Aging Neuropsychology and Cognition
2002, Vol. 9, No. 4, pp. 298–333

1382-5585/02/0904-298$16.00

Aging, Retrograde Amnesia, and the Binding Problem for Phonology and Orthography: A Longitudinal Study of "Hippocampal Amnesic" H.M.

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

ABSTRACT

This study develops and tests a theory of aging and long-term retrograde amnesia (RA) that extends to word retrieval, including the seemingly simple retrieval task of reading isolated words. Under the theory, transmission deficits due to aging, nonrecent use of connections, and infrequent use of connections over the lifespan cause mild and reversible RA in normals but severe and irreversible RA in amnesics who cannot readily form new connections to replace nonfunctioning ones. Consistent with this theory, "hippocampal amnesic" H.M. exhibited little or no retrieval deficit relative to memory-normal controls in reading short, moderately high frequency words at ages 60 or 71 but exhibited accelerated age-linked declines for low frequency words that were unrelated to cerebellar function, working memory capacity, practice effects, speed-accuracy trade-off, and sensory or attentional deficits.

This article reports three studies. Study 1 develops and tests a new theory of aging and retrograde amnesia (RA), defined here as long-term deficits in representing and/or retrieving information acquired years and sometimes decades before the amnesia-causing trauma (for other types of RA, see Hodges, 1995; Kapur, 1993). This RA theory is a derivative of Node Structure Theory (NST; MacKay, 1987, 1990) known as the Node Structure Aging Theory (NSAT). NSAT encompasses aging, memory, and language more generally, but we apply it to word retrieval here since tests of episodic memory in RA research have often involved specific proper names for people, places, and events. The General Discussion section outlines other theories relevant to more limited aspects of results in Studies 1–3.

To test interactions between aging and amnesia predicted under NSAT, Study 1 examined the longitudinal performance of "hippocampal amnesic" H.M. at ages 60 and 71 in the seemingly simple retrieval task of reading isolated words. Study 2 then adopts the NST framework to develop a detailed theory of reading that served to guide more detailed analyses of H.M.'s age 71 reading performance for pseudo-words and high- versus low-frequency words. Using various control procedures, Study 3 reran Study 2 with H.M. and 3 cerebellar patients to test for possible relations between H.M.'s deficits in Studies 1–2 and his cerebellar damage. Study 3 also exam-ined effects of repetition on H.M.'s reading, including the test-retest consistency of his errors.

NSAT: A THEORY OF AGING AND RA AS APPLIED TO WORD-RETRIEVAL

A major challenge facing theories of RA is to explain the complex relations between RA and

---

*Lori E. James is now at University of Colorado, Colorado Springs, USA.

Address correspondence to: Dr. Don MacKay, Psychology Department, UCLA, Los Angeles, CA 90095-1563, USA. Tel.: +1-310-825-8465, +1-909-607-4158. Fax: +1-310-206-5895. E-mail: mackay@psych.ucla.edu

Accepted for publication: July 16, 2002.

anterograde amnesia (AA), that is, postmorbid deficits in representing, retrieving, and/or consolidating new information. Low correlations between the severity of AA and RA have been reported, and whereas all amnesics exhibit AA, only some amnesics exhibit the poorly understood phenomenon of RA. Moreover, unlike AA, RA is often ‘‘patchy’’ (with ‘‘islands’’ of intact memory for salient episodes), temporally-graded (with greater sparing of remote than recent episodes), and variable across different amnesics, types of memory, and times-since-trauma (Hunkin et al., 1995; Kapur, 1993; Levin et al., 1985; Morris, 1999; Murre, 1996). Such findings are problematic for theories that directly attribute both AA and RA to hippocampal damage and ascribe two roles to the hippocampus, namely to encode concepts and events as they unfold in the fleeting present, and (simultaneously) to consolidate thousands and perhaps millions of episodic memories from the past.

Unlike other RA theories, NSAT is an interactive activation model that assigns a single role to hippocampal systems: binding or the formation of new connections when encoding novel aspects of sentences and visual figures, and novel experiences or episodic memories (see MacKay & Burke, 1990). We illustrate this binding function here for the learning and retrieval of words. The nodes or representational units for words are organized into three systems in NSAT (following MacKay, 1987): a *semantic system* (representing, e.g., word-meanings), a *phonological system* (representing hierarchically organized syllables, phonological clusters, and speech sounds), and an *orthographic system* (representing hierarchically organized letters and letter clusters). Word retrieval within these systems presupposes three fundamentally different processes (see, e.g., MacKay, 1990): binding (to initially form the nodes and connections that represent the word), priming (to prepare preformed nodes for activation), and node activation (to actually retrieve the word). To illustrate priming and activation processes in NSAT, Figure 1 depicts preformed nodes representing semantic and phonological aspects of the low frequency (LF) word *consultant*. Connections from lexical to proposition nodes representing, for example, *gives expert advice*, embody the meaning of *consultant*, and connections to phonological nodes representing syllables and other hierarchically organized components embody its phonology. Unique to NSAT, the same node connects bottom-up (enabling meaning retrieval given activation of lexical phonology) and top-down (enabling phonological retrieval given meaning activation). To illustrate the top-down processes for retrieving the phonology of *consultant*, one or more of its meaning nodes, for example, *gives expert advice*, is normally activated first, which primes (top-down) and enables activation of the *consultant* node as the most primed node in the domain of noun nodes (the most-primed-wins activation principle). Activation of *consultant* in turn primes (top-down) and indirectly (via the most-primed-wins principle) enables activation of the syllable nodes representing *con*, *sul*, and *tant*, and so on, until the full phonology for *consultant* has been activated/retrieved in sequence (see MacKay, 1987, pp. 39–89 for further details).

Under NSAT, the basic cause of RA is transmission deficits, which increase with aging, and nonrecent and infrequent use of connections over the life span. Transmission deficits reduce connection strength, which determines the rate and amount of priming that connections throughout the network can transmit. For example, transmission deficits can reduce the top-down priming transmitted from *consultant* to its syllable nodes, thereby preventing activation/retrieval of its phonology, an instance of phonological RA. To illustrate transmission deficits, consider the mild and reversible form of phonological RA known as the tip-of-the-tongue (TOT) phenomenon. People in the TOT state typically know the meaning of a not-recently-used LF word they are trying to retrieve, and although they can sometimes retrieve part of the word, for example, its initial sound, they cannot retrieve its full phonology, sometimes for many days (Burke, MacKay, Worthley, & Wade, 1991). Under NSAT, TOTs occur in older adults when transmission deficits prevent the activation/retrieval of specific phonological units with infrequently and not-recently-used connections. However, TOTs are temporary: Perceiving or producing phonologically related words can strongly prime the critical phonological units

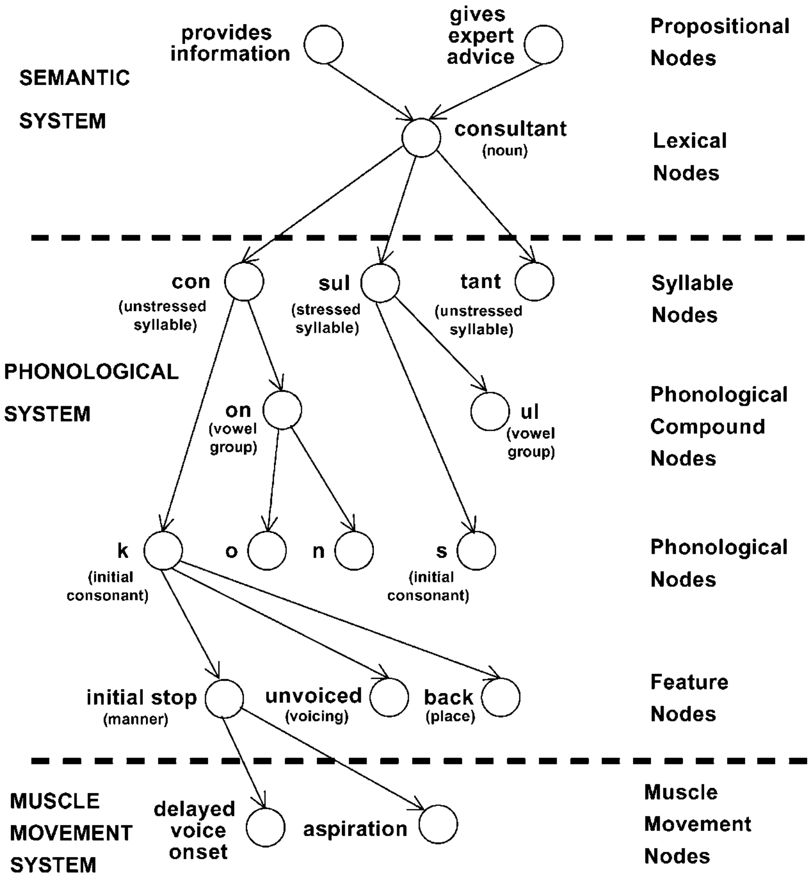

Fig. 1. Selected top-down connections in the semantic, phonological and muscle movement systems for producing the word *consultant* in NSAT.

suffering transmission deficits, enabling the target word to pop into mind, seemingly spontaneously during other activities (James & Burke, 2000).

More severe RA occurs with extreme disuse and old age. The resulting transmission deficits cause not-recently-used connections to become defunct, that is, permanently nonfunctional, and a node with no functional connections is impossible to activate. For example, if the node for the syllable *sul* in *consultant* has become defunct due to old age and nonuse over an extended time period, presenting words such as *sultan*, *insult*, and *desultory* will no longer enable activation of the *sul* node (see White & Abrams, 2002, for relevant evidence). Nonetheless, normal older adults can readily replace this defunct *sul* node with a new one. When next they encounter *sul*, they simply form new connections to re-represent *sul* using intact hippocampal-binding mechanisms. However, defunct information remains defunct and irretrievable in hippocampal amnesics such as H.M. who cannot readily form new connections either to represent novel information or to re-represent old information (see e.g., MacKay & James, 2001).

This may explain why correlations between AA and RA are weak but positive: By preventing the re-representation of defunct information, AA exacerbates but does not directly cause the RA that normally accompanies aging. This may also explain why RA often exhibits temporal gradients. Newly formed connections are fragile and unstable (they rapidly lose connection strength over time) so that memories immediately preceding onset of AA quickly tend to become (and remain) defunct in amnesics. However, remote memories (formed long before the amnesia-causing trauma) are more resistant to RA: Either they are frequently-used, with connections that are strong and relatively stable over time, or they are rarely used, with connections formed anew via binding mechanisms that were intact at that time. Differences across amnesics in age-at-test and time-since-trauma-when-tested may therefore explain the variable nature of RA across studies, and the great variability in how frequently and how recently different premorbid memories have been used may explain the patchy nature of RA.

PARTICIPANTS IN STUDIES 1–2: BACKGROUND INFORMATION

H.M.'s extremely severe AA originated when his hippocampal systems were bilaterally lesioned (with little neo-cortical damage) in 1953 (Corkin, Amaral, González, Johnson, & Hyman, 1997), and for decades researchers believed that AA for post-1953 events was H.M.'s only deficit. However, when tested at ages 44–47, H.M. also exhibited relatively mild RA for names learned 3–13 years before his operation (Marslen-Wilson & Teuber, 1975), and consistent with NSAT, H.M.'s RA had increased in severity when tested at age 57 (Corkin, 1984). At ages 41, 44, and 47, H.M. also exhibited selective deficits in the immediate encoding of unfamiliar (but not familiar) sentence-concepts. To illustrate, a 1973 ambiguity-detection experiment (Lackner, 1974) indicated that H.M. could discriminate between ambiguous versus unambiguous sentences no better than unbiased coin tosses, and a 1967 experiment indicated that H.M. could detect the two meanings of short sentences that he knew were ambiguous less often than memory-normal controls of similar intelligence (IQ), age, work background, and education (MacKay, Stewart, & Burke, 1998). Controls born the same year as H.M. and a patient with bilateral frontal lobe damage also outperformed H.M. in ambiguity detection, demonstrating that H.M.'s sentence-meaning deficits are related to his particular lesion, rather than to cohort effects or general effects of brain damage. These (age 41–47) comprehension deficits provided the initial support for the present theoretical framework, and subsequent tasks involving picture-description, conversational speech, sentence-reading, and description of sentence-meanings clarified that H.M.'s comprehension and production deficits extend to unambiguous sentences, continue to the present day, and apply to reading high frequency words in unfamiliar (but not familiar) phrases in sentences (MacKay, Burke, & Stewart, 1998; MacKay & James, 2001). Studies 1–3 go beyond earlier work by testing H.M.'s ability to read moderately high- and low-frequency words and pseudo-words presented in isolation.

Table 1 summarizes the 1997 background characteristics of H.M. and normal controls in Studies 1–2. The normal controls received $10/h for participating and were 6 native English speakers matched as closely as possible with H.M. for highest educational degree (high school), pre-trauma employment, age, and IQ (especially in the case of control 3; see Table 1). H.M.'s

Table 1. Background Characteristics of H.M. and Memory-Normal Controls: Studies 1–3.

Participants	Age in 1997	Highest educational degree	Mean IQ
H.M.	71.75	High School	112.00
Control 1	73	High School	122.00
Control 2	74	High School	121.00
Control 3	74	High School	111.50
Control 4	70	High School	116.50
Control 5	70	High School	123.50
Control 6	67	High School	122.00
Mean for Controls (*SD*)	71.33 (2.80)	High School	119.42 (4.55)

performance on our IQ test (the Wechsler-Bellevue: Form I) has remained remarkably stable across time. For example, H.M.'s verbal IQ was identical when tested in 1997 and 1977 (see MacKay, Burke, et al., 1998), ruling out general cognitive decline and progressive semantic dementia during the period 1977–1997.

STUDY 1: EFFECTS OF AGING ON H.M.'S ABILITY TO READ ISOLATED WORDS AND PSEUDO-WORDS

Study 1 compared the ability of H.M. and memory-normal controls to read pseudo-words, moderately high frequency (MHF) words (between 39 and 100 instances per million in Francis & Kucera, 1982), and LF words (less than four instances per million) at ages 60 and 71. A note on the relevance of memory retrieval to reading is therefore necessary because many researchers view reading as a fundamentally perceptual task with no relevance to memory (including word retrieval; see Bock, 1996), and they view the production side of reading aloud as requiring only simple links from perceived letters to low-level production mechanisms for pseudo-words and "regularly-spelled" words, that is, ones that follow the dominant or most common orthography-to-pronunciation pattern in English. Within this framework, letter units for regularly spelled words such as *consultant* and regularly spelled pseudo-words such as *quintity* are linked 1-to-1 with phoneme-level production units via "grapheme-phoneme-conversion processes," so that activating letter units in left-to-right order in perception causes phoneme units to become activated in appropriate sequence on the output side of reading aloud (see, e.g., Coltheart, Curtis, Atkins, & Haller, 1993).

However, *a priori* considerations indicate that grapheme-phoneme-conversion processes are insufficient for reading many regularly spelled words and pseudo-words. Consider the process of stressing the appropriate syllable in the regularly spelled word *consultant*. English orthography does not directly specify syllabic stress, and given only grapheme-phoneme-conversion processes, readers could follow several possible stress patterns in reading *consultant*, for example, either con.SUL.tant (the correct pattern), or CON.sul.tant with the stress pattern of *Protestant* (PRO.tes.tant) and *consulship* (CON.sul.ship), and the same is true of regularly spelled pseudo-words such as *quintity.*[1] However, normal readers consistently pronounce *quintity* with stress on the first syllable and current theories cannot explain this consistency without postulating retrieval processes and rules that span the entire pseudo-word. How readers translate letter strings into syllables raises similar problems for grapheme-phoneme-conversion theories because, like stress, English orthography does not directly and unambiguously constrain syllable structure to ensure, for example, that *quintity* is pronounced quin.ti.ty and not quint.i.ty or quin.tit.y.

Another problem for grapheme-phoneme-conversion theories concerns multiple grapheme-phoneme correspondences. For example, consider the pseudo-word *consultment*. Normal readers invariably pronounce the [E] in *consultment* as /e/, but why do they not sometimes follow other common grapheme-phoneme patterns such as [ə]-/i/ (as in *media*) or [E]-/e/ (as in *mesa* and *Megan*)? Lacking memory retrieval processes for morphological units that are pronounced the same throughout the lexicon (here, the suffix *ment*), grapheme-phoneme-conversion theories cannot answer such questions.[2]

Study 2 develops a Node Structure Reading theory (NSRT) that directly addresses such questions under the assumption that during childhood,

---

[1] In these and other examples, capitalization indicates a syllable with primary stress and the periods indicate syllable boundaries. To subsequently distinguish orthographic from phonological units, we will follow the generally accepted convention of capitalizing orthographic units within square brackets, e.g., [-ER], and placing phonological units between slashes, e.g., /-êr/. Our phonological transcriptions follow current IPA rules (available at www.arts.gla.ac.uk/IPA/fullchart.html), with ′ indicating that the subsequent syllable carries primary stress. In addition, [. .] will indicate a brief pause, with [. . .] and [. . . .] indicating progressively longer pauses, and [-] will indicate a glottal stop or perhaps stuttering block.

[2] We note that Rastle and Coltheart (2000) do recognize syllabic stress as a serious problem for word learning in grapheme-phoneme-conversion theories, but the "multiple correspondences" and syllable structure problems are so far neither recognized nor solved.

reading becomes grafted onto the phonological and semantic node structures for retrieving and producing speech (illustrated in Fig. 1). However, details of NSRT are not relevant to Study 1 except to note that NSRT functions via the same basic processing principles as word retrieval/production and that the present aging and RA principles apply to reading as well as to retrieving spoken words. Since the main measure in Study 1 was reading time (from initial presentation to final correct pronunciation of a word or pseudo-word), these principles predict little or no deficits in H.M.'s correct response times for MHF words at age 60 or 71, insofar as word frequency implies recent and frequent use over the course of a lifetime. The reason is that H.M. formed the connections for MHF words prior to his operation, and since recent and frequent use offsets the transmission deficits that result from normal aging, nodes for H.M.'s MHF words will currently be functional due to everyday use since 1953. However, our aging and RA principles predict large deficits for LF words at age 60 and 71, with exaggerated or greater-than-normal age-linked declines between those ages, that is, faster declines with aging than for same-age memory-normal controls. The reason is that aging, nonrecent use and infrequent use over the lifetimes of H.M. and same-age controls causes LF words to become defunct, but unlike same-age controls, H.M. cannot form the new connections required to re-represent defunct LF words. As a result, H.M.'s ability to read LF words will decline with aging at an exaggerated rate relative to same-age controls. By contrast, a ''pure memory deficit'' predicts no difference in correct response times for H.M. versus controls at any age since H.M. learned to read prior to his operation, and because pronouncing isolated, visual words aloud does not require memory storage as traditionally defined. Study 2 will discuss H.M.'s predicted performance for two types of pseudo-words.

METHOD

Participants

The 1986 participants were H.M. and 1 male control matched with H.M. for age and educational level in Gabrieli, Cohen, and Corkin (1988, Experiment 2b). The 1997 participants were H.M. and the 6 controls in Table 1.

Materials

The 1986 stimuli were 54 MHF words and 54 LF words that were virtually identical to the 1997 words in length, frequency and other characteristics, plus 108 pronounceable pseudo-words formed by substituting two letters in each of the 108 words. The 1997 stimuli were 23 pseudo-words and 52 words which were either MHF ($N = 26$; mean frequency 84 per million in Francis & Kucera, 1982), or LF ($N = 26$; mean frequency 0.92 per million). To ensure that H.M. knew our LF words before his age 26 lesion, we chose LF words from Webster's (1949) New Collegiate Dictionary that had entered English at least 3 years before H.M.'s (1953) lesion, and had a mean age of acquisition of 13 or less in three sources: Stratton, Jacobus, and Brinley (1975), http://www.psy.uwa.edu.au/MCRDataBase/mrc2.htm, and http://allserv.rug.ac.be/∼hnaessen/vakgroep/AoA.html. Since the mean age of acquisition for these words was 9.7 (range = 5–13), H.M. almost certainly knew them at age 25 (1953). The pseudo-words were easily pronounced nonwords created by substituting a single letter in a MHF base-word, for example, *quantity* became *quintity*, and *trash* became *drash*. Mean frequency of the base words was 69 per million in Francis and Kucera.

Procedures

The 1986 stimuli were presented via computer one at a time overlaid with a ''blanket'' of random visual noise. The noise initially obscured 70% of a stimulus and diminished over time in small increments to 0% obstruction after 15 s. The participant's task was to correctly pronounce each stimulus as quickly as possible after initial presentation. The main dependent variable in Gabrieli et al. was the time required to pronounce each stimulus correctly, as signaled via experimenter key press. No errors were reported at any noise level. The 1997 stimuli are shown in Appendix A together with current International Phonetic Alphabet (IPA) transcriptions of how pilot participants most commonly pronounced them. The stimuli were presented in blocked order (MHF words, LF words, then pseudo-words) to enable more detailed condition-specific measures of reading time in Study 2. The 1997 stimuli were typed in lower-case, 12 point Courier font and in a single, double-spaced column on a separate 8.5 × 11″ page for each condition. Participants were instructed to read each list orally from top to bottom as quickly as possible without making errors, guessing at the pronunciation of ''unfamiliar items.'' Instructions did not mention words, nonwords or pseudo-words and were summarized on a prominently displayed card. To facilitate timing, the experimenter said "OK" as soon as the participant could see a list, and we measured the time for each correct response from a digitized version of the tape-recorded responses (44,100 samples per second using SoundEdit 16 v. 2).

RESULTS AND DISCUSSION

Table 2 shows mean correct response times and errors in 1986 and 1997 for H.M. and controls (with standard deviations; *SD*). Although conceptually analogous, absolute response times in 1986 and 1997 differed due to just summarized procedural differences between the two reading tasks. To enable direct comparison, we therefore computed percent deficit scores for the identical conditions in the two experiments, calculated as correct response time for H.M. divided by correct response time for controls × 100 minus 100%. Table 2 shows H.M.'s deficit scores at age 60 (1986) and age 71 (1997), together with deficit scores for errors (defined as the % incorrect responses for H.M. minus controls). H.M. had no response time or error deficits for MHF words at age 60 (see Table 2), but a large response time deficit for pseudo-words, and a smaller, but nonetheless sizeable (51%) response time deficit for LF words (unnoted in Gabrieli et al., 1988). H.M.'s substantial response time deficit for LF words at the relatively young age of 60 (1986) is consistent with NSAT which predicts continuous, progressive and irreversible effects of aging for LF words since H.M.'s 1953 lesion at age 26 (see Burke et al., 1991).

By age 71, H.M. also had a large response time deficit for MHF words (see Table 2), larger response time and error deficits for pseudo-words, and extremely large deficits for LF words (see Study 2 for statistics). H.M.'s age 71 deficits for LF words indicate an accelerated or greater-than-normal aging effect (relative to controls) when compared with his age 60 deficits, as in Figure 2, which plots age-linked deficit increase (i.e., the difference in deficit scores for H.M. and controls at age 60 vs. 71) for response times (left ordinate) and errors (right ordinate). Because age-linked deficit increases for normal controls are by definition 0% across all conditions, H.M.'s age-linked deficit increases between age 60 and 71 were greater-than-normal in all three conditions, but especially greater for LF words (response times = 160%; errors = 57%;

Table 2. Mean Correct Response Times in Seconds for H.M. and Controls by Condition in 1997 (Study 2, With *SD* in Parentheses) and 1986 (From Gabrieli et al., 1988).

Measure	MHF words	LF words	Pseudo-words
Mean correct response times in 1997 (s)	(*N* = 18)	(*N* = 6)	(*N* = 8)
Participants			
H.M. (*SD*)	1.383 (.592)	2.831 (1.033)	3.025 (1.255)
Controls (*SD*)	.650 (.086)	.909 (.154)	1.118 (.210)
H.M.'s 1997 percent deficit scores			
For response times (%)	113	211	171
For errors (%)	18	57	44
Mean correct response times in 1986 (s)	(*N* = 54)	(*N* = 54)	(*N* = 108)
Participants			
H.M.	5.1	9.2	13.6
Control	5.4	6.1	7.8
H.M.'s 1986 percent deficit scores			
For response times (%)	−6	51	74
For errors (%)	0	0	0

see Fig. 2). These extreme age-linked deficit increases for LF words comport with predictions of NSAT, as does the relatively small (18%) age-linked deficit increase for errors on MHF words.[3] The sizeable age-linked deficit increases for pseudo-words (response times = 97%; errors = 44%) and for response times involving MHF words (119%) will be discussed in Study 2.

How do we know that H.M.'s heightened 1997 deficits for LF words were due to aging rather than differing stimuli, stimulus presentation details, or criteria for acceptable pronunciation in the two studies? The near identity of LF stimuli in 1986 and 1997 (see also footnote 3) argues against this procedural-difference hypothesis, as does the selective nature of H.M.'s age-linked deficits: 1997 procedures would have to selectively affect LF more so than MHF stimuli, an unlikely scenario. Other evidence against the procedural-difference hypothesis is that H.M. also exhibited exaggerated aging effects for virtually identical LF words in another longitudinal study involving closely comparable stimuli, procedures, and scoring criteria: James and MacKay (2001) compared lexical decision performance of H.M. and controls at age 57 (1983) versus 71 (1997) for MHF and LF words of comparable length and difficulty to the 1997 stimuli, and from 1983 to 1997 H.M.'s lexical decision accuracy declined 36% (6.0 *SD*) for LF words, but only 4% for MHF words, consistent with present results. By contrast, lexical decision accuracy and *SD*s for LF words were identical for control participants at ages 57 and 71, indicating no effect of procedural differences on controls and comporting with results in many other studies of normal aging and lexical decision (for reviews, see Burke & MacKay, 1997; Laver & Burke, 1993).

THEORETICAL ISSUES FOR STUDIES 2 AND 3

This section develops Node Structure Reading Theory (NSRT), a theory for guiding more detailed analyses of H.M.'s age 71–72 reading performance in Studies 2–3. These analyses focused on specific types of errors and two additional response time measures: planning times (the duration of silence immediately preceding correctly- and fluently-read stimuli) and production times (the acoustic duration of correct and fluent responses). Because age-linked deficits are more severe on the output-side than the input-side of perception-production tasks such as reading aloud (see MacKay, Abrams, & Pedroza, 1999), we apply this principle in developing NSRT: We describe the output-side before addressing the input-side of reading aloud, first for familiar words and then for pseudo-words.

NSRT and the Ability to Read Familiar Words

Under NSRT, normal adults *produce* a familiar word such as *consultant* in the same way when reading aloud as when producing *consultant* in everyday speech: via top-down hierarchies such as the one in Figure 1. What makes reading complex and unlike speech in NSRT is that many bottom-up links contribute priming in parallel to determine what lexical node becomes most primed and therefore activated for production under the most-primed-wins activation principle: There are many redundant "routes" from

---

[3] It is unlikely that the greater relative deficit for PS pseudo-words than LF words in 1986 but vice versa in 1997 reflect the fact that PS pseudo-words altered only one letter in 1997 words, versus two letters in 1986 words: Under NSRT, a one letter change is more likely than a two letter change to interfere with H.M.'s reading (see Study 2). However, despite several requests, we were unable to obtain the data, stimuli, stimulus presentation details, or criteria for acceptable pronunciation in Gabrieli et al. (1988) so as to examine this and other issues, for example, why H.M. made no errors in reading pseudo-words at 0% obstruction in 1986, a surprising result under NSRT (see also Study 2). We were also unable to obtain crucial data on how H.M. read comparable MHF and LF words in 1993 (Postle & Corkin, 1998, Experiment 2), again despite many requests over several years. Finally, three 1997 pseudo-words seemed to differ more fundamentally from the 1986 pseudo-words but removing these stimuli from the 1997 data caused an increase rather than a decrease in H.M.'s 1997 pseudo-word deficit (212%), again ruling out a stimulus-difficulty account of the relative magnitude of deficits for PS pseudo-words versus LF words in 1986 versus 1997.

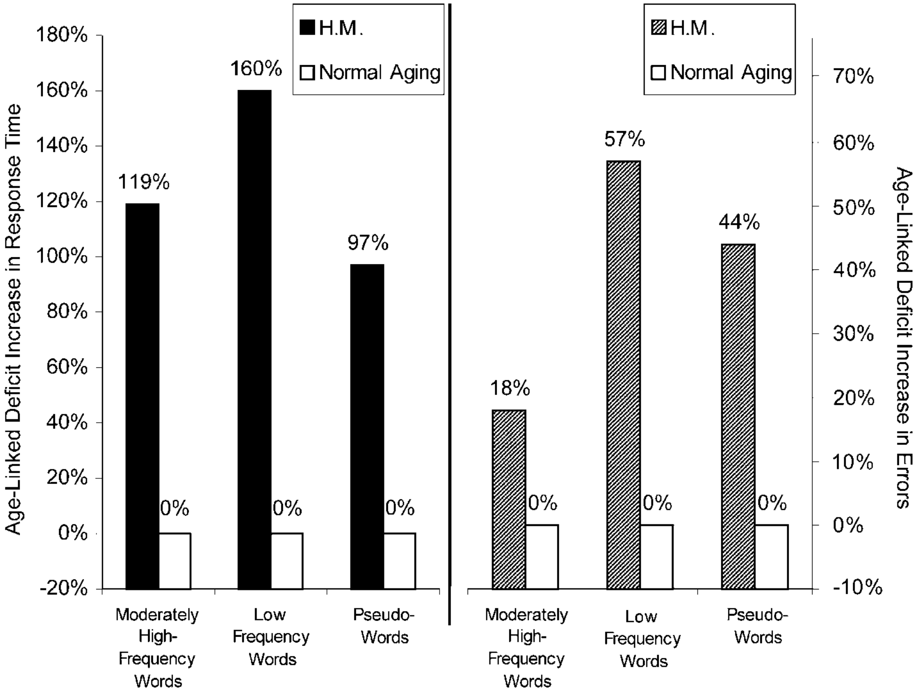

Fig. 2. H.M.'s age-linked deficit increase (his deficit scores for 1997 minus 1986) for response times (left ordinate) and errors (right ordinate) in producing MHF words, LF words, and phoneme substitution pseudo-words. Normal aging predicts 0% in all conditions.

orthography to semantics and phonology for reading familiar words in NSRT.

To illustrate the role of these multiply-redundant links for reading, consider the preformed bottom-up connections that represent the suffix *ment* in words such as *agreement* and *government* in orthographic, phonological and semantic systems in NSRT. Figure 3 shows two of these redundant routes: One route links orthographic units directly to semantic-system units: Letter nodes representing common suffixes such as [MENT] link directly to a suffix node in the orthographic system (see Fig. 3) that links directly to a semantic-system node, here *ment* (noun suffix). The second, more complex "route" from letters to meaning involves "lateral links" (illustrated with dotted lines in Fig. 3) that connect orthographic system nodes representing, for example, letters, to phonological system nodes representing, for example, segments. Segments in turn link bottom-up to phonological cluster and syllable nodes in the phonological system, which link to nodes in the semantic system such as *ment* (noun suffix).

Two fundamental factors distinguish NSRT from dual route theories such as Coltheart et al. (1993). First, the direct and indirect routes to semantic system nodes in NSRT are "both-and" rather than "either-or": Higher-level orthographic and phonological units contribute priming simultaneously and in parallel to determine the most-primed node in the semantic system. Second, as discussed next, lateral links differ in important respects from the grapheme-phoneme-conversion processes in dual route theories.

Differences Between Lateral Links and Grapheme-Phoneme-Conversion Rules

Lateral links in NSRT do not just represent the dominant or most common orthography-to-pronunciation pattern for "regularly-spelled letters," for example, the [E] in *ment*: They simultaneously represent less common or "irregular" orthography-to-pronunciation patterns, e.g., the [E]-/i/

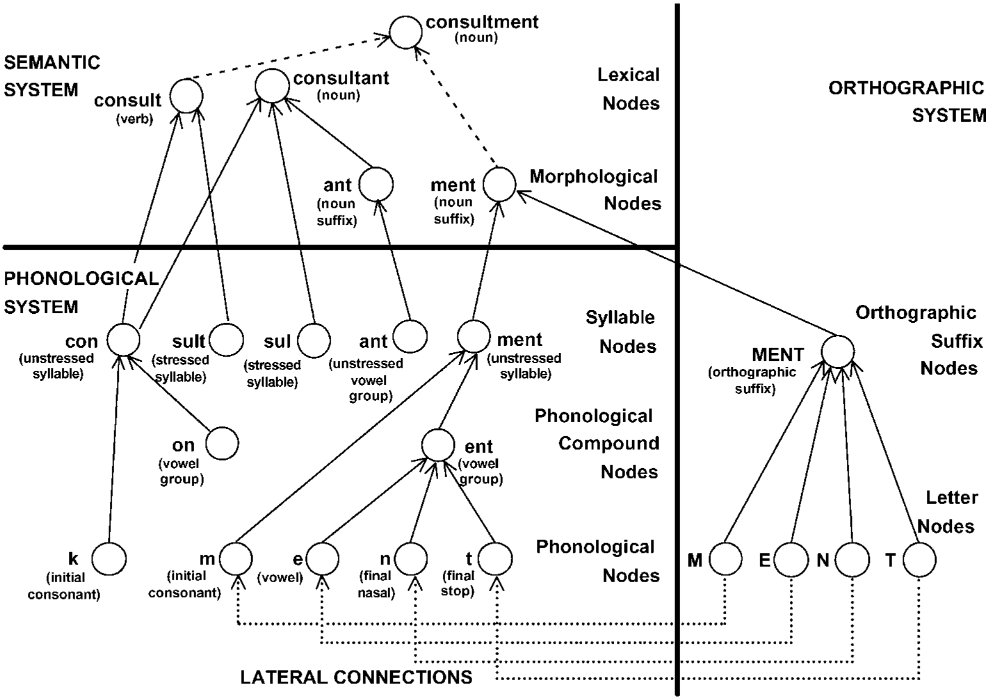

Fig. 3. Bottom-up connections in the semantic, phonological and orthographic systems for representing the noun *consultant* and the pseudo-word *consultment* in NSRT. Solid links and dotted links indicate preformed connections. Broken links indicate new connections. Binding nodes for forming new connections are omitted.

pattern in *media*, *menial*, and *mediocre* (not shown in Fig. 3). The only difference between regular versus irregular orthography-to-pronunciation patterns in NSRT is that lateral links representing more common patterns receive greater frequency of use.

Also unlike grapheme-phoneme-conversion rules, lateral links do not just connect the lowest level letter units to phonological segment units. Orthographic nodes representing common letter clusters also connect laterally to phonological nodes representing consonant clusters, vowel groups or rhymes (e.g., -*ent* in Fig. 3) and entire syllables (e.g., *ment* in Fig. 3). Such "multi-layer redundancy" is possible because lateral links do not directly determine pronunciation of familiar "regularly-spelled" words and morphological units in NSRT. Rather, lateral links influence pronunciation only indirectly by helping to determine which lexical node receives greatest bottom-up priming and becomes activated for production. For example, when reading the word *government*, summated priming from nodes in the phonological and orthographic systems (see Fig. 3) enables *ment* (noun suffix) to become activated as the most primed node in the (noun suffix) domain. For the [E] in *ment*, lateral links representing "irregular" patterns such as [E]-/i/ cannot predominate in the most-primed-wins process for three reasons: their priming transmission is relatively weak since irregular patterns are by definition infrequent across the lexicon; there exists no preformed competitor representing, for example, -*meent* (noun suffix); and *ment* (noun suffix) receives additional priming via the direct connection from the orthographic suffix node for [MENT] (see Fig. 3). However, lateral links representing the irregular [E]-/i/ pattern can and usually do predominate in reading the word *media* because there exists a preformed lexical node representing /midiʌ/ but not /mɛdiʌ/. As we will see, the simultaneous functioning of multi-layered lateral links for regular and irregular spelling patterns and the indirect relation between lateral links and pronunciation in NSRT will clarify some otherwise bizarre aspects of H.M.'s reading errors.

NSRT and the Ability to Read Pseudo-Words

Under NSRT, normal speakers must form new bottom-up and top-down connections to represent and produce multi-syllabic pseudo-words *as coherent units*. We therefore discuss the mechanisms for forming new connections in NSRT as a basis for generating predictions regarding H.M.'s ability to read never previously encountered pseudo-words.

Forming New Connections and Committing New Nodes

Any discussion of connection formation processes must begin by examining what preestablished nodes require connection. To illustrate, consider the preestablished nodes in Figure 3 that become conjoined for reading the never previously encountered pseudo-word *consultment*: *consult* (verb) and *ment* (noun suffix). To conjoin these nodes, *consult* (verb) and *ment* (noun suffix) must first become activated, a process discussed earlier. These simultaneously activated nodes are then conjoined to an ''uncommitted'' or ''chunk node'' in the appropriate domain, first bottom-up, and then top-down (see MacKay, 1990). For example, to represent *consultment* as a unit, the bottom-up connections must link the simultaneously activated nodes for *consult* and *ment* to a chunk node in the noun domain, because the suffix *ment* is a noun marker. This ''domain-selection step'' is essential for representing the possible meanings of *consultment* in subsequently encountered propositions, and for forming top-down connections to syllables representing the stress pattern appropriate to *verb* + *noun suffix* combinations. (This stress pattern is seen in, e.g., *achievement*, *nourishment*, and *measurement*, and differs from the nonderivational stress patterns for other three-syllable nouns ending in *ment*, e.g., *element*, *ligament*, and *monument*.) Preformed top-down connections from syllables with this derivational stress pattern then enable retrieval of the appropriate phonological segments during recall and production of *consultment*, including in principle, everyday conversation (see MacKay & Burke, 1990).

Special mechanisms known as binding nodes are normally engaged to facilitate the formation of new bottom-up and top-down connections (see MacKay, 1990). Binding nodes are located in binding systems, here, the phonological, orthographic, and semantic binding systems, and different binding nodes within these systems specialize in conjoining nodes in different domains.[4] For example, to represent the pseudo-noun *consultment*, the semantic binding node for conjoining a verb and a noun suffix becomes engaged. Similarly, when children initially learn and produce the verb *consult*, specific types of phonological binding nodes are engaged for conjoining an initial unstressed syllable (*con*) and a final stressed syllable (*sult*) and for conjoining the onset or initial consonant group (*k*-) and the rhyme or vowel group (-*on*) in the unstressed syllable *con* (see Fig. 3). Selective damage to phonological binding nodes and no other binding nodes is therefore possible in principle and will only impair the ability to learn newly encountered syllables, phonological clusters and speech sounds when learning, for example, words in a foreign language. Selective damage to orthographic binding nodes is also possible in principle and will only impair the mapping of new or never

---

[4] See MacKay (1990, 1992) for a discussion of conditions that trigger binding nodes and constrain what domains of committed nodes can be directly conjoined. Note that NSAT and NSRT are committed to binding nodes as a theoretical construct, but are not committed to a specific brain location for this theoretical construct. What precise brain areas house different types of binding nodes is an empirical question that does not affect NSAT principles. It is nonetheless likely that hippocampal systems, and specifically, bilateral hippocampal structures and associated entorhinal, perirhinal, and parahippocampal cortices, are the likely locus for most binding nodes given current information (see, e.g., Milner, 1975; O'Keefe & Nadel, 1978). However, it is possible in principle under NSAT that hippocampal system damage may diminish the efficacy of binding nodes located outside hippocampal systems.

previously encountered letter combinations onto phonological and lexical nodes.

By hypothesis, H.M.'s bilateral lesion has damaged all three binding systems: Some (but perhaps not all) of the binding nodes required to efficiently establish new connections in semantic, orthographic and phonological systems have been impaired. As a result, H.M. can represent new semantic, orthographic, and phonological units only with great difficulty. This important "great difficulty" caveat is discussed next.

Relations Between Binding and Engrainment Learning

Activation of a preformed node is normally a brief and self-terminating process that increases by a trivial amount the strength of connections to other preformed or committed nodes, a process known as engrainment learning. Binding mechanisms essentially enhance engrainment processes by inhibiting the self-terminating process, causing simultaneously activated nodes to prolong their activation for an extended period. This increases to a nontrivial degree the strength of connections to other nodes, especially the weak connections to new or uncommitted nodes. Assume, for example, that the noun binding node prolongs activation of semantic-system nodes representing *consult* and *ment*. This will essentially speed up the normal engrainment learning process, causing a rapid boost in connection strength for the uncommitted node within the noun domain that receives joint connections from both *consult* and *ment*. With this rapid boost in connection strength, this uncommitted noun node can amass sufficient priming to become activated as the most primed node in its domain. This and subsequent activations will further strengthen these connections, enabling this new noun node to provide an enduring representation of the conjunction *consult* + *ment*. Once formed, a new node must become activated again within a critical period (several days) or its extremely fragile connections will decay to strength zero and it will revert to uncommitted status. This has negative consequences (the node no longer represents a specific unit and cannot be activated without a new binding process) and positive consequences (the node can potentially represent another novel conjunction, say, *preventment*, *reversement*, or *discussment* rather than *consultment* because each uncommitted node receives hundreds of redundant and initially nonfunctional connections). However, if a newly committed node is used repeatedly during the critical period since last activation, engrainment learning will eventually increase its connection strengths to a durable or "commitment" level, with positive consequences (automatic activation without binding system input during normal language use), and negative consequences (it is committed to representing *a single concept*).

To summarize, binding system input initially helps to form new connections by enhancing the engrainment learning process, and subsequent use without binding system input causes further engrainment learning that ensures durability of newly formed connections. Engrainment learning alone can even serve to form new connections, but without supplementary input this is a slow and inefficient process: Nevertheless, sufficient use or rehearsal will eventually commit new connections at any level in the network. For example, consider the bottom-up phonological connections from the syllables *con* and *sult* to the lexical node for *consult* in Figure 3. When a normal child first encounters *consult*, input from binding nodes helps to form the new connections between this lexical node and its syllables *con* and *sult* in a single trial (without rehearsal). However, hundreds of repetitions of *con* + *sult*, each during the critical period, would also suffice to commit these new connections *without input from binding nodes*. Each subsequent use will then further strengthen these committed connections as long as the *consult* node remains functional.

NSRT Predictions for H.M.'s Age 71–72 Planning Times, Production Times, and Errors

Basic assumptions for relating NSRT to data in Study 2 are that production times mainly reflect the time required to activate preformed and functional nodes, whereas planning times mainly reflect the time required to organize a response for production, including the time to form new connections to represent that response. Under these assumptions, NSRT predicts that

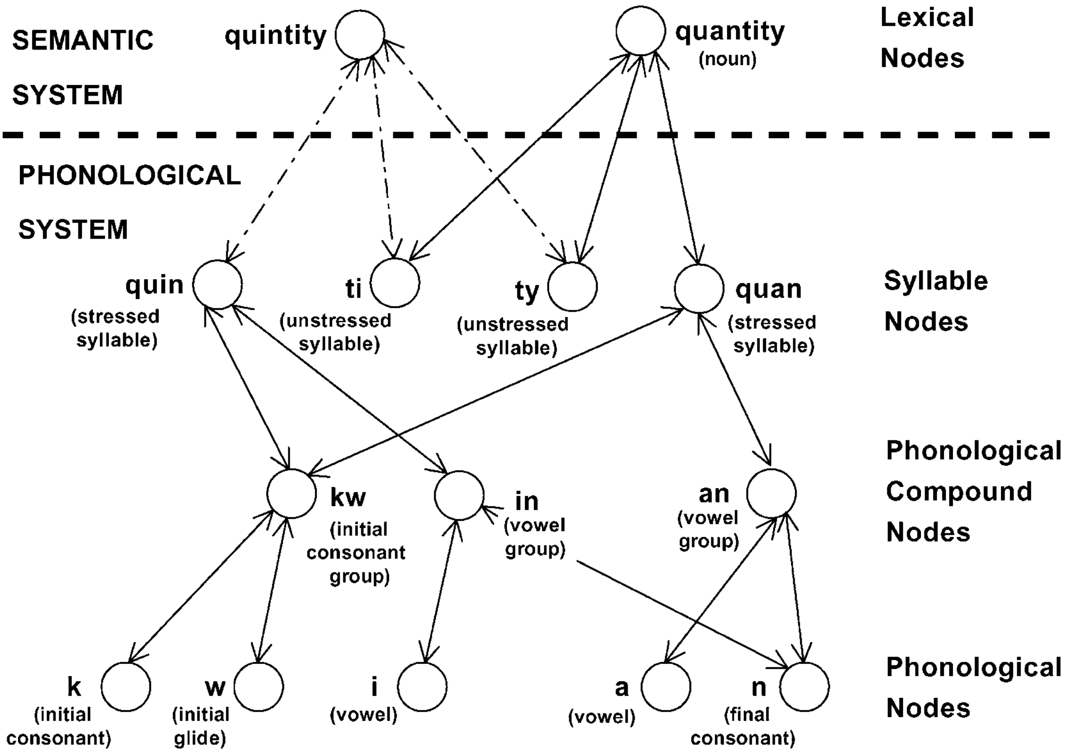

Fig. 4. Bottom-up and top-down connections in NSRT for producing selected aspects of the pseudo-word *quintity*. Solid links indicate preformed connections. Broken links indicate new connections. Binding nodes for facilitating new connection formation are omitted.

H.M.'s errors will parallel his planning times rather than his production times since activating frequently used nodes is not a problem for H.M. under NSRT. Under the assumption that H.M. has used MHF words such as *consult* and *quantity* recently and frequently since 1953, NSRT also predicts no impairment in H.M.'s planning and production times for MHF words as all of the nodes and connections for such words were formed prior to H.M.'s 1953 operation and can be activated without binding system input (see Figs. 3 and 4).

However, NSRT predicts impairment in H.M.'s planning and production times for LF words such as *adumbrate* because transmission deficits due to normal aging, nonrecent use, and infrequent use over his lifetime have rendered H.M.'s previously functional phonological and orthographic nodes for LF words difficult or impossible to activate: The difficult-to-activate nodes predict increased production times to correctly produce LF words (see Study 1). The defunct or impossible-to-activate nodes predict planning time deficits for correctly produced LF words. The reason is that H.M. can only form new connections for representing defunct information via the inefficient connection formation process of engrainment learning, that is, multiple repetitions.[5] Covert or internal repetitions will increase the time it takes H.M. to begin to pronounce LF words, but he may occasionally repeat the onsets of these words

---

[5] Freed, Corkin, and Cohen (1987) present data that are relevant to the hypothesis that H.M. can form weak or fragile new connections, but requires extensive repetition for these connections to achieve normal levels of strength. For H.M., Freed et al. had to multiply the time available for rehearsal by a factor of 20 in order to equate initial performance of H.M. and memory-normal controls in a yes-no recognition memory task.

aloud, an error type involving repetition that reflects deliberate attempts to form and strengthen new connections via rehearsal, that is, engrainment learning.

NSRT predicts similar repetition tendencies and deficits in H.M.'s planning times for pseudo-words because representing pseudo-words, like re-representing defunct LF words, requires new connection formation. However, NSRT predicts greater impact of aging for H.M.'s reading of LF words than pseudo-words, with precipitously deteriorating performance in the period 1986–1997 for LF words but not for pseudo-words *relative to same-age controls*. The reason for H.M.'s precipitous age-linked declines for LF words is that forming new connections only becomes necessary for re-representing defunct LF words beginning around age 65 since age-linked transmission deficits for preformed connections only become severe around that age (see, e.g., MacKay & Abrams, 1998). Since same-age controls have normal binding mechanisms for re-representing and successfully producing defunct LF words whereas H.M. does not, this means that H.M.'s performance for LF words will decline precipitously after age 65 *relative to same-age controls*. By contrast, forming new connections is necessary for representing pseudo-words *at any age*, which means that H.M. has had an (undetected) deficit for pseudo-words that has remained constant ever since his 1953 lesion, with little age-linked change relative to same-age controls with intact binding mechanisms.

NSRT predicts lower probabilities of error correction for H.M. than controls since H.M.'s errors derive from unformed or degraded internal representations that provide no basis for correction (see also MacKay & James, 2001). However, NSRT predicts few errors for MHF words and qualitatively different errors for LF words and morpheme replacement (MR) pseudo-words (which combine real English morphemes in creative and/or anomalous ways, e.g., *consultment*) versus phonological substitution (PS) pseudo-words (which substitute a single letter in a real word, e.g., *quintity* for *quantity*). Specifically, wrong-word errors (i.e., words phonologically similar to the target) will be more common for PS pseudo-words than MR pseudo-words and LF words: When H.M.'s binding deficit prevents efficient formation of new connections for accurately representing and producing a PS pseudo-word, its nearest real-word approximation will receive most priming via already formed bottom-up connections (such as the ones in Fig. 3) and become activated, causing a wrong-word error. For example, presenting *quintity* will strongly prime *quantity* via bottom-up connections established during H.M.'s childhood, causing a wrong-word error. To read *quintity* without error, a new chunk node must become activated via three new bottom-up and top-down connections conjoining the syllable nodes *quin* (as in *quintet*), *ti* (as in *citizen*) and *ty* (as in *pretty*) (see Fig. 4). Because H.M.'s connection formation deficit prevents him-from quickly representing *quintity* as a coherent unit, strong priming from bottom-up connections formed prior to H.M.'s operation will therefore tend to trigger the wrong-word error *quantity*.[6] By contrast, few wrong-word errors are predicted for MR pseudo-words such as *consultment* and LF words such as *akimbo*. For example, if *akimbo* becomes defunct, there is no phonologically similar MHF word that could replace it in error due to bottom-up priming.

To summarize, NSRT predicts that H.M.'s errors will parallel his planning times rather than his production times, with no errors and no planning or production time deficits for MHF words, and large planning time deficits and

---

[6] Close inspection of Figures 3 and 4 allows several valid inferences. One is that cortical nodes for comprehending and producing units are identical at semantic and phonological levels in NSRT, unlike other distributed-memory theories (see MacKay, 1987, pp. 14–194). A second inference from Figures 3 and 4 is that top-down and bottom-up connections are not entirely symmetrical, despite the shared semantic and phonological nodes for comprehension and production. A third inference is that different affixes call for different patterns of top-down connections to achieve the appropriate syllabic segmentation within the phonological system. For example, top-down connections for producing *consultment* connect with the syllables *con*, *sult*, and *ment*, whereas top-down connections for producing *consultant* connect with the syllables for *con*, *sul*, and *tant* (see Figs. 1 and 3).

frequent errors involving repetition for LF words and pseudo-words. NSRT also predicts that relative to controls, H.M. will produce fewer error corrections than controls, and fewer wrong-word errors for LF words and MR pseudo-words than for PS pseudo-words, for example, *quintity*. NSRT nevertheless predicts that at age 71, H.M. will exhibit more errors and longer planning and production times for LF words than for PS and MR pseudo-words.

STUDY 2: DETAILED ANALYSES OF H.M.'S 1997 RESPONSE TIMES AND ERRORS

Study 2 had four goals: to rule out alternate accounts of H.M.'s exaggerated age-linked declines for LF words in Study 1; to analyze H.M.'s age 71 data statistically; to compare H.M.'s reading performance for PS pseudo-words, for example, *quintity*, with MR pseudo-words that miscombine real English morphemes, for example, *consultment*; and to provide preliminary tests of NSRT predictions. However, because Study 2 yielded some unexpected results that are replicated in Study 3, we defer overall evaluation of NSRT predictions to the General Discussion section.

METHOD

Participants

The participants were H.M. and the 6 controls in Table 1.

Materials

Materials were H.M.'s 1997 stimuli, plus 16 MR pseudo-words that followed common English grapheme-to-phoneme pronunciation rules (see Table 3) and recombined a high frequency base word (mean frequency 119 per million in Francis & Kucera, 1982) with a common suffix in either legal or illegal ways. Morphemes combine in the same way as real words for legal MR pseudo-words ($N = 9$), for example, *undoctoral* (which combines a negative prefix with a base adjective in the same way as, e.g., *unable* and *untrue*), but not for illegal MR pseudo-words ($N = 7$; *duskly*, *slumply*, *friendlyhood*, *eventment*, *peoplement*, *untale*, and *retrend*). For example, *dusk* (noun)+*ly* (adverb suffix) represents an illegal morpheme combination because adverb suffixes normally combine with adjectives and not nouns.

Procedures

Procedures were as described in Study 1 with MR pseudo-words presented last.

RESULTS AND DISCUSSION

We separately report results for response times versus errors since these variables were independent: H.M. both produced more errors and had longer correct response times than controls. Error types appear with examples in Table 4 for H.M. and controls.

Response Time Analyses

Table 5 shows overall reading times by condition, which were over 8.81 *SD* longer for H.M. than controls in every condition. However, these overall reading times included pauses between responses (planning time), response duration (production time), and a wide variety of errors and self-corrections (see Table 4), fillers (e.g., um, er), false starts, and interjections (some quite lengthy, e.g., "Well I say it that way, but it's...."). To overcome these confounding

Table 3. Mean Frequency and Length in Syllables and Letters for Stimuli in Studies 2–3.

	MHF words ($N = 26$)	LF words ($N = 26$)	PS pseudo-words ($N = 26$)	MR pseudo-words ($N = 16$)
Mean frequency	83.96	0.92	N/A	N/A
Mean length in syllables	2.04	2.31	2.08	2.75
Mean length in letters	6.50	6.46	6.23	9.00

Table 4. Example Error Characteristics by Type in Study 2, Together With Separate Relative Frequencies for H.M. and Controls (Means With *SD* in parentheses; see Text for Explanation).

Error characteristics	H.M	Typical H.M. example	Controls (*SD*)	Typical control example
Wrong-word errors	22%	"sanctify" for *satisfy*	2% (1%)	"lapse" for *lapte*
Successful self-corrections	19%	"guest.. no that's uh.. crypt"	28% (14%)	"muddle, no… mundle"
Successive approximations	17%	"abicurgle .. duh .. abidackle.. abedickle" for *abdicate*	4% (2%)	"papaya…pap-papyruse" for *papyrus*
Segmentation errors	22%	"consultant" for *consultment*	0% (0%)	N/A
Unshared segments	38	"international" for *internal*	1.7 (1.9)	"muddle" for *mundle*
Repetitions of earlier units	10%	"ambryite" after "papryism" after "embryism"	0% (0%)	N/A
Order errors	7%	"barshite" for *boshertin*	0% (0%)	N/A
Suffix errors	19%	"metalness" for *metalousness*	0% (0%)	"joined" for *join*
Stress shifts	20%	"labrinth" (laBRINTH) for *labyrinth*(LAbyrinth)	0% (0%)	N/A

Note. The error characteristics are not mutually exclusive and do not sum to 100%.

factors, our subsequent response time measures only included correctly- and fluently-read stimuli.

Production Times

Table 5 shows the mean per item production time by condition for H.M. and controls (with *SD*), together with the number of stimuli that H.M. and all controls read correctly and fluently in each condition. These mean per item production times were problematic from a theoretical point of view since word length (in letters) covaries with production time and varied across our conditions (see Table 3). To factor out stimulus length we therefore divided per item production time by number of letters for each item, and computed H.M.'s mean deficits relative to controls in *SD*, with results in Figure 5. H.M.'s mean production time deficits were large for LF words (8.24 *SD*), intermediate for MR pseudo-words (4.84 *SD*), and relatively small for MHF words (2.21 *SD*) and PS pseudo-words (1.77 *SD*; see Fig. 5).

Planning Times

Table 5 shows mean per item planning times by condition for stimuli read correctly and fluently by H.M. and all controls (with *SD*s). Per item planning times were over 8.40 *SD* longer for H.M. than controls in every condition but required adjustment since word length differed across conditions (see Table 3) and covaries with planning time (see, e.g., Santiago, MacKay, Palma, & Rho, 1999). To factor out stimulus length we divided per item planning time by number of letters per item, and computed H.M.'s mean deficits relative to controls in *SD*, with results in Figure 6. H.M.'s mean planning time deficits were larger for LF words (about 10 *SD*) and MR pseudo-words (19 *SD*) than MHF words (7.77 *SD*) and PS pseudo-words (8.00 *SD*).

Table 5. Overall Reading Times (s), Mean Correct Production, Planning and Production Times (ms), and Percent Deficit for H.M., by Condition for H.M. and Controls in Study 2 (With *SD* in Parentheses).

Measure	MHF words	LF words	PS pseudo-words	MR pseudo-words
Overall reading time (s)	(*N* = 26)	(*N* = 26)	(*N* = 26)	(*N* = 16)
Participants				
H.M.	40.61	94.41	115.09	79.40
Controls	16.75	32.97	35.76	22.19
(*SD*)	(1.73)	(6.97)	(5.74)	(2.73)
H.M.'s percent deficit in overall reading time (%)	142	186	222	258
Mean correct planning time (ms)	(*N* = 18)	(*N* = 6)	(*N* = 8)	(*N* = 8)
Participants				
H.M. (*SD*)	686 (580)	1550 (1091)	2225 (1185)	2210 (953)
Controls (*SD*)	156 (63)	293 (95)	480 (167)	329 (72)
H.M.'s percent deficit in correct planning time (%)	340	429	364	572
Mean correct production time (ms)	(*N* = 18)	(*N* = 6)	(*N* = 8)	(*N* = 8)
Participants				
H.M. (*SD*)	699 (110)	1261 (109)	800 (206)	1335 (391)
Controls (*SD*)	494 (65)	589 (79)	638 (121)	756 (99)
H.M.'s Percent Deficit in Correct Production Time (%)	41	114	25	77

Percent Deficit Scores for Planning Versus Production Time

To compare H.M.'s deficits for planning versus production times, we calculated H.M.'s percent deficit scores by condition as in Study 1, with results in Table 5. Overall, H.M. had much larger deficit scores for planning times (425%) than production times (48%), a pattern that held for every condition. This pattern comports with the hypothesis that planning times primarily reflect the process of forming new connections (H.M.'s basic problem under NSRT), whereas production times primarily reflect the activation of already formed connections (not a problem for H.M. under NSRT).[7] Also consistent with this NSRT connection formation hypothesis are H.M.'s larger planning time deficits for LF words and pseudo-words than for MHF words: Unlike MHF words, both LF words and pseudo-words require new connection formation under NSRT. However, H.M.'s planning and production time deficits and his greater deficit scores for planning than production contradict a "pure memory deficit."

Main Error Measures

Overall Errors

Overall errors included true reading errors, self-corrected errors, and successive approximations (for examples, see Table 4), together with general-fluency errors, that is, other deviations from the phonemic transcriptions in Appendix A (which represent the most common ways that pilot participants pronounced the stimuli). Table 6 shows the mean frequency of overall errors by condition for H.M. and controls (with *SD*s). Overall errors occurred on 50% of the stimuli for H.M. versus a mean of 11% for controls (*SD* = 4%), a difference in excess of 9.75 *SD*. This difference contradicts a "pure memory deficit," and cannot be attributed to IQ differences

---

[7] Note that we do not claim that new connection formation is the sole contributor to planning time, or that no connection formation can occur following onset of production.

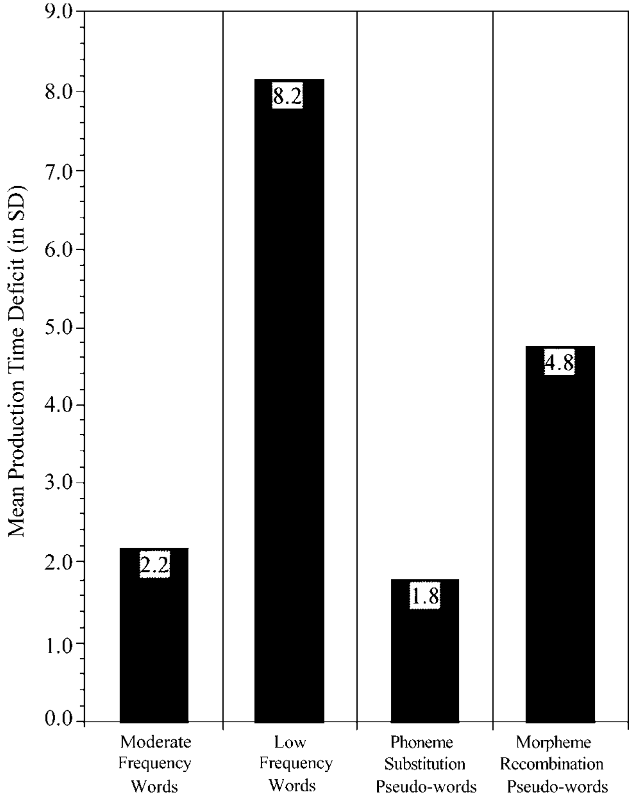

Fig. 5. H.M.'s mean production time deficits relative to controls in *SD* (per condition in ms per letter per item).

between H.M. and controls because Control 3 (whose IQ was slightly lower than H.M.'s) performed similarly to other controls and made at least 4.7 *SD* fewer overall errors than H.M. for all but the MHF condition.

Preplanned sign tests with stimuli as unit of analysis indicated that overall errors were reliably more common for H.M. than controls reading LF words, $p < .001$, PS pseudo-words, $p < .001$, and MR pseudo-words, $p < .05$, but not MHF words ($p > .13$). This pattern comports with deficits predicted under NSRT for LF words and pseudo-words but not MHF words. Relative to controls, H.M. also made more overall errors when reading PS pseudo-words (a 7.14 *SD* increase relative to controls; see Table 6) than MR pseudo-words (a 4.86 *SD* increase relative to controls; see Table 6), a remarkable result in view of length effects discussed later since MR pseudo-words were 38% longer than PS pseudo-words (see Table 3).

True Misreadings

True misreadings, for example, *mundle* misread as "muddle" without correction, excluded general-fluency errors, self-corrected misreadings and repair signals, for example, "muddle, no... mundle," and acceptable pronunciations of pseudo-words that happened to deviate from the "most common pronunciation(s)" prespecified as correct in Appendix A. Acceptable pronunciations had the same grapheme-to-phoneme pattern as one or more English words in the X-words program of the American Heritage dictionary. For example, pronouncing the [OU] in *metalousness* as /aw/ was acceptable because [OU] is pronounced /aw/ in several English words, for example, *lousy* and *blouse*. Similarly, pronouncing the [ER] in *merling* as /eyr/ was acceptable because [ER] is pronounced /eyr/ in several English words, for example, *merit*. Controls produced more acceptable misreadings than

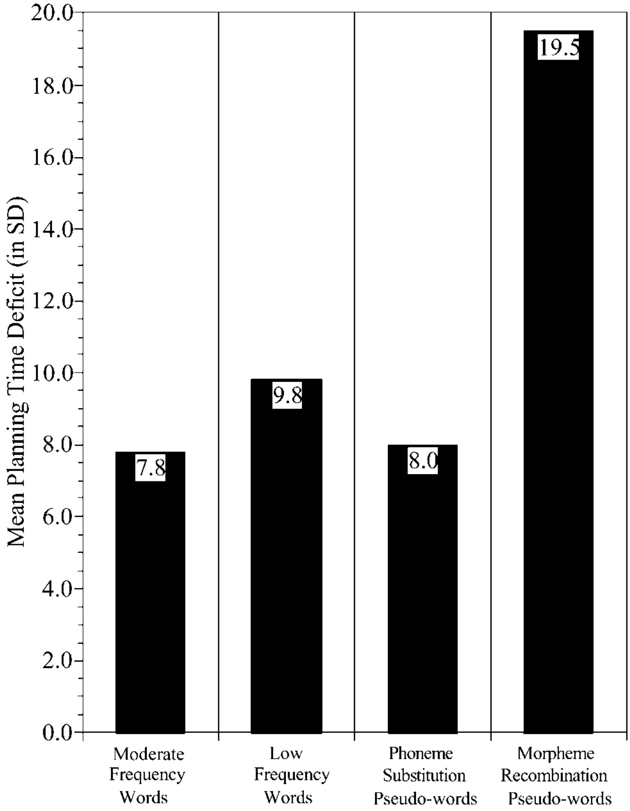

Fig. 6. H.M.'s mean planning time deficits relative to controls in SD (per condition in ms per letter per item).

Table 6. Overall Errors, True Misreadings, and General-Fluency Errors by Condition for H.M. and Controls (Means in % with *SD*) in Study 2 (H.M.[a]) and Study 3 (H.M.[b]).

Error Category	Participants	MHF words (*N* = 26)	LF words (*N* = 26)	PS pseudo-words (*N* = 26)	MR pseudo-words (*N* = 16)
Overall errors(%)					
	H.M.[a]	23	69	62	44
	H.M.[b]	8	58	58	31
	Controls	2	22	12	10
	(*SD*)	(2)	(9)	(7)	(7)
"True" misreadings (%)					
	H.M.[a]	19	62	50	19
	H.M.[b]	0	46	42	25
	Controls	1	5	6	2
	(*SD*)	(2)	(3)	(4)	(3)
"General-fluency" errors (%)					
	H.M.[a]	0	0	0	0
	H.M.[b]	0	0	0	0
	Controls	1	3	0	1
	(*SD*)	(2)	(5)	(0)	(2)

H.M., whose misreadings usually violated all grapheme-to-phoneme relations in English. For example, "ambryite" is an unacceptable misreading of *ampetite* because [PET] is never pronounced /briy/ in English.

As a type of overall error, true misreadings occurred relatively more often for H.M. than controls: 79% of H.M.'s overall errors were true misreadings, but only 27% of controls' overall errors were true misreadings ($SD = 17\%$), a difference in excess of 3 *SD*. In absolute number, true misreadings were also more common for H.M. than controls by 17 *SD*, and exhibited the same pattern across conditions as overall errors and planning times for H.M. and controls (see Table 6). Preplanned sign tests with stimuli as unit of analysis indicated that true misreadings were reliably more common for H.M. than controls for LF words ($p < .001$) and PS pseudo-words ($p < .01$), but not MHF words ($p > .11$). True misreadings were also over 5 *SD* more common for H.M. than controls for MR pseudo-words (see Table 6).

Wrong-Word Errors

Table 7 shows the frequency of wrong-word errors (i.e., other real English words) by condition for controls and H.M., who misread, for example, *building* as "build" (MHF condition), *ampetite* as "appetite" (PS condition), and *consultment* as "consultant" (MR condition). Consistent with NSRT, wrong-word errors were at least 3 *SD* more common for H.M. than for controls in each condition. For example, H.M. misread 25% of the MR pseudo-words as real English words, versus 0% for control participants. Also consistent with NSRT, H.M. produced more wrong-word errors in the PS condition than any other condition, for example, twice as many than for LF words (a difference that exceeds 3 *SD*s relative to controls).

Repetition of Phonological Units

As predicted under NSRT, H.M. tended to repeat phonological units that he had perceived or produced earlier in the experiment. For example, H.M. first misread the LF word *euphemism* as "embryism," and 15 items later on the next page, he misread *papyrus* as "papryism," and then, on the subsequent page, he misread the pseudo-word *ampetite* as "ambryite," again repeating the [Y.I] of "embryism" and "papryism." Similarly, H.M. misread *mundle* as "muddily" and *minate* as "mumby" as if repeating the suffixes [-Y] and [-LY] from, for example, *yelkey*, and he misread *slumply* as "supplement," as if repeating [-MENT] from *accusement* (see Appendix A). H.M. even repeated an entire word from a prior list, misreading the pseudo-word *lapte* as "labyrinth." Repetition errors are shown (as percentage of total items) for H.M. and controls in Table 4, and occurred for 10% of items for H.M. but 0% for controls ($SD = 0\%$).

Similar repetition tendencies are apparent in H.M.'s successive approximations, defined as two

Table 7. Wrong-word Errors and Successful Self-correction of Errors (in %) by Condition for H.M. and Controls (means with *SD*) in Study 2 (H.M.[a]) and Study 3 (H.M.[b]).

	Participants	MHF words ($N = 26$)	LF words ($N = 26$)	PS pseudo-words ($N = 26$)	MR pseudo-words ($N = 16$)
Wrong-word errors (%)					
	H.M.[a]	23	19	31	25
	H.M.[b]	8	12	27	19
	Controls	1	4	6	0
	(*SD*)	(2)	(4)	(6)	(0)
Successful self-corrections (%)					
	H.M.[a]	17	6	19	57
	H.M.[b]	50	13	13	0
	Controls	0	21	33	31
	(*SD*)	(0)	(17)	(29)	(40)

or more attempts to read an item, regardless of whether the item was ever read correctly. For example, H.M. produced three successive approximations in misreading *akimbo* as ''ak. . .akibo'' and finally ''akbo,'' and in misreading *abdicate* as ''abiCURgle'' (stress on CUR) and then ''abidackle'' (stress on initial A), followed by ''abeDIckle'' (stress on DI). As shown in Table 4, successive approximations occurred on 17% of all items for H.M. versus 4% for controls. This 6.5 *SD* difference is consistent with the NSRT hypothesis that H.M. tried deliberately to form the new connections required to represent and produce LF words and pseudo-words via rehearsal or engrainment-learning processes (see MacKay, Burke, et al., 1998, for other unusual repetition strategies that H.M. adopted during sentence production and memory tasks). Four aspects of the data contradicted the alternate hypothesis that H.M.'s successive approximations represent normal error correction processes: Unlike normal error corrections, H.M.'s successive approximations did not always achieve the correct form, and sometimes continued even after he produced the correct form. For example, H.M. read the pseudo-word *lapte* as ''la .. laptey .. laptley,'' and proceeded to the next stimulus, as if his second approximation was correct by accident rather than by design. H.M.'s successive approximations also lacked the normal ''repair signals'' that usually accompany error correction, e.g., ''no,'' ''er,'' or ''rather.'' Finally, unlike successive approximations, error corrections were more common for controls than H.M., as discussed next.

Error-Corrections

Error-corrections excluded general-fluency errors and involved the following typical sequence: an error followed by the correct, spontaneously-produced form and no additional response, e.g., *mundle* misread as ''muddle,'' followed immediately by self-correction, ''no. . . mundle,'' and the next item on the list. As in this example, controls usually marked error-corrections with a repair signal, but H.M.'s sole use of a repair signal was inappropriate: He misread *serrated* as ''san.GRATE'' and then claimed to correct himself with ''SEE.grated.'' Controls produced more error-corrections (28%) than H.M. (19%) (see Table 7 for *SD*s), a finding consistent with MacKay and James (2001) and MacKay, Burke, et al. (1998), where H.M. corrected fewer errors than controls during sentence-production. These results comport with the NSRT hypothesis that H.M.'s errors derive from unformed or degraded internal representations that do not provide the basis for correction.

Control Procedure Results

Controls for Regular Versus Irregular Spelling

With spelling-regularity defined as in Balota and Ferraro (1993), regular spelling was about equally common for our MHF ($N = 16$) and LF ($N = 15$) stimuli, and so was irregular spelling (10 MHF words and 11 LF words), which rules out irregular spelling as an account of our main results for MHF versus LF words. Not only did H.M.'s exaggerated LF deficits hold for true misreadings of regularly spelled words (a 19.3 *SD* deficit for LF words relative to controls vs. a 1.67 *SD* deficit for MHF words), irregular spelling reduced rather than increased H.M.'s LF word deficits. H.M. had larger deficits for regularly than irregularly spelled LF words (19.3 vs. 8 *SD*s). This last finding contradicts SDA (senile dementia of the Alzheimer's type) as an account of present results since SDA patients show the opposite pattern, with larger deficits for reading irregularly than regularly spelled LF words (see Balota, & Ferraro).

Controls for Stimulus-Blocking Interactions

Although blocking by condition was essential for developing condition-specific measures of reading time, control participants may have quickly learned the general nature of stimuli within a list condition, thereby facilitating responses to subsequent stimuli in the list. This strategy predicts reduced errors with practice or item number in a list for controls, but no relation between errors and practice for H.M. whose memory deficits would preclude this learning-strategy. However, errors for controls were evenly distributed across the first and second half of MHF, MR and PS pseudo-word lists. Controls did make somewhat more errors on the first than second half for LF words (3.30 vs. 2.33 errors), but H.M. exhibited the same pattern: H.M. made 11 errors on the first half of the LF word list versus 7 errors on the second half. These results contradict the hypothesis that our by-condition blocking procedure supported special strategies for controls that were unavailable to H.M.

General-Fluency Errors

General-fluency errors included false starts or stutters, for example, ''pap.. papyrus'' and miscellaneous exclamations, filler words, and intrusions, as in, ''uh… a… boy.. adumbrate,'' but excluded meta-comments, such as ''had to correct myself.'' Only controls produced general-fluency errors, never H.M. (see Table 6). This finding indicates that H.M.'s deficits are selective rather than general in nature: H.M. produced certain types of errors more often than controls, but was not more error-prone than controls for all error types.

Subsidiary Results Consistent With NSRT

Segmentation Errors

We defined segmentation errors as production of an appropriate speech sound in the wrong syllabic position. An example is H.M.'s misreading of *papyrus* (paPYrus) as ''PAPryism'' where the second [P] is syllable-initial in the stimulus, but syllable-final in H.M.'s response. Segmentation errors occurred in reading 22% of the items for H.M. versus 0% for controls ($SD = 0\%$; see Table 4).[8] This difference comports with the NSRT hypothesis that some of the semantic- and phonological-system connections required for appropriate syllabic segmentation have become permanently defunct in H.M.

Stress Shifts

We defined stress shifts as production of primary stress on one or more inappropriate segments in a word or pseudo-word. For example, when H.M. misread *akimbo* as ''ak…akibo'' and finally ''akbo,'' he produced major stress on the vowel /a/ each time, which is inappropriate since /a/ is unstressed in *akimbo* (see Table 4). Controls produced no stress shifts in their true misreadings ($SD = 0\%$), but H.M. produced stress shifts in 20% of all items, a conservative estimate because H.M. sometimes produced several stress shifts on a single trial.[9] This difference suggests that H.M. often guessed at stress patterns due to his inability to form new connections to represent pseudo-words and to re-represent degraded representations for LF words.

Suffix Errors

Suffixes posed special difficulty for H.M. in all four conditions, and his suffix errors fell into three categories: substitutions, omissions and additions. Examples are *consultment* misread as ''consultant'' (substitution) and *remarkable* misread as ''remark,'' *serrated* misread as ''sangrate,'' and *metalousness* misread as ''metalness'' (omissions). H.M.'s least common type of suffix error involved addition, for example, *internal* misread as ''international.'' One control participant made one suffix error for an overall mean of 0% ($SD = 0\%$), whereas H.M. made suffix errors on 21% of all items (see Table 4), a conservative estimate because only the first error was counted, and H.M. sometimes produced several suffix

---

[8] Although controls sometimes segmented units differently from Appendix A, these were acceptable errors (as defined earlier, based on grapheme-to-phoneme relations in the X-words program of the American Heritage Dictionary) that did not fit the definition of ''true'' reading errors. Note that since H.M.'s responses often contained several errors, any given response could enter several different error categories in Table 4, a conservative procedure since choosing a single error category for each response raised controversial issues.

[9] Although controls sometimes stressed different units from Appendix A, these were acceptable errors (as defined earlier) that did not fit the definition of ''true'' reading errors. For example, one control read *internal* with contrastive stress as INternal, acceptable in, for example, ''not EXternal, INternal.'' By contrast, H.M. misread *internal* as ''international,'' which counts as a stress shift under our definition since neither [IN] nor [TER] are stressed in ''international.'' However, it might be argued that word errors take precedence over stress shifts: If H.M. thought the word was ''international,'' he produced the correct stress pattern for ''international'' without stress shift. However, only two of H.M.'s stress shifts could be explained in this way.

errors on a single trial. Under NSRT, H.M.'s special difficulty with suffixes reflects three factors: his inability to form new connections, his tendency to substitute simpler and more frequent word forms (wrong-word errors), and age-linked degradation of connections representing real-word morpheme combinations.

Legal Versus Illegal MR Pseudo-Words

H.M. made relatively more overall errors on illegal MR pseudo-words (*slumply*, *friendlyhood*) than legal ones (e.g., *consultment*, *undoctoral*). H.M. misread 57% of the illegal MR pseudo-words versus 5% for controls (*SD* = 7%, a difference in excess of 7.4 *SD*), but H.M. only misread 33% of the legal MR pseudo-words versus 13% for controls (*SD* = 8%; a 2.5 *SD* difference). This relatively greater deficit for illegal than legal MR pseudo-words suggests that H.M. was sensitive to how English morphemes normally combine and is consistent with NSRT. Irregular MR pseudo-words exacerbate H.M.'s connection formation problem by obscuring the "domain-selection step" in NSRT, that is, the process of selecting the domain or syntactic category for the chunk node representing an MR pseudo-word (see Introduction).

Unexpected Subsidiary Effects

This section reports unpredicted effects that are replicated in Study 3 and discussed in the General Discussion section.

Unshared Segments

Unshared segments were phonemes added to or subtracted from the original stimulus (as per the IPA transcriptions in Appendix A), and H.M.'s true misreadings usually contained many unshared segments, for example, *slumply* misread as "supplement" whereas controls' true misreadings usually contained only one, for example, *slumply* misread as "slumpily" and "sumply." Across all conditions, unshared segments were over 19 *SD* more common in errors of H.M. than controls (see Table 4), with omissions exceeding additions: H.M. omitted 24 segments (vs. a mean of 1.0 for controls, a difference in excess of 14 *SD*), and added 14 new segments (vs. a mean of 0.7 for controls, a difference in excess of 11 *SD*).

Order Errors

Close inspection indicated that H.M.'s true misreadings often contained correct target phonemes in the wrong order. For example, H.M. misread the pseudo-word *boshertin* as "barshite," misordering the /r/ with respect to the /sh/. Similarly, H.M. misread *thrish* as "thirst," misordering the /r/ with respect to the /i/. Overall order errors are shown (as percentage of total items) in Table 4 for H.M. and controls (means and *SD*), and occurred for 7% of the stimuli for H.M. but 0% for controls (*SD* = 0%).

Stimulus Length Effects

Reliable correlations between stimulus length (in letters) and H.M.'s overall errors, for example, $r_s = 0.93$ for LF words ($p < .01$) motivated our per letter per item measures of H.M.'s reading time deficits in Figures 5 and 6. To further explore these stimulus length effects, we divided our full set of stimuli (mean length 6.8 letters) into the categories "short" (6 letters or fewer) versus "long" (7 letters or more). Table 8 shows the number of items in each condition in the long versus short categories, together with overall reading errors by condition for H.M. and controls. Overall errors occurred on 42% of short stimuli for H.M. versus 8% for controls on average (*SD* = 4%), a difference of 8.5 *SD*. Long stimuli increased this difference between H.M. and controls by an additional 10%: H.M. misread 59% of the long stimuli whereas controls misread 15% on average (see Table 8). Of additional interest in Table 8 is the large difference between H.M.'s deficit for long MHF words (12.3 *SD*) versus short MHF words, where H.M. fell within 1.3 *SD* of the mean for controls. H.M. also produced no stress shifts, no successive approximations, and no order errors in reading short MHF words, and he produced no errors whatsoever in reading short MR pseudo-words, which consist of short, familiar morphemes that resemble very short MHF words.

Table 8. Overall Reading Errors (in %) for Short Words (Top Panel) and Long Words (Bottom Panel) by Condition for H.M. and Controls (Means With *SD* in Parentheses) in Study 2.

Word length	MHF words	LF words	PS pseudo-words	MR pseudo-words	Overall
Short words (M = 5.3 letters)	(*N* = 13)	(*N* = 15)	(*N* = 19)	(*N* = 3)	(*N* = 50)
Participants					
H.M.	8%	53%	63%	0%	42%
Controls	3%	17%	6%	0%	8%
(*SD*)	(4%)	(9%)	(6%)	(0%)	(4%)
Long Words (M = 8.6 letters)	(*N* = 13)	(*N* = 11)	(*N* = 7)	(*N* = 13)	(*N* = 44)
Participants					
H.M.	38%	91%	57%	54%	59%
Controls	1%	28%	26%	11%	15%
(*SD*)	(3%)	(11%)	(14%)	(8%)	(5%)

STUDY 3: A REPLICATION AND EXTENSION OF STUDY 2

Study 3 followed Study 2 by about 15 months and addressed five issues. Issue I was whether we could replicate the unexpected results in Study 2. Issue II was whether H.M. would make similar errors in reading LF words that he definitely knew prior to 1983. Issue III concerned test-retest consistency and effects of repetition, that is, whether H.M. would make the same or similar errors but with lower probability when re-reading the same stimuli. Issue IV was whether sensory and/or attentional factors influenced H.M.'s ability to read normal-font stimuli in list format in Studies 1–2: Study 3 stimuli were typed on individual index cards in large font, resembling the half-inch graphic font of Gabrieli et al. (1988). In another (separate) test for sensory deficits, H.M. named isolated letters in this half-inch font. Issue V was whether H.M.'s bilateral cerebellar damage (due to long-term use of epilepsy-controlling drugs; Corkin et al., 1997) caused his deficits in Studies 1–2: In Study 3 H.M. and patients with bilateral cerebellar damage read identical stimuli, an important control since neuroimaging and neuropsychological research (e.g., Helmuth, Ivry, & Shimizu, 1997; Ivry & Keele, 1989) indicates unexpected cerebellar involvement in several cognitive tasks (albeit not word reading).

METHOD

Participants

Participants were H.M. and 4 patients who were diagnosed with bilateral cerebellar damage in the absence of other brain damage or disorders, for example, dementia. Their mean age was 55 (range 31–79) and mean years of education 16.5 (range 14–18 years).

Materials

Materials were the Appendix A stimuli, plus 22 LF words that H.M. definitely knew in 1970 because he produced them spontaneously and appropriately in Marslen-Wilson's (1970) study (as cited Marslen-Wilson & Teuber, 1975). All stimuli were typed in upper-case, 24 point Courier font centered on individual, unlined 4 × 6″ index cards. To test for sensory deficits, H.M. also read the 26 letters of the alphabet typed in 24 point font on randomly ordered index cards.

Procedure

Basic procedures and instructions were identical to Study 2. The experimenter presented the Appendix A stimuli in the same order as in Study 2 on 12/15/98, and presented the remaining 22 words on 12/12/99. H.M. read the 26 cards containing individual letters on 3/15/99 and made no true errors on this letter-naming task.

RESULTS AND DISCUSSION

H.M.'s overall errors, true reading errors, and general-fluency errors for Appendix A stimuli were scored as in Study 2 and appear in Table 6. Despite the large font and isolated stimulus-presentation procedure, these error types were quantitatively similar in Studies 2 and 3 and exhibited the same pattern: H.M. misread LF words, PS pseudo-words, and MR pseudo-words more often than MHF words (see Table 6). H.M.'s misreadings were also qualitatively similar to Study 2: they frequently violated grapheme-to-phoneme relations in English, included many wrong-word errors, and were rarely corrected. H.M. corrected only 14% of his Study 3 errors, half as many as controls in Study 2. Basic Study 3 results therefore replicated Study 2 and ruled out sensory and attentional deficits as possible accounts of Study 2 results.

Turning to the issue of test-retest consistency, 89% of H.M.'s Study 3 errors involved items misread in Study 2 and H.M. made 10 errors that were identical (e.g., *drash* twice misread as "dash") or remarkably similar, for example, *adumbrate* as "embryate" (Study 2) and "emberate" (Study 3). In short, H.M.'s errors were very similar on retest, despite procedural differences and the 13-month lapse between the two experiments.

However, there were several differences. One was that H.M.'s successive approximations tended to be shorter in Study 3 (with one notable exception: H.M. misread *biltow* as "billow" in Study 2, but as "billow .. biTIllow .. 'Huh? (LJ)' .. biTIllow . . .bituh . . . . biTIL" in Study 3). Difference two was that H.M. made fewer errors in Study 3 than Study 2 for words (MHF and LF combined; $p < .05$, sign test with stimuli as unit of analysis), but not for pseudo-words ($p > .2$, same test). Both findings are consistent with NSRT: Under NSRT, word repetition improved H.M.'s performance by automatically strengthening pre-formed connections via engrainment learning, but repetition cannot improve H.M.'s pseudo-word reading because H.M. was incapable of forming the new connections for representing pseudo-words as coherent units.

Difference three concerned a remarkable error correction in Study 3. H.M. first misread *governor* as "govern," and two trials later, with *governor* no longer in view, H.M. spontaneously said, "It was governor." This long-delayed, memory-based error correction illustrates H.M.'s strong desire to succeed in this task but also points to an unusual underlying process not seen in controls. To make this correction H.M. must have been internally reiterating his earlier response over this period, a repetition process that comports with a wide range of other repetition tendencies documented in MacKay, Burke, et al. (1998), with his successive approximations, and with his tendency to repeat units that he had perceived or produced earlier in Studies 2 and 3.[10] Whether such a covert repetition process also contributed to H.M.'s planning time deficits for LF words and MR pseudo-words in Study 2 is currently unknown, however.

Turning to the 22 LF words that H.M. had used spontaneously and appropriately in 1970, H.M. in 1999 correctly read these words 7.33 *SD* less often than 1997 controls, and his 1999 reading errors were similar to his 1997 errors, for example, *undeclared* misread as "declaration." This replication of Study 2 results using LF words that H.M. definitely knew and used without error in 1970 (age 44) strengthens the case for exaggerated effects of aging on amnesia-linked RA.

Finally, the cerebellars fell within the range of 1997 controls but they outperformed H.M. in both errors and response times. Mean correct production times for cerebellars (measured as in Study 1) were 494 ms for MHF words ($SD = 69$; 3.0 *SD* shorter than H.M.), 511 ms for LF words ($SD = 84$; 8.9 *SD* shorter than H.M.), 516 ms for PS pseudo-words ($SD = 119$; 2.4 *SD* shorter than H.M.), and 844 ms for MR pseudo-words ($SD = 189$; 2.6 *SD* shorter than H.M.). Per-item planning times for cerebellars (measured as in Study 2) were 337 ms for MHF words ($SD = 154$; 2.3 *SD* faster than H.M.), 584 ms for LF words ($SD = 138$; 7.0 *SD* faster than H.M.), 783 ms for PS pseudo-words ($SD = 156$; 9.2 *SD* faster than H.M.), and 774 ms for MR pseudo-words ($SD = 143$; 10.0 *SD* faster than H.M.). However, mean correct planning times for cerebellars fell within the control range (despite being longer on average), and mean correct production

---

[10] For a detailed discussion of perseveration and data indicating that H.M. does not suffer from a general perseverative tendency, see MacKay, Burke, et al. (1998).

times also fell within the control range (mean overall difference, 0.35 *SD*).

Turning to errors, cerebellars produced 6% overall errors on MHF words (*SD* = 7%; 2.4 *SD* fewer than H.M.), 22% on LF words (*SD* = 11%; 4.3 *SD* fewer than H.M.), 22% on PS pseudo-words (*SD* = 12%; 3.3 *SD* fewer than H.M.), and 19% on MR pseudo-words (*SD* = 10%; 2.5 *SD* fewer than H.M.). However, cerebellar errors fell within the frequency range for controls (0.6 *SD* different overall), and like controls (but unlike H.M.), cerebellars never produced errors completely unlike the target in phonology, they virtually never produced unusual within-word pauses, and they usually corrected their errors. These findings suggest that H.M.'s deficits in Studies 1–2 are not attributable to his cerebellar damage, and render the cerebellum an unlikely locus for orthographic and phonological binding processes.

GENERAL DISCUSSION

This section summarizes H.M.'s current reading deficits and relates them to NSAT and NSRT. We then note problematic aspects of present results for NSRT and other major theories and hypotheses applicable to H.M.'s memory and reading abilities.

H.M.'s Selective Deficits in Reading Words and Pseudo-Words

Eleven aspects of present results indicate that in addition to his other deficits, H.M. currently suffers from deficits in reading LF words and pseudo-words aloud. These deficits involve planning and production time; a tendency to repeat units perceived or produced earlier; H.M.'s error rate, and the unusual nature of H.M.'s errors, for example, the minimal overlap in phonology between H.M.'s responses and the actual stimuli; H.M.'s successive approximations; his tendency to produce wrong-word errors; his suffix errors; his tendency to leave errors uncorrected; and his stress-shift errors, order errors, and segmentation errors.

Control procedures indicated that H.M.'s reading deficits were not due to experimental artifact, sensory problems, speed-accuracy-trade-off, practice effects, or attentional deficits. Moreover, Study 1 results indicated that H.M. had reading deficits for LF words and PS pseudo-words at age 60 or earlier, and relative to age-matched controls, these deficits have increased at an accelerated rate from age 60 to 71, especially for LF words. By extrapolation, then, H.M.'s ability to read LF words may have been deficit-free at some time between 1953–1986, which may explain why researchers failed to detect reading problems soon after H.M.'s operation.

However, it is important to note that H.M. has selective rather than across-the-board reading deficits, and viable theoretical accounts must explain why H.M. is currently as good or better than controls for some stimuli and some measures. For example, H.M. produced *fewer* general-fluency errors than controls, ruling out across-the-board factors such as reduced motivation or diffuse brain damage as accounts of H.M.'s reading deficits. H.M. also had much larger deficits for LF words and pseudo-words than MHF words, especially short MHF words (where H.M. closely resembled controls). Effects of practice or repetition on H.M.'s reading were also selective, with improvement from Studies 2 to 3 for words but not pseudo-words.

Aging and H.M.'s Selective Reading Deficits: NSAT and Other Theories

Our main aging results in Studies 1–3 strongly supported NSAT predictions. Under NSAT, transmission deficits interact with H.M.'s binding deficits to cause the exaggerated deterioration from age 60 to 72 in how H.M. reads LF words. Under NSAT, transmission deficits due to normal aging, nonrecent use, and infrequent use over his lifetime have rendered H.M.'s previously-functional phonological and orthographic nodes for LF words difficult or impossible to activate: The difficult-to-activate nodes explain H.M.'s greatly increased production times to correctly produce LF words. The impossible-to-activate nodes explain H.M.'s large planning time deficits for correctly produced LF words: Due to his binding deficit, H.M. could only re-re-establish the connections required to correctly produce defunct LF words via the inefficient process of repetition, which lengthened his planning times for LF words. H.M.'s intact engrainment learning mechanisms explain his tendency to repeat irrelevant units from earlier in the experiment, and also explain why H.M.'s reading improved from Studies 2 to 3 for words but not pseudo-words: Since engrainment learning is age-constant and can only strengthen preformed connections for words, internal and overt repetition improved H.M.'s reading for words but not pseudo-words, which under NSAT *always* require new connection formation, a process that has been difficult for H.M. ever since his age 26 lesion (including the period from age 60–72). Intact engrainment learning also explains why H.M.'s performance improved with repetition in priming tasks involving MHF words but not pseudo-words in Gabrieli et al. (1988).

In summary, a small number of theoretical constructs (intact engrainment learning with binding deficits that interact with frequency, recency, lexical status and aging) can readily account for H.M.'s selective age-linked deficits in reading isolated LF words and pseudo-words. Moreover, these same constructs also account for H.M.'s selective deficits in recalling recent episodes (see the review in MacKay, Burke, et al., 1998), in discovering hidden figures and errors in visual scenes (see MacKay & James, 2000), in comprehending novel sentences (see MacKay, Stewart, et al., 1998), in comprehending isolated words and pseudo-words (see James & MacKay, 2001), in producing novel sentences (see MacKay, Burke, et al., 1998), and in reading sentences containing MHF words (see MacKay & James, 2001). In short, H.M.'s spared and impaired abilities in a wide range of language and memory abilities illustrate a ‘‘binding syndrome’’ with a common cause or coherent set of causes under NSAT.

Whether H.M. exhibits planning time deficits in other behaviors besides reading is currently unknown, but many studies have already demonstrated close temporal links between hippocampal system activity and various types of action planning in humans (Halgren, 1991) and other species (e.g., Eichenbaum, Otto, & Cohen, 1994; Muller & Kubie, 1987; O'Mara, Rolls, Berthoz, & Kesner, 1994; Ringo, Sobotka, Diltz, & Bunce, 1994; Wilson, 1994, p. 499; Wilson, Riches, & Brown, 1990). Although Eichenbaum et al. (1994, p. 507) dismissed these close temporal links as irrelevant to ‘‘the more ‘central’ functions of the hippocampus,’’ these links make sense within NSAT because binding also plays a role in planning actions, including seemingly simple actions such as reading aloud.

Other Theories of RA and Aging

Present data are problematic for the theory that progressive age-linked anterograde deficits preceding detection and diagnosis of Alzheimer's and Korsakoff's amnesia cause RA (e.g., Albert, Butters, & Levin, 1979): Progressive AA cannot explain present data since H.M. had sudden-onset AA that has remained unchanged since 1953. H.M.'s selective age-linked RA (e.g., large deficits for LF but not MHF words) is also problematic for the theory that RA reflects across-the-board damage to retrieval mechanisms (see, e.g., Kopelman, Stanhope, & Kingsley, 1999). Also ruled out for present data are theories of RA requiring frontal lobe damage to disrupt retrieval processes (Kopelman et al., 1999): H.M. has no known frontal damage (Corkin et al., 1997).

Single-Factor Theories

The selective nature of H.M.'s reading deficits present a challenge for theories that ascribe H.M.'s age-linked deficits to a single, nonselective factor, e.g., reduced working memory capacity, diffuse brain damage, reduced motivation, and ‘‘general-cognitive-decline.’’ These aging theories have difficulty explaining H.M.'s disproportionately greater age-linked decline in reading LF words than PS pseudo-words between age 60 and 71, his disproportionately greater planning time deficits for pseudo-words than MHF words, and his disproportionately greater deficits for planning times than production times. H.M's greater deficits for planning than production times also contradict the hypothesis of Postle and Corkin (1998) that H.M. exhibits exaggerated age-linked slowing (which implies proportionate declines across all conditions and all time measures for H.M. relative to age-matched controls). The stability in H.M.'s IQ scores over the last 20 years (see the Introduction) is also problematic for a general-cognitive-decline hypothesis.

NSRT Predictions and H.M.'s Detailed Reading Deficits

The next section addresses two NSRT predictions that failed. However, most results were consistent with NSRT predictions: the parallels between the relative frequency of H.M.'s errors across conditions and his planning times but not his production times; his tendency to leave errors uncorrected (see also MacKay, 1992; and MacKay, Burke, et al., 1998); his frequent errors involving repetition and his long planning times for LF words and pseudo-words; his increased errors for illegal versus legal MR pseudo-words; his high rate of wrong-word errors, especially for PS pseudo-words; and his larger error and production time deficits for LF words than PS and MR pseudo-words; his normal planning times, production times and errors for short MHF words; and his smaller deficits for long MHF words than LF words and pseudo-words. Also consistent with NSRT were H.M.'s relatively greater deficits for illegal than legal MR pseudo-words, and his unusual errors involving syllabic segmentation, stress, and suffixes.

Problematic Results for NSRT and Other Major Theories and Hypotheses

NSRT

Two results were inconsistent with NSRT predictions. One was H.M.'s longer planning time deficits for MR pseudo-words than LF words (see Fig. 6). The small number of correct responses that H.M. produced for LF words ($N=6$) and MR pseudo-words ($N=8$) may account for this unexpected result, which suggests that errors may represent a better index of H.M.'s deficits in these conditions than correct response times in these conditions. However, the second inconsistent result cannot be explained in terms of stimulus factors: H.M.'s reliable planning and production time deficits for long MHF words. The connections for activating long MHF words should be strong and functional if H.M. has used these words recently and frequently since his 1953 lesion. However, this precondition may not hold if H.M's main source of language input has been watching television (see Gabrieli et al., 1988): H.M. may have read, comprehended, and produced very few long MHF words from 1953 to 1997, and those especially rarely and nonrecently.

Under a second hypothesis, some of H.M.'s MHF words are no longer deficit-free after almost 50 post-lesion years of progressive and irreversible transmission deficits: H.M. may now be completely deficit-free only for words with very high frequency, for example, over 100 instances per million. A third hypothesis not only explains H.M.'s difficulties with MHF words but also specifically addresses the word length effect and H.M.'s problems with syllable segmentation, stress location, and segment order. We develop this hypothesis in the next section.

The "Try-Then-Test" Hypothesis

Some aspects of H.M.'s segmentation, stress-shift, and order errors are readily explained via his tendency to produce phonologically similar wrong-word errors that by chance involve re-segmentation, stress-shifts and re-ordering. However, other aspects of these complex and unusual errors call for a more complex account. Consider H.M.'s misreading of *thrish* as "thirst," followed by the successive approximation "thrust" which both corrects the misordered [R] in *thirst* and increases phonological similarity to *thrish*, as if H.M. had compared his initial response with the stimulus [THRISH], detected some of his errors, and corrected them on his second attempt. By this "try-then-check" strategy, H.M. may have initially produced the first word resembling *thrish* that came to mind, but subsequent approximations benefited from active comparison of the stimulus with this initial response.

Consistent with an active try-then-check strategy, H.M. seemed to experiment with differing stress patterns and grapheme-phoneme correspondences to maximize fit. For example, consider H.M.'s approximations to the LF word *abdicate*: "abiCURgle .. duh .. Abidackle .. abeDIckle." Here, H.M.'s "duh" suggests that he noted the missing [D] in his first approximation (abiCURgle) and therefore included [D] in subsequent approximations, "Abidackle" and "abeDIckle." The differing segments, syllabic groupings, and stress patterns in these approximations suggest an active process of experimenting with various ways of combining or parsing components and stressing syllables to better match the target string. Such a process may help explain the role of stimulus length in H.M.'s errors because short items have higher *a priori* probability of success via try-then-check than long items. For example, the chance of success on the first pass of a try-then-check strategy for producing primary stress is 50% for two-syllable words versus 33% for three-syllable words, and *internal* reiterations of try-then-check would increase the overall hit rate while maintaining the effect of length. Internal reiterations of the try-then-check strategy may also explain the absence of successive approximations for short MHF words and short MR pseudo-words (which themselves consist of short, highly familiar subunits).

However, another aspect of the *abdicate* example suggests the possibility of an additional, more fundamental problem. Note that H.M. produced the correct order for [I] relative to [A] in his second approximation to *abdicate* (“Abidackle”), but not in his third approximation (“abeDIckle”). The fact that H.M.'s order approximations sometimes proceeded in the wrong direction suggests a problem in representing orthographic sequences and/or visual sequences in general, and further research into H.M.'s perception of letter sequences seems warranted.

Procedural Memory Hypothesis

H.M.'s reading deficits contradict the hypothesis that H.M. has intact procedural memory for skills learned before his operation (Cohen & Eichenbaum, 1993, pp. 49–219; Kesner, 1998; Squire, 1987, pp. 151–169; see also Shanks, 1996, for a review of other problems with this procedural-memory hypothesis), a conclusion that will probably generalize to other amnesics since “the purity of the preservation of procedural memory in H.M. has rarely been equaled in other amnesic patients” (Parkin, 1996, p. 342).

Pure Storage Deficit Hypothesis

H.M.'s reading deficits contradict the hypothesis that H.M. suffers from a pure storage deficit that has left his ability to read intact. It might nevertheless be claimed that H.M.'s storage deficit is earlier in origin and independent from his problems in reading and other aspects of language use (including anatomically independent; see MacKay, 2001). However, this claim is problematic: After all, H.M.'s language comprehension deficits were discovered in 1966 (see MacKay, Stewart et al., 1998), the same year as H.M.'s “hippocampal amnesia” (see Milner, Corkin, & Teuber, 1968). Moreover, H.M. might also have exhibited reading deficits in 1966 had reading been tested then: All we currently know is that the first reported data on H.M.'s reading (1986, see Study 1) indicated reading deficits.

Stages-Of-Processing Theories

The procedural-memory and pure-storage-deficit hypotheses are part of a more general theoretical framework describing relations between perception, memory, and response production (including language). This general stages-of-processing framework originated with Descartes (see MacKay, Burke, et al., 1998) and comes in many explicit and implicit forms (see, e.g., Massaro, 1978; Sternberg, 1985). Figure 7 illustrates the six stages postulated in Pashler and Carrier (1996, pp. 15–20): stimulus recognition, sensory memory, short-term memory storage, long-term memory storage, retrieval from long-term memory, and response production. However, stage modules have also been postulated for retrieval from short-term memory (see, e.g., Atkinson & Shiffrin, 1968; Pashler &

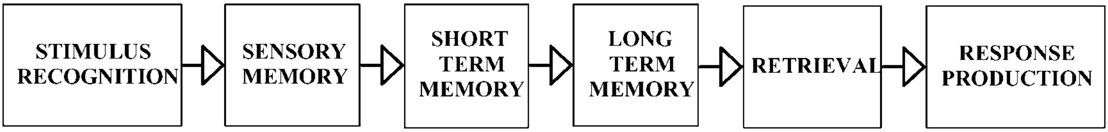

Fig. 7. A standard stage-theory flow chart with sequentially ordered stages for perception, storage, retrieval, and production (as described in Pashler & Carrier, 1996, pp. 15–20).

Carrier, 1996), for episodic versus semantic memory (e.g., Tulving, 1983), for explicit versus implicit memory (e.g., Graf & Schachter, 1985), for sensory and perceptual versus modality-independent memory (e.g., Schachter, 1990), for procedural versus declarative memory (e.g., Squire, 1987, pp. 151–169), for reference versus working memory (Cohen & Eichenbaum, 1993, pp. 49–219), for short-term working memory versus long-term working memory (Kintsch, 1998), and for long-term versus very long-term memory (Squire, 1987, pp. 204–214).

Although H.M. has been cited as support for most of these stages, only four basic stages, implicit in all stage-theories, suffice to illustrate the historic importance of H.M. within the stage-theory framework: perception/comprehension, memory storage, memory retrieval, and response production (see e.g., Gordon, 1989, pp. 196–216). These four stages are assumed to contain entirely separate units, mechanisms, or processes, to perform distinct functions, and to run off sequentially, that is, people first comprehend verbal inputs, next store and consolidate the comprehended inputs in memory, and then during recall, retrieve the stored memory, and finally, express the memory in language. Since the 1960s, leading memory researchers (e.g., Atkinson & Shiffrin, 1968) have suggested that H.M.'s condition dissociates all four stages (with damage to the storage stage, but not to the comprehension, production, or retrieval stages for memories preceding H.M.'s operation), thereby providing the strongest possible support for stage-theories (Wickelgren, 1975).

However, H.M.'s deficits in reading isolated words and pseudo-words contradict the stage-theory assumptions that perception and production involve entirely separate units and processes from memory storage and retrieval, at least for lower-level stages involving orthography, phonology, and lexical meaning. H.M.'s deficits in comprehending and producing novel sentences contradict stage-theory assumptions at higher levels: H.M.'s comprehension deficits in MacKay, Stewart, et al. (1998) have ruled out the postulated dissociation between stage 1 (language comprehension) and stage 2 (memory storage) at sentential-semantic levels, and H.M.'s production deficits in MacKay, Burke, et al. (1998) have ruled out postulated dissociations between stage 2 (memory storage), stage 3 (memory retrieval), and stage 4 (language production) at sentential-semantic levels.

However, H.M.'s language deficits are just one of many empirical challenges facing the stage-theory framework (see e.g., Bock, 1996; Burke et al., 1991; MacKay & Abrams, 1996). Another challenge for stage-theories is to establish clear theoretical dividing lines between where comprehension, storage and retrieval of verbal information ends and where language production (including reading aloud) begins (see MacKay, Burke, et al., 1998). In this regard, Goodglass and Wingfield (1997, p. 18) note that stage-theorists often take "observations … as correct because they fit the logic of the model, (while) equally common observations that do not fit the model" are ignored (parentheses ours; see also MacKay & James, 2001).

Despite these serious problems, however, stage-theories are likely to remain popular because of their flexibility (see MacKay, Burke, et al., 1998). For example, *ad hoc* modifications for increasing the ostensible fit of stage-theories with present data are readily imagined: One could assume that H.M.'s procedural memory is intact except as it relates to reading LF words and pseudo-words. Or that H.M.'s brain damage has affected some new systems for memory storage that are essential for exactly those aspects of reading that are problematic for H.M., let us say a "kinetic intelligence" system for reading LF words, and an "adaptive memory" system for reading pseudo-words.

Other Distributed Memory Theories

NSAT and NSRT are distributed-memory theories within a family of recently developed distributed-memory theories that make reference to hippocampal systems and/or binding mechanisms to explain both verbal memory phenomena and detailed aspects of language processing (see, e.g., McClelland, 1985; also Carpenter & Grossberg, 1993; Grafman & Weingartner, 1996; MacKay, 1990; McClelland, McNaughton, & O'Reilly, 1995; Metcalfe, Cottrell, & Mencl, 1992; Saffran, 1990; and Wickelgren, 1979). Unlike stage theories, distributed-memory theories do not represent memory for verbal materials via discrete serial stages involving differing units and processes, but via varying connection strengths for millions of neural units or nodes distributed throughout a vast interactive network that plays a role in both perception and production of language.

Because most distributed-memory theories view the formation of new cortical connections as necessary for producing pseudo-words (including, in H.M.'s case, words that entered English after his 1953 operation) but not for producing familiar words (see MacKay, Burke, et al., 1998), H.M.'s deficits for pseudo-words are consistent with most distributed memory theories (with one notable exception; see MacKay, Burke, et al., 1998; MacKay, Stewart, et al., 1998). Nevertheless, distributed-memory theories have been faulted because "their primitive elements … do not have names and clearly diagrammable effects on other named systems" and they "lack the transparent logic of linear stage models" (Goodglass & Wingfield, 1997, p. 24). However, these criticisms do not apply to NSAT where nodes do have names and clearly diagrammable relations to other nodes and node systems, and where logical operations of the theory are transparent (although some aspects of this logic have been developed in detail elsewhere; see MacKay, 1987, 1990). Moreover, NSAT is the only distributed-memory theory that takes aging into account and postulates fine-grained relations between reading aloud (including errors and rate in perceiving, planning, and producing words and pseudo-words), sentence comprehension, sentence production, and other distributed-memory functions such as episodic memory in order to explain the full spectrum of H.M.'s (highly selective) deficits.

Theoretical Extrapolation, Limitations and Caveats

Further research is needed to test whether other amnesic patients with hippocampal damage resembling H.M.'s (and no damage to temporal neocortex) will show the same pattern of age-linked reading deficits as H.M. However, replicating the present results with other hippocampal amnesics will not be as simple as the Study 3 replication of Study 2. As Corkin (1984, p. 258) notes, "the patient group to which H.M. belongs (is) N = 1 (parentheses ours)." For example, no otherwise comparable amnesics are similar to H.M. in time-since-lesion: All have had less time to develop or use compensatory strategies (e.g., engrainment learning), or to develop transmission deficits due to aging, and nonrecent and infrequent use of information since their lesion and over the course of their lifetimes. This means that replicating H.M.'s results with other amnesics will require theoretical extrapolation.

Theoretical extrapolation will also be necessary to compare the reading deficits of H.M. and patients classified as dyslexic. However, such comparisons raise some interesting questions: For example, do some dyslexics exhibit H.M.'s pattern of impaired planning times with relatively unimpaired production times? Do some dyslexics tend to repeat units perceived or produced earlier? Do some dyslexics have damage to hippocampal and/or cerebellar structures? Do dyslexics with damaged hippocampal and/or cerebellar structures often produce suffix errors, for example, *internal* misread as "international" (an error typical of deep dyslexics; see Shallice, 1988, p. 99)? Answers to these questions are currently unknown.

We conclude with some caveats for present hypotheses and conclusions. One set of caveats concerns the brain areas responsible for H.M.'s age-linked reading deficits. Results of Study 3 did rule out the cerebellum, but damage to H.M.'s parahippocampal structures and to the tips of his temporal poles are other possible candidates (but see footnote 8). A related caveat concerns H.M.'s brain damage as a function of time since 1953: Hippocampal lesions can trigger transneuronal degeneration (involving, for example, dendrites in the medial mammillary nucleus; Loftus, Knight, & Amaral, 2000), especially in older adults (see Wilkinson & Davies, 1978) so that H.M.'s 1953 lesion may have triggered age-linked damage elsewhere since then. A final caveat concerns the functional status of H.M.'s posterior hippocampus. Under NSAT, thousands of binding nodes play a role in memory as traditionally defined and in the normal comprehension, acquisition, and production of words and phrases (see MacKay & Burke, 1990). There also exists a surplus of binding nodes for binding entire domains of new nodes, so that, for example, any semantic-system node can connect with virtually any other semantic-system node with the help of binding nodes. If left-hippocampal structures contain thousands of binding nodes for conjoining different classes of never previously linked units during normal use and acquisition of language (at all ages; MacKay, 1990), and if H.M.'s posterior hippocampus is intact and functional (see Corkin et al., 1997), then some of H.M.'s language-specific binding nodes may have been spared, but we cannot be certain how many or which have been destroyed versus spared. Similar caveats regarding damaged versus intact binding nodes also apply to H.M.'s impaired episodic memory capacities and to his residual capacities in other areas, for example, motor learning.

Another caveat concerns MacKay's (1990) simplifying assumptions that all possible binding nodes for language exist at birth as part of the innate basis for language acquisition, and that binding nodes damaged after some critical period cannot be replaced. These simplifying assumptions are not central to NSAT, and we hope to test and perhaps reject them in future experiments with H.M., consistent with recent neuro-anatomical evidence indicating hippocampal neurogenesis in older adults (see, e.g., Eriksson et al., 1998; Kempermann, Kuhn, Winkler, & Gage, 1998).

Our final caveats concern the limits of our hypothesized relations between processes for episodic memory and reading LF words and pseudowords aloud: A sufficiently precise lesion in another patient could in principle dissociate these processes by damaging high-level episodic binding nodes without damaging low-level phonological and orthographic binding nodes, or vice versa. We also wish to stress the limits of our claim that H.M.'s binding deficits contribute to his deficits in both reading and episodic memory. We are not arguing that H.M.'s reading deficits caused his episodic memory deficits in the many previous studies involving recall of written words, an argument that follows logically within stage-theories, but not within distributed-memory theories. Nor are we arguing that aging has made H.M. an impure case that no longer applies to stage-theories. Case H.M. remains as relevant as ever to both distributed-memory and stage-theories.

ACKNOWLEDGEMENTS

Don MacKay gratefully acknowledges support of this research from the 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), and Dr. William Marslen-Wilson for providing his unpublished transcript of H.M.'s conversational speech. We also thank Deborah Burke, Angela Kessell, Barbara Knowlton, Kay Livesay, William Marslen-Wilson, Michelle Miller, Christine Rho, Indre Viskontas, and Nancy Woolf for helpful comments on an earlier version of this paper, Irwin Stein for recruiting participants, and Ibby Ambrose, Bryan Chae, Helen Chon, Deirdre Milligan, and especially Angela Kessell for research assistance.

REFERENCES

Albert, M.S., Butters, N., & Levin, J. (1979). Temporal gradients in the retrograde amnesia of patients with alcoholic Korsakoff's disease. *Archives of Neurology*, *36*, 211–216.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation*: *Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.

Balota, D.A., & Ferraro, F.R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer type. *Journal of Memory & Language*, *32*, 573–592.

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

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

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

Carpenter, G., & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in the Neurosciences*, *16*, 131–37.

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

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*, 589–608.

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.

Eichenbaum, H.B., Otto, T., & Cohen, N. (1994). Two functional components of the hippocampal memory system. *Behavioral & Brain Sciences*, *17*, 449–518.

Eriksson, P.S., Perfilieva, E., Bjork-Eriksson, T., Alborn, A.M., Nordborg, C., Peterson, D.A., & Gage, F.H. (1998). Neurogenesis in the adult human hippocampus. *Nature Medicine*, *4*, 1313–1317.

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

Freed, D.M., Corkin, S., & Cohen, N.J. (1987). Forgetting in H.M.: A second look. *Neuropsychologia*, *25*, 461–471.

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

Goodglass, H., & Wingfield, A. (1997). Word-finding deficits in aphasia: Brain-behavior relations and clinical symptomatology. In H. Goodglass & A. Wingfield (Eds.), *Anomia: Neuroanatomical and cognitive correlates* (pp. 3–27). New York: Academic Press.

Gordon, W.C. (1989). *Learning and memory.* Pacific Grove, CA: Brooks/Cole.

Graf, P., & Schachter, D. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning*, *Memory & Cognition*, *11*, 501–518.

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

Halgren, E. (1991). Firing of human hippocampal units in relation to voluntary movements. *Hippocampus*, *1*, 153–162.

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.

Hodges, J.R. (1995). Retrograde amnesia. In A.D. Baddeley, B.A. Wilson, & F.N. Watts (Eds.), *Handbook of memory disorders* (pp. 81–108). New York: Wiley.

Hunkin, N.M., Parkin, A.J., Bradley, V.A., Burrows, E.H., Aldrich, F.K., Jansari, A., & Burdon-Cooper, C. (1995). Focal retrograde amnesia following closed head injury: A case study and theoretical account. *Neuropsychologia*, *33*, 509–523.

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

James, L.E., & Burke, D.M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1378–1391.

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

Kapur, N. (1993). Focal retrograde amnesia in neurological disease: A critical review. *Cortex*, *29*, 217–234.

Kempermann, G., Kuhn, H.G., Winkler, J., & Gage, F.H. (1998). New nerve cells for the adult brain. Adult neurogenesis and stem cell concepts in neurological research. *Nervenartz*, *69*, 851–857.

Kesner, R.P. (1998). Neural mediation of memory for time: Role of the hippocampus and medial prefrontal cortex. *Psychonomic Bulletin & Review*, *5*, 585–596.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Kopelman, M.D., Stanhope, N., & Kingsley, D. (1999). Retrograde amnesia in patients with diencephalic, temporal lobe or frontal lesions. *Neuropsychologia*, *27*, 939–958.

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.

Levin, H.S., High, W.M., Meyers, C.A., von Laufen, A., Hayden, M.E., & Eisenberg, H.M. (1985). Impairment of remote memory after closed head injury. *Journal of Neurology, Neurosurgery, and Psychiatry*, *48*, 556–563.

Loftus, M., Knight, R.T., & Amaral, D.G. (2000). An analysis of atrophy in the medial mammilary nucleus following hippocampal and fornix lesions in humans and nonhuman primates. *Experimental Neurology*, *163*, 180–190.

MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, *89*, 483–506.

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. (1992). Awareness and error detection: New theories and paradigms. *Consciousness and Cognition*, *1*, 199–225.

MacKay, D.G. (2001). A tale of two paradigms or metatheoretical approaches to cognitive neuropsychology. Did Schmolck, Stefanacci, and Squire demonstrate that "Detection and explanation of sentence ambiguity are unaffected by hippocampal lesions but are impaired by larger temporal lobe lesions"? *Brain and Language*, *78*, 265–272.

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

MacKay, D.G., & Abrams, L. (1998). Age-linked declines in retrieving orthographic knowledge: Empirical, practical, and theoretical implications. *Psychology and Aging*, *13*, 647–662.

MacKay, D.G., Abrams, L., & Pedroza, M.J. (1999). Aging on the input versus output side: Age-linked asymmetries between detecting versus retrieving orthographic information. *Psychology and Aging*, *14*, 3–17.

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

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

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

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

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

Marslen-Wilson, W. (1970). *Biographical interviews with H.M*. Unpublished transcript (pp. 1–144). Referenced in Marslen-Wilson & Teuber (1975).

Marslen-Wilson, W., & Teuber, H.-L. (1975). Memory for remote events in anterograde amnesia: Recognition of public figures from news photographs. *Neuropsychologia*, *13*, 353–364.

Massaro, D.W. (1978). A stage model of reading and listening. *Visible Language*, *12*, 3–26.

McClelland, J.L. (1985). Distributed models of cognitive processes: Applications to learning and memory. In D.S. Olton, E. Gamzu, & S. Corkin (Eds.), *Memory dysfunctions: An integration of animal and human research from preclinical and clinical perspectives* (Vol. 444, pp. 1–9). New York: New York Academy of Sciences.

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

Metcalfe, J., Cottrell, G.W., & Mencl, W.E. (1992). Cognitive binding: A computational-modeling analysis of a distinction between implicit and explicit memory. *Journal of Cognitive Neuroscience*, *4*, 289–298.

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.

Morris, R.G.M. (1999). Memory: Topographical knowledge survives hippocampal damage. *Current Biology*, *9*, 890–892.

Muller, R.U., & Kubie, J.L. (1987). The firing of hippocampal place cells predicts the future position of freely moving rats. *Journal of Neuroscience*, *9*, 4101–4110.

Murre, J.M.J. (1996). TraceLink: A model of amnesia and consolidation memory. *Hippocampus*, *6*, 675–684.

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

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

O'Mara, S.M., Rolls, E.T., Berthoz, A., & Kesner, R.P. (1994). Neurons responding to whole-body motion in the primate hippocampus. *Journal of Neuroscience*, *14*, 6511–6523.

Parkin, A.J. (1996). H.M.: The medial temporal lobes and memory. In C. Code, C.W. Wallesch, Y. Joanette, & A.R. Lecours (Eds.), *Classic cases in neuropsychology*. Hove, UK: Erlbaum, Taylor, & Francis.

Pashler, H., & Carrier, M. (1996). Structures, processes, and the flow of information. In E.L. Bjork & R.A. Bjork (Eds.), *Memory* (pp. 4–29). San Diego: Academic 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.

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory & Language*, *42*, 342–364.

Ringo, J.L., Sobotka, S., Diltz, M.D., & Bunce, M.C. (1994). Eye movements modulate activity in hippocampal, parahippocampal, and inferotemporal neurons. *Journal of Neurophysiology*, *71*, 1285–1288.

Saffran, E.M. (1990). Short-term memory impairment and language processing. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics: Advances in models of cognitive function and impairment* (pp. 137–168). Hillsdale, NJ: Lawrence Erlbaum.

Santiago, J., MacKay, D.G., Palma, A., & Rho, C. (1999). Picture naming latencies vary with sequential activation processes in hierarchic syllable and word structures. *Language and Cognitive Processes*, *15*, 1–44.

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

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

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

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

Sternberg, R.J. (Ed.). (1985). *Human abilities: An information processing approach* (259pp.). New York: Freeman.

Stratton, R.P., Jacobus, K.A., & Brinley, B. (1975). Age-of-acquisition, imagery, familiarity and meaningfulness norms for 543 words. *Behavior Research Methods & Instrumentation*, *7*, 1–6.

Tulving, E. (1983). *Elements of episodic memory.* Oxford: Clarendon Press.

Webster's new collegiate dictionary. (1949). Springfield, MA: G. & C. Merriam.

White, K.K., & Abrams, L. (2002). Does priming specific syllables during tip-of-the-tongue states facilitate word retrieval in older adults? *Psychology and Aging*, *17*, 226–235.

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

Wickelgren, W.A. (1975). The long and the short of memory. In D. Deutsch, & J.A. Deutsch (Eds.), *Short-term memory* (pp. 41–63). New York: Academic Press.

Wickelgren, W.A. (1979). Chunking and consolidation: A theoretical synthesis of semantic networks, configuring in conditioning, S-R versus cognitive learning, normal forgetting, the amnesic syndrome and the hippocampus arousal system. *Psychological Review*, *86*, 44–60.

Wilkinson, A., & Davies, I. (1978). The influence of age and dementia on the neurone population of the mammillary bodies. *Age Aging*, *7*, 151–160.

Wilson, F.A. (1994). Hippocampal neuronal activity in rat and primate: Memory and movement. *Behavioral & Brain Sciences*, *17*, 499–500.

Wilson, F.A., Riches, I.P., & Brown, M.W. (1990). Hippocampus and medial temporal cortex: Neuronal activity related to behavioral responses during performance of memory tasks by primates. *Behavioral & Brain Research*, *40*, 7–28.

APPENDIX A

Stimuli by Condition, With Their Most Common Pronunciation(s) in IPA Transcription

HF Words	IPA Transcription	LF Words	IPA Transcription	PS Pseudoword	IPA Transcription	MR Pseudowords	IPA Transcription
escape	ɛsˈkep	serrated	ˈsɛretɪd, sɛˈretɪd	quintity	ˈkwɪntəti	eventment	ɪˈvɛntmənt
payment	ˈpeymənt	grovel	ˈgrɑvl	astery	ˈæstəri	unmelt	ʌnˈmɛlt
crowd	ˈkraʊd	adumbrate	ædˈʌmbret	jat	ˈdʒæt	safetyhood	ˈseftihʊd
internal	ɪnˈtərnl	crypt	ˈkrɪpt	bondit	ˈbɑndɪt	duskly	ˈdʌskli
attract	əˈtrækt	zealot	ˈzɛlət	pediodical	pɛdiˈɑdɪkl, pidiˈɑdɪkl	reversement	rɪˈvərsmənt
remarkable	riˈmɑrkəbl	akimbo	əˈkɪmbo	biltow	ˈbɪltaʊ, ˈbɪltoʷ	friendlyhood	ˈfrɛndlihʊd
driver	ˈdraɪvər	chameleon	kəˈmiliən	etual	ˈɛtʊəl, ˈɛtʃʊəl	undoctoral	ʌnˈdɑktərəl
literature	ˈlɪtərətʃʊr	jettison	ˈdʒɛtəsn	kovernor	ˈkovənər, ˈkʌvənər	metalousness	məˈtæləsnəs
satisfy	ˈsætɪsfaɪ	efface	ɪfˈes	lapte	ˈlæpt, ˈlæpti	consultment	kənˈsʌltmənt
forget	fərˈgɛt	euphemism	ˈjufənɪzm	bosh	ˈbɑʃ, ˈboʃ	mortalness	ˈmɔrtəlnəs
governor	ˈgʌvənər	gulp	ˈgʌlp	merling	ˈmərliŋ	peoplement	ˈpiplmənt
mistake	mɪsˈtek	abacus	ˈæbəkəs	ampetite	ˈæmpətaɪt	retrend	riˈtrɛnd, rɪˈtrɛnd
guest	ˈgɛst	sulk	ˈsʌlk	yelkey	ˈjɛlki	accusement	əˈkjuzmənt
hence	ˈhɛns	inane	ɪnˈen	minate	ˈmɪnet	slumply	ˈslʌmpli
worry	ˈwʊri	zodiac	ˈzodiæk	rulber	ˈrulbɑr, ˈrʌlbɑr	discage	dɪsˈkedʒ
kitchen	ˈkɪtʃən	labyrinth	ˈlæbərɪnθ	drash	ˈdræʃ	untale	ʌnˈtel
marine	məˈrin	abdicate	ˈæbdəket	monder	ˈmɑndər		
noise	nɔɪz	wretch	ˈrɛtʃ	rittlefin	ˈrɪtlfɪn		
protect	prəˈtɛkt	lentil	ˈlɛntl	boshertin	ˈboʃərtɪn, ˈbɑʃərtɪn		
building	ˈbɪldiŋ	ellipsis	ɪˈlɪpsɪs	chim	ˈtʃɪm		
orchestra	ˈɔrkəstrʌ	lob	ˈlɑb	thrish	ˈθrɪʃ		
rural	ˈrʊrəl	primp	ˈprɪmp	dunger	ˈdʌngər, ˈdʌndʒər		
scene	ˈsin	squander	ˈskwɑndər	mundle	ˈmʌndl		
join	ˈdʒɔɪn	stamina	ˈstæmənə	prain	ˈpren		
theater	ˈθiətər	papyrus	pəˈpaɪrəs	strome	ˈstrom		
warn	ˈwɔrn	yolk	ˈjok	vithon	ˈvɪθɑn		