errors on our measures of production-, planning-, and pause-time. Second, we wanted to observe errors in some (e.g., syntactically complex) sentences to test NST predictions regarding errors. Third, we wanted to determine whether sentence length is an especially important determinant of errors for H.M. Note, however, that none of our sentences are anomalous and only comparisons between H.M. and controls reading identical sentences are of interest in Experiment 1 because no between-sentence factors were systematically controlled and manipulated.

Method

Participants. Participants were H.M. and six memory-normal controls matched with H.M. for highest educational level, previous employment, age, and overall IQ (see Table 1 for a summary of background characteristics). Corkin et al. (1997) review H.M.'s surgical and clinical history. H.M.'s bilaterally symmetrical midbrain surgery mainly affected "hippocampal systems" (MacKay et al., 1998a), removing virtually the entire amygdaloid complex, some entorhinal cortex, part of the dentate gyrus and subicular complex, and all but the caudal 2 cm of the hippocampal body, although this spared 2 cm has unknown functional

TABLE 1
Characteristics of H.M. and normal controls in Experiments 1–4

Participants	*Age in 1997*	*Highest educational degree*	*Verbal IQ*	*Performance IQ*	*Mean IQ*
H.M.	71.75	High school	107	117	112.00
Control 1	73	High school	115	129	122.00
Control 2	74	High school	114	128	121.00
Control 3	74	High school	110	113	111.50
Control 4	70	High school	115	118	116.50
Control 5	70	High school	117	130	123.50
Control 6	67	High school	120	124	122.00
Mean for controls (*SD*)	71.33 (2.80)	High school	115.17 (3.31)	123.67 (6.83)	119.42 (4.55)

status. Spared were the temporal stem, the collateral sulcus, including portions of the ventral perirhinal cortex, parahippocampal cortex and all neocortex except for small areas at the tips of his temporal poles where thin metal suction tubes passed during H.M.'s sub-orbital surgery (Scoville & Milner, 1957). H.M. also exhibits bilateral cerebellar damage due to his long-term doses of dilantin and other drugs for controlling epilepsy since age 15.

Control participants were native English speakers (five female, one male) who received $10/hour for participating. Like H.M., controls' highest educational degree was high school. Regarding background/work history, H.M. once held a skilled labour position, and controls once held skilled labour, manual labour, clerical, sales, and administrative positions. Like H.M., controls were 71 years old when tested (age range 67–74) and their IQ matched H.M.'s as closely as possible (see Table 1). H.M.'s most recent IQ test (1997, S. Corkin, personal communication) was the Wechsler–Bellevue (Form I), and controls took this test less than 6 months before the present studies. H.M.'s performance on this test has remained remarkably stable between 1977 and 1997 (see MacKay et al., 1998a), counter-indicating semantic dementia or accelerated decline in general cognitive function.

Materials. Materials were 11 sentences, typed with normal punctuation (commas and full stops) on separate 4″ × 6″ index cards in large (18 point, mainly lower case) Courier font. Words in the sentences were HF (mean Francis & Kucera frequency = 12 486 tokens per million) and short (mean length = 4.5 letters). The sentences were 7 to 23 words long (M = 15.3 words), and most contained one or more complex syntactic constructions, e.g., subordinate clauses, pronouns or other anaphora, passive constructions, left-branching structures, prepositional phrases, and consistent co-ordination of various types. Three sentences (2, 9, and 11 in the Appendix) were semantically less coherent than the others.

Procedure. Instructions made no mention of prosody, were repeated orally, and appeared on a continuously displayed card: ''Read each sentence as quickly as possible without making errors''. To present each sentence, the experimenter turned over a stimulus card while saying ''OK''. Reading responses were taped, and because initial recording quality seemed less than satisfactory for prosody measures, the experimenter had H.M. read the sentences again (about 26 hours later). Although H.M. made more errors on the first than second reading, results of the two tests were basically similar and are averaged together here. Reading errors were scored from the tapes, as was overall reading time from the experimenter's ''OK'' to the end of a sentence as measured via

stopwatch. Other more precise time measures were derived from a digitised version of the tapes recorded at 44 100 samples per second using SoundEdit 16 v. 2 running on a Power Macintosh G3 computer.

Results and discussion

Response time analyses

Mean overall reading time per sentence (see Table 2) was 4.6 standard deviations (*SD*) longer for H.M. ($M = 11.9$ s) than controls ($M = 5.4$ s; $SD = 1.4$), a difference due in part to H.M.'s errors, but also to some abnormally long pauses at unusual points in H.M.'s sentences. For example, both H.M. and controls produced *dogs* and *got* without error in *Although the boys who were fed hot dogs got stomach aches*, (sentence 9),[3] but H.M. paused for 1356 ms between these words, over 43 *SD*s longer than controls ($M = 50$ ms; $SD = 30$).

To eliminate effects of errors on temporal measures, most subsequent analyses involved the three sentences that both H.M. and all six controls read without error (see Appendix). Five times were computed from the digital record of these sentences: overall correct response time, correct planning time (the period of silence between the experimenter's "OK" and onset of phonation), correct production time (the duration of phonation, excluding all measurable pauses), pause time (the duration of each measurable between-word pause), and overall pause time (the

TABLE 2
Mean overall response times (in s) per sentence, and mean response, planning and production times for H.M. and controls for correctly-produced sentences in Experiment 1 (with *SD*s in parentheses)

Measures (Number of sentences)	*Participants*	
	H.M.	*Controls (SD)*
Overall response time ($N = 11$)	11.9	5.4 (1.4)
Overall correct response time ($N = 3$)	10.8	4.2 (0.4)
Correct planning time ($N = 3$)	1.7	0.5 (0.1)
Correct production time ($N = 3$)	5.4	3.3 (0.3)
Pause time within correctly produced sentences ($N = 3$)	3.7	0.4 (0.2)

[3] Numbered sentences in the text appear in full with corresponding numbers in the Appendix.

summed duration of all pauses inside sentences). For these error-free sentences, overall correct response times were 16.5 *SD*s longer for H.M. ($M = 10.8$) than controls ($M = 4.2$; $SD = 0.4$; see Table 2), overall pause times were 16.5 *SD*s longer for H.M. ($M = 3.7$ s) than controls ($M = 0.4$ s; $SD = 0.2$), and correct production times (acoustic durations) were 7 *SD*s longer for H.M. ($M = 5.4$ s) than controls ($M = 3.3$ s; $SD = 0.3$; see Table 2). Correct planning times were 12 *SD*s longer for H.M. ($M = 1.7$ s) than controls ($M = 0.5$ s; $SD = 0.1$; see Table 2), and similar planning time effects obtained for the 10 sentences with initial words correctly produced by all participants: Planning times were 17 *SD*s longer for H.M. ($M = 2.3$ s) than controls ($M = 0.6$ s; $SD = 0.1$) reading these sentences. In short, H.M.'s deficits were relatively smaller for production times (7 *SD*s greater than controls) than for planning times (16.5 *SD*s greater than controls) and pause times within sentences (17 *SD*s greater than controls), consistent with NST predictions.

Effects of commas. Consistent with NST predictions, when all participants correctly produced words on either side of major constituent boundaries, pause durations at boundaries with commas differed little for H.M. ($M = 245$ ms) versus controls ($M = 120$ ms; $SD = 78$), but were over 26 *SD*s longer for H.M. ($M = 767$ ms) than controls ($M = 79$ ms; $SD = 26$) at boundaries without commas.

Familiar versus unfamiliar phrases. To test NST predictions regarding familiar versus unfamiliar phrases, we selected 16 two-or-three-word phrases from error-free sentences that intuitively seemed high or low in familiarity, and eight from sentences containing errors that intuitively seemed intermediate in familiarity (as foils). We then presented these phrases to eight native English speakers (college graduates or graduate students conducting research on language-related issues; mean age = 28) who estimated their familiarity with each phrase using 5-point scales (1 = "I've never heard it before", and 5 = "I hear it very often"). The phrases were typed on individual index cards, with presentation order randomised across judges under two constraints: Phrases from the same sentence never occurred consecutively, and four or more cards separated phrases containing an identical word. Judges were instructed to ignore between-phrase overlap and familiarity of individual words in judging phrase-familiarity.

Two pairs of phrases contained overlapping words, e.g., *cotton farmers* and *the cotton farmers*, under the following rationale: If *the cotton farmers* receives higher familiarity ratings than *cotton farmers* then such ratings are invalid and/or insensitive measures of prior encounters with these word combinations (unconditional familiarity of a or b in isolation must exceed

familiarity of conjoint event a + b). Both overlapping phrase pairs passed these validity/sensitivity tests, e.g., *the cotton farmers* had lower mean familiarity ratings (2.1) than *cotton farmers* (2.3), enabling confidence in the validity and sensitivity of the ratings.

To establish our final familiar- versus unfamiliar-phrase categories, we categorised phrases with a mean rating above 4.0 (''I hear this phrase often'') as familiar (N = 6; e.g., *for lunch*) and phrases with a mean rating below 3.0 (''I've occasionally heard this phrase'') as unfamiliar (N = 3; e.g., *cotton farmers*). Mean familiarity ratings were 4.7 for familiar phrases (range 4.1–5.0) versus 2.4 for unfamiliar phrases (range 2.1–2.8).

Consistent with the NST prediction that H.M. will pause longer between words in unfamiliar phrases (which require new connections) than between words in familiar phrases (which don't require new connections), H.M.'s pause times within unfamiliar phrases (M = 185 ms) were over 5.8 *SD*s longer than controls' (M = 69 ms, SD = 20), but within familiar phrases (M = 42 ms) were only 1.9 *SD*s longer than controls' (M = 17 ms; SD = 13).

To illustrate these phrase-familiarity effects in greater detail, Figure 2a shows the mean pause times (in ms) for H.M. and controls between each word in error-free sentence 6: *For lunch Bill is having either pork or chicken and fries.* H.M. and controls produced pauses of approximately equal duration within the familiar phrases *for lunch* and *is having* in this sentence (see Figure 2a), but H.M. produced a longer pause within the unfamiliar phrase *or chicken* (M = 1580 ms) than controls, who produced no pauses between *or* and *chicken* (see Figure 2a).[4] Figure 2b transforms Figure 2a data into proportions, i.e., per cent of their overall pause times that H.M. and controls spent between each word in this sentence. These proportions matched for only one word-pair, and varied across a 47% range, from 20% greater for controls than H.M. (between *lunch* and *Bill*), to 27% greater for H.M. than controls (between *or* and *chicken*), a proportion difference that exceeds 10 *SD*s. The remaining two error-free sentences exhibited a similar pattern, indicating that pauses were not simply longer overall for H.M. than controls: H.M. paused disproportionately longer than controls at some points (predicted under NST), and

[4] Repeated replaying indicated that H.M. and controls produced *familiar phrases* in error-free sentences with perceptually indistinguishable patterns of stress and pitch, but H.M. produced *unfamiliar phrases* with pitch and stress patterns that were unusual and variable from reading to reading. For example, on his first reading, H.M. produced all four words of the unfamiliar phrase *either pork or chicken* with perceptually equal stress and pitch, i.e., [EITHER] [PORK] [OR] [CHICKEN], but on his second reading, H.M. produced a pitch drop on [CHICKEN] that was much stronger than his (perceptually normal) pitch drop at the end of the sentence. By contrast, controls consistently produced this phrase with modulated pitch and stress, i.e., [either] [PORK] [or] [CHICKEN], and with a stronger pitch drop on sentence-final *and fries* than on [CHICKEN].

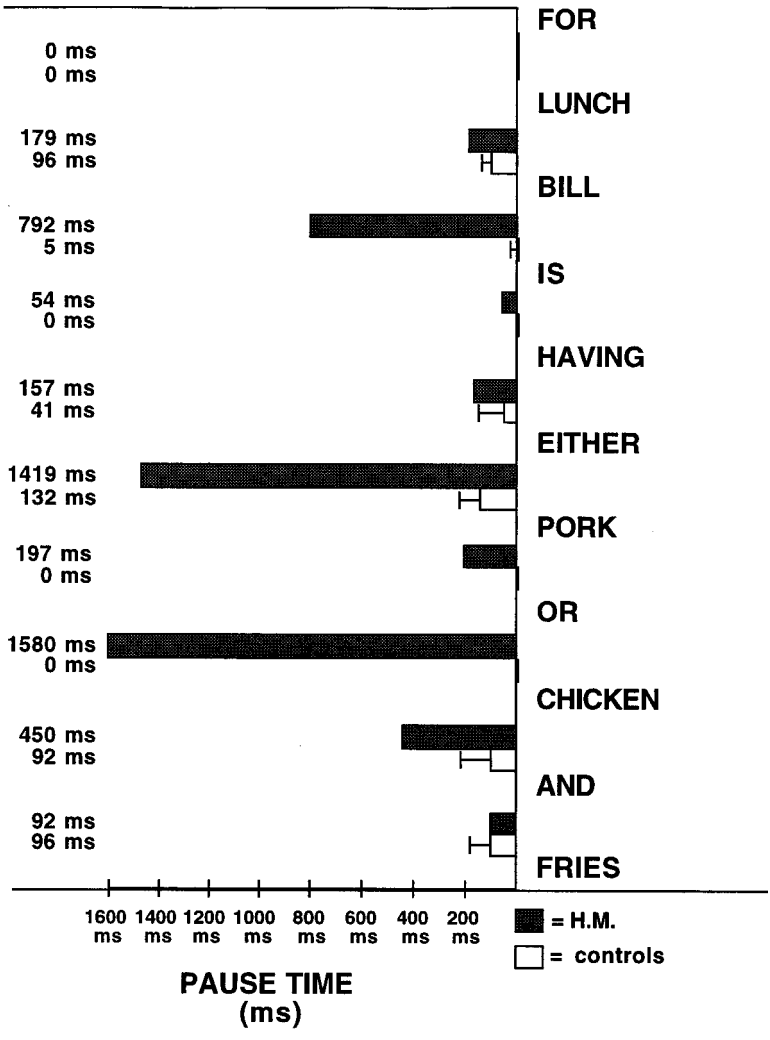

Figure 2a. Pauses in ms between each word in (error-free) sentence 6 for H.M. and controls (note right-to-left scale).

disproportionately shorter than controls at other points (predicted under NST).

Error frequency analyses

Errors were scored conservatively: When H.M.'s output was difficult to decipher, up to five judges listened to the tape, and if more than one judge considered H.M.'s output error-free, H.M.'s output was scored as error-

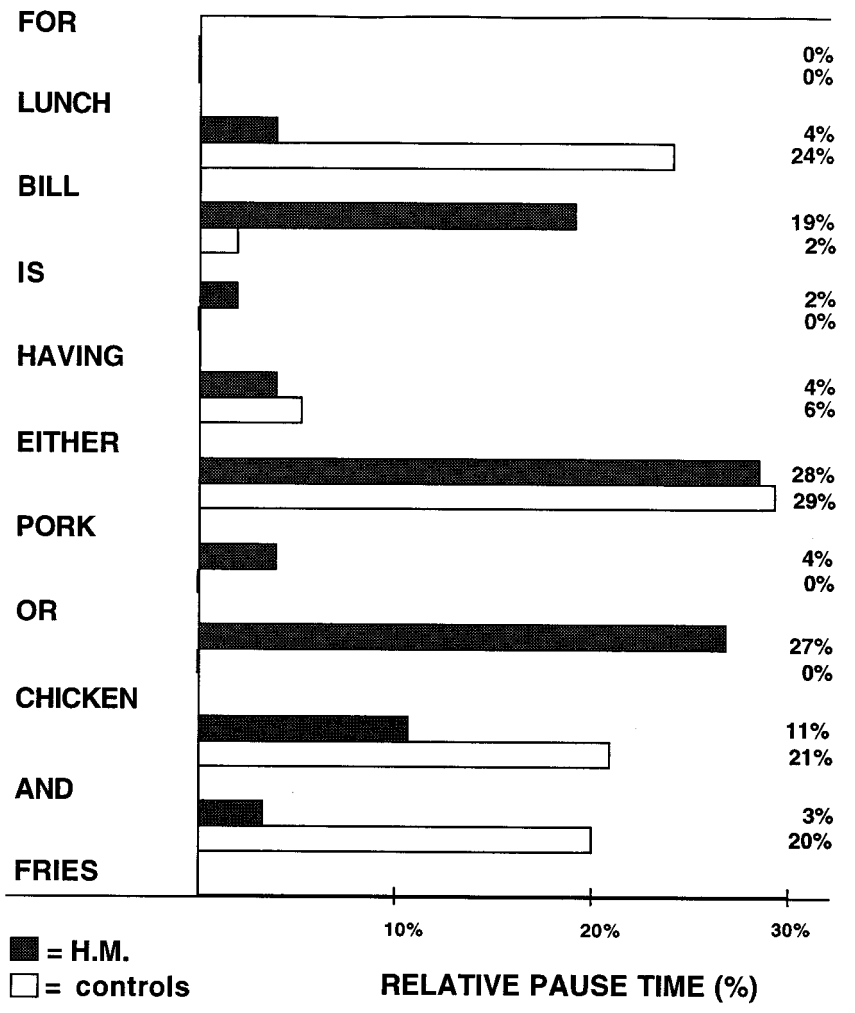

Figure 2b. Proportion of total pause time (in %) between words in (error-free) sentence 6 for H.M. and controls.

free. We initially scored overall reading errors using sentences rather than error instances (words) as unit of analysis because determining where one error left off and another began in H.M.'s responses was often difficult. Overall errors included false starts, repetitions, anticipations of phonological units, and substitutions, omissions and additions of words, both corrected and uncorrected. For example, H.M. misread *Although the boys who were fed hot dogs got stomach aches, the genie ate the golden figs* (in sentence 9) as "Although the boys _____ were fed hot dogs got stomach

aches, *and* the genie ate the _____ figs'' (underscores within quotes indicate omitted and added words). H.M. also left out *who* on his second reading, again rendering sentence 9 ungrammatical in the manner of semantic-level omissions discussed in the introduction.

Overall errors occurred in 55% of H.M.'s sentences versus 14% of controls' sentences on average ($SD = 21\%$), a difference reliable at $p < .01$ using a sign test with sentences as unit of analysis. This difference is not attributable to speed-accuracy trade-off because H.M. produced each word more slowly, paused longer between words, and paused much longer before beginning to produce a sentence than controls (see Response Time results).

Target words that were omitted or mispronounced were comparable in mean length for H.M. (5.18 letters, range 2–9) and controls (5.20 letters, range 3–9). However, frequency of these target words in Francis and Kucera (1982) was higher for H.M. (mean frequency 7560 per million, range 2–28 543) than controls (1826 per million, range 2–19 427).

To test whether sentence length contributed to H.M.'s overall errors, a second set of analyses equated opportunities for errors across sentences of different lengths by calculating error words per word per sentence, i.e., how many words a participant misproduced, omitted, or added divided by number of words per sentence and transformed into per cent. Error words per word per sentence were 4 *SD* greater for H.M. ($M = 5\%$) than controls ($M = 1\%$; $SD = 1\%$; see Table 3). However, mean number of error words per word per sentence correlated weakly and unreliably with sentence length (range 7 to 23 words) for H.M. ($r_s = .31$, $p > .46$) and controls ($r_s = .05$, $p > .90$), indicating that misreadings were unrelated to sentence length.

Interestingly, semantic complexity increased misreadings relatively more so for controls than H.M. Controls made 3.5 times as many errors on the semantically less-coherent sentences than they made on the remaining sentences whereas H.M. made only 2.6 times as many errors on the semantically less-coherent sentences than on the remaining sentences.

Qualitative differences between errors of H.M. versus controls

Reading errors of H.M. and controls differed in several qualitative ways summarised in Table 3. We used words as unit of analysis to evaluate word-level qualitative factors, e.g., phonological overlap of substituted and target words, and sentences as unit of analysis to evaluate sentence-level qualitative factors, e.g., grammaticality.

TABLE 3
Example error characteristics with frequency (in %*) for H.M. and controls in Experiment 1 (with *SD*s in parentheses)

Characteristics	*Example error*	*H.M.*	*Controls* (*SD*)
Overall errors per word per sentence	(see text)	5%	1% (1%)
Non-corrections	(see text)	73%	48% (50%)
Meaning-changing errors	"the figs" for *the golden figs*	46%	8% (11%)
Ungrammaticality	"they were train" for *they were training*	27%	2% (4%)
Syntax-altering errors	"went to the movie" for *went to see the movie*	36%	2% (4%)
Syntax-altering errors that decreased syntactic complexity	"Although the boys_ were fed hot dogs ..." for *Although the boys who were fed hot dogs ...*	50%	0% (0%)
Syntax-altering errors that increased syntactic complexity	"Kevin was frightened ... but he knew" for *Kevin was frightened ... but knew*	0%	17% (41%)

* Note: Error characteristics sum over 100% as not mutually exclusive.

Phonological overlap. H.M.'s misreadings shared few phonological segments with the original stimulus, e.g., complete omissions of *golden* (in sentence 9), whereas misreadings of controls usually added or subtracted only a single segment, e.g., *boys* misread as "boy". With unshared segments in uncorrected reading errors subcategorised as omissions versus additions, H.M. omitted 26 segments (versus a mean of 0.3 for controls), and added 10 new segments (versus a mean of 0.3 for controls), an omission-to-addition ratio that resembles H.M.'s additions and omissions of isolated segments in reading pseudo-words and LF words (see MacKay & James, 1999). Omissions of entire words or morphemes (e.g., *training* misread as "train" in sentence 4) were also more common for H.M. (6) than for controls ($M = 0.17$, $SD = 0.41$), and made up 55% of H.M.'s morpheme errors, which all occurred in rarely used phrases. An example is *giggled in the hallway* misread as "juggling in the hallway", an error that involves substitution of two morphemes.

Meaning-altering misreadings. Meaning-altering misreadings, e.g., omission of *golden* in *the golden figs* (sentence 9) changed the meaning

of the original sentence, and occurred 3.5 *SD*s more often for H.M. (M = 46%) than controls (M = 8%; SD = 11%; see Table 3).

Ungrammatical misreadings. Ungrammatical misreadings, e.g., H.M.'s "they were train for the Olympics" (sentence 4) rendered sentences ungrammatical, and occurred 6.3 *SD*s more often for H.M. (M = 27%) than controls (M = 2%; SD = 4%).

Error correction. H.M. usually failed to correct even ungrammatical misreadings such as, "they were train for the Olympics". However, controls usually corrected their errors immediately, as in "Kevin was frightened by the bat in the bushes but grew—knew that—that when you learn gradually you worry less" (sentence 2). Error words were uncorrected more often for H.M. (M = 73%) than controls (M = 48%; SD = 50%).

Syntax-altering misreadings. Misreadings changed the syntax of the original sentence 8.5 *SD*s more often for H.M. (M = 36%) than controls (M = 2%; SD = 4%; see Table 3). For example, H.M.'s omission of *see* in *went to see the movie* (sentence 3) eliminates the infinitive phrase *to see*, and transforms its infinitive marker (*to*) into a preposition in the prepositional phrase, "to the movie".

Considering only syntax-altering errors, misreadings reduced the syntactic complexity of a sentence many *SD*s more often for H.M. (M = 50%) than controls (M = 0%; SD = 0%). For example, H.M. misread *Although the boys who were fed hot dogs* (in sentence 9) as "Although the boys were fed hot dogs", a syntax-altering error that reduces syntactic complexity by eliminating the subordinate *who*-clause. Conversely, syntax-altering errors increased syntactic complexity more often for controls (M = 17%; SD = 41%) than H.M. (0%; see Table 3). For example, *Kevin was frightened ... but knew* (in sentence 2) misread as "Kevin was frightened ... but *he* knew" is a syntax-altering error that increases syntactic complexity by adding the pronoun *he*.

To summarise, H.M. doesn't just misread long and LF words (see MacKay & James, 1999): As predicted under NST, H.M. also made more errors than memory-normal controls when reading short, HF words in novel sentences. The syntax- and meaning-distorting nature of H.M.'s misreadings comports with the hypothesis that H.M.'s binding deficits made it difficult for him to form the new connections required to represent the syntax and meaning of unfamiliar phrases in novel sentences, even when the phrases contained HF words and were typed on a card before him. Also consistent with this hypothesis was the high proportion of word omissions, and the fact that H.M.'s errors only simplified the syntax of a

sentence, occurred mainly in unfamiliar phrases, and involved very HF connectors (such as *but* and *who*) that *only* linked to sentence structures via new connections (see Figure 1).

EXPERIMENT 2: CONTROLS FOR LEXICAL FACTORS AND CEREBELLAR DAMAGE

Experiment 2 had two goals. One was to test lexical-level versus NST accounts of H.M.'s sentence-reading deficits by having H.M. read isolated content- and function-words taken from Experiment 1 sentences. Under NST, H.M. should read these HF words without difficulty when presented in isolation rather than in sentences. However, under lexical-level accounts, H.M. should mainly misread isolated words that he misread in Experiment 1 (and not ones he read correctly), and of these, he should misread more function-words than content-words, as in Experiment 1.

Goal two was to determine whether H.M.'s sentence reading deficits are attributable to his bilateral cerebellar damage. To test this hypothesis, Experiment 2 compared the sentence reading of H.M. and patients with bilateral cerebellar damage. This control seemed important because recent neuroanatomic, neuroimaging, and neuropsychological data (see e.g., Ivry & Keele, 1989; Helmuth, Ivry, & Shimizu, 1997) indicate unexpected cerebellar activity during several cognitive functions, including timing (albeit not in sentence reading tasks).

Method

Participants. Participants were H.M. and four patients who were diagnosed with bilateral cerebellar damage and/or degeneration in the absence of other brain damage or disorders, e.g., dementia (mean age = 54.75, range 31–79; mean years of education = 16.5, range 14–18 years).

Materials and procedure. H.M. read 24 isolated words using procedures resembling Experiment 1. The isolated words came from sentences in Experiment 1, and included 12 short function-words (e.g., *in*, *that*, *who*) and 12 content-words matched as closely as possible for mean length and frequency (e.g., *bat*, *see*, *spy*). H.M. had misread half of the content- and function-words in Experiment 1, but had correctly read the remainder. The words were typed in 18 point Courier upper case font on individual index cards and presented in intermixed order. Sentences and procedures for the cerebellar patients were identical to Experiment 1.

Results and discussion

For isolated words, H.M. misread one word (17%) that he had read correctly in Experiment 1, and he misread no words (0%) that he had misread in Experiment 1. Both findings contradict lexical-level accounts of H.M.'s sentence-reading deficits, which predicted more errors on isolated words that H.M. misread in Experiment 1, and of these, more errors on function- than content-words. However, both findings comport with results of MacKay & James (1999) and with NST: H.M. has little difficulty reading HF words in isolation, and his special difficulty with function-words in sentences lies in quickly linking such words to novel syntactic and semantic structures in particular sentence contexts. Both NST hypotheses also comport with the fact that H.M. correctly read in other sentences in Experiment 1 all of the function-words that he misread in Experiment 1.

The cerebellar patients did not differ in either errors or onset times from controls in Experiment 1. However, cerebellars produced fewer sentences containing errors (25%, $SD = 15\%$) than H.M. (55%, a 2.0 *SD* difference), usually corrected their misreadings, virtually never produced meaning- and syntax-altering or ungrammatical sentences, and virtually never produced unusual pauses resembling H.M.'s. These data render H.M.'s bilateral cerebellar damage an unlikely cause of his sentence reading deficits, and render the cerebellum an unlikely locus for the semantic binding nodes underlying sentence reading.

EXPERIMENT 3: AN INCIDENTAL TEST OF H.M.'S SENTENCE-READING ABILITIES

Experiment 3 had three goals. One was to replicate Experiment 1 results using a large number of short sentences. Goal two was to rule out ambiguity as an explanatory factor in Experiment 1 (where many sentences were ambiguous): Experiment 3 tested whether H.M. also alters the meaning and syntax of sentences specifically constructed to be unambiguous. Goal three was to determine whether H.M. exhibits sentence-reading deficits when fast and accurate reading is not the primary task: Experiment 3 participants understood that their main task was to specify whether sentences were ambiguous or unambiguous, following the incidental task of reading the sentence aloud.

Method

Participants. The seven participants were identical to Experiment 1 (see Table 1).

Materials. Stimuli were 48 ambiguous sentences modified from Lackner (1974) and 24 unambiguous[5] sentences modified from MacKay et al. (1998a). The ambiguous sentences were either structurally ambiguous ($N = 36$) or lexically ambiguous ($N = 12$; see the Appendix for examples). Each sentence was typed in large (18 point Courier, mainly lower case) font on a separate $4'' \times 6''$ index card. Mean sentence length was 9.2 words overall (range 5–14 words), with 10.4 words in lexically ambiguous sentences, 9.3 words in structurally ambiguous sentences, and 8.5 words in unambiguous sentences.

Procedure. The experimenter verbally instructed participants that the main task was to determine whether sentences had one or more than one meaning, and to describe the one or two meanings that they detected.[6] First, however, participants were to read the sentence aloud to ensure that they had registered it accurately. The instructions were summarised on a prominently displayed card, and the sentences were presented in the same "semi-random" order to each participant. When misreadings noticeably distorted the meaning of a word or sentence, the experimenter asked participants to read the word or sentence again, and these experimenter prompts were 29 times more common for H.M. ($M = 5$) than controls ($M = 0.17$).

Results and discussion

Overall errors were defined and scored as in Experiment 1, and included false starts, repetitions, anticipations and omissions of phonological units, and omissions, additions, and substitutions of words, e.g., *hunters* misread as "hunter", and *like* misread as "want". H.M. also produced unusually long pauses within words that we excluded from our error analyses. Also excluded were three unusual types of omissions that we attributed to the

[5] By "unambiguous" we mean containing no obvious ambiguities. The ubiquity of ambiguity in natural language makes complete absence of ambiguity impossible to guarantee (see e.g., MacKay & Bever, 1967).

[6] This primary task replicated and extended results for auditory sentences reported in MacKay et al. (1998b, Experiment 1), i.e., H.M. was impaired relative to memory-normal controls in discriminating ambiguous from unambiguous sentences. One novel result was that, unlike controls, H.M. often changed his responses from "two meanings" to "one meaning" and vice versa, sometimes several times, as if he was basing his responses on subtle, evoked cues from the experimenter (L.J.) rather than comprehension per se. Counting only his initial responses, H.M. correctly responded "one meaning" for only 58% of the 24 unambiguous sentences, versus 87% for controls on average ($SD = 12\%$), a 2.4 *SD* difference. For ambiguous sentences that participants correctly identified as ambiguous, controls successfully described both meanings for more ambiguous sentences ($M = 78\%$; $SD = 11\%$) than H.M. ($M = 29\%$), a 4.5 *SD* difference.

primary task in Experiment 3 (ambiguity detection): omissions of sentence-initial *the* and failures to read or finish reading a sentence following discovery of the ambiguity.

Sentences contained errors 5.2 *SD* more often for H.M. ($M = 38\%$) than controls ($M = 12\%$; $SD = 5\%$; see Table 4, top panel). Analyses excluding self-corrected errors gave similar results, as did error words per word per sentence (see Table 4, bottom panel). As in Experiment 1, H.M. primarily omitted or misread short, HF words, e.g., *a*, *the*, *are*, and *after*, with overall mean length 4.35 letters (range 1–11) and mean frequency 10 741 per million (range 4–69 975) in Francis and Kucera (1982). No word that H.M. misread in Experiment 3 fit the LF category in MacKay and James (1999). Also as in Experiment 1, controls misread words with similar mean length but much lower mean frequency than words H.M. misread.

Misreadings changed the meaning of a sentence, e.g., *Her only choice* misread as "Her choice", 3.5 *SD*s more often for H.M. ($M = 8\%$ of sentences) than controls ($M = 1\%$ of sentences, $SD = 2\%$). Also as in Experiment 1, H.M.'s misreadings tended to alter syntax and grammaticality, and to reduce syntactic complexity more often than misreadings of controls. For example, H.M. misread, *The spy put out a torch which was our signal to attack* (sentence 12) as "The spy out of touch which was our signal to attack", where substituting *of* for *a* alters both meaning and syntax, substituting *touch* for *torch* alters meaning but not syntax, and omitting *put* and transforming its particle *out* into a preposition alters and simplifies syntax.

H.M.'s error corrections generally followed experimenter prompts after he finished reading a sentence. For example, following a specific experi-

TABLE 4

Misreadings for unambiguous, structurally ambiguous, lexically ambiguous, and across all sentences in Experiment 3 for H.M. and controls (means with *SD*s) using two units of analysis: sentences (top panel) and error words per word per sentence (bottom panel)

Participants	*Unambiguous sentences* ($N = 24$)	*Structurally ambiguous sentences* ($N = 36$)	*Lexically ambiguous sentences* ($N = 12$)	*Across all sentence types* ($N = 72$)
Sentences containing errors				
H.M.	33%	39%	42%	38%
Controls	9%	11%	22%	12%
(*SD*)	(9%)	(5%)	(12%)	(5%)
Error words per word per sentence				
H.M.	4.9%	6.7%	6.5%	5.8%
Controls	1.1%	1.3%	2.9%	1.5%
(*SD*)	(1.0%)	0.7%	(2.9%)	(0.9%)

menter request, H.M. corrected his ''spy-sentence'' errors (above) as follows: ''The spy put out of touch ... put out a torch which, ... it says 'torch,' which ___ our, was our signal to attack''. As in this example, H.M.'s error-corrections were laborious and not always coherent or error-free, whereas controls usually corrected their errors concisely, without prompts, and quickly, sometimes before the end of word, and always before they finished reading the sentence. H.M. corrected 50% of his errors whereas controls corrected 67% of their errors (SD = 21%) even though the experimenter asked controls for error corrections 29 times less often than H.M.

H.M. nonetheless initiated one error correction that is worthy of note. H.M. misread *The cat always hid under the couch when dogs were in the room* (sentence 21) as ''The cat always hides, no, the cat always hide, under the couch when dogs were in the room. And that wouldn't be hide, it'd be hid''. Here H.M.'s multiple self-corrections (''hides'', ''hide'', and ''hid'') are reminiscent of his successive approximations in reading isolated LF words, e.g., *akimbo* misread as ''ak ... akibo'' and finally ''akbo'' with the vowel /a/ receiving inappropriate stress on all three attempts (MacKay & James, 1999). However, the ungrammatical nature of H.M.'s initial ''hides''-to-''hide'' correction is remarkable, as is his final (accurate) ''hide''-to-''hid'' correction, which occurred after he finished reading the sentence and required recall of what he had said earlier. Controls often made error corrections, but never after so long a delay, and never in a way that rendered a grammatical utterance ungrammatical.

For all three sentence types (unambiguous, structurally ambiguous, and lexically ambiguous), H.M. misread at least 1.9 times as many sentences as the mean for controls (see Table 4, top panel). For unambiguous sentences, H.M. misread 2.67 SDs more sentences than controls, indicating that H.M.'s sentence-reading deficits are not solely attributable to ambiguity. This finding represents an important extension of previous research on H.M.'s language deficits which has relied heavily on data from ambiguous sentences (see MacKay et al., 1998a).

Interestingly, the difference between H.M. and controls was greater for unambiguous than lexically ambiguous sentences, even though unambiguous sentences were shorter (mean length = 8.5 words) and offered less opportunity for error than lexically ambiguous sentences (mean length = 10.4 words). Using words as unit of analysis to correct for differences in sentence length across the three sentence types (see Appendix), H.M. produced 3.8 SDs more error words per word per sentence than controls reading unambiguous sentences, but only 1.2 SDs more error words per word per sentence than controls reading lexically ambiguous sentences (see Table 4, bottom panel). However, H.M. produced 7.7 SDs more error words per word per sentence than controls reading structurally ambiguous

sentences, suggesting greater difficulty with structurally ambiguous than lexically ambiguous and unambiguous sentences: However, close inspection of the transcripts revealed no direct link between H.M.'s errors and ambiguities in structurally ambiguous sentences.

To summarise, Experiment 3 both replicated and extended the basic error results of Experiment 1, indicating that reading HF words in novel sentences is problematic for H.M. even for sentences that are unambiguous and short. H.M.'s sentence-reading deficits in the incidental reading task in Experiment 3 also indicate that H.M.'s deficits are not limited to tasks whose primary goal is fast and accurate reading.

EXPERIMENT 4: DOES AMBIGUITY DISRUPT H.M.'S SENTENCE-READING?

Experiment 4 definitively tested whether ambiguity per se influences H.M.'s reading. This issue is important because structural ambiguities can be expected to cause production interference during sentence reading (see the introduction; also MacKay et al., 1998a) and because several researchers have suggested that amnesics are especially sensitive to retrieval/production interference (Nadel, 1994; Warrington & Weiskrantz, 1974, 1978; see Shapiro & Olton, 1994; and Hayman, MacDonald, & Tulving, 1993, for reviews).

Experiment 4 procedures resembled MacKay (1966): H.M. read on different days two intermixed sets of ambiguous and unambiguous sentences that were identical except for a single word change that transformed each sentence from ambiguous (e.g., *The United States sent troops over a week ago*) to unambiguous (e.g., *The United States sent troops almost a week ago*) or vice versa, thereby allowing within-subject comparisons that controlled for length, and syntactic- and semantic-complexity of ambiguous versus unambiguous sentences.

Materials and procedure

All procedures resembled Experiment 1. Stimuli were 40 sentences and 24 isolated words. The sentences were syntactically simple and short ($M = 7.2$ words), contained relatively short HF words, and included 10 lexically ambiguous and 10 structurally ambiguous versions adapted from MacKay (1966), plus 20 unambiguous versions created by replacing a single word in ambiguous versions with a syntactically appropriate and semantically similar word of similar length and frequency. Each sentence was typed on a separate $4'' \times 6''$ index card in (mainly) lower case, 24 point bold Courier font. On day 1 (December 11, 1999), H.M. read a randomly selected and shuffled set of 5 lexically ambiguous versions, 5 structurally ambiguous versions, and the 10 remaining (non-overlapping) unambiguous versions,

and he read the 20 residual (counterbalanced) versions on day 2 (December 12, 1999).

Results and discussion

H.M. correctly read 45% of the unambiguous sentences versus 60% of the ambiguous sentences. The direction of this difference replicates Experiment 3, and contradicts three hypotheses: that ambiguity exacerbates H.M.'s reading deficits, that ambiguity causes special retrieval or production interference when amnesics such as H.M. read sentences aloud, and that ambiguity was responsible for H.M.'s sentence reading deficits in Experiment 1. Focusing on ambiguity types, H.M. correctly read 50% of the lexically ambiguous sentences versus 70% of the structurally ambiguous sentences. This finding comports with MacKay et al. (1998a) and indicates that structural ambiguities do not exacerbate H.M.'s reading deficits given appropriate controls for sentence length, and syntactic and semantic complexity.

GENERAL DISCUSSION

The General Discussion first summarises H.M.'s selective sentence-reading deficits and their implications for NST and other theoretical frameworks. We then discuss how present results bear on the three general issues that introduced this paper.

H.M.'s selective deficits in sentence-reading

More than ten aspects of the present results indicate that in addition to his other deficits, H.M. currently suffers from selective deficits in reading HF words in novel sentences: H.M. misread ambiguous and unambiguous sentences more often than controls, usually without self-correction, and his misreadings simplified meaning and syntax, and rendered sentences ungrammatical more than misreadings of controls. H.M. also produced each word more slowly than controls, and he paused longer than controls before beginning to produce a sentence, but his planning time deficits greatly exceeded his production time deficits (see Figure 3a). H.M. also introduced unusual pauses between words in unfamiliar phrases, and he failed to produce appropriate prosodic pauses at major constituent boundaries that were unmarked by punctuation.

However, H.M.'s prosody deficits were *selective* rather than *universal*. H.M. read familiar phrases with appropriate prosody, and he produced appropriate prosodic responses to commas and full stops. These selective deficits both replicate and extend the transcript-based observations of

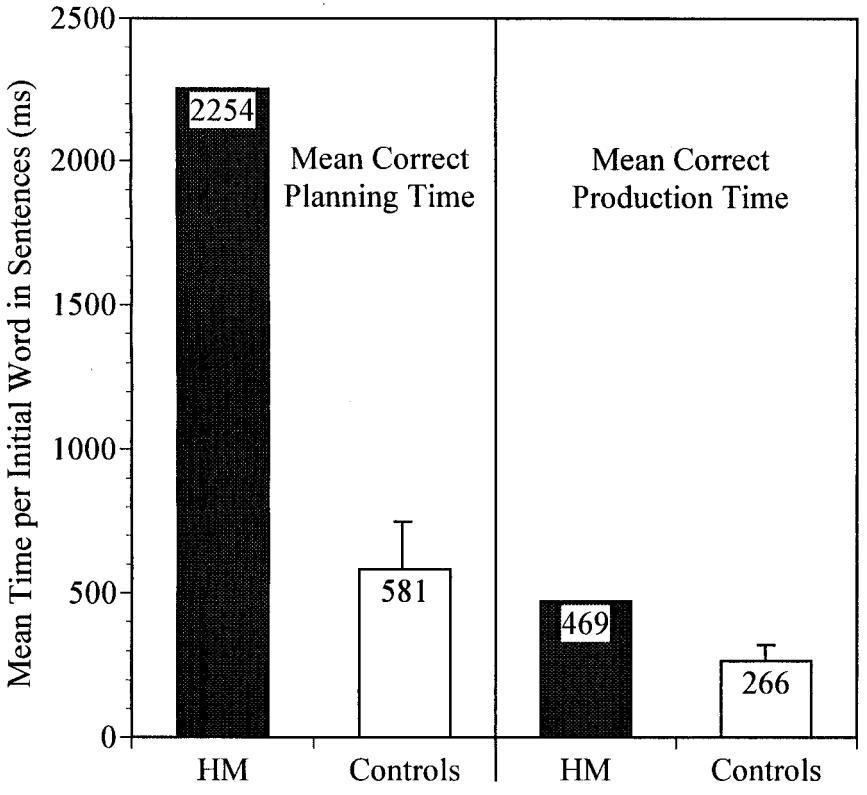

Figure 3a. Mean planning times (left panel) and production times (right panel) for Experiment 1 sentences with initial words correctly produced by H.M. and controls (error bars represent 1 *SD*).

MacKay et al. (1998a) that H.M. tends to misread short ambiguous sentences in ways that alter both meaning and syntax. For example, in Corkin's (1973) transcript, H.M. misread *John is the one to help today* as "John is the one that helped today", changing the verb meaning from future to past tense and changing the infinitive *to help* into a subordinate clause ("that helped"). However, present experiments go beyond MacKay et al. (1998a) by systematically comparing H.M. and controls reading a large number of sentences under different task demands, by examining prosody, phrase-familiarity and effects of punctuation, by demonstrating that H.M. misreads unambiguous sentences as often as he misreads ambiguous sentences, and in the same meaning- and syntax-altering way, and by introducing controls for effects of cerebellar damage and lexical-level factors.

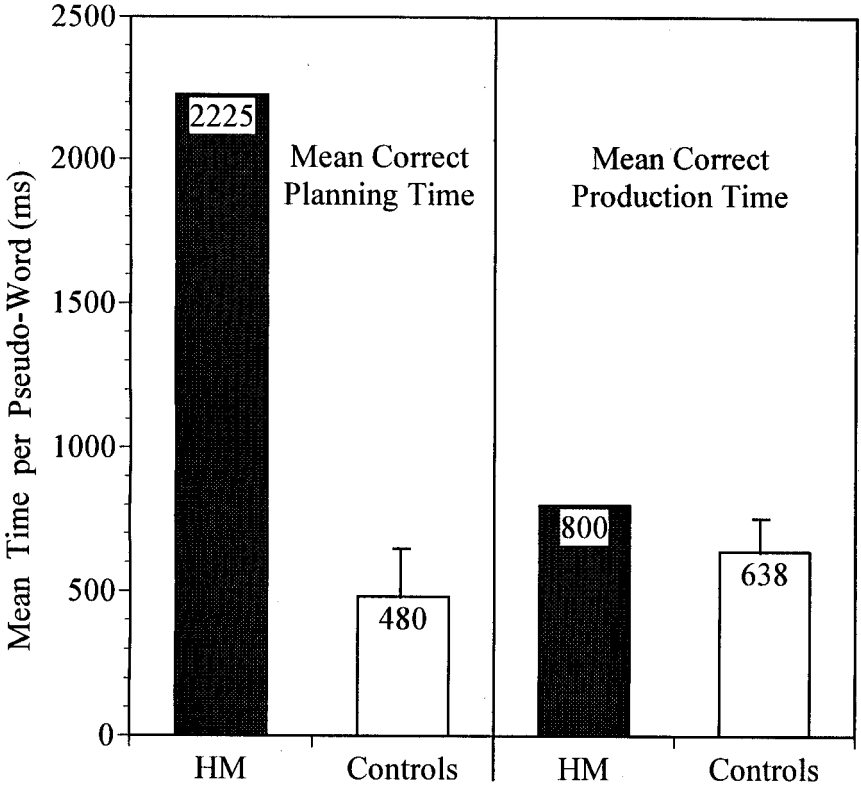

Figure 3b. Mean planning times (left panel) and production times (right panel) for pseudo-words correctly produced by H.M. and controls in MacKay and James (1999; Experiment 1).

Implications of H.M.'s sentence-reading deficits under NST

NST explains H.M.'s sentence-reading errors and abnormal pause patterns as reflecting failures to form or adequately strengthen the new top-down connections required to represent the plan for producing unfamiliar phrases and novel prosody during and prior to onset of production. However, H.M. could use old top-down connections to produce appropriate prosodic pauses within familiar phrases under NST. By way of illustration, Figure 4a shows the new versus old top-down connections that memory-normal controls (in theory) formed and used when reading aloud the proposition, "The boys who were fed hot dogs got stomach aches" (from sentence 9). For comparison, Figure 4b shows the limited number of old top-down connections that H.M. (in theory) used when he twice

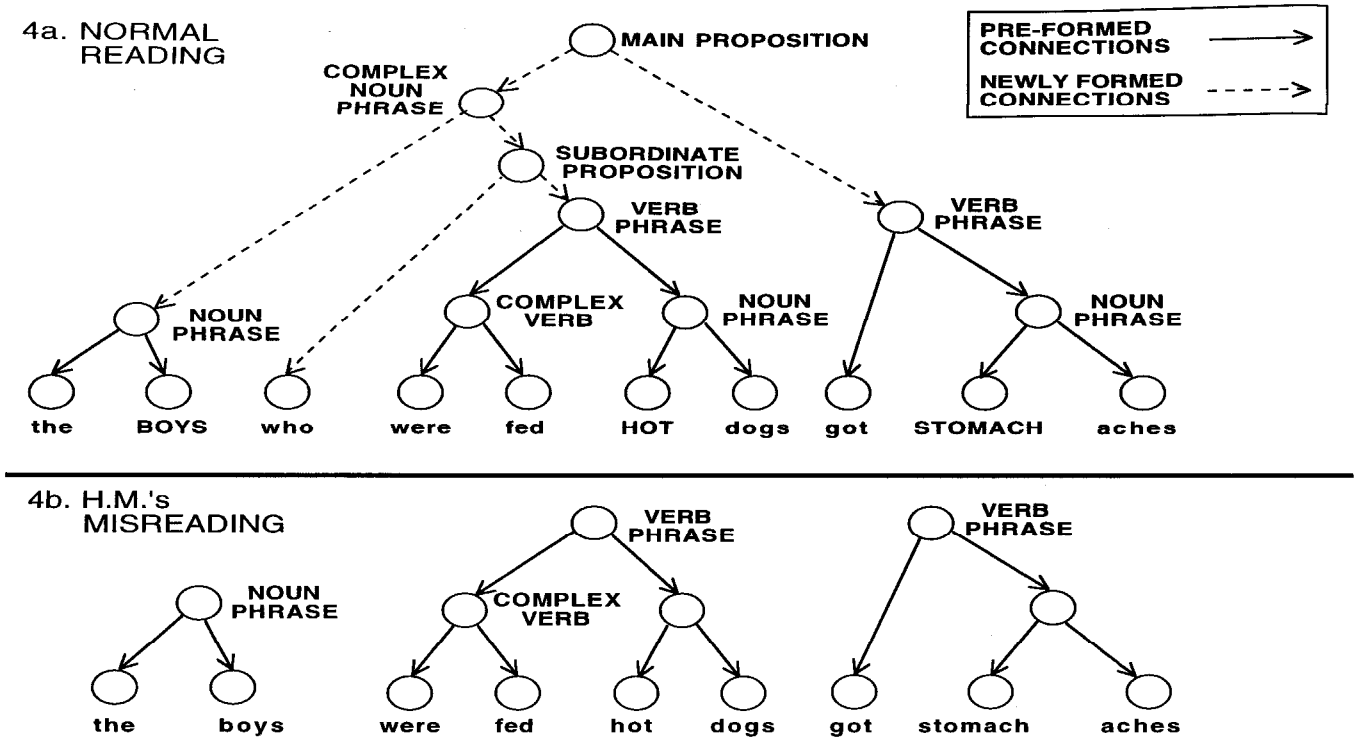

Figure 4a. Hypothetical top-down connections in NST's semantic system underlying normal production of, "The boys who were fed hot dogs got stomach aches" (from sentence 9; broken lines, solid lines, and capitalisation as in Figure 1).

Figure 4b. Top-down connections underlying H.M.'s sentence 9 misreading, "The boys were fed hot dogs got stomach aches".

misread this same proposition as, ''The boys were fed hot dogs got stomach aches''.

These NST hypotheses explain H.M.'s deficits in producing the prosody for unfamiliar but not familiar phrases and at major constituent boundaries unmarked by commas but not at ones marked by commas, and his greater deficits for planning times than for production times. NST also explains why H.M. tended to produce ungrammatical sentences with altered meaning and simplified syntax due to omission of HF connectors such as *but* and *who*. Even the fact that controls made relatively more errors than H.M. on semantically less-coherent vs. more-coherent sentences makes sense under NST: If H.M. is reading word-by-word or phrase-by-familiar phrase (see Figure 4b), unlike controls (see Figure 4a), then sentence-level semantic incoherence should disrupt controls relatively more so than H.M.

However, NST did *not* predict that production times for HF words in sentences would be 7 *SD*s longer for H.M. than controls (see Table 2; also MacKay & James, 1999). This finding may nonetheless comport with NST if production times or acoustic durations don't provide a pure measure of activation processes during sentence production: If semantic-level planning and phonological-level activation processes are interwoven during production of sentences (but not isolated HF words), then H.M.'s semantic-level binding deficits could slow down his production of HF words in sentences.

Implications of H.M.'s sentence-reading deficits for other major theories and hypotheses

Left-to-right reading. Under the left-to-right reading assumption, sentence reading involves production of each word from left to right in the coherent, grammatical and comprehensible order provided on the page. This being the case, H.M.'s ''semantic binding deficits'' might apply to extemporaneous speech where new connections are required for assembling ideas into a coherent and grammatical plan when communicating some novel state of the world (see MacKay et al., 1998a), but not to sentence reading, where readers simply produce each word in the pre-assembled left-to-right order on the page.

Left-to-right reading is the foundation assumption in a complex, three-syndrome theory of selective reading deficits that closely resemble H.M.'s (Friedman, 1995; 1996a; discussed shortly). However, many aspects of the present data contradict the left-to-right reading assumption. An example is the fact that H.M. correctly read words presented in isolation that he misread in sentences (see Experiment 2; also MacKay & James, 1999): Given left-to-right reading, H.M. should read a sentence correctly if he can read its constituent words in isolation.

General cognitive decline. Under the "general cognitive decline" hypothesis, brain damage is an information-loss factor that causes "general slowing" or proportionately longer response times across the board for all units and for all cognitive tasks (see e.g., Myerson, Hale, Wagstaff, Poon, & Smith, 1990). This "general cognitive decline" hypothesis predicts proportionately longer pauses and response times in many tasks for H.M. relative to age-matched controls without brain damage (see Postle & Corkin, 1998), and predicts that all units in H.M.'s reading should be slowed in equal proportion relative to controls, e.g., H.M. should pause after the same units in a sentence as controls, but for proportionately longer times.

Present results contradicted all of these "general cognitive decline" predictions (for other contradictory results, see Laver & Burke, 1993, and MacKay & James, 1999): Relative to controls, H.M.'s pauses were disproportionately longer after some units (unfamiliar phrases, and at major constituent boundaries unmarked by commas) but not others, and H.M.'s planning times were slowed disproportionately relative to his production times. The stability of H.M.'s IQ from 1977–97 also contradicts a general cognitive decline hypothesis, as does H.M.'s selective difficulty in reading HF words in sentences but not in isolation (Experiment 2; see also MacKay & James, 1999).

Working-memory capacity. According to Just and Carpenter (1992), working-memory capacity influences reading comprehension, which suggests the hypothesis that H.M.'s reduced working-memory capacity may cause his sentence-reading deficits. However, a factor such as working-memory capacity that is fixed and general in nature might explain across-the-board declines, but cannot explain the highly selective deficits we identified in H.M. Working-memory limitations also fails to explain why H.M.'s sentence-reading deficits were relatively greater for semantically more- vs. less-coherent sentences, and did not differ for short vs. long sentences, or for syntactically complex vs. syntactically simple sentences.

Stages-of-processing framework. H.M.'s sentence-reading deficits contradict the widely accepted stages-of-processing framework (see MacKay et al., 1998a) wherein H.M. suffers from a pure memory deficit that has spared his language production abilities (including reading aloud).

The procedural-memory hypothesis. H.M.'s sentence-reading deficits contradict the hypothesis of Squire (1987, pp. 151–169) and Cohen and Eichenbaum (1993, pp. 49–219) that procedural-memory (a theoretically distinct and separate store for generating cognitive skills such as reading) is spared in amnesics (including H.M.). Other parallels between H.M.'s

spared versus impaired performance in reading and memory (discussed shortly) are also problematic for the assumption that declarative-, episodic- and procedural-memories involving language engage entirely separate stores (see Shanks, 1996, for a review of other problems with the procedural-memory hypothesis).

THREE GENERAL ISSUES REVISITED

Issue I: Does H.M. pause normally in speech and reading?

Present results indicate that H.M. produces normal-duration pauses in reading familiar phrases and at major constituent boundaries marked by commas, but abnormal-duration pauses elsewhere and prior to sentence onset, an indication of planning trouble. However, further research is needed to determine whether H.M. exhibits similar prosodic deficits in spontaneous speech. Milner et al.'s (1968) observations focused on the ''monotonic'' nature of H.M.'s speech, and may apply only to emotional tone or prosodic intonation rather than to prosodic pauses (Levelt, 1989, p. 364). Given established relations between amygdala and emotional tone (see e.g., Gray, 1982, p. 5), further research on relations between H.M.'s amygdala damage and his prosodic intonation in both reading and speaking seems warranted.

Issue II: Does H.M. produce analogous errors in speech and reading?

For both ambiguous and unambiguous sentences, H.M. produces remarkably similar omission errors in reading and spoken sentence production (see Figure 4b and the introduction for typical examples), and NST explains these analogous omissions as due to semantic-level binding deficits, i.e., failures to form or adequately strengthen the new connections required to represent the plan for producing novel aspects of a sentence prior to onset of production (see MacKay et al., 1998a).

However, H.M.'s errors in reading versus speech also exhibited differences. One difference concerned extensiveness: Unlike H.M.'s reading errors, H.M.'s errors in novel spoken discourse were so extensive as to render his output incoherent and incomprehensible (see MacKay et al., 1998a). This difference seems attributable to the differing representations of H.M.'s intended output in reading versus speaking: When reading, the sentence on the page provides a strong and enduring representation that is absent when speaking spontaneously, but could in principle guide all aspects of H.M.'s sentence-reading except for (unpunctuated) prosody. If in the present experiments H.M. had *solely* read word-by-word and

extremely slowly with no attempts at normal prosody, his output probably would have been error-free (see MacKay & James, 1999). From this perspective, what is interesting is not that H.M. made fewer errors in reading than extemporaneous speech, but that he made any non-prosodic (e.g., omission) errors whatsoever in reading HF words in sentences.

Another difference is that H.M. produced more error corrections in reading (27% in Experiment 1) than speech (0% in MacKay et al., 1998a). This difference also may reflect the strong and enduring visual representation of sentences available during reading. When reading, H.M. (and the experimenter) can immediately compare the words on the page with his spoken output to enable error correction (and experimenter prompts regarding errors). However, H.M. cannot correct his novel spoken discourse in this way because coherent or ordered representations of his intended output are either absent (see Figure 4b) or too weak and short-lived to compare with his spoken output, and the experimenter cannot call for specific error-corrections in spoken output that is incoherent and incomprehensible (see MacKay et al., 1998a).

Issue III: Does a coherent theoretical syndrome underlie all of H.M.'s deficits?

Assigning patients to syndromes represents a common and longstanding practice in neuropsychology (see e.g., Shallice, 1988), and Issue III was whether H.M.'s "hippocampal amnesic syndrome" is part of a much more general, theoretically coherent syndrome that runs through all of his deficits: in reading sentences, in reading isolated LF words and pseudo-words, in producing spoken sentences, in visual cognition, and in recall from episodic memory. To address Issue III, we examine three contrasting approaches for establishing syndromes: the surface-, anatomical-, and theoretical-syndrome approaches.[7] We illustrate some limitations of the surface- and anatomical-syndrome approaches wherein H.M.'s deficits represent neither a surface-syndrome nor an anatomical-syndrome that is coherent, generalisable, and explanatory in nature. However, under the theoretical-syndrome approach H.M. does exhibit a coherent syndrome that carries explanatory power, generalises to other patients, and suggests interesting new questions for future test.

The surface-syndrome approach and its limitations. Under the surface-syndrome approach the basic goal is to search for similarities in the observed deficits of various patients, who are then grouped into syndromes

[7] Although the three approaches can each be illustrated via at least one pure example-study, example-studies representing mixtures of these approaches can also be found.

or similarity categories in the hopes of discovering a common (explanatory) cause for their shared deficit-pattern. We illustrate this approach and its limitations via a minor sub-aspect of Issue III: Do H.M.'s deficits in reading sentences and isolated words fit one or more of the many already-postulated dyslexic syndromes (see e.g., Ellis & Young, 1988, p. 199–221)?

Table 5 summarises some of the surface similarities between H.M.'s deficits in reading sentences (present study) and isolated words/pseudo-words (MacKay & James, 1999) that might suggest a common (explanatory) cause or set of causes underlying H.M.'s reading deficits under the surface-syndrome approach. H.M. tended to produce morphological and wrong-word errors for both sentences and isolated words, and he tended to leave both types of errors uncorrected. Reduced phonological overlap with target words also characterised H.M.'s reading errors for sentences and isolated words, with a similar omission-to-addition ratio for non-overlapping segments. Stress or emphasis shifts also characterised H.M.'s reading of both sentences and isolated words, although stress involved entire words in sentences versus syllables in isolated words. H.M.'s reading also exhibited remarkably similar temporal characteristics for words in sentences versus pseudo-words presented in isolation: H.M.'s planning

TABLE 5
Empirical similarities between H.M.'s reading deficits for isolated words (MacKay & James, 1999) versus sentences (present experiments) and their theoretical bases under NST

	Empirical similarities		
Characteristic	*Isolated words and pseudo-words*	*Sentences containing HF words*	*Theoretical bases under NST*
Correct production times	Relatively small deficit for HF words (only)	Relatively small deficit	Relatively intact activation processes
Error correction	Rare for H.M.	Rare for H.M.	Damaged binding nodes
Abnormal pauses	Mainly within LF words and pseudo-words	Mainly within unfamiliar phrases	Damaged binding nodes
Planning times	Large deficit for LF words and pseudo-words	Large deficit	Damaged binding nodes
Phonological similarity	Low similarity	Low similarity	Damaged binding nodes
Inappropriate stress	Very common (on syllables in LF words)	Very common (on words)	Accidental similarity

times revealed large deficits for correctly produced pseudo-words and for sentences containing HF words (compare Figures 3ab, left panel), but H.M.'s production times revealed small deficits for correctly-produced pseudo-words and sentences (compare Figures 3ab, right panel).

Such similarities suggest that H.M.'s sentence and isolated-word deficits may reflect a coherent syndrome under the surface-syndrome approach. However, H.M.'s sentence and isolated-word deficits also exhibited differences (summarised in Table 6). One set of differences concerned word frequency and word length. H.M. only reliably misread LF words in isolation, but he misread extremely HF words in sentences. Indeed, H.M. misread words with much higher mean frequency in sentences (7560 per million in Experiment 1) than he read *correctly* in isolation (60 per million in MacKay & James, 1999, Study 1). H.M. also misread shorter words in sentences (M = 5.2 letters in Experiment 1) than in isolation (M = 8.2 letters in MacKay & James).

A second set of surface differences concerned some unusual ways that H.M. misread words/pseudo-words, classified as follows in MacKay and James (1999): wrong-word misreadings, e.g., *satisfy* misread as "sanctify"; repetition of previously produced or perceived segments or letter units; segment order errors, e.g., the pseudo-word *boshertin* /'bɑʃərtɪn/ misread as "barshite" /'bɑrʃaɪt/ with the SH and R produced in the wrong order; successive approximations, e.g., *abdicate* misread in sequence, "abicurgle ... duh ... abidackle ... abidickle"; /'æbɪkərgɥl ... 'dɥ ... æbɪ'dækl ... æbɪ'dɪkl/; stress-shift errors, e.g., *labyrinth* (LAbyrinth with capitalisation indicating primary stress /'læbərɪn' /) misread as "labrinth" (laBRINTH /

TABLE 6
Empirical differences between H.M.'s reading deficits for isolated words and pseudo-words (MacKay & James, 1999) versus sentences (Experiments 1 and 2)

Empirical dimension	*Isolated words and pseudo-words (MacKay & James, 1999)*	*Sentences containing HF words (present experiments)*
Word frequency	LF	Extremely HF
Word length in letters	Relatively long	Relatively short (and not different from controls)
Repetitions of earlier units	Very common	Virtually 0%
Successive approximations	Very common	Virtually 0%
Syllabic segmentation errors	Very common	0%
Wrong-word misreadings	Very common	Not different from controls
Order errors	Very common	0%
Typical morpheme error	Substitution/misproduction	Omission
How errors relate to meaning	Not related (under the surface-syndrome approach)	Simplify sentence meaning
How errors relate to syntax	Not relevant (under the surface-syndrome approach)	Simplify sentence syntax

lə'brɪn' /), and syllabic segmentation errors, i.e., production of an appropriate speech sound in the wrong syllabic position, e.g., *papyrus* /pə'paɪrəs/ misread as "PAPryism" /'pæpriɪzm/, where the second [P] is syllable-initial in the stimulus, but syllable-final in H.M.'s response. By contrast, when H.M. read HF words in sentences, his wrong-word misreadings did not differ from controls', he virtually never produced successive approximations, he never produced order errors or syllabic segmentation errors, and he never repeated units that he had perceived or produced in earlier sentences (see Table 6).

A third set of surface differences concerned the pattern of H.M.'s morpheme errors for words in sentences (e.g., *training* misread as "train") versus in isolation, as when H.M. misread the isolated pseudo-word *metalousness* /mə'tæləsnəs/ as "metalness" /'mɛtəlnəs/, omitting the morpheme [OUS] /əs/. H.M.'s typical morpheme error involved omissions in sentences, but segment errors and morpheme substitutions in isolated words/pseudo-words (see Table 6). Overall (excluding added morphemes), morpheme omissions made up 50% of H.M.'s morpheme errors in sentences (Experiment 1), but only 25% of H.M.'s morpheme errors in isolated words (MacKay & James, 1999).

Under the surface-syndrome approach, such differences suggest that H.M.'s reading deficits cannot in principle reflect a single, coherent dyslexic syndrome, but may reflect a syndrome-mixture, i.e., one dyslexic syndrome for isolated words/pseudo-words (say Type II surface dyslexia; see e.g., Behrmann & Bub, 1992) and another dyslexic syndrome for HF words in sentences (say phonological text alexia; see Friedman, 1996a). However, syndrome-mixtures raise the more general issue of whether syndromes established under the surface-syndrome approach carry causal or explanatory significance (see also Caramazza & McCloskey, 1988): Applying one diagnostic label to H.M.'s sentence-reading deficits and another to his reading deficits for isolated words/pseudo-words explains nothing and would only serve to downplay the empirical or surface similarities (reviewed earlier) between H.M.'s reading deficits for sentences versus isolated words.

The final limitation of the surface-syndrome approach discussed here (for additional limitations, see Caramazza & McCloskey, 1988; Caramazza & Badecker, 1991) concerns the degree of similarity necessary for assigning patients to a syndrome. Consider H.M.'s order errors for LF words and pseudo-words, e.g., production of SH and R in the wrong order in the pseudo-word boshertin (misread as "barshite" /'bɑʃaɪt/). Do such order errors suffice to classify H.M. as an "attentional dyslexic", a category of dyslexics who also make order errors involving single segments (see e.g., Allport, 1977; Shallice & Warrington, 1980)? Or do differences between H.M.'s order errors and those of attentional dyslexics (variously known as

"visual segmentation" or "migration" errors) indicate a different syndrome-assignment? For example, attentional dyslexics generally only misorder or intrude letters that share identical syllabic position, whereas H.M. always misordered letters from different syllabic positions in isolated words/pseudo-words in MacKay and James (1999). Similarly, attentional dyslexics misorder letters between adjacent HF words in word strings, e.g., *win fed* misread as "fin fed", and *live lone* misread as "love lone" (see Ellis & Young, 1988, p. 198), whereas H.M. never misordered segments in HF words in sentences (see Table 6).

The anatomical-syndrome approach and its limitations. Under the anatomical-syndrome approach, the basic goal is to examine patients with lesions restricted to particular anatomical areas or structures, e.g., the hippocampus, and to specify the deficits common to this category of patients in the hopes of discovering the functions of that anatomical area or structure. To illustrate this approach and its limitations, we examine the hypothesis of Schmolck and Squire (1999) that H.M.'s damage is not as clean as was once thought, so that H.M.'s memory deficits or "hippocampal amnesia" arise solely from his damaged hippocampus, whereas his other, newly discovered deficits involving language and visual cognition are attributable either to aspects of H.M.'s personal history, including especially his pre-surgical seizure history and his life-history post-surgery, or to brain damage elsewhere, including *potential* damage or atrophy to as-yet-undiscovered areas or microstructures.

The anatomical-syndrome approach underlying the Schmolck–Squire hypothesis carries all of the limitations of the surface-syndrome approach discussed earlier, plus some new ones. One is whether deficits such as "hippocampal amnesia" can *in principle* be considered "pure". For example, no one has succeeded in establishing a dividing line, either empirically or theoretically, between where memory storage and retrieval involving verbal materials ends and where language production begins (see Bock, 1996; MacKay & Abrams, 1996; MacKay et al., 1998a), and until someone does so, neither pure memory deficits nor "pure hippocampal amnesia" are possible in principle (see also MacKay et al., 1998b, for analogous dividing-line problems involving comprehension and memory).

Another limitation of the anatomical-syndrome approach concerns between-patient differences in anatomical damage. Because no two patients have ever experienced precisely the same brain damage, one must again ask the dividing-line question: How similar must their brain damage be to assign two patients to an anatomical syndrome? Lack of correspondence between anatomical structures and actual or potential damage raises a similar issue. For example, no patient has ever experienced actual or potential damage to the entire hippocampus and

only to the entire hippocampus (including H.M.). So determining whether some function is attributable to one observed (let alone potential) area of brain damage rather than another is extremely difficult *for a single patient*, let alone for a category of patients (and perhaps impossible; see Caramazza & McCloskey, 1988).

A final limitation of the anatomical-syndrome approach concerns explanatory power, the ultimate goal of all syndrome-categories. For example, assigning H.M.'s selective reading deficits to his epilepsy, his medication history, his slightly damaged polar tips, or to some other as-yet-undiscovered atrophy or micro-structural damage explains nothing. Explanatory power concerning deficits traceable to epilepsy, the polar tips, or any other anatomical structure requires answers to questions such as these: How does epilepsy and/or damaged polar tips enable H.M. to read HF words in isolation but not in novel sentences? How does epilepsy and/or damaged polar tips enable H.M. to produce appropriate pauses for major constituent boundaries marked with commas, but not for ones unmarked with commas? How does epilepsy and/or damaged polar tips cause H.M. to produce unfamiliar but not familiar phrases with long pauses and unusual patterns of pitch and stress? Why does epilepsy and/or damaged polar tips cause greater deficits in planning times than production times? The anatomical-syndrome approach is fundamentally correlational in nature and unlikely to raise such questions let alone answer them.

The theoretical-syndrome approach. The basic goal of the theoretical-syndrome approach is to develop a theory with explanatory mechanisms that accurately predict specific patterns of spared and impaired function in one or more patients who can be said to exhibit a syndrome with causal relations to those mechanisms. Four recent papers have adopted this theoretical-syndrome approach to H.M.'s selective deficits in reading sentences (present research), in explicit versus implicit memory for novel versus familiar verbal information and in producing spoken sentences (MacKay et al., 1998a), in reading isolated words and pseudo-words (MacKay & James, 1999), and in analysing hidden figures and comprehending spoken and written sentences (MacKay et al., 1998b). Because all four papers postulated a single explanatory mechanism (damaged binding nodes) that predicted H.M.'s spared and impaired abilities with remarkable accuracy, H.M.'s deficits form a coherent pattern that we call the binding syndrome. No other established syndrome or syndrome-mixture captures these strong theoretical parallels spanning memory, visual cognition, reading, and spoken language.

How do the theoretical- and surface-syndrome approaches differ? One difference is that empirical similarities between deficit-patterns within a patient or between different patients do not necessarily indicate a common

cause under the theoretical-syndrome approach. What is needed is a theory of the processes underlying these empirical similarities. By way of illustration, NST specifies a common underlying cause for some empirical similarities in Table 5, but specifies different underlying causes for others, and no causal relation whatsoever for still others. For example, the first empirical similarity between reading sentences versus isolated words in Table 5 (H.M.'s relatively small production time deficits for HF words in sentences and in isolation) reflects an underlying cause (relatively intact activation processes) that is unrelated to H.M.'s binding syndrome under NST. However, the next four empirical similarities in Table 5 reflect H.M.'s damaged binding nodes, and are therefore part of his binding syndrome under NST, even though the syntactic, semantic, and prosodic binding nodes for reading sentences differ from the orthographic and phonological binding nodes for reading words and pseudo-words (see MacKay & James, 1999). And the final empirical similarity in Table 5 (H.M.'s stress errors for words in sentences versus syllables in words) is accidental or has two unrelated causes.[8]

A second difference between the theoretical- and surface-syndrome approaches is that some empirical differences (either between patients or between deficit-patterns within a patient) are large and important under the surface-syndrome approach, but unimportant or irrelevant under the theoretical-syndrome approach. For example, the length and frequency differences between words that H.M. misread in sentences versus in isolation were extremely large and therefore important under the surface-syndrome approach, but represent artifacts under the theoretical-syndrome approach: H.M.'s syntactic binding deficits especially impact words linked via new connections to the structure of a sentence (see the Introduction), and HF function-words often happen to exhibit this characteristic in sentences (but not in isolation).

Conversely, important differences under the theoretical-syndrome approach can be trivial or extraneous under the surface-syndrome approach. For example, the most important differences between H.M.'s reading errors involving sentences versus isolated words under the theoretical-syndrome approach were the context-dependent nature of his sentence errors and the way they simplified sentence meaning and syntax: All three characteristics follow from H.M.'s semantic binding deficits (see

[8] This similarity is accidental under NST because totally differing causes underlie H.M.'s stress assignment errors for syllables in LF words versus HF words in sentences. For sentences, H.M. had difficulty computing novel syntax (essential for assigning appropriate sentential stress), but for LF words, H.M. was trying out different patterns of syllabic stress in an attempt to activate the non-functional but "old" connections for once familiar LF words (see MacKay & James, 1999).

the Introduction). However, meaning and syntax in sentences might seem so unrelated to reading isolated words under the surface-syndrome approach as to render these differences inconsequential or meaningless (see Table 6). Indeed, the surface-syndrome approach might consider context-effects a *similarity* rather than a difference between H.M.'s reading of isolated words vs. sentences: After all, when H.M. repeated units previously perceived or produced in isolated words (see MacKay & James, 1999), these too were context-effects.

How do the theoretical- and anatomical-syndrome approaches differ? One difference is that under the theoretical-syndrome approach, theories can be committed to a particular theoretical construct or causal mechanism without being committed to a specific brain locus for that causal mechanism. For example, many NST binding nodes are probably located in bilateral hippocampal structures and connected entorhinal, perirhinal, and parahippocampal areas given current information (see e.g., Milner, 1975; O'Keefe & Nadel, 1978). However, there is no *theoretical* reason in NST or any other theory for localising a particular theoretical construct to a particular brain structure such as the hippocampus. What precise brain areas house different types of NST binding nodes is clearly an empirical question, and an especially complex and challenging one since it is possible in principle for hippocampal system damage to diminish the efficacy of binding nodes located *outside* hippocampal systems and vice versa (see MacKay, 1990).

Does H.M.'s binding syndrome generalise to other patients? It turns out that Friedman (1995; 1996a, b) has described several dyslexics with reading deficits remarkably similar to H.M.'s. For example, Friedman-dyslexics misread isolated pseudo-words, and HF words in sentences (especially short function-words) but not in isolation, and some Friedman-dyslexics also misread isolated LF words. Under the theoretical-syndrome approach, such similarities suggest that Friedman-dyslexics share H.M.'s binding syndrome because NST postulates causal relations between this specific pattern of deficits.

Assigning H.M. and Friedman-dyslexics to the same syndrome under the theoretical-syndrome approach suggests many interesting questions for further research. For example, do Friedman-dyslexics exhibit H.M.'s pattern of greatly impaired planning times versus relatively unimpaired production times for isolated pseudo-words and sentences containing HF words? Do Friedman-dyslexics tend to repeat units perceived or produced earlier in isolated LF words and pseudo-words? Do Friedman-dyslexics have damaged hippocampal systems? Do other dyslexics with damaged hippocampal systems produce morpheme errors, e.g., *internal* misread as ''international'', an error typical of H.M., deep dyslexics (see Shallice, 1988, p. 99), and some Friedman-dyslexics (Friedman, 1996b)? Do deep

dyslexics have damaged hippocampal systems? Answers to these questions are currently unknown, but the balanced focus on theory as well as detailed similarities and differences between empirical deficits in the theoretical-syndrome approach clearly suggests fruitful directions for further research.

Manuscript received October 1999
Revised manuscript received June 2000

REFERENCES

Allport, D.A. (1977). On knowing the meaning of words we are unable to report: The effects of visual masking. In S. Dornic (Ed.), *Attention and performance* (Vol. 6). Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.

Behrmann, M., & Bub, D. (1992). Surface dyslexia and dysgraphia: Dual routes, single lexicon. *Cognitive Neuropsychology*, *3*, 209–251.

Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review*, *3*, 395–421.

Caramazza, A., & Badecker, W. (1991). Clinical syndromes are not God's gift to cognitive neuropsychology: A reply to a rebuttal to an answer to a response to the case against syndrome-based research. *Brain & Cognition*, *16*, 211–227.

Caramazza, A., & McCloskey, M. (1988). The case for single-patient studies. *Cognitive Neuropsychology*, *5*, 517–528.

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

Cooper, W.E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Cambridge, MA: Harvard University Press.

Corkin, S. (July, 1973). H.M.'s comprehension and description of ambiguous meanings. Unpublished transcript, Dept. of Brain and Cognitive Sciences, MIT, 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.

Cutler, A. (1987). Speaking for listening. In A. Allport, D.G. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading, and writing*, pp. 23–40. London: Academic Press.

Deese, J. (1984). *Thought into speech: The psychology of a language.* Englewood Cliffs, NJ: Prentice-Hall.

Ellis, A.W., & Young, A.W. (1988). *Human cognitive neuropsychology.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Francis, W.N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Friedman, R.B. (1995). Two types of phonological text alexia. *Cortex*, *31*, 397–403.

Friedman, R.B. (1996a). Phonological text alexia: Poor pseudo-word reading plus difficulty reading functors and affixes in text. *Cognitive Neuropsychology*, *13*, 869–885.

Friedman, R.B. (1996b). Recovery from deep alexia to phonological alexia: Points on a continuum. *Brain and Language*, *52*, 114–128.

Gee, J.P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*, 411–458.

Gray, J.A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system.* New York: Oxford University Press.

Hayman, C.A.G., MacDonald, C.A., & Tulving, E. (1993). The role of repetition and associative interference in new semantic learning in amnesia. *Journal of Cognitive Neuroscience*, *5*, 375–389.

Helmuth, L.L., Ivry, R.B., & Shimizu, N. (1997). Preserved performance by cerebellar patients on tests of word generation, discrimination learning, and attention. *Learning and Memory*, *3*, 456–474.

Ivry, R.B., & Keele, S.W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*, 136–152.

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122–149.

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.

Laver, G.D., & Burke, D.M. (1993). Why do semantic priming effects increase in old age? A meta-analysis. *Psychology & Aging*, *8*, 34–43.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

MacKay, D.G. (1966). To end ambiguous sentences. *Perception and Psychophysics*, *1*, 426–435.

MacKay, D.G. (1987). *The organization of perception and action: A theory for language and other cognitive skills.* New York: Springer-Verlag.

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., & 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, Fourth Edition*, pp. 251–265. San Diego: Academic Press.

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

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. (1999). *H.M.'s selective reading deficits for isolated words and pseudo-words: Relations between aging and the binding problem for phonology and orthography*. Manuscript submitted for publication.

MacKay, D.G., & Miller, M. (1996). Can cognitive aging contribute to fundamental psychological theory? Repetition deafness as a test case. *Aging, Neuropsychology, and Cognition*, *3*, 1–18.

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.

Miller, M.D., & MacKay, D.G. (1994). Repetition deafness: Repeated words in computer-compressed speech are difficult to encode and recall. *Psychological Science*, *5*, 47–51.

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

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.

Myerson, J., Hale, S., Wagstaff, D., Poon, L.W., & Smith, G.A. (1990). The information-loss model: A mathematical theory of age-related cognitive slowing. *Psychological Review*, *97*, 475–487.

Nadel, L. (1994). Multiple memory systems: What and why; an update. In D.L. Schachter & E. Tulving (Eds.), *Memory systems 1994*, pp. 39–63. Cambridge, MA: MIT Press.

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 Inc.

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

Postle, B.R., & Corkin, S. (1998). Impaired word-stem completion priming but intact perceptual identification priming with novel words: Evidence from the amnesic patient H.M. *Neuropsychologia*, *36*, 421–440.

Schmolck, H., & Squire, L. (1999, April). *Amnesic patients can detect and explain ambiguous sentences.* Poster presented at the annual meeting of the Cognitive Neuroscience Society, Washington, DC.

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

Shallice, T. (1988). *From neuropsychology to mental structure.* New York: Cambridge University Press.

Shallice, T., & Warrington, E.K. (1980). Single and multiple component central dyslexic syndromes. In M. Coltheart, K.E. Patterson, & J.C. Marshall (Eds.), *Deep dyslexia.* London: Routledge.

Shanks, D. (1996). *Learning and memory.* In D.W. Green et al. (Eds.), *Cognitive Science: An introduction*, pp. 276–309. Oxford: Blackwell.

Shapiro, M.L., & Olton, D.S. (1994). Hippocampal function and interference. In D.L. Schachter, & 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.

Warrington, E.K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia*, *12*, 419–428.

Warrington, E.K., & Weiskrantz, L. (1978). Further analysis of the prior learning effect in amnesic patients. *Neuropsychologia*, *16*, 169–177.

Wingfield, A., & Butterfield, B. (1984). Running memory for sentences and parts of sentences: Syntactic parsing as a control function in working memory. In H. Bouma & D.G. Bouhuis (Eds.), *Attention and performance X: Control of language processes*, pp. 351–363. Hove, UK: Lawrence Erlbaum Associates Ltd.

APPENDIX

Sentences for Experiment 1 (1–11) and example sentences (12–21) for ambiguity categories in Experiment 3

Sentences for Experiment 1

1. I usually wore sandals, even though the sand on the beach in Hawaii was fine.
2. Kevin was frightened by the bat in the bushes but knew that when you learn gradually you worry less.
3. After eating lunch she walked to the post office, mailed some letters, and then went to see the movie at the new theater.
4. The superb athletes rose early in May because they were training for the Olympics.
5. The cotton farmers spoke about bad floods just before harvest time. (no errors)
6. For lunch Bill is having either pork or chicken and fries. (no errors)
7. David read the review literally learning nothing.
8. Lisa remarked that the stars are quite visible in Los Angeles. (no errors)
9. Although the boys who were fed hot dogs got stomach aches, the genie ate the golden figs in the ancient temple.
10. The police officer, watching the cat run through the dark alley, asked the witness to describe the thief.
11. The young girls' yellow ball fell from the roof onto the lawn while they giggled in the hallway.

Structurally ambiguous sentences (Experiment 3)

12. The spy put out a torch which was our signal to attack.
13. The men with the women who were complaining were told to move along.
14. Her only choice was to throw it up immediately or die.
15. He threw out the suggestion that the trouble was due to a communist plot.

Lexically ambiguous sentences (Experiment 3)

16. When a strike was called it surprised everyone.
17. He was not able to handle the case by himself.

Unambiguous sentences (Experiment 3)

18. My friend gave me a necklace for my birthday.
19. Jennifer lacked motivation to finish reading that story.
20. She did not like to walk in the rain.
21. The cat always hid under the couch when dogs were in the room.