
Donald G. MacKay
Psychology Department, University of California, Los Angeles
Lori E. James
Psychology Department, University of Colorado, Colorado Springs
Christopher B. Hadley
Psychology Department, University of California, Los Angeles
and
Kethera A. Fogler
Psychology Department, Saint Louis University

Grant sponsor: The Samuel A. MacKay Memorial Research Fund.

Correspondence to: Donald G. MacKay, Psychology Dept., University of California, Los Angeles, Los Angeles, CA 90095-1563.

Email: mackay@ucla.edu; Phones: 310-825-8465        FAX: 310-206-5895
Abstract
Three language production studies indicate that amnesic H.M. produces speech errors unlike everyday slips-of-the-tongue. Study 1 was a naturalistic task: H.M. and 6 controls closely matched for age, education, background and IQ described what makes captioned cartoons funny. Nine judges rated the descriptions blind to speaker identity and gave reliably more negative ratings for coherence, vagueness, comprehensibility, grammaticality, and adequacy of humor-description for H.M. than the controls. Study 2 examined “major errors”, a novel type of speech error that is uncorrected and reduces the coherence, grammaticality, accuracy and/or comprehensibility of an utterance. The results indicated that H.M. produced seven types of major errors reliably more often than controls: substitutions, omissions, additions, transpositions, reading errors, free associations, and accuracy errors. These results contradict recent claims that H.M. retains unconscious or implicit language abilities and produces spoken discourse that is “sophisticated,” “intact” and “without major errors.” Study 3 examined whether three classical types of errors (omissions, additions, and substitutions of words and phrases) differed for H.M. versus controls in basic nature and relative frequency by error type. The results indicated that omissions, and especially multi-word omissions, were relatively more common for H.M. than the controls; and substitutions violated the syntactic class regularity (whereby, e.g., nouns substitute with nouns but not verbs) relatively more often for H.M. than the controls. These results suggest that H.M.’s MTL damage impaired his ability to rapidly form new connections between units in the cortex, a process necessary to form complete and coherent internal representations for novel sentence-level plans. In short, different brain mechanisms underlie H.M.’s major errors (which reflect incomplete and incoherent sentence-level plans) vs. everyday slips-of-the tongue (which reflect errors in activating pre-planned units in fully intact sentence-level plans). Implications of Study 1-3 results are discussed for systems theory, binding theory and relational memory theories.

Keywords: speech errors, amnesic H.M., language production, comprehensibility, grammaticality

Thousands of people have heard the famous amnesic H.M. produce error-free sentences on National Public Radio (NPR; Newhouse, 2007). These and many other people have the impression that H.M.’s language production is normal, intact, artful, “sophisticated” (Kolb & Whishaw, 2003, p. 500) and “without major errors” (Skotko, Andrews & Einstein, 2005, p. 406). The present paper reports experimental data that contradict this impression and indicate that H.M. produces new types of speech errors unlike normal, everyday slips-of-the-tongue.

First some background information. Following a highly localized medial temporal lobe (MTL) lesion in 1953, H.M. has exhibited selective memory deficits, with impaired recall of new or never-before-encountered semantic and episodic information, but spared recall of semantic information that H.M. encountered before and after 1953 (see e.g., Gabrieli, Cohen & Corkin, 1988; James & MacKay, 2001). Under a hypothesis that has profoundly influenced theories in psychology and the brain sciences over the past 50 years (see e.g., MacKay, Burke, & Stewart, 1998), H.M.‘s selective memory deficits reflect separate systems for storing new information (damaged in H.M.) versus retrieving already-stored information (undamaged in H.M.).

Milner, Corkin, and Teuber (1968) proposed a related hypothesis with equally profound impact on current theories of the relation between language and memory. Under this Milner et al. hypothesis, language-linked processes are intact and normal in H.M. If correct, this “intact language hypothesis” indicates dissociations between memory storage systems (damaged in H.M.) and the systems for comprehending and producing sentences (by hypothesis undamaged in H.M.). These hypothesized dissociations have motivated the independent modules for processing memory versus language in current systems theories: Under these theories, a language comprehension system processes words and sentences, and transmits the products of comprehension to an entirely separate system for long term memory storage. A retrieval system later retrieves these stored memories for transmission to a language production system for verbally expressing the retrieved memory (see e.g., MacKay, James, Taylor, & Marian, 2007).

The Present Research: Studies 1-3

The present research followed the standard convention of describing differences between patient and controls that exceed two standard deviations (SDs) as deficits and characterizing indefinitely large differences (as can occur when a control group outperforms a patient with $SD = 0$) as 6 SD deficits. The research consisted of three studies. Studies 1-2 compared language production in H.M. and memory-normal controls matched as closely as possible with H.M. on five dimensions: age, education, IQ, background and native language. The task was to describe captioned cartoons so that a listener could understand what made them funny, with no constraints on description length or duration. This task can be considered naturalistic and ecologically valid (see Benuzzi, Nocetti, Baraldi & Nichelli, 2006, and Giora, 2003), involving implicit (but not explicit) production of coherent, easy-to-understand, and grammatical sentences, an important feature because H.M. may retain unconscious, implicit or off-line language abilities (see Knott & Marslen-Wilson, 2001), but not conscious, explicit or on-line language abilities (see Caplan & Waters, 2006, for an analogous phenomenon in aphasic patients).

Cartoon description requires two everyday skills: the ability to comprehend and appreciate the humor in cartoons, and the ability to effectively communicate that comprehension and appreciation. Reports in the literature suggest that H.M.’s ability to comprehend and appreciate humor is intact: According to Carlson (2004, p. 452; see also Kolb & Whishaw, 2003, p. 447), H.M. both comprehends and enjoys humor, laughing “endlessly … at the same jokes,
finding them fresh and new each time.” Studies 1-2 therefore addressed the second skill: Can H.M. describe captioned cartoons and communicate why they are funny as readily as controls? The only currently available data on H.M.’s ability to explain cartoon humor appears in Marslen-Wilson (1970). Marslen-Wilson showed H.M. (then age 44) a cartoon and asked him why it might be funny. We describe the cartoon in [1], and quote H.M.’s answer in [1a].

[1]. The cartoon: A distraught woman is saying, “The Pill. The Pill.” [like someone dying of thirst in the desert, crying “Water! Water!”] as she crawls out of a messy kitchen containing dirty laundry, unwashed pots and dishes, toys strewn over the floor, and five young children, one crying, one quizzical, and three squabbling with each other.

[1a]. H.M. (answering the question, “Why do you think that’s funny?”): “Well ... ... it’s a wonder of the .. uh .. the mother of course going out of the room .. but seeing “The Pill, the Pill” and all the .. like soap suds in a way that there’s been raised there ... she can’t do anything, however, she has to do everything .. [emphasis in the original] she ... both ways of looking at it ... as you could say .. because the pots and everything. (WM-W: Why .. What’s the ... why’s she saying “The Pill, The Pill”?) She isn’t saying “The Pill, The Pill” ... it’s the little girl that’s saying to the boy. (WM-W: Oh, I see, yes .... that’s right ... why’s she saying it to the little girl .. little boy?) Well ... to .. point out to the boy that that’s what it was that .. the little pill that the mother possibly had dropped in to make the soapsuds and .. and maybe .. she thought maybe well, it was more than one pill that she had put in, and that got .. that’s why she’d got so many.” (Dot strings indicate hesitations of varying lengths; from Marslen-Wilson, 1970)

Close inspection of [1a] indicates that H.M.’s description of [1] was in places incoherent, ungrammatical and difficult to understand, e.g., “it’s a wonder of the mother”. However, Marslen-Wilson’s (1970) methods were inadequate for addressing the intact language hypothesis. First, Marslen-Wilson ran no memory-normal controls for comparison with H.M. Second, this “pill cartoon” confounds H.M.’s language production with his memory deficits because oral contraception pills only emerged as a concept after his 1953 lesion. To overcome these problems, the present research incorporated memory-normal controls and tested humor description abilities for cartoons containing words and concepts familiar to H.M. prior to 1953.

Study 1 tested the Milner et al. (1968) intact language hypothesis: H.M. and six memory-normal controls described a set of cartoons and explained why they might be funny. Then judges blind to speaker identity rated a transcript of each description on global dimensions such as vagueness, comprehensibility, grammaticality, coherence, and repetitiveness. Under the intact language hypothesis, blind ratings on these dimensions will not differ for H.M. vs. the controls.

Study 2 tested the Skotko et al. (2005) hypothesis that H.M.’s language production is “without major errors.” To our knowledge, no prior study has distinguished between major versus minor speech errors: Studies of speech errors to date (e.g., Baars, 1980; Dell. 1990; Fromkin, 1971; Garnham, Shillcock, Brown, Mill & Cutler, 1982; MacKay, 1969, 1970, 1972, 1973, MacKay & James, 2004) have focused exclusively on minor errors, which by definition do not impair the communication process because the speaker’s original intent is obvious from the context or from error correction processes. By contrast, major errors impair the communication process because they are uncorrected and the speaker’s original intent is neither obvious nor readily determined. To determine the frequency of major errors for H.M. versus normal controls, Study 2 judges with the definition of major errors in hand and blind to speaker identity estimated the number of major errors in the cartoon descriptions from Study 1. Under the intact language
H.M.’s Language Production Errors

(Milner et al.) and no-major-errors (Skotko et al.) hypotheses, the estimated frequencies of major errors should not differ for H.M. versus the memory-normal controls.

To determine whether H.M.’s errors differ in nature from those of normal speakers, Study 3 conducted three relative frequency analyses that bear on the cognitive and neural mechanisms underlying language. For example, one analysis examined the relative frequency of three classical types of major and minor errors that together include about 90% of normal errors involving words and phrases (see Garnham et al., 1982, p. 254): omissions, additions, and substitutions (see Table 2 for definitions). If H.M.’s speech errors reflect the same processes or brain mechanisms as normal errors, these classical error types should not differ in relative frequency for H.M. vs. memory-normal controls.

Study 1: A Test of the Intact Language Hypothesis

In Study 1, H.M. and six controls matched for age, education, IQ, background and native language described three cartoons and what made them funny. Three listeners then transcribed the descriptions and three readers analyzed their content to determine whether H.M. and the controls were describing similar aspects of the cartoons. Finally, nine judges blind to speaker identity rated the descriptions on six dimensions: vagueness, comprehensibility, grammaticality, coherence, repetitiveness, and adequacy of the humor descriptions. Although some of these dimensions have been rated in other tasks, Study 1 is the first to rate H.M.’s output on all six dimensions in the same task. Under the intact language hypothesis, blind ratings of the cartoon descriptions should not differ for H.M. versus the controls on any of these dimensions.

Method

Participants

The participants were H.M. and six healthy memory-normal controls. All responded “no” or “I don’t remember” when asked post hoc about prior exposure to the present cartoons.

H.M. We tested H.M. in 1997 at age 71 when his most recent mean IQ score on the Verbal and Performance subtests of the Wechsler Bellevue I (W-B I) was 112. MRI data from about 1997 suggested (without data from same-age memory normal controls) “possible” and at most “minimal” damage to lateral temporal neocortex that was not due to H.M.’s original 1953 MTL ablation (Corkin, Amaral, González, Johnson, & Hyman, 1997). Follow-up MRI data from a decade later (i.e., long after the present study) discounted Alzheimer-related degeneration in H.M. relative to four memory-normal controls (unmatched with H.M. for IQ, education or background) but indicated vascular changes and cortical thinning of unknown etiology, unknown time of onset, and unknown relations to behavior (Salat et al., 2006).

Controls. The controls reported an absence of neurological problems and participated for $10/hr. We selected the controls from the combined participant pools (N=750) of the UCLA Cognition and Aging Laboratory and the Claremont Project on Memory and Aging to match H.M. as closely as possible on five factors: highest educational degree (high school diploma), background (unskilled or semi-skilled labor), mean age at time of test, mean IQ score on the Verbal and Performance subtests of the W-B I, and first language (English).

We tested the controls from 1999-2003 when their mean age was 71.00 (SD = 2.53) and their mean IQ score on the W-B I Verbal and Performance subtests was 119.42 (SD = 4.55). Neither mean age nor mean IQ differed reliably for the 6 controls versus H.M. Correlations between the controls’ data in Studies 1 and 2 and their Verbal and Performance IQ were also non-reliable, indicating that Verbal and Performance IQ cannot account for the present results.

Materials
The materials were three cartoons with captions containing words that entered English dictionaries no later than 1949. Cartoons 1 and 2 came from a New Yorker (1951) collection published before 1949, and Cartoon 3 was published more recently (see Figure 1, from Larson, 1993). However, the general scenarios depicted in all three cartoons would have been familiar to H.M. before his 1953 lesion. To ensure ease of perceptual processing, we photocopied each cartoon with enlargement to fill an 8” x 11” page.

Experimental Procedures

We presented the instructions verbally (with some aspects repeated mid-trial for H.M. but not the controls; see Appendices V-X) and visually on a continuously displayed card: Describe each cartoon and explain why it is supposed to be funny in sufficient detail that someone could understand the humor based solely on your description. We presented the cartoons in the order 1, 2, 3, and to help H.M. remember already described aspects of a cartoon, all participants received a marker for use if necessary to indicate already described aspects.

Transcription procedures

All trials were tape recorded for later transcription. We established the final transcript in three steps: Step 1). A primary listener first transcribed the tapes word-for-word, using the label (?) when confidence in her transcription of a particular word was less than 100%, and using the label “inaudible” when she could not decipher one or more words after repeated replay. Step 2). A second listener read the Step 1 transcript while listening to the tapes and either agreed with the transcription or suggested a change. Step 3). A third listener read conflicting (Step 1 versus 2) transcriptions and either chose one after listening to the relevant section of the tapes, or rejected them both and added the label “inaudible” in the final transcript.

Content analyses: Preliminary procedures

To prepare the transcripts for content analyses, we edited out errors and other irrelevant aspects of the cartoon descriptions in three stages: 1). We eliminated experimenter comments (e.g., “OK”, “Good” and “Mm hm”), participant pauses, extraneous comments (e.g., about a noticeable burp), interjections (e.g., “um” or “uh”), false starts, and other common dysfluencies; 2). We divided the cartoon descriptions into conceptually coherent segments containing a relatively small number of errors to allow close inspection, one error at a time. The boundaries between segments (up to 21 per cartoon description) always occurred at natural points in the descriptions (see Table 1 for the segmented descriptions of H.M. and a representative control participant); 3). We edited out or corrected the errors, with the results shown in Table 1.

Five pre-specified rules governed our choice of best possible correction (BPC) for errors in Study 1. Rule 1 trumped all the others: If a participant corrected an error either on their own or following an experimenter prompt, that correction was the BPC. Although some speech error studies have relied on less stringent “contextual criteria” to specify the BPC or intention underlying a speaker’s error (see Fromkin, 1971, 1973; and Garnham et al., 1982), Rule 1 has a long history in speech error studies, represents the “gold standard” for specifying a speaker’s intention following an error (see Meringer & Mayer, 1896), and is usually easy to apply because normal speakers spontaneously correct most of their substitutions (as in Put it on the table… I mean, chair; see Boomer & Laver, 1968) and omissions (as in, too South, I mean, too far South; from Garnham et al.), they readily produce a correction when listeners point out or otherwise react to uncorrected errors, and when asked directly, they effortlessly indicate their original intent (albeit sometimes begrudgingly due to the interruption; see e.g., Meringer, 1923).

However, Rule 1 was insufficient for present analyses because, as will be seen, H.M. fails to correct most of his errors, either on his own or when asked directly, even for errors that render
his utterances ungrammatical, incoherent, or difficult to understand. We therefore created four additional rules to overcome these problems. Rule 2 was a “best fit” rule that gave priority to whatever correction added the fewest words and retained the most words from what participants actually said. Rule 3 gave priority to whatever correction was most consistent with the cartoon and its caption and the prosody and content of the utterance context. When prosody seemed critical for selecting a BPC under Rule 3, a phonologist (M.P.) repeatedly listened to the relevant sections of the tapes and provided a detailed prosodic analysis blind to speaker identity. Rule 4 gave priority to the most coherent, grammatical and readily understood correction. Rule 5 (applicable only to H.M.) gave priority to the correction most consistent with H.M.’s use of words, prosody and syntax in prior studies (see, e.g., MacKay, Burke et al., 1998, and MacKay & James, 2001, for data on H.M.’s prosodic idiosyncrasies). Rule 5 often proved essential for resolving ties between the other rules (see Appendix Y, which illustrates how we applied Rules 1-5 to H.M.’s utterances) and for suggesting possible BPCs in the first place because most observers unfamiliar with the nature of H.M.’s errors can neither comprehend H.M.’s utterances nor suggest BPCs (see Appendices V-Y). Rule 5 might nonetheless be questioned because coders familiar with H.M.’s linguistic idiosyncrasies (here, the experimenter, L.J., and the senior author) might be open to hypothesis-linked coding bias. However, four factors argued against coding bias. First, no hypotheses guided our content analyses of the cartoon descriptions, the original goal of our BPC coding. Second, being blind to speaker identity and unaware of the hypotheses under test, the raters in Studies 1 and 2 were immune to hypothesis-linked bias. Third, because we developed the hypotheses for Study 3 many months after establishing the BPCs in Table 1, Study 3 results were immune to hypothesis-linked bias. Fourth, hypothesis-linked coding bias would have called for different BPC rules. For example, eliminating “add as few words as possible” from (best fit) Rule 2 would have favored the main prediction in Study 3 (relatively more omissions and fewer additions for H.M. than controls), as in [5ab], where modified Rule 2 now favors [5a] over the originally selected BPC for [5], namely [5b], which yields fewer omissions and more additions than [5a].

In short, we believe that our transcription and scoring procedures were unbiased with respect to the hypotheses tested in Studies 1-3, and that no other reasonable rules or procedures would change our results.

[5]. H.M. (responding to the experimenter question, “What do you make of that?”): “And that’s um, mixed up in a way, because one person is talking, and (pause) and one person is, is uh, saying what the- another person is. Causing a lot of trouble, each crisis and everything” (Study 3; from Appendix X)

[5a]. one person is talking, and that one person is saying that another person is causing a lot of trouble and crises (possible correction of [5]; added words are in italics).

[5b]. one person is talking, and is saying that another person is causing a lot of trouble or crises (BPC for [5]; added words are in italics).

Content analysis procedures

The edited transcripts for H.M. and the six memory-normal controls were analyzed for content in three pre-planned steps: Step 1). A primary reader applied three possible labels to each segment in the transcript: humor-irrelevant versus humor-relevant (i.e., it described why the cartoon was supposed to be funny, however inadequately), context-irrelevant versus context-relevant (i.e., it described the physical surroundings of protagonists in the cartoons, however inadequately), and “neither” (discussed shortly) versus “indeterminate” (if she was unsure). Step 2). A secondary reader read the labeled segments and either agreed with the primary reader’s label or suggested an alternate label. Step 3). A third reader examined conflicting (primary
versus secondary) labels, and after listening to the relevant section of the tape, chose between the two or indicated “indeterminate” as the final content category.

**Rating Procedures**

Nine judges rated the unedited cartoon descriptions of H.M. and the controls on six dimensions: vagueness, repetitiveness, comprehensibility, grammaticality, coherence, and adequacy of the humor explanations. The judges were graduate or undergraduate students at the University of Colorado, Colorado Springs (3 males, 6 females; mean age 25, range = 21-39) who were unaware that amnesic H.M. was one of the speakers or participants in the study.

Each judge received an initial instruction page followed by three rating booklets, one for each cartoon. Each rating booklet began with a full page copy of a cartoon, followed by seven separate pages, one for each rating dimension. The rating pages for each dimension provided a five point rating scale with the labels 0 (not at all), 2 (somewhat) and 4 (extremely), plus example ratings for statements unlike the target descriptions in meaning. The judges used these examples whenever necessary to anchor the low and high ends of the scales.

Under each rating scale were the seven randomly ordered cartoon descriptions of the participants, labeled Speaker 1, 2 etc. to mask speaker identity, with different speaker-linked numbers across the three cartoons so that judges could not tag any particular speaker number as unusual. To render the mean number of words per speaker approximately equal for H.M. and the controls, we divided H.M.’s descriptions into sections at natural points (see Appendices V-X) and averaged his final ratings across sections.

The instruction page asked the judges to examine the cartoon and its caption and read the responses of seven different participants in a prior experiment where the task was to describe the cartoon and explain why it was funny. The judges were to focus exclusively on a single rating dimension for each section, simultaneously comparing all sections on that rating dimension while ignoring all other features such as description length, utterance completeness and accuracy, corrected errors, stutters, and intrusions such as "um", "uh", "no" or "I mean." The instructions explained why some descriptions were divided into sections, and asked the judges to imagine, if possible, a word that could replace each word labeled with a question mark (indicating “difficult to hear”) and yield a better sentence. The judges then rated the seven descriptions on the dimensions shown on subsequent pages of the rating booklets.

**Results**

**Transcription Results**

The final (complete and unedited) transcripts contained two words labeled “difficult to hear” (see Appendix X), with one “inaudible” label for a control participant but none for H.M. It is noteworthy that the experimenter frequently asked H.M. to clarify difficult to understand phrases but her requests received no response (see Appendices V-X). For example, when asked what he meant by the phrase “double correction” in [6], H.M. simply repeated this phrase without elaboration in [7].

[6]. H.M.: “And in, and- he’s making double correction.” (from Appendix V)
[7]. H.M. (responding to the experimenter’s request to clarify “double correction”): “He’s making a double correction” (from Appendix V)

**Task-relevant Content Results**

Task-relevant content was either context-relevant or humor-relevant because no segment in any cartoon received the content label indeterminate for H.M. or the controls. To control for H.M.’s larger number of segments relative to the controls (see Table 1), our analyses compared mean percentages of segments labeled context-relevant or humor-relevant per participant.
Context-relevant content described the physical surroundings of protagonists in a cartoon. Context-relevant examples from H.M.'s descriptions included: the location of a window in a wall behind a protagonist, the relative location of a pen and notebook on the protagonist’s desk, the length of the extension cord for the telephone receiver in the protagonist’s hand, the way the artist drew the windows on skyscrapers visible through a window behind the protagonist (see Appendix W), a picture or window behind the main protagonist (see [8]), the kind of writing instrument used by a minor protagonist (see Appendix W), and contents of a picture on the wall (see Appendix X; also Figure 1). Overall, 57% of H.M.’s segments were context-relevant versus a mean of 48% for the controls ($SD = 13\%$), a non-reliable 0.69 $SD$ difference.

[8]. H.M.): “And, uh, that picture they've got in the background there, it's, um, just a picture. (EXP: Mmhm.) But they're um, a business, in a way. In the area. Or in… maybe a distant view, because the mountain area in the back.” (here H.M.’s “mountain” reflects misinterpretation of a business performance chart central to Cartoon 2; see Appendix W)

Humor-relevant content (as in [9]) indicated why a cartoon was supposed to be funny (no matter how inadequately). Overall, 4% of H.M.’s segments were humor-relevant versus a mean of 21% for the controls ($SD = 11\%$), a non-reliable 1.55 $SD$ difference. Extending this result, a separate analysis indicated that 6% of task-relevant segments were humor-relevant for H.M. versus a mean of 29% for the controls ($SD = 15\%$), a non-reliable 1.53 $SD$ difference.

[9]. H.M.: “He's making a double correction. Because, 'it's never good for you,' means that, that never been good for the person he's talking- person he's talking to.” (humor-relevant content; from Appendix V)

Task-irrelevant Content Results

Task-irrelevant content fell into six categories discussed next: speculations about details not shown in a cartoon, complaints about cartoons, references to the creator of a cartoon, references to what a protagonist was not doing in a cartoon, meta-comments about task difficulty, and self-referential expressions of puzzlement.

Speculation about unobservable details. Example [10] illustrates task-irrelevant speculations about Cartoon 2: After misinterpreting a business performance chart as a window, H.M. speculated that the window enabled inspection of “black stuff” that might be crawling up the unseen exterior wall of the room. Example [11] illustrates task-irrelevant speculations about a “blackening rule” that the creator of Figure 1 (Cartoon 3) might have used. Other Cartoon 3 speculations included the possible color of a protagonist’s shoes, whether a protagonist was “afraid of falling off her chair” or wanted “to have it her way,” and whether “no one will talk to” one of the protagonists or “give her a hand” (see Appendix X). Overall, 11% of H.M.’s segments contained task-irrelevant speculations versus a mean of 0% for the controls ($SD = 0\%$), a reliable difference in excess of 6.00 $SD$s.

[10]. H.M. (the slanted window refers to a business performance chart): “Looking out a window in uh, place that they have it slanted that way, or is people supposed ... to be shaped f- in any- black stuff crawled under there. (EXP: That's true.) On the wall.” (example illustrating task-irrelevant speculation; from Appendix W)

[11]. H.M.: “And you can tell that that’s a… the base of the… chair that she just bl- the-blackens the whole way, and everything, because that was one of the rules.” (example illustrating task-irrelevant speculation; from Appendix X)

Complaints about cartoon art, depictions or captions. H.M.’s task-irrelevant complaints included “drawn wrong” (see [12]) and “mixed up in a way.” He also complained that a protagonist in Cartoon 3 was misnamed (see Appendix X), and that the caption for Cartoon 2 contained a...
spelling error (which it did not; see Appendix W). Excluding semi-relevant complaints about how un-amusing a cartoon was, 7% of H.M.’s segments contained task-irrelevant complaints versus a mean of 1% for the controls ($SD = 2\%$), a reliable $3.00$ $SD$ difference.

   [12]. H.M.: “And this one, is- is drawn wrong, because that woman sittin’ there, you see the top of the chair there. And not in- her hair also. And you can’t tell just what her hair color is.” (task-irrelevant complaints about Cartoon 3; from Appendix X)

References to the artist’s signature. Example [13] illustrates a task-irrelevant reference to the signature of a cartoon’s creator. Overall, 4% of H.M.’s segments contained task-irrelevant references to the artist’s signature versus a mean of 0% for the controls ($SD = 0\%$), a reliable difference in excess of 6.00 $SD$s.

   [13]. H.M.: “And he has a, a couple(?) down there, and two, and then of course the signature of the person who k- made the cartoon.” (task-irrelevant signature reference; see Appendix V)

References to what protagonists were not doing. Example [14] illustrates task-irrelevant description of what the Cartoon 1 protagonist was not doing. Overall, 2% of H.M.’s segments contained task-irrelevant references to non-behavior versus a mean of 0% for the controls ($SD = 0\%$), a reliable difference in excess of 6.00 $SD$s.

   [14]. H.M.: “And he’s not, looking out the window (burps), excuse me, (EXP: Mmhm.) or anything like that, he’s just looking at- at his date book” (task-irrelevant reference to what a protagonist was not doing; from Appendix V).

Expressions of puzzlement. Examples [15]-[16] illustrate task-irrelevant expressions of puzzlement. Overall, 13% of H.M.’s segments contained such self-referential expressions versus a mean of 1% for the controls ($SD = 2\%$), a reliable 6.00 $SD$ difference.

   [15]. H.M.: “I wonder about that” (self-referential expression of puzzlement; see Appendix X)
   [16]. H.M.: “you can’t tell if she is the boss …. or just what she is.” (self-referential expression of puzzlement; from Appendix W)

Meta-comments about task difficulty. Examples [17]-[18] illustrate meta-comments about task difficulty, e.g., “you can’t tell…”; “can’t see…”; and “can’t read…”. Overall, 13% of H.M.’s segments contained task-irrelevant meta-comments about task difficulty versus a mean of 3% for the controls ($SD = 3\%$), a reliable 3.33 $SD$ difference.

   [17]. H.M.: “you can’t tell if that is uh, ghosts of what someone- some some, because it’s white over here, and the- don’t have any eyes on it.” (meta-comment about task difficulty; see Appendix X)
   [18]. H.M.: “Can’t read the name, it starts with an “O.”” (task difficulty meta-comment; see Appendix X)

Rating Results

We report mean ratings of the nine judges separately for the six rating dimensions: humor description (how adequately participants described what made a cartoon funny), vagueness, comprehensibility, grammaticality, coherence, and repetitiveness.

Humor description dimension. Categories for rating the adequacy of a humor description ranged from 0 (not at all adequate), to 2 (somewhat adequate), to 4 (extremely adequate). The mean rating for humor description adequacy was $1.86$ per section for the controls ($SD = 0.51$) versus $0.11$ for H.M., a reliable $3.43$ $SD$ deficit. Reinforcing this result, 90% of H.M.’s humor description ratings were “not at all adequate” versus a mean of 31% for the controls ($SD = 12\%$), a reliable $4.92$ $SD$ deficit.
Vagueness dimension. Categories for rating vagueness of the descriptions ranged from 0 (not at all vague), to 2 (somewhat vague), to 4 (extremely vague). The mean vagueness rating was 1.45 per section for the controls ($SD = 0.25$) versus 3.04 for H.M., a reliable 6.36 $SD$ deficit. Reinforcing this result, the mean percentage of ratings in the "extremely vague" category was 36% for H.M. versus 8% for the controls ($SD = 4\%$), a reliable 7.00 $SD$ difference.

Comprehensibility dimension. Comprehensibility ratings ranged from 0 (not at all difficult to comprehend), to 2 (somewhat difficult), to 4 (extremely difficult). The mean comprehensibility rating was 1.02 per section for the controls ($SD = 0.45$) vs. 3.11 for H.M., a reliable 4.64 $SD$ deficit. Reinforcing this result, the mean percentage of "not at all difficult to comprehend" ratings was 42% for the controls ($SD = 17\%$) vs. 3% for H.M., a reliable 2.29 $SD$ deficit.

Grammaticality dimension. Grammaticality ratings ranged from 0 (not at all ungrammatical), to 2 (somewhat ungrammatical), to 4 (extremely ungrammatical). The mean ungrammaticality rating was 1.38 per section for the controls ($SD = 0.33$) versus 2.88 for H.M., a reliable 4.55 $SD$ deficit. Reinforcing this result, the mean percentage of "not at all ungrammatical" ratings was 31% for the controls ($SD = 10\%$) versus 0% for H.M., a reliable 3.10 $SD$ deficit.

Coherence dimension. Coherence ratings ranged from 0 (not at all incoherent), to 2 (somewhat), to 4 (extremely). The mean coherence rating was 1.09 per section for the controls ($SD = 0.32$) vs. 3.17 for H.M., a reliable 6.50 $SD$ deficit. Reinforcing this result, the mean percentage of "not at all incoherent" ratings was 43% for the controls ($SD = 16\%$) vs. 3% for H.M., a reliable 2.50 $SD$ deficit.

Repetitiveness dimension. Repetitiveness ratings ranged from 0 (not at all repetitive), to 2 (somewhat), to 4 (extremely). The mean repetitiveness rating was 1.41 per section for the controls ($SD = 0.49$) versus 1.93 for H.M., a non-reliable 1.06 $SD$ difference. Reinforcing this result, the mean percentage of "not at all repetitive" ratings was 27% for the controls ($SD = 13\%$) versus 12% for H.M., a non-reliable 1.15 $SD$ deficit.

Discussion

Content analysis results: Task-relevant content

Descriptions of H.M. and the controls did not differ reliably in context-relevant (physical surroundings of the cartoon protagonists) or humor-relevant (why a cartoon was supposed to be funny) content. These data carry two implications: that H.M. was motivated and able to follow the instructions; and that Study 1 rating results did not reflect differences in task-relevant content between H.M. versus the controls.

Nonetheless, H.M. often tried to avoid the humor-description and caption-reading tasks, as in [19] and [20]. Moreover, when H.M. stopped avoiding and did try to explain why the cartoon was funny in [20], he was especially vague, incoherent and difficult to understand, and produced numerous false starts, dysfluencies, minor errors (word repetitions) and major errors, e.g., the inaccurate claim that the “other person” on the phone said “it was never good” when in reality, this person said nothing (see Appendix V).

[19]. H.M. (responding to, “Why is that supposed to be a funny cartoon?” in Appendix V): “And, this is, um, be- naturally the building’s behind him, and he's talking on the, to somebody, the gentleman who wrote him. (EXP: Mmhm.) H.M.: And he’s not, looking out the window (burps), excuse me, (EXP: Mmhm ..) or anything like that, he’s just looking at- at his date book. (EXP: Yeah.) And he has a, a couple(?) down there, and two, and then of course the signature of the person who k- made the cartoon.”.

[20]. H.M. (responding to the experimenter’s request to read what it says at the bottom): “Yeah… uh, and so, just what he said, over the telephone. To the person he's talking to ...
And in, and— he’s making double correction. (EXP: He’s doing what?) He’s making a
double correction. Because, “it’s never good for you,” means that, that never been good
for the person he’s talking— person he’s talking to. (EXP: Mhm.) And, he has stated
something, he stated it about, person who’s out, and he’s just statin’ it, the other person
always ma— said that, uh… (EXP: The other person always what?) Said it was never
good. And he’s just repeatin’ something.” (from Appendix V).

Content analysis results: Task-irrelevant content

Descriptions of H.M. and the controls differed reliably on six task-irrelevant dimensions:
H.M., significantly more so than the controls, speculated about details not shown in the cartoons,
complained about cartoon captions or depictions, discussed the signature of the artist and what
the cartoon protagonists were not doing, expressed puzzlement, and produced meta-comments
about task difficulty. Under one hypothesis, H.M. raised these task-irrelevant issues as a way to
avoid the more difficult task of describing what made the cartoons funny. Consistent with this
hypothesis, 90% of the judges rated H.M.’s humor descriptions as “not at all adequate” and H.M.
avoided requests to explain what made the cartoons funny or to read the cartoon captions aloud.

Also consistent with avoidance of a difficult task were H.M.’s meta-comments,
complaints about the cartoons, and expressions of puzzlement. H.M. may have been puzzled
because he did not fully comprehend the cartoon captions, because he did not recognize some of
the objects in the cartoons, and because he did not comprehend relationships between the
depicted protagonists. Consistent with non-comprehension of the cartoon captions, H.M.
produced major errors when asked to read captions aloud: Someone who insists that a caption
about morale (The beatings will continue until morale improves) is really about “morality” (see
Appendix W) cannot be said to comprehend the caption. Consistent with failure to recognize
objects depicted in the cartoons, H.M. misconstrued a performance chart central to a business
meeting as an outwardly slanting window affording a mountain view. Consistent with failure to
comprehend relations between the protagonists depicted in Figure 1, H.M. also misconstrued the
speaker to be discussing some third person rather than Edith, the person addressed.

Rating results

Judges blind to speaker identity rated the cartoon descriptions as reliably more vague and
reliably less grammatical, coherent, comprehensible and adequate in humor description for H.M.
than the controls, findings that replicated and extended earlier results at age 44 and 72.5. In
MacKay, Burke, et al. (1998), judges blind to speaker identity rated answers to conversational
questions concerning common childhood experiences (e.g., kindergarten) as reliably less
grammatical, less coherent, and less comprehensible for H.M. (age 44) than memory-normal
controls closely matched for age, IQ, and education. In MacKay et al. (2007), judges blind to
speaker identity rated descriptions of the two meanings of familiar ambiguous words, e.g., lots,
and phrases e.g., run out of, as reliably less grammatical, less coherent, and less comprehensible
for H.M. (age 72.5) than closely matched memory-normal controls.

However, two seemingly identical dimensions (redundancy versus repetitiveness) yielded
inconsistent rating results in MacKay et al. (2007) vs. the present study. Study 2 will address
these inconsistent results (reliably higher redundancy ratings for H.M. than the controls in
MacKay et al. but not for repetitiveness ratings in Study 1).

Study 2: The No-major Errors Hypothesis

Study 2 re-examined the hypothesis that H.M.’s language production is “without major
errors” (Skotko et al., 2005, p. 406-407). Four factors motivated this re-examination: 1).
problems with the Skotko et al. evidence for the no-major-errors hypothesis; 2). H.M.’s language

**Factor 1: Problematic Evidence favoring the No-major-errors Hypothesis**

One reason to re-examine the no-major-errors hypothesis is the inconclusive evidence in its favor. To support the no-major errors hypothesis, Skotko et al. (2005) simply cited error-free examples during 5-6 hours of conversational interviews with H.M. However, a selective focus on examples favoring the no-major-errors hypothesis is problematic because science can only progress as an empirical enterprise by seeking counterexamples and analyzing them in detail (see e.g., Popper, 1959; see MacKay et al., 2008, for other procedural flaws in Skotko et al., e.g., statistical errors, an inappropriate control group, inadequate control procedures).

**Factor 2: Extensive Deficits in H.M.’s Language Production**

A second reason to re-examine the no-major-errors hypothesis is that H.M. has exhibited language production deficits from 1970-1999 on the Test of Language Competence (MacKay et al., 2008), on the Reporters Test of Language Production (Corkin, 1984), and in 12 experiments on the spoken production of novel words, phrases and sentences, with more than 25 statistically reliable deficits relative to carefully matched controls in one experiment alone (MacKay, Burke et al., 1998; MacKay & James, 2001, 2002; MacKay et al., 2007; MacKay, Stewart et al., 1998, and MacKay et al., 2008). Given this extensive evidence for production deficits, it would be remarkable if H.M. produced fewer major errors than carefully matched controls. However, this possibility warrants test because many secondary sources continue to ignore H.M.’s production deficits and to describe H.M.’s language production as “sophisticated” (e.g., Kolb & Whishaw, 2003, p. 500) and “intact” (e.g., Eichenbaum, 2008, p. 277; Klein & Thorne, 2007, p. 522).

**Factor 3: Unusual Aspects of H.M.’s Utterances in Skotko et al. (2005)**


**Abnormal self-repetition.** H.M.’s answer to the question, “Would you like to tell us anything?” in Skotko et al. (2005, Appendix B) was that research on him will help others. However, H.M. has produced this same “help others” answer to many, quite different questions in the past: Would you like to tell us anything?; Are you happy?; How are you feeling?; Where do you think you are?; What aspect of remembering are you wondering about?; and Don’t you remember things quite well before the operation? (see [21]-[25]). Repeating the same response may be normal for the same question but seems abnormal for quite different questions.

[21]. H.M. (responding to the question, “Are you happy?”): “Yes, well, the way I figure it is, they- what they find out about me helps them to help other people.” (response selected for broadcast on NPR; from Newhouse, 2007)

[22]. H.M. (responding to the question “do you know where we are, where you are at the moment?”): “……I think of Lahey right off… [W.M-W: Lahey? You think we’re in the Lahey clinic?] “And… I have….yes, that’s what I thought of right off…and then I have a mental reservation right there……because I’m not sure, can’t put my finger definitely on it… but I know it’s some hospital .. and .. whatever’s being done …. well .. to me, is good because it will not alone help me, helps others, and that’s the important thing ………………… “ (from Marslen-Wilson, 1970)

[23]. H.M. (responding to the question, “how are you, how are you feeling? … too many questions?): “…. And ….. I wasn’t thinking just ... uh ... me giving the answers or asking the questions or anything like that .......... uh ..... I was thinking of other ...... other
people.” [W.M-W.: Yes ... how do you mean, ‘other people’?] “Because possibly it can help you to help others.” (from Marslen-Wilson, 1970)

[24]. H.M.: “...Gee...well...that's one thing right there .. you got that .. because a .. but the .. the remembering part is what I'm...I've wondered and wondered and wondered....” [W.M-W.:About ..? What .. wondered about what?] “Well, me .. just wondered just .. uh ... well .. me and alone in a way but I know in a way .. being ... at MIT, that something is being learned ..” [W.M-W: Oh yes.] “. And what is being learned ... is going to help plenty of others.” (from Marslen-Wilson, 1970)

[25]. H.M. (responding to the question, “You remember things from before the operation ... quite well, don’t you?”): “Yes ... before that, yes (said in a whisper, almost in tears) ... I do remember them .... in a way...was that .. uh .. was ... the ... worrying part .... and ..... always the wondering...myself of...just how things are...and...well, not just the effect on me but others......” [W.M-W.: How do you mean?] “Well ... how it affects me, affects others ... two ways ..... my parents, affects them naturally .... either way in a way as you could say ... but the medical profession too ... would be affected because ... they ... would ... uh ... find ... probably that they maybe had done something a little different ...... and .. just the correction they could make ... themselves will help others.” (from Marslen-Wilson, 1970)

**Topic shifts.** H.M.’s topic shifts in [22]-[25] are also noteworthy. H.M. shifted the topic from “Where are we?” to helping others in [22], from “How do you feel?” to helping others in [23], from “What aspect of remembering are you wondering about?” to helping others in [24], and from his long term memory to helping others in [25]. Study 2 will investigate two possible accounts of such topic shifts: artfulness in conversational control vs. inability to coherently discuss unfamiliar topics not discussed repeatedly in the past.

**Non-answers.** Examples [21], [23] and [25] illustrate non-answers to experimenter questions. In [21], H.M. suggested that research discoveries will help other people without answering the experimenter’s question: Are you happy? Similarly in [23], H.M. discussed what he was and wasn’t thinking about without answering the experimenter’s question: How are you feeling? Likewise in [27], H.M. discussed becoming less afraid of the nuns around the school without answering the experimenter’s question: Why did the lay teacher take over the class? Perhaps H.M.’s non-answers are normal, artful and “sophisticated” (Kolb & Whishaw, 2003, p. 500), reflecting a skill common among politicians. Or perhaps H.M.’s non-answers, false starts, and topic shifts reflect a way of hiding his communication difficulties, an issue addressed in Study 2.

**Factor 4: Major Errors in H.M.’s Utterances in Skotko et al. (2005)**

A fourth reason to re-examine the no-major-errors hypothesis concerns previously un-noted errors that H.M. produced in Skotko et al. (2005). The next sections argue that these errors can be considered major because they were uncorrected and impair three dimensions essential to effective communication: utterance comprehensibility, coherence, and grammaticality.

**Errors yielding ungrammatical utterances.** Skotko et al. (2005, p. 401) specifically cited [26] as evidence for the intact language hypothesis, but [26] is in fact ungrammatical and contradicts this hypothesis because it contains a pronoun (it) with no discernable referent anywhere in the conversational context (five speaker turns that ended with the interviewer comment, “Not everyone gets to be famous, sir, but you are!”). Only if modified as in [26a], [26b], or [26-c] does [26] become grammatical and readily understood.

[26]. H.M.: “Well, you come to a realization that maybe you are because it helps others.”
Was the referential indeterminacy involving the pronoun *it* in [26] a momentary lapse attributable to H.M.’s age (74)? Contrary to this hypothesis, H.M. also produced large numbers of difficult-to-discern referents at age 44 (see [27], where the label [REFERENT?] follows words with difficult-to-discern referents).

[26a]. (possible correction of [26]) “Well, you come to a realization that maybe you are because your fame helps others” (words in italics were substituted for *it* in [26]).

[26b]. (possible correction of [26]) ‘Well, you come to a realization that maybe you are because your being famous helps others” (words in italics were substituted for *it* in [26]).

[26c]. (possible correction of [26]) “Well, you come to a realization that maybe you are because you are helping others.” (words in italics were substituted for *it* in [26]).

Errors yielding incoherent utterances. Example [28] illustrates an error that impairs conceptual coherence because the verb *be* is incoherent with the conversational context and with the remainder of [28]. Perhaps H.M. had intended some sort of continuation, but a coherent and grammatical continuation of [28] is difficult to imagine, and the Skotko et al. (2005) transcript contained no indication that [28] was incomplete. Moreover, H.M. has elsewhere produced many similar errors involving *to be* that rendered his utterances incoherent, ungrammatical and/or difficult to understand, e.g., H.M.’s “more eased” instead of *be more at ease* in [27].

[28]. H.M. (replying to the question, “Would you like to tell us anything?”): “What’s found out about me will help others be.” (from Skotko et al., 2005, Appendix B)

Errors yielding difficult-to-understand utterances. Example [29] contains omission errors that reduced H.M.’s comprehensibility, coherence and grammaticality in Skotko et al. (2005) and call for elaborations to address issues such as; What did H.M. say “No” to?; Who does “they” refer to?; How do *they* relate to the small ones?; What do the “small ones” refer to (lesions?; petit mal seizures?)?; and How do the small ones relate to someone learning things that would help others? Researchers familiar with H.M.’s background and speech patterns might be able to imagine plausible answers to these questions and suggest that H.M. was trying to say something like [29a]. However, our ability to correct H.M. does not indicate that H.M. can do so, and in fact, H.M. corrected none of his errors in Skotko et al. (2005) and he consistently fails to correct his errors in many domains (see MacKay, 2010).

[29]. H.M. (replying to a question about his alleged intention to become a doctor): “Yeah, and I said, “No,” because if I have one of these small ones <<points to head>> but they can learn more about me and others. And would help others.” (from Skotko et al., 2005, Appendix C)

[29a]. (possible correction of [29]): H.M.: “I said, “No” to that desire because if I have one of these small seizures <<points to head>> I can become incapacitated. As it is, other doctors can learn more by studying me. And what they learn can help others” (hypothesized elaborations in italics).
In summary, H.M. produced errors in Skotko et al. (2005) that can be considered major because they were uncorrected and rendered his utterances ungrammatical, incoherent, and difficult-to-understand. However, normal speakers also sometimes produce major errors that they fail to correct and the critical question is: How often does H.M. produce major and minor errors relative to closely matched controls?

To address this critical question, Study 2 adopted two procedures for estimating the frequency of major versus and errors in Study 1 cartoon descriptions: standard versus rating procedures. Most speech error studies (e.g., MacKay & James, 2004) have adopted standard procedures, where errors are classified into standard categories, such as word omissions or word transpositions, by comparing a speaker’s actual versus intended or corrected utterances. Frequencies of such standard error types are then computed for experimental versus control groups, as in Study 2, where the mean number of immediately repeated words was compared for similar length outputs of H.M. versus the controls in Study 1.

As our second procedure, judges blind to speaker identity received a definition of major errors and estimated their frequency in the transcripts for H.M. and the controls. These blind ratings enabled the first ever comparison between standard versus rating procedures in speech error studies (see the Discussion section). Predictions for both measures were identical under the intact language and no-major-errors hypotheses: no more major errors for H.M. than the memory-normal controls.

### Method

#### Rating procedures

The same judges as in Study 1 received the same rating booklets except for two changes: 1). Study 2 instructions called for a focus on major errors, defined as uncorrected errors that rendered the speaker’s utterance ungrammatical, incoherent, or difficult to comprehend; 2). The rating scale was labeled 0 (none whatsoever), 1 (1 major error), 2-3 (2-3 major errors), 4-5 (4-5 major errors), and 6+ (6 or more major errors).

#### Standard frequency procedures

Two new judges compared the unedited transcripts in Appendix V-X with the edited (corrected) transcripts in Table 1 and scored seven types of major errors and four types of minor errors by consensus, as discussed next.

**Major error types.** Major errors fell into seven categories: substitutions, omissions, additions, transpositions, reading errors, free associations, and accuracy errors (see Table 2 for definitions). Major substitutions were scored when speakers substituted one or more words without correction, rendering an utterance inaccurate, ungrammatical, incoherent, difficult to comprehend or some combination of these. To illustrate a major substitution resulting in inaccuracy, H.M. substituted “wrote” for “phoned” without correction in describing the ongoing telephone conversation in Cartoon 1 (see Appendix V). To illustrate a major substitution resulting in ungrammaticality, H.M. substituted “Is” for “Are” without correction in Are people supposed to see if any black insects crawled under there on the wall (corrected description), yielding the ungrammatical question: “Is people supposed…?” (see Appendix W).

Major omissions were scored when speakers omitted without correction one or more critical words e.g., major constituents such as the verb or subject of a sentence, thereby rendering their utterance inaccurate, ungrammatical, incoherent, difficult-to-comprehend or some combination of these, as in this uncorrected utterance from Appendix V: “Then of course the signature of the person who made the cartoon”. Major additions were scored when participants added one or more contiguous words without correction, rendering an utterance inaccurate,
H.M.’s Language Production Errors

ungrammatical, incoherent, difficult to comprehend or some combination of these. An example is H.M.’s added phrase, *one person*, in this uncorrected and difficult-to-comprehend utterance: “one person is talking, and one person is saying that another person is causing a lot of trouble.” Here H.M.’s second (added) “one person” implies inaccurately that two people in the cartoon are talking rather than one (see Figure 1). Major *free associations* were scored when speakers produced uncorrected words or phrases that were closely related to each other but not to the conceptual context, rendering the overall utterance inaccurate, ungrammatical, incoherent, difficult to comprehend or some combination of these. An example is the phrase *in her way* in [30a], which inaccurately describes Cartoon 3 (see Figure 1) and shares the phonological form *way* with H.M.’s prior “make it *her way*” but caries a different and contextually inappropriate meaning. By hypothesis, H.M. produced “They’re in her way” via free association with the prior phonological forms, *her way* and *only her way*, rendering [30a] inaccurate, incoherent, and difficult to comprehend. Moreover, H.M. often free associates from the word *way* to *someone’s way is the only way*, consistent with his abnormal self-repetitions discussed earlier. For example, in [30b], twenty-seven years prior to [30a], H.M. free associated from *way* to “his way was the way” in discussing *Martin Luther King*.

[30a]. H.M.: “And, uh, I can’t tell just what- she possibly wants to make it her way, only her way. They’re in her way.” (example illustrating major free association; from Appendix X)

[30b]. H.M. [responding to the question *Have you ever heard of anybody called Martin Luther King?*]: “Well, in a way that he .. well .. everything was, I guess .. we .. er .. better explain it .. the way…everything was OK for everyone else but .. er … just what he’s done, it’s got to be just right …. their .. they can do anything, it doesn’t make any difference, but what I do is right, that’s .. it ..” [W.M-W.: “I’m not…so what was he saying, what was he doing?”] “Well, in a way, he was just .. telling the people in a way that no matter they could think of things they wanted to and everything but .. er .. his way was the way.” (example illustrating a topic shift to “someone’s way was the only way,” from Marslen-Wilson, 1970).

Major *accuracy errors* were scored when speakers inaccurately described some aspect of a cartoon without correction, as in [31], where H.M. inaccurately described the young boy (ghost) in Figure 1 as an older man who couldn’t be in school or named *Billy*.

[31]. H.M. (after misreading *Billy* as “Gus” in example [32]): “Gus been having trouble in school. And that’s a very young kid. See, Billy. (EXP: Mmhm.) And in school ... And this is a- an older man.” (from Appendix X)

Major *reading errors* were scored when speakers misread a cartoon caption without correction, making it ungrammatical, difficult-to-understand, inaccurate or some combination of these, as in [32], where H.M. without correction misread *Billy* as “Gus” and omitted two instances of the auxillary verb *has*, rendering [32] ungrammatical and difficult-to-understand. Similarly in [33], H.M. omitted the *you* in *I tell you Edith it’s not easy raising the dead*, as if the main protagonist in Figure 1 was talking about someone named *Edith* rather than to *Edith*. Similarly in [34], H.M. misread *morale* as “morality”, changing what the cartoon is about (see Appendix W). Interestingly, H.M. subsequently noticed the word *morale* in this caption, but rather than correcting his reading error, insisted that the caption contained a misprint (see [35]), so that the cartoon dealt with *morality* rather than *morale*.

[32]. H.M. (misreading the caption: *Oh, I don’t know. Billy’s been having trouble in school and Sally’s always having some sort of crisis.*): “Oh. (slowly) I don’t know. Gus
(pronounced ‘Guzz’) having been having trouble in school and Sally having sh- always having some sort of crisis” (a major reading error that rendered a caption ungrammatical and difficult-to-understand; from Appendix X)

[33]. H.M. (misreading the caption, I tell you, Edith, it’s not easy raising the dead): “I tell Edith, it’s not… easy, the- raising the dead.” (a misreading that rendered a caption inaccurate relative to its cartoon; from Appendix X).

[34]. H.M. (reading the caption The beatings will continue until morale improves): “The beatings will continue until morality improves” (from Appendix W).

[35]. H.M.: “Well … she’s making a comment there that the beatings will continue … Until… morality improves. And then- she said, "moral improves." (EXP: Good.) Morality, in a way. (EXP: OK.) Instead of "I" it’s a "t". It should be a "t" (an illustration of erroneous correction; from Appendix W).

Major transpositions were scored when speakers produced words in transposed order without correction, rendering an utterance inaccurate or difficult to comprehend, as in [36], where H.M. transposed the words extension and cord in an intended utterance or BPC such as the telephone, which has got a long extension cord, rendering this uncorrected utterance inaccurate as well as difficult-to-comprehend (see Appendix V).

[36]. H.M. (describing Cartoon 1, which depicts a telephone with a long cord on the receiver or handset but not on the telephone itself): “Well… it’s a bye- about a person talking on the telephone, got a long cord on the extension.” (uncorrected transposition of the words extension and cord in the intention of BPC: the telephone which has a long extension cord; from Appendix V).

False starts and dysfluencies. False starts were scored when speakers began with one syntactic structure and then shifted to another, as in [37], where H.M. began with a passive (BPC: Half the wall is taken up by the window) and shifted to an active structure (BPC: The window is right in the middle of the wall). Dysfluencies involved use of “uh” and “um” as in “Um, you can't tell.”

[37]. H.M.: “Well half the, you could say, um, right in the middle of the wall” (false start; from Appendix V)

Minute error types. The scorers classified four types of minor errors by consensus: word repetitions, stuttering, corrected speech errors, and corrected reading errors (see Table 3 for definitions). Word repetitions were scored when speakers immediately repeated a word or phrase without otherwise altering the syntax or meaning of a sentence, as in [38], where H.M. immediately repeated the words these here are. Stutters were scored when speakers repeated initial phonological segments in a word without repeating the entire word, as in [39]. Corrected speech errors were scored when speakers corrected an error, as in [40]. Corrected reading errors were scored when speakers corrected one or more misread words, as in [41], where a control participant initially misread the word morale as moral, then corrected herself.

[38]. H.M.: “And he's, um, these here are, these here are writin’, and the note was up there.” (immediately repeated words are in italics; from Appendix V)

[39]. Control participant: “m- mixture.” (stuttering)

[40]. Control participant: “a doorway behind her, or… a graph behind her.” (corrected error in italics)

[41]. Control participant (reading the caption The beatings will continue until morale improves): “until moral, morale improves” (corrected reading error in italics)

Results

Major errors: Overall rating results
To estimate the mean number of major errors per section from the rating data, we entered 2.5 for the rating category 2-3 major errors, 4.5 for the rating category 4-5 major errors, and 6 for the rating category 6 or more major errors. Using this procedure, the mean number of major errors per section was 1.55 for the controls (SD = 0.63) versus 3.62 for H.M., a reliable 3.29 SD deficit. Reinforcing this result, 83% of the judges indicated 2.5 or fewer major errors per section for the controls (SD = 6%) versus only 50% for H.M., a reliable 5.50 SD deficit.

Major errors: Standard frequency analyses
Using standard frequency analyses, the overall mean number of major errors per section was 1.17 for the controls (SD = 0.89) vs. 28.82 for H.M., a reliable 31.07 SD deficit. By major error subtypes, the mean number of major substitutions per section was 0.22 for the controls (SD = 0.27) vs. 10.83 for H.M., a reliable 39.30 SD deficit; the mean number of major omissions per section was 0.67 for the controls (SD = 0.67) vs. 10.00 for H.M., a reliable 13.93 SD deficit; the mean number of major additions per section was 0.00 for the controls (SD = 0.00) vs. 2.83 for H.M., a reliable deficit in excess of 6.00 SDs; the mean number of major free associations per section was 0.00 for the controls (SD = 0.00) vs. 0.33 for H.M., a reliable deficit in excess of 6.00 SDs; the mean number of major inaccuracies per section was 0.17 for the controls (SD = 0.18) vs. 1.00 for H.M., a reliable 4.61 SD deficit; the mean number of major reading errors per section was 0.11 for the controls (SD = 0.17) vs. 1.33 for H.M., a reliable 7.18 SD deficit; the mean number of major transpositions per section was 0.00 for the controls (SD = 0.00) vs. 2.50 for H.M., a reliable deficit in excess of 6.00 SDs.

Minor errors, false starts and dysfluencies: Standard frequency analyses
The overall mean number of minor errors, false starts and dysfluencies per section was 5.06 for the controls (SD = 2.59) vs. 15.17 for H.M., a reliable 3.90 SD deficit. However, only false starts and word repetitions contributed to this deficit: The mean number of false starts per section was 1.44 for the controls (SD = 0.96) vs. 5.00 for H.M., a reliable 3.71 SD difference; and the mean number of word repetitions per section was 1.22 for the controls (SD = 0.91) vs. 4.83 for H.M., a reliable 3.97 SD difference. By contrast, the mean number of stutters per section was 0.28 for the controls (SD = 0.25) vs. 0.17 for H.M., a non-reliable 0.44 SD difference; the mean number of corrected speech errors was 0.33 per section for the controls (SD = 0.37) vs. 0.17 for H.M., a non-reliable 0.43 SD difference; the mean number of corrected reading errors was 0.11 per reading trial for the controls (SD = 0.17) vs. 0.00 for H.M., a non-reliable 0.65 SD difference; the mean number of dysfluencies (“uh”s and “um”s) per section was 1.61 for the controls (SD = 2.07) vs. 5.00 for H.M., a non-reliable 1.64 SD deficit.

Discussion
Overall major errors: Rating versus standard analysis results
Results of our rating and standard analysis procedures contradicted claims that H.M. may retain implicit language abilities and produces spoken discourse that is “sophisticated,” “intact” and “without major errors.” Judges blind to speaker identity rated cartoon descriptions as containing reliably more major errors (defined as uncorrected errors that reduced utterance comprehensibility, coherence, and grammaticality) for H.M. than controls closely matched for age, IQ, background and education. This result comports with Study 1 rating results for comprehensibility, grammaticality, and coherence and with the results of our standard error analyses: reliably more major errors per section for H.M. than the controls.

Moreover, our rating and standard analysis procedures yielded virtually identical overall means for the controls (1.17 versus 1.55 major errors), and this congruence rules out coding and scoring biases in our standard error analyses because our raters were blind to H.M.’s identity. This congruence also offsets four possible sources of error in our rating procedures: scaling issues (with "none whatsoever" as the low-end category for major errors, the raters may have...
avoided the high-end category, "6 or more major errors"), time constraints (in the brief time available, raters may have overlooked some of the errors in H.M.'s difficult-to-understand utterances), inexperience (unfamiliar with errors resembling H.M.'s, the raters may have underestimated their frequency), and ceiling effects (with "6 or more" as the high-end category, 6.00 was the maximum possible average score for major errors4). However, none of these sources of error applied to our standard error analyses, which also contributed a degree of precision not possible with rating procedures, e.g., 0.22 major substitutions, 0.67 major omissions, 0.00 major additions, 0.00 major free associations, 0.17 major accuracy errors, 0.11 major reading errors and 0.00 major transpositions per section for the controls.

**Major errors: Error type results**

H.M. produced reliably more major substitutions, major omissions, major additions, major reading errors and major free associations than controls in Study 2. These findings replicate and extend less systematic observations in five earlier studies that tested H.M.'s language production from age 44 to 72.5 using a wide variety of procedures. However, H.M.'s major errors were remarkably similar in the present and earlier studies (illustrated in Table 4, cross-classified by task and H.M.'s age at time of test, and numbered [42]-[49] for expository purposes). In [42] (see Table 4), H.M. omitted the word on without correction in his BPC or intended utterance, thereby rendering [42] ungrammatical, difficult-to-understand, and inaccurate: One’s “position” can be on a passenger line but cannot be a passenger line, a structure that equates the concept “position of employment” with “a passenger line” (see MacKay, Stewart et al., 1998). In [43], H.M. added the words “out,” “in,” and “a person,” additions that rendered [43] ungrammatical, incoherent and difficult to understand (see MacKay, Burke et al. (1998). Such major omissions and additions almost certainly contributed to H.M.'s grammaticality and comprehensibility deficits in Study 2 and these earlier studies.

H.M. also produced remarkably similar major reading errors in Study 2 and earlier studies, e.g., [44] (see Table 4), a uncorrected reading error in MacKay and James (2001), where the task was to read a set of isolated sentences aloud (see also MacKay, Burke et al., 1998, and MacKay, Stewart, et al., 1998). H.M.’s major free associations were likewise remarkably similar in Study 2 and these earlier studies, e.g., [45], where H.M.’s “most of all” is a free associative phrase phonologically related to his just produced “of all” but semantically unrelated to the topic of his sentence (being above them). Similarly in [46], H.M. inaccurately defined the word lots as “more” via free association with his immediately prior “many,” yielding the familiar phrase “many more” (MacKay et al., 2007). Then in [47], H.M. responded to the experimenter’s immediately subsequent request for another meaning for lots with several additional free associations that rendered [47] incoherent and difficult-to-comprehend relative to the dictionary definition that he clearly wanted to convey: Lots can refer to pieces of straw of various lengths that people use to choose between alternatives by chance.

In all, H.M. produced reliably more major free associations than memory-normal controls in five experiments (Study 2; MacKay, Burke et al., 1998; MacKay et al., 2007; MacKay et al., 2008; MacKay, Stewart, et al., 1998) and a naturalistic study (Marslen-Wilson, 1970), a remarkable pattern because major free associations have never been reported in the speech of normal speakers (see e.g., Garnham et al., 1982). Under a hypothesis proposed in MacKay, Burke, et al., H.M. produces uncorrected free associations when either the situation or his own self-produced output triggers familiar phrases stored in long term memory. This free association process enables H.M. to say something (however incomplete, inaccurate, ungrammatical, vague, incomprehensible or irrelevant), despite his lesion-induced inability to form complete and
coherent plans or internal representations of what he wants to say in the current task or situation. H.M.’s inability to clarify difficult-to-understand phrases that he himself produces, e.g., “double correction” (see examples [6] and [7]) is also consistent with the MacKay-Burke free association hypothesis, as are Study 2 rating results. If no preformed, complete, accurate, grammatical, coherent and task-relevant sentence plan guides H.M.’s utterances, then judges are likely to rate his utterances as more incomplete, inaccurate, ungrammatical, vague, irrelevant and incoherent than normal (preplanned) utterances.

However, H.M.’s incoherent and difficult-to-comprehend utterances might impress naïve listeners as artful and “sophisticated” (unlike raters blind to speaker identity). For example, listeners unaware that H.M. routinely shifts topics to “I’m helpful” (as in [22]-[25]) may misconstrue such non-sequiturs as artful and “sophisticated.” H.M.’s free associative topic shifts, e.g., from “Are you happy?” to “I’m helpful” in [21], may likewise have impressed NPR listeners who mistakenly believed that H.M. only had memory problems (see Newhouse, 2007).

Error type results: False starts and minor errors

H.M. produced reliably more false starts than memory-normal controls in Study 2, perhaps, like his non-answers and topic shifts, as a means of avoiding major errors. However, H.M.’s false starts probably had no effect on his grammaticality, comprehensibility and coherence ratings in Study 1 because: 1). Study 1 raters were instructed to ignore false starts; and 2). Like dysfluencies, false starts are routinely ignored due to their minor status and high frequency (see Lickley, 1996; Lickley & Bard, 1996; see also Erard, 2007, pp. 78-110).

H.M. produced reliably more minor word repetitions than memory-normal controls in Study 2, a result that extends earlier observations, e.g., [48] in Table 4, where H.M. repeated the phrase on the bus and the proposition it's crowded, and [49], where H.M. produced seven repetitions of the words bus, drive and it (or their variants) in a 17 word statement (see MacKay et al., 2008). These observations in turn comport with earlier observations that H.M. repeats cliché phrases and propositions more often than memory-normal controls, e.g., repeating the exact phrase in a way 174 times in Marslen-Wilson (1970) (for instances, see [50]; also Appendix V-X), and repeating variants of in a way, I thought of, right off, I wonder, you’d call it, I guess, and I have an argument with myself reliably more often than controls in MacKay, Burke et al. (1998; see also Skotko et al., 2004). Moreover, H.M. has also been known to spontaneously repeat the same complex sentence many times a day (Hilts, 1995, p. 136), to repeat whole stories nearly verbatim on different occasions (see Ogden & Corkin, 1991), and to deliberately repeat a string of unrelated words during a long (unplanned) interruption in a memory test (see Ogden & Corkin). Repetition clearly represents a general characteristic of H.M.’s behavior as well as a common type of minor word-level error in his speech.

[50]. H.M.: “Yes .... in a way, it was .... free .... uh ... it was open that(?) we we re able to move into it then .... because .... and .. uh ... the Capellas .... liked Daddy in a way, and ..... liked mother ........ they could do things ..... uh ..... doing things for them ..... when .... their relations came and visited them, they were able to use part of our place.” (repeated use of in a way; from Marslen-Wilson, 1970)

H.M.’s selective deficits and sparing

It is important to note that in general, H.M. exhibited selective rather than across-the-board deficits in Study 1 and 2. In Study 1, H.M. did not differ from the controls in humor-relevant or context-relevant content, or in the ratio of humor-relevant to task-relevant segments. Similarly, in Study 2, H.M. produced some types of errors reliably more often than controls, e.g., major omissions, but not others, e.g., stutters. Likewise in Study 2, H.M. failed to correct major
errors that rendered his utterances ungrammatical, incoherent or difficult-to-comprehend, but he corrected his minor speech and reading errors no less often than the controls. Although this non-difference may reflect a floor effect (with only 0.11 corrected errors per reading trial for the controls), it nonetheless contradicts the hypothesis that a low error correction criterion explains H.M.’s error correction deficits: Contrary to the present results, this hypothesis predicts a higher correction rate for errors that render utterances ungrammatical, incoherent or difficult-to-comprehend for H.M. and normal speakers alike (see MacKay, 1992).

H.M.’s selective deficits and sparing also extend to other aspects of sentence production, as shown in MacKay et al. (2007; see also MacKay, Burke et al., 1998; MacKay & James, 2001, 2002; MacKay, Stewart et al., 1998). The MacKay et al. task was adapted from a standardized language production test involving constrained picture description (Wiig & Secord, 1988): On each trial, H.M. and eight carefully matched controls attempted to create a single grammatical sentence that included two or three target words and accurately described a picture. The results indicated that H.M. described the pictures reliably less accurately and less completely, included reliably fewer target words, and produced more agreement rule violations, more non-sequiturs, more run-on sentences, and more incomplete sentences (e.g., lacking a subject or verb) than the controls. Examples of these problems can be seen in [51-52ab], where the target words and general descriptions for two stimulus pictures appear in [51] and [52]. H.M.’s “single-sentence” descriptions appear in [51a] and [52a], and those of a typical control participant appear in [51b] and [52b]. H.M.’s inaccurate, incomplete, ungrammatical, and incoherent utterances in [51a] and [52a] seem to reflect concatenation of familiar units in the complete absence of an overall sentence plan. However, H.M.’s production problems were selective, and his error-free production of frequently-used words, phrases and propositions, e.g., “it’s wrong”, “to be”, “the same way”, “some of that,” “some pie” and “I’ll have some” in [51a] and [52a], may underlie the common but mistaken impression that H.M.’s language production is “normal” or “intact.”

[51]. TARGET WORDS: although wrong

PICTURE STIMULUS: Scene: A clothing store with several racks of jogging suits. Protagonists: A male customer, a female customer, and a male clerk. Contrary to [51a], the clerk and customers are not dressed “the same way.”

Action: The male customer looks on as the female customer addresses the clerk while pointing at a suit that the clerk is holding.

[51a]. H.M. description: “Because it’s wrong for her to be and he’s dressed just as this that he’s dressed and the same way”. (target words are in italics)

[51b]. Control description: “Well, I think I’ll take that one although it looks wrong.” (target words are in italics)

[52]. TARGET WORDS: pie either have

PICTURE: Scene: A check out line at a cafeteria with desserts on display. Protagonists: A male customer, a female customer, and a female clerk (behind the counter).

Action: As the female customer looks on, the male customer points while addressing the clerk.

[52a]. H.M. description: “I want some of that pie either some pie and I’ll have some.” (target words are in italics)

[52b]. Control description: “Uh, there are two people getting pie, but there’s only one piece of blueberry pie left, and so, either one of them will have to have it.” (target words are in italics)
Study 3: Relative Frequency Analyses of Classical Error Types

Study 2 conducted absolute frequency analyses of H.M.’s major versus minor errors defined in terms of correction status (uncorrected versus usually corrected) and their effects on utterance comprehensibility, coherence, and grammaticality. By contrast, Study 3 conducted relative frequency analyses to determine whether errors differ in nature for H.M. versus controls, without regard to absolute error frequencies or the distinction between major versus minor errors. One Study 3 analysis examined the relative frequency of three classical error types to test the hypothesis that H.M. and normal speakers produce basically similar errors with the same relative frequency. The classical types were omissions, additions and substitutions of words and phrases (see Garnham et al., 1982), where substitutions include both paradigmatic substitutions, as in [53], and syntagmatic substitutions, which involve anticipation, perseveration or transposition of already produced or about-to-be produced words or phrases in a intended utterance or BPC, as in [54] (see MacKay, 1987, p. 59-120). All Study 2 errors fell into these three categories, with none in the remaining two categories in Garnham et al.: word blends (where words with the same meaning or pragmatic import in some context become blended at the phonological level, see, e.g., MacKay, 1972) and phrase blends (where phrases with the same meaning or pragmatic import in some context become blended at the sentential level; see e.g., MacKay, 1973).

[53]. “…American expression…” [intention or BPC, “Australian expression”]
(paradigmatic substitution error; from Garnham et al., 1982).

[54]. “threw the window through the clock…” [BPC: “threw the clock through the window”] (syntagmatic substitution error involving transposition; from Fromkin, 1973).

Achieving power sufficient to detect relative frequency differences between sentence-level omissions, additions, and substitutions required a group sample of 60 or more, an unachievable criterion for the control sample in Study 2 (N=5). We therefore replaced the Study 2 control sample with the London-Lund corpus, perhaps the largest published collection of speech errors in the spontaneous conversations of normal adults. Svartvik and Quirk (1980) originally recorded and transcribed the speakers’ conversations, and Garnham et al. (1982) carefully classified their errors to create the London-Lund corpus, which contained 73 sentence-level substitutions, omissions and additions. This was less than the 158 corresponding errors for H.M. in Study 2 but this absolute frequency difference was irrelevant for the relative frequency analyses in Study 3. Of more concern were the variable topics of discussion, conversational settings, and speakers in the London-Lund corpus, but such variability could only favor the null hypothesis in Study 3, thereby rendering reliable differences all the more remarkable.

Our second relative frequency analysis compared exceptions versus non-exceptions to the syntactic class regularity for H.M. versus controls. The syntactic class regularity applies to single-word substitution errors and occurs when the substituted and intended words belong to the same syntactic category, e.g., common noun substituted for common noun, proper noun substituted for proper noun, adjective substituted for adjective, verb substituted for verb, and preposition substituted for preposition. This regularity applies to both paradigmatic substitutions (as in [53], where the adjective American substituted the adjective Australian) and syntagmatic substitutions (as in [54], where the transposed words, clock and window, are both common nouns) and represents the most general phenomenon established to date in studies of normal speech errors. However, exceptions do sometimes occur, as in [55], where the speaker substituted a common noun (government) for a proper noun (Obama), and Study 3 tested the hypothesis that H.M. and controls will produce basically similar substitution errors, with no difference in the relative frequency of exceptions to the syntactic class regularity.
Senator John McCain (during an internationally televised political debate, Oct. 15, 2008): “Senator government … I mean, Senator Obama” (a substitution error that violates syntactic class regularity)

Our third relative frequency analysis examined two types of substitution errors: single-word substitutions as in [55], versus multi-word substitutions as in [56]-[57]. This third analysis tested the hypothesis that H.M. and controls produce basically similar substitution errors, with no difference in the relative frequency of single- versus multi-word substitutions.

[56]. “If you’ll meet him you’ll stick around…” [intention or BPC: “If you’ll stick around you’ll meet him”] (multi-word transposition error; from Fromkin, 1973).

[57]. “extinguish your seat belts…” [intention or BPC: “extinguish your cigarettes and fasten your seat belts”] (multi-word omission error; from Fromkin, 1973).

The three relative frequency analyses in Study 3 also addressed two contrasting accounts of H.M.’s error deficits in Study 2: general process vs. specific process hypotheses. Under general process hypotheses, H.M. produces more errors than controls due to extreme values for one or more general processes, say tolerance for errors. This and other general process hypotheses predicted different absolute error frequencies but the same relative frequencies for omissions vs. additions vs. substitutions, for exceptions to the syntactic class regularity, and for multi- vs. single-word substitutions for H.M. and the controls. By contrast, specific process hypotheses predicted reliably different relative frequencies for omissions vs. additions vs. substitutions, for exceptions to the syntactic class regularity, and for multi- versus single-word substitutions for H.M. versus controls because processes specific to H.M.’s MTL and cerebellar damage cause abnormal speech errors under specific process hypotheses. A link between speech errors and cerebellar lesions is especially plausible under specific process hypotheses because Ackermann, Mathiak, & Riecker (2007) and Ackermann, Mathiak & Ivry (2004) reported “agrammatic” errors involving substitution and omission of grammatical morphemes in at least two patients with cerebellar damage (for a recent review, see Ackermann, 2008).

Method

Participants

H.M. H.M.’s background information is outlined in Study 1. We analyzed all of H.M.’s omission, addition and substitution errors involving words and phrases in Appendix V-X regardless of error status (major versus minor) or follow-up (corrected versus uncorrected). Counted as substitution errors were paradigmatic errors involving substitution of a word from outside an intended utterance or BPC (as in [55]), and three types of syntagmatic errors: anticipations involving the substitution of an upcoming word or phrase in an intended utterance or BPC, perseverations involving the substitution of an already produced word or phrase in an intended utterance or BPC, and transpositions involving mutual substitution of two words or phrases in an intended utterance or BPC (as in [36]). We ignored all other error types in Study 3.

Controls. Our controls were the normal adults in Svartvik and Quirk (1980). Their errors were classified as omissions, additions or substitutions by consensus of five expert scorers in Garnham et al. (1982) who were careful to exclude ambiguous examples and non-errors. Because neither H.M. nor the controls produced tag questions in Study 2, we excluded London-Lund errors involving tag questions (N=5) to ensure comparability with H.M.’s corpus, yielding an overall control corpus of 68 word and phrase omissions, additions and substitutions.

Procedures: Relative frequency analyses by error type

We conducted three relative frequency analyses by error type. One was an overall analysis comparing the relative frequency (in %) of substitutions versus omissions versus...
additions produced by H.M. versus the controls. Our second relative frequency analysis compared single- versus multi-word substitutions, omissions and additions for H.M. versus the controls. Our third relative frequency analysis compared how often the substituted and substituting words in single-word substitutions of H.M. versus the controls matched versus mismatched in lexical class, using two reference sources (www.open-dictionary.com and eslus.com/LESSONS/GRAMMAR/POS/pos.htm) to determine lexical class.

Results

Overall relative frequencies of substitution, omission and addition errors

Table 5 shows the absolute number of substitutions, omissions and additions, together with their relative frequencies for H.M. vs. the controls (in %). Additions constituted 6% of the target errors for the controls vs. 11% for H.M. (see Table 5), a non-reliable difference in relative frequency, $X^2(1) = 1.46, p > .20$. Substitutions constituted 84% of the target errors for the controls vs. 51% for H.M. (see Table 5), a reliable difference in relative frequency, $X^2(1) = 8.73, p < .01$. Omissions constituted 10% of the target errors for the controls vs. 38% for H.M. (see Table 5), a reliable difference in relative frequency, $X^2(1) = 17.99, p < .001$. H.M. therefore produced substitutions relatively less often and omissions relatively more often than the controls, with no reliable difference in the relative frequency of additions for H.M. vs. the controls.

Relative Frequencies of Single- versus Multi-word Errors by error type

Single-word omissions occurred when speakers omitted a single word in their intended utterance or BPC, as in [58], where the omitted word *phone* rendered [58] ungrammatical.

Single-word additions occurred when speakers added a single word to their intended utterance, as in [59], where the added word *because* rendered [59] ungrammatical. Single-word substitutions occurred when speakers substituted one word for another in their intended utterance, as in [60], where *wrote* substituted *phoned*, yielding an inaccurate utterance.

[58]. H.M: “He’s talking on the to somebody.” [intention or BPC: *He’s talking on the phone to somebody*] (example word omission; from Appendix V).

[59]. H.M.: “Because the guy is partially down on the floor.” [intention or BPC: *The guy is partially down on the floor*] (example word addition; from Appendix X).

[60]. H.M.: “he’s talking on the .. to somebody, the gentleman who wrote him” [intention or BPC: *he’s talking on the phone to the gentleman who phoned him*] (example single-word substitution; from Appendix V).

Multi-word omissions occurred when speakers omitted several contiguous words in their intended utterance or BPC, as in [61], where the omitted words *those buildings and his building* rendered [61] incomplete, incoherent and difficult-to-comprehend. Multi-word additions occurred when speakers added several contiguous words to their intended utterance, as in [62], where the added words *and just* rendered [62] ungrammatical. Multi-word substitutions occurred when speakers substituted several contiguous words in their intended utterance, as in [63], where substitution of *Because* for *Note that* rendered [63] incomplete and ungrammatical.

[61]. H.M: “there must be a street in between. Because he’s in his office” [intention or BPC: *There must be a street in between those buildings and his building because he’s in his office*] (example multi-word omission; from Appendix V).

[62]. H.M.: “the guys are all out there and just at their table.” [intention or BPC: *those guys are all out there at their table*] (example multi-word addition; from Appendix V).

[63]. H.M.: “Because she’s afraid of falling off her chair” [intention or BPC: *Note that she’s afraid of falling off her chair*] (multi-word substitution; from Appendix X).
Figure 2 shows the absolute number of single- versus multi-word omissions, substitutions and additions for H.M. versus the controls. As can be seen there, H.M. produced more multi-word errors of every type and more single-word omissions and additions, but somewhat fewer single-word substitutions than the controls.

Relative Frequency of Single- versus Multi-word Errors

The relative frequencies of single- versus multi-word omission, substitution and addition errors are shown for H.M. versus the controls in Table 6. As a proportion of all additions, the relative frequency of single-word additions (see Table 6) did not differ reliably for H.M. (94%) versus the controls (100%), $X^2(1) < 1$, and neither did multi-word additions (H.M., 6%; versus the controls, 0%; $X^2(1) < 1$; see Table 6). However, multi-word omissions were relatively more common for H.M. (53%) than the controls (0%), a reliable relative frequency difference, $X^2(1) = 18.16$, $p < .001$, whereas single-word omissions were relatively more common for the controls (100%) than H.M. (47%), a reliable relative frequency difference, $X^2(1) = 6.54$, $p < .05$. Similarly, multi-word substitutions were relatively more common for H.M. (24%) than the controls (2%; see Table 6), a reliable relative frequency difference, $X^2(1) = 10.28$, $p < .01$, but single-word substitutions did not differ reliably in relative frequency for H.M. (76%) versus the controls (98%; $X^2(1) = 1.46$, $p > .20$; see Table 6).

Multi-word substitution error sub-types. The controls produced only one (two-word) multi-word substitution (see Table 6), versus 14 for H.M., with relatively more two-word (79%, N=11) than three-word (14%, N=2) or four-word (7%, N=1) substitutions.

Relative frequency of syntactic class regularity

For this analysis, we first coded the syntactic class of the substituted and substituting words in single-word substitutions into twenty distinct categories: proper nouns (e.g., Mexico, Susan), common nouns (e.g., friend, aunt, idea), transitive verbs (e.g., throw, love), intransitive verbs (e.g., go, listen), auxiliary verbs (e.g., could, should), regular adjectives (e.g., lazy, brown, tall), demonstrative adjectives (e.g., this, those), possessive adjectives (e.g., my, their), adverbs of time (e.g., tomorrow, recently), adverbs of frequency (e.g., always, usually), personal pronouns (e.g., I, he), possessive pronouns (e.g., hers, mine), intensive pronouns (e.g., "myself" in "I myself saw it"), reflexive pronouns (e.g., "myself" in "I saw it myself"), prepositions of time (e.g., at 2:00, for a day), prepositions of place (e.g., at my house, in my hand), coordinating conjunctions (e.g., and, or, but), subordinating conjunctions (e.g., after, if, when), correlative conjunctions (e.g., either/or, both/and), indefinite articles (i.e., a/an), and definite articles (i.e., the). We then coded the substituted and substituting words as matching or mismatching in syntactic class. The results indicated that single-word substitutions obeyed the syntactic class regularity relatively more often for the controls than for H.M.: Substituted and substituting words had matching syntactic categories in 51% of H.M.’s single-word substitutions versus 96% for the controls (see Table 6), a reliable relative frequency difference, $X^2(1) = 6.83$, $p < .01$, and had mismatching syntactic categories in 49% of H.M.’s single-word substitutions versus 4% for the controls (see Table 6), a reliable relative frequency difference, $X^2(1) = 18.93$, $p < .001$. Moreover, the mismatches or violations of the syntactic class regularity, e.g., “out” [preposition] substituted for calling [verb], invariably rendered utterances ungrammatical.

Discussion

General process hypotheses

General process hypotheses, e.g., an especially high error tolerance for H.M. relative to controls, predicted no difference in the relative frequency of omissions vs. additions vs. substitutions for H.M. versus the controls, contrary to Study 3 data, where omissions were
relatively more common and substitutions relatively less common for H.M. than the controls. General process hypotheses likewise predicted no difference in the relative frequency of single-versus multi-word omissions and substitutions for H.M. versus the controls, contrary to Study 3 data, where single-word omissions were relatively less common, and multi-word omissions, substitutions and additions were relatively more common for H.M. than the controls. General process hypotheses also predicted same relative frequency of exceptions to the syntactic class regularity for H.M. and the controls, contrary to Study 3 data, where exceptions to the syntactic class regularity were relatively more common for H.M. than the controls.

Specific process hypotheses

Study 3 data were broadly consistent with specific process hypotheses: that specific brain mechanisms or processes associated with H.M.'s cerebellar and/or MTL damage causes abnormal speech errors in H.M. relative to controls.

Cerebellar damage hypothesis. Under the cerebellar damage hypothesis, H.M.'s abnormal speech errors reflect his cerebellar damage, consistent with data indicating “agrammatic” speech errors in cerebellar patients (see Ackermann et al., 2007; and Ackermann et al., 2004). However, H.M.'s errors differed in etiology and nature from those of the Ackermann et al. patients. The two agrammatic patients in Ackermann et al. had transient agrammatic symptoms (lasting several weeks) whereas H.M.’s cerebellar symptoms were chronic (lasting decades); since 1970, H.M. has exhibited the usual pattern of symptoms of (non-agrammatic) cerebellar patients: reduced speech rate and abnormal hesitations or pauses (see H.M.’s unusual hesitations in [64] and [65] from Marslen-Wilson, 1970; see also Appendices V-X); even their lesion cites differed: unilateral in the Ackermann et al. patients (reflecting ischemic infarctions involving the right posterior inferior cerebellar artery) but a bilateral in H.M. (reflecting chronic use of dilantin for controlling epilepsy; see Corkin et al., 1997). The cerebellar damage hypothesis therefore faces four challenges as an account of H.M.’s abnormal errors: the fact that most cerebellar patients (including H.M.) exhibit “classical cerebellar symptoms” (abnormal hesitations or pauses, and reduced speech rate) rather than agrammatic symptoms, the differing nature of H.M.’s cerebellar symptoms versus his sentence-level errors, which were chronic and non-progressive since his 1953 lesion, and the differing origins and symptoms for H.M. (dilantin-linked, longstanding, progressive and classical symptoms) versus the Ackermann et al. “agrammatics” (short-lived and unusual symptoms).

[64]. H.M. (p. 25): “… uh … thank you …. it’s funny … the way … what I always wanted to be …. though I know I couldn’t … uh …. well ….. put it down as the …. well ….. why I couldn’t be it … uh … naturally with the … epileptic seizures of any kind ….. but I thought of …. the wearing the glasses.” [W.M-W.: “of .. of what?”] The wearing the glasses. [W.M-W.: “Yes ....”] “….. brain surgeon ….” [W.M-W.: Yes ... that’s what you wanted to be ? ....... oh, well ....] “…… because I know, in brain surgery …. that wearing glasses …. these little bits(???) (makes gesture showing hand slipping slightly) …. that person is gone …..” [W.M-W.: “Yes .... got to be ...”] “It has to be just right …. ” [W.M-W.: “This is what you wanted to be when you were young?”] “I wanted to be a doctor ….. before … I thought of a lawyer too ….. but .. seeing that … well …. law … well .. came along .. it was more of an excuse in a way …. because I thought of being a doctor.” [W.M-W.: “Yes. …”] “….. and …. in …….. naturally(???) in mechanics … mechanical way …. uh …. but .. uh .. helping people …. and … helping, and doing, and knowing …. because I wore the glasses then ….. that … one thing I could not be …. and I didn’t want to trust myself to ………….. as …. as … well …. uh …. I could say trust myself to, in a way.” (from Marslen-Wilson, 1970)
H.M.: “That’s what they’re there for, yes, but the important thing is the other people….and the doctors too, how it helps them….because little things that they...learn themselves, they can pass on to others...and they can pass them, and collect it and group them, that way...” (from Marslen-Wilson, 1970)

MTL damage hypotheses. Under the MTL damage hypothesis, H.M.’s MTL damage causes speech errors that differ in basic nature from the everyday slips of normal speakers. Consistent with this hypothesis, omission and substitution errors differed in relative frequency for H.M. versus the controls in Study 3. Also consistent with this hypothesis, H.M. has produced omission errors and incomplete sentences resembling [64-65] from 1970 to 2000. For example, like [29] from Skotko et al. (2005), [64] from Marslen-Wilson (1970) requires elaborations that address issues such as: What does it refer to in H.M.’s “couldn’t be it” in [64]? Who was wearing the glasses in H.M.’s “the wearing the glasses” in [64]? What was being excused and why in H.M.’s “more of an excuse in a way” in [64]? What was being done “in mechanics” or in a “mechanical way in [64]”? What was being done and by whom in H.M.’s “and doing” in [64]?

However, the MTL damage hypothesis must also explain detailed causes of H.M.’s error patterns in Studies 1-3: Why did H.M. produce major but not minor errors more often than the controls in Study 2? Why were omissions relatively more common but substitutions relatively less common for H.M. than the controls in Study 3? Why were single-word omissions relatively less frequent and multi-word omissions relatively more frequent for H.M. than the controls? Why did H.M. correct major errors (resulting in ungrammatical, incoherent and difficult-to-understand utterances) but not minor errors less often than the controls? Why did single-word substitutions violate the syntactic class regularity relatively more often for H.M. than the controls? Why does H.M. frequently shift topics of conversation to familiar themes that he has frequently discussed in the past? Why does H.M. often repeat familiar words, phrases, propositions, sentences and stories than controls?

Only one version of the MTL damage hypothesis addresses all of these why questions: the original MacKay, Burke et al. (1998) hypothesis (see also MacKay & James, 2001, 2002; MacKay et al., 2007, 2008; and MacKay, Stewart, et al., 1998). Under this hypothesis, H.M.’s 1953 lesion impaired his ability to rapidly and without rehearsal form new connections between units in the cortex. H.M. therefore produces incomplete and incoherent utterances lacking critical elements such as a main verb and object because new cortical connections are necessary to form novel internal representations for never-previously encountered conjunctions of word and phrase concepts in sentence-level plans (see MacKay, Burke et al.).

The MacKay-Burke hypothesis also explains why omissions, and especially multi-word omissions, were relatively more common for H.M. than the controls in Study 3: Conjunctions of word and phrase concepts by definition involve more than one word. The MacKay-Burke hypothesis likewise explains why H.M. produced relatively more free associations and multi-word substitutions than the controls: H.M. produced free associations as multi-word gap fillers for completing his sentence-level plans. Because free associations in an ongoing utterance are by definition unrelated to sentence syntax, H.M.’s free associative substitutions may also explain why his major substitution errors often violated the syntactic class regularity.

The MacKay-Burke hypothesis also explains the selective nature of H.M.’s sentence-production deficits. Under this hypothesis, H.M. produces familiar words, phrases and propositions without deficits because his mechanisms for activating internal representations formed prior to his lesion and used frequently since then have been spared. To favor his spared activation processes, H.M. also shifts conversational topics to familiar themes that he has
frequently discussed in the past, as in [64] and [65], where H.M. in 1970 switched to the same topics (becoming a medical doctor, little one, and helping other people) as in [28] and [29] from Skotko et al. (2005). H.M.’s spared activation processes also enable engrainment learning, the primitive and relatively inefficient process whereby rehearsal or repeated activation facilitates performance. A repetition-based engrainment learning strategy therefore enables H.M. to form sentence- and paragraph-level internal representations or plans by repeating words, phrases, propositions, sentences and stories (see Study 2; also MacKay et al., 2008).

The MacKay-Burke hypothesis also explains H.M.’s deficits in detecting and correcting major but not minor errors. By hypothesis, minor errors are normal and occur during the activation of familiar or pre-formed phrase- and sentence-level plans, so that H.M. activates and mis-activates pre-formed plans in the same way as normal speakers. By contrast, major errors are abnormal and occur when unfamiliar or novel phrase- and sentence-level plans are incomplete or incoherent: Lacking well-formed sentence-level plans or intentions, H.M. cannot compare his actual output with his intended output, an essential step in error detection and correction. Consistent with Study 2 results, H.M. will therefore fail to detect major but not minor errors more often than normal speakers under the MacKay-Burke hypothesis.

Also consistent with the MacKay-Burke hypothesis were the experimental data of MacKay et al. (2007, Experiment 2), where H.M. neither detected nor corrected major substitution errors deliberately planted within simple sentences. Participants in MacKay et al. had three goals: 1). to identify whether sentences such as, The boy make a cake, were grammatical or ungrammatical; 2). to specify what word made an ungrammatical sentence ungrammatical, e.g., make in The boy make a cake; and 3), and to replace the erroneous word with a new word that makes the ungrammatical sentence grammatical, e.g., The boy made a cake. The results indicated that: 1). H.M. correctly identified grammatical sentences as grammatical and ungrammatical sentences as ungrammatical reliably less often than the controls; 2). H.M. correctly specified the erroneous word in sentences he identified as ungrammatical relatively less often than the controls; and 3). H.M. successfully corrected words that he identified as erroneous relatively less often than the controls.

General Conclusions

In summary, major errors (which disrupt communication by definition) were more common for H.M. than closely matched memory-normal controls. These errors differed from everyday slips-of-the-tongue in effects on communication (reduced accuracy, grammaticality, coherence and comprehensibility) and follow-up (never corrected versus usually corrected). Turning to minor errors (which do not disrupt communication), only word repetitions were reliably more common for H.M. than memory-normal controls: H.M. was deficit-free for stutters, dysfluencies and for corrected errors in speech and reading.

Classical error types also differed in nature for H.M. versus controls: Omission, substitution and addition errors involved multiple rather than single words relatively more often for H.M. than the controls; substitution errors violated the syntactic class regularity relatively more often for H.M. than the controls; and omissions were relatively more common and substitutions relatively less common for H.M. than normal speakers;

Theories of speech production must eventually explain these fundamental differences between the speech errors of normals versus H.M. and other “amnesic-aphasics” (see MacKay, in press; MacKay et al., 2007; and MacKay, Burke, et al., 1998). Just as the backfiring of a car engine instantiates general principles that theories of internal combustion must account for, the speech errors of normal and brain damaged speakers instantiate general principles that language
production theories must account for (see MacKay, 1973). However, further research on the relation between speech errors and the MTL is needed to determine whether other patients with MTL damage similar to H.M.’s exhibit the same pattern of speech errors as H.M. The same also applies to reading errors: Although H.M.’s reading errors in MacKay and James (2001, 2002) closely resembled those of amnesic-dyslexics in Friedman (1996), the locus of brain damage in Friedman-dyslexics and other amnesic-dyslexics is currently unknown.

Speech production theories must also explain why H.M. exhibits selective rather than across-the-board speech error deficits. Theories that ascribe H.M.’s deficits to a single non-selective factor, e.g., reduced working memory capacity, diffuse brain damage, reduced motivation, perseverative tendencies, or general cognitive decline are inadequate (see e.g., MacKay & James, 2002). Hypotheses based solely on H.M.’s spared functions are likewise inadequate, e.g., the intact language hypothesis, which ignores H.M.’s well-established sentence production deficits in a wide variety of tasks, as does the hypothesis that H.M.’s language production is artful and “sophisticated” (Kolb & Whishaw, 2003). The present research alone indicated 15 reliable deficits relative to carefully matched controls for standard error frequency analyses (nine significant differences between H.M. and controls for seven types of major errors and two types of minor errors) and blind rating data (six significant differences involving more negative ratings for vagueness, comprehensibility, grammaticality, coherence, humor-description adequacy, and number of major errors). H.M.’s reliable deficits for major errors likewise contradict the Skotko et al. (2005) hypothesis that H.M.’s spoken discourse is without major errors (see also H.M.’s major errors in MacKay, Burke et al., 1998; MacKay & James, 2001, 2002; MacKay et al., 2007; MacKay, Stewart et al., 1998; Marslen-Wilson, 1970; and the Skotko et al. Appendices).

Theories applicable to H.M.’s condition must also address one final class of constraints: the parallels between H.M.’s deficits and sparing in spoken sentence production, reading aloud, visual cognition, language comprehension, and episodic, semantic, procedural and implicit memory (for two such theories, see the comparison in Appendix Z). These parallels fall into two categories: general performance parallels and error-related parallels. MacKay et al. (2007), MacKay and Hadley (2009), and MacKay and James (2002, 2009) have discussed general performance parallels in detail. Here is a small sample: H.M. misproduces unfamiliar but not familiar phrases when producing sentences in conversational contexts, and he likewise misreads unfamiliar but not familiar phrases when reading visually presented sentences aloud (MacKay & James, 2001); H.M. understands familiar words in familiar but not unfamiliar contexts when comprehending sentences, and he likewise readily detects familiar figures in familiar but not unfamiliar contexts in visual cognition tests such as the What’s-wrong-here task (MacKay & James, 2009; see also H.M.’s window-for-performance-chart error in the unfamiliar cartoon scene in Study 2); H.M. can remember familiar information (e.g., familiar semantic memories) but not novel information (including all episodic memories). Theories that do not address these and other general performance parallels or apply to only a single domain without addressing H.M.’s parallel deficits and sparing in language comprehension, reading aloud, visual cognition, spoken sentence production, and episodic, semantic, procedural and implicit memory can be considered inadequate.

We turn now to error-related parallels or consistent idiosyncrasies in H.M.’s errors in spoken speech, word retrieval, visual object identification, reading isolated words aloud, and reading sentences aloud. These consistent idiosyncrasies include non-correction (H.M. corrects his errors reliably less often than controls), anomaly (H.M.’s errors yield anomalous outcomes reliably more often than controls’) and omissions (omission errors are relatively more common
for H.M. than controls). All three idiosyncrasies (non-correction, anomaly and omissions) characterize H.M.’s errors in spoken speech when describing cartoons (Studies 1-3), when answering questions about common childhood experiences (MacKay, Burke et al., 1998; see also Marslen-Wilson, 1970), when describing two meanings in short ambiguous sentences (MacKay, Stewart et al., 1998), when describing two meanings for isolated ambiguous words and phrases (MacKay et al., 2007), and when using two or three pre-specified words in a grammatical sentence that accurately describes a picture (MacKay et al., 2008). Moreover, the same three idiosyncrasies also characterize H.M.’s errors when reading isolated low frequency words aloud (see Appendix Z; also MacKay & James, 2002; and MacKay & Hadley, 2009), when retrieving low frequency words on the Boston Naming Test (MacKay & Hadley, 2009), when identifying objects in unfamiliar visual scenes (see MacKay & James, 2009; see also H.M.’s uncorrected window-for-easel error in Study 2, where he omitted the easel from his internal representation of the cartoon scene in Appendix W, an uncorrected error that yielded anomaly because easels do not support windows in common experience), and when reading isolated sentences aloud (see MacKay & James, 2001, and MacKay, Stewart et al., 1998; see also H.M.’s omission of the word you without correction in reading the cartoon caption in Figure 1, I tell you, Edith, it’s not easy raising the dead, an uncorrected omission that yielded anomaly because you is essential to accurately indentify Edith as the listener depicted in the cartoon rather than an unknown third person named Edith). Parsimony therefore demands the same explanation for H.M.’s parallel idiosyncrasies (non-correction, anomaly and omissions) in all five domains. Theories that cannot explain these error-related parallels in H.M.’s spoken speech, word retrieval, visual object identification, word reading, and sentence reading can be considered inadequate (for a detailed application of this important theoretical constraint, see Appendix Z).

Theory-linked errors: One Theory (or Less) is not Enough

H.M.’s major errors in this and other studies (MacKay, Burke et al., 1998; MacKay & James, 2001, 2002; MacKay et al., 2007; MacKay, Stewart et al., 1998; Skotko et al., 2005) can be considered theory-linked because they contradict the intact language and no-major-errors hypotheses, and are difficult to explain under current systems theories (see MacKay et al., 2007). We turn now to another (quite different) type of theory-linked errors, namely inaccuracies in how Skotko et al. (2005) described published statements about procedures, data, results and theoretical implications in other studies of H.M. These numerous inaccuracies are shown with readily verified corrections in Table 7 and can be considered theory-linked because they all favor the intact language hypothesis over alternate hypotheses.5

Why are such theory-linked errors important to note and correct in this case? One reason concerns the enormous impact that data on H.M. have had and continue to have on theories in psychology and the brain sciences. A second reason is that H.M. is the most studied patient in the history of the field: If the field cannot get the procedures, results, and theoretical implications right with H.M., what hope is there of getting them right with other less-studied patients or groups of patients? A third reason for correcting theory-linked errors is that accuracy is of the essence in science (unlike, say, politics), an issue that applies also to the secondary sources that continue to describe H.M.’s language production with the labels “intact” (despite extensive data to the contrary) and “relatively intact” (presumably relative to H.M.’s episodic memory deficits rather than relative to memory-normal controls). Although presented as fact, this “relatively intact” label represents an interesting hypothesis that can only be tested via comparisons of H.M.’s episodic memory and language production under similar conditions, with controls for, say, stimulus novelty and complexity, and for the time between stimulus presentation and recall.
or production of stimulus descriptors. Because no such experiment has been conducted with H.M., the “relatively intact” label inaccurately characterizes an untested hypothesis as fact.

A related issue concerns the process whereby the editor and reviewers of the claim that H.M.’s spoken discourse is “intact” and without major errors came to miss H.M.’s major errors and ungrammatical, incoherent, and difficult-to-understand utterances in Skotko et al. (2005). Under one hypothesis, such oversights are both “normal” and readily understood. The history of science contains many examples where “seeing” or attending to existing data that falsify or delimit a generally accepted account has required the adoption, consideration, or development of alternative theoretical frameworks. For example, the biological sciences ignored centuries of facts on selective breeding of domesticated animals until Darwin developed a theoretical alternative that made sense of these facts (see MacKay, 1993). The moral for systems theory and the H.M. story is clear: One theory or less is not enough in psychology and the brain sciences. To “see” or register H.M.’s language production deficits and their implications may require the acquisition and use of at least one alternative theory or “theoretical language” that makes sense of H.M.’s major errors, false starts, and repetitions, and their effects on the comprehensibility, grammaticality, coherence, adequacy, vagueness, relevance and accuracy of his utterances (for one such theory, see MacKay, Burke et al., 1998; MacKay & James, 2001, 2002; MacKay et al., 2007, 2008; and MacKay, Stewart, et al, 1998).
References


H.M.’s Language Production Errors


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


Author Notes

The authors gratefully acknowledge support from the Samuel A. MacKay Memorial Research Fund and NIH grant R01 AG 09755. We thank Dr. Suzanne Corkin for permission to study H.M. from 1997-1999, and Kataryna Boese, Alicia Chang, Meghan Gould, Jennifer Taylor, and Diane Marian for general research assistance. We also thank Dr. Arthur Shimamura for suggesting the relational memory theory account of H.M.’s language production deficits discussed in Appendix Z. This article grew out of a reply to Skotko et al. (2005) that was originally intended for publication in the same journal as Skotko et al. However, that journal has a policy of not accepting replies. Queries should be addressed to Donald G. MacKay, mackay@ucla.edu, or Psychology Dept., University of California, Los Angeles, Los Angeles, CA 90095-1563.
Footnotes

1 Our use of present tense describes H.M.’s behavior 11 to 38 years before his death.
2 We could not obtain permission to reprint Cartoons 1 and 2. Researchers can nevertheless view these cartoons in The New Yorker (1951) using the descriptions in Table 1.
3 The present appendices are labeled Appendix V-Z to avoid confusion with subsequent references to Appendix A-K in Skotko et al. (2005, pp. 409-413). Context resolved virtually all of H.M.’s “difficult to hear” words, which reflected two factors: poor articulatory quality (perhaps related to his recent-origin cerebellar damage; see Corkin et al., 1997), and his sentence-level incoherence (discussed shortly).
4 In combination, these limitations of our rating procedures may explain the almost ten-fold difference in H.M.’s estimated mean number of major errors in our rating results (3.62) versus our overall standard error frequencies (28.82), computed as the sum of seven major error types: 10.83 substitutions, 10.00 omissions, 2.83 additions, 0.33 free associations, 1.00 accuracy errors, 2.67 reading errors, and 2.50 transposition errors (per section).
5 H.M.-linked inaccuracies are not unique to Skotko et al. (2005). For example, Martin and Freedman (2001) inaccurately claimed that H.M.’s surgery took place in 1957 and involved complete removal of his perirhinal cortex, his entorhinal cortex, and his hippocampi; that H.M.’s pre-1953 epilepsy differentially affected his left hemisphere; and that all patients with episodic memory deficits necessarily show language processing difficulties according to MacKay, Burke et al. (1998) and MacKay, Stewart et al. (1998) (a claim explicitly denied in MacKay, Burke et al.). Although H.M.’s episodic memory and language was unexamined prior to his lesion, Martin and Freedman also speculated that H.M.’s language deficits may have originated prior to his lesion (without suggesting the same for his episodic memory deficits; see MacKay, 2001).
Table 1

Edited descriptions of H.M. (right panel) and a representative control (left panel) for cartoons 1-3 (captions indicated in quotes). [SEGMENT] indicates a boundary used in content analyses.

<table>
<thead>
<tr>
<th>Cartoon 1 (business man on the phone): “No, Thursday’s out. How about never – is never good for you?”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control Description</strong> (Edited)</td>
</tr>
<tr>
<td>C 12: A fellow is on the phone in his office looking in his date book and somebody is obviously on the phone asking him to come on Thursday. [SEGMENT] He says, “No, Thursday’s out.” And now he’s saying “How about never?” [SEGMENT] He doesn’t want to see him. [SEGMENT] He says, “Is never good for you?” [SEGMENT] And, of course, behind him is Wall Street and everything like that. [SEGMENT]</td>
</tr>
<tr>
<td>H.M.: Well, it's a about a person talking on the telephone, which has got a long extension cord. [SEGMENT] And he is looking at his notebook, besides. [SEGMENT] And naturally, the window is behind him and shows other buildings over there, see. [SEGMENT] And there must be a street in between those buildings and his building because he's in his office. [SEGMENT] And this here is a pen beside the notebook up there. [SEGMENT] And these lines right here just represent the edge of the window [SEGMENT] which you could say is right in the middle of the wall. [SEGMENT] And this building’s behind him, and he's talking on the phone to the gentleman who phoned him. [SEGMENT] And he’s not looking out the window, or anything like that; he’s just looking at his date book, and he has two dates down there. [SEGMENT] And then of course this is the signature of the person who made the cartoon. [SEGMENT] Yeah (in response to the experimenter question: Can you read that part down there?): It says just what he said over the telephone to the person he's talking to. [SEGMENT] And he's making a double negation: no and never. [SEGMENT] Because to ask, “Is never good for you?” means that never is good for the person talking. [SEGMENT] And, he has stated something about the person calling. [SEGMENT] This person said that never was good. And I’m just repeating what’s there. [SEGMENT] And the buildings in the background, some of them are just being constructed in a way, because they aren't finished.</td>
</tr>
<tr>
<td>Control Description (Edited)</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td><strong>C 5:</strong> “The beatings will continue until morale improves.”</td>
</tr>
<tr>
<td>[SEGMENT] There’s an executive meeting, or a board meeting, or it’s some sort of a meeting. [SEGMENT] And there is a mixture of men and women and they all have dour looks on their faces [SEGMENT] and this woman is standing there and she says, “The beatings will continue until morale improves.” [SEGMENT] And looking at these faces, you know that, forget it, nothing’s going to improve, except the beatings. [SEGMENT]</td>
</tr>
</tbody>
</table>
Cartoon 3 (see Figure 1): “Oh, I don’t know. Billy’s been having trouble in school and Sally’s always having some sort of crisis. I tell you, Edith, it’s not easy raising the dead!”

<table>
<thead>
<tr>
<th>Control Description (Edited)</th>
<th>H.M.’s Description (Edited)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C 14: “Oh, I don’t know. Billy’s been having trouble in school and Sally’s always having some sort of crisis. I tell you Edith, it’s not easy raising the dead!” [SEGMENT] There are two women sitting in a corner. [SEGMENT] They look like women but I think they’re women ghosts. [SEGMENT] And you see two ghosts, one sitting on the floor and one falling down the steps [SEGMENT] and there’s a picture of a bat and a picture of a ghost [SEGMENT] and there’s two ghosts discussing their kids it seems. [SEGMENT] “It’s not easy raising</td>
<td>H.M.: Well, this woman ghost is talking to this woman ghost. And this woman ghost is falling down the stairs. [SEGMENT] And you can’t tell but she struck me more as wanting to have it her way, only her way. [SEGMENT] The children are in her way. Especially this one over here. This one over here (in response to the experimenter question: Mm, which one?) Yeah, she's the one that’s talking to this woman. [SEGMENT] This guy is partially on the floor. And you can't tell exactly what it is she's telling her, whether because of the picture or what. [SEGMENT] Note that she's afraid of falling off her chair: You can see both sides, the front, and the inside of the chair. [SEGMENT] And he's down on the floor. And she's falling down the stairs. There in the picture- I wonder about that railing they have there for the stairway. [SEGMENT] And look, what is wrong with his braces here, in between his nose and chin? [SEGMENT] His mouth is just that single tooth and he can't talk. [SEGMENT] No one will talk to him but they all stare. Well, they’re not willing to give a hand. [SEGMENT] And that is a picture of a ghost. And that's a picture of a bat. And this picture is drawn wrong, because that woman sitting there, you can see the top of the chair right through her face there and through her hair also. (Seeing the chair right through the ghost apparently puzzles H.M.) And you can't tell just what her hair color is. [SEGMENT] And this dress of hers is dark on this side but you can't tell what these dots are on this side. [SEGMENT] There is a continuation of the dots over here. And this one you naturally can’t see the color of her shoes. She’s just got transparent shoes on. [SEGMENT] And you can tell that that’s the base of the chair that the artist just blackens the whole way down, and everything, because</td>
</tr>
<tr>
<td>that was one of her rules. [SEGMENT] Then this looks like a shoe but a shoe would be narrower. That really illustrates why there’s a rule. [SEGMENT] I can’t read the name (in response to the experimenter question: Can you read me the part on the bottom? Can you read me what it says?). It starts with “O.” [SEGMENT] (Here H.M. is reading the fine print signature of the artist rather than the caption at the bottom and the experimenter asks: But what about the typing down at the bottom). It says, “Oh, I don’t know. Billy’s been having trouble in school and Sally’s always having some sort of crisis. I tell you, Edith, it’s not easy raising the dead.” [SEGMENT] That’s mixed up in a way, because one person is talking, and is saying that another person is causing a lot of trouble or crises and everything. [SEGMENT] This man has been having trouble in school, but Billy, that’s the name of a very young kid in school. And this is an older man in the picture. [SEGMENT] (H.M. is confused because he considers the boy ghost in the picture to be an older man who couldn’t have a boy’s name like Billy). And you can’t tell if they are ghosts of some dead people [SEGMENT] because it’s white over here, but her face doesn’t have any eyes on it. [SEGMENT] And it’s sorta whitish on both her arm and her wrist and there are white parts up there on her face. [SEGMENT]</td>
<td></td>
</tr>
</tbody>
</table>
Table 2

*Definitions for six types of major errors (center panel) with examples (right panel)*

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
</table>
| Minor substitution errors | Minor substitutions occur when speakers substitute one word or more words in sequence for another word or words, *paradigmatic* substitutions *syntagmatic* substitutions (anticipations, perseverations and transpositions or exchanges of upcoming and/or already produced words and phrases in an intended utterance). | Normal speaker: “Put it on the table, I mean, chair”
|                  |                                                                           | Normal speaker: “Put it on the table, I mean, chair”                   |
| Major Substitution Errors | Major substitutions [S] occur when speakers substitute without correction one word or more words in sequence for another word or words, yielding an utterance that is inaccurate, ungrammatical or infelicitous (incoherent, vague, repetitious, or difficult to comprehend). | H.M.: No one will take [S] him…
|                  |                                                                           | {BPC: No one will *talk to* him…}                                       |
| Major Omission Errors | Major omissions [O] occur when speakers omit without correction one or more adjoining words in a sentence, yielding an utterance that is inaccurate, ungrammatical (when major constituents such as the verb or subject are omitted) or infelicitous (incoherent, vague, or difficult-to-comprehend). | H.M.: And you can't tell exactly what it is she's telling him [S] …
<p>|                  |                                                                           | {O} picture or what. {BPC: And you can't tell exactly what it is she's telling <em>her</em>, <em>whether because of the</em> picture or what.} |
| Major additions [A] | Major additions [A] occur when speakers                                                                           | H.M.: (those) guys are all out                                           |</p>
<table>
<thead>
<tr>
<th>Type of Error</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Addition Errors</strong></td>
<td>add words or phrases that render an ongoing sentence ungrammatical and have no discernable function in the sentence.</td>
<td>there and [A] just at their table. {BPC: (those) guys are all out there at their table.}</td>
</tr>
<tr>
<td><strong>Major Accuracy Errors</strong></td>
<td>Major inaccuracies [IA] occur when speakers inaccurately describe some feature without correction.</td>
<td>H.M.: because the mountain area in the back. {H.M. has confused the chart with a window showing a view of mountains}</td>
</tr>
<tr>
<td><strong>Major Free Association Errors</strong></td>
<td>Major free associations [FA] occur when speakers produce without correction words or phrases that are strongly linked in semantic memory but task-irrelevant or unrelated to an ongoing utterance.</td>
<td>H.M.: She possibly wants to make it her way, only her way. They're in her way [FA]. (Note: “in her way” is strongly linked in semantic memory to H.M.’s prior “her way” but unrelated to the remainder of his utterance.)</td>
</tr>
<tr>
<td><strong>Major Transposition Errors</strong></td>
<td>Major transposition errors occur when speakers produce words in inappropriate order within a sentence without correction.</td>
<td>H.M.: telephone, got a long cord on the extension {BPC: the telephone has a long extension cord}.</td>
</tr>
<tr>
<td><strong>Major Reading Errors</strong></td>
<td>Reading errors [MR] occur when a caption cartoon is misread without correction.</td>
<td>H.M.: it's [MR] never good for you. {correction: “Is never good for you?”}</td>
</tr>
</tbody>
</table>
Table 3

Types of minor errors and error-related events, with definitions (center panel) and examples with corrections (right panel).

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minor syntagmatic substitution errors</td>
<td>Minor <em>syntagmatic</em> substitutions occur when speakers anticipate, perseverate, transpose or exchange one or more words with one or more upcoming or already produced words in their intended utterance.</td>
<td>Normal speaker: “He threw the window through a clock, I mean, threw a clock through the window”</td>
</tr>
<tr>
<td>Minor paradigmatic substitutions</td>
<td>Minor paradigmatic substitutions occur when speakers substitute one or more words for a word or words in the same syntactic slot or position in their intended utterance.</td>
<td>Normal speaker: “Put it on the table, I mean, chair”</td>
</tr>
<tr>
<td>Minor omission errors</td>
<td>Minor omissions occur when speakers omit one or more words in executing their sentence plan.</td>
<td>Normal speaker: “As Morton and Broadbent out, I mean, point out.”</td>
</tr>
<tr>
<td>Minor addition errors</td>
<td>Minor additions occur when speakers add one or more unintended words to their sentence plan.</td>
<td>Normal speaker: “you’re in a more better position, I mean, you’re in a better position.”</td>
</tr>
<tr>
<td>Minor Corrected Speech Errors</td>
<td>Corrected errors [CE] include all speech errors that are subsequently corrected.</td>
<td>C14: “has this appointment desk on hi- [CE] uh- book on his desk.” {correction: appointment book on his desk}</td>
</tr>
<tr>
<td>Minor Corrected Reading Errors</td>
<td>Corrected reading errors [CRE] occur when a cartoon caption is misread followed by a correction.</td>
<td>C7: “until moral, morale [CRE] improves.”</td>
</tr>
</tbody>
</table>
Table 4

Errors by Task, H.M.’s Age when tested and Error Type, with [O] for omissions, [A] for additions, [S] for stuttering, and [R] for repetitions. Following standard format, arrows representing “was misproduced as” link H.M.’s BPC or intended output (in italics) with his actual output (in quotes).

<table>
<thead>
<tr>
<th>Task</th>
<th>H.M.’s Age at Test</th>
<th>Type of Errors</th>
<th>Typical Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describing two meanings in short ambiguous sentences</td>
<td>47 (data from MacKay, Stewart et al., 1998)</td>
<td>Major Spoken Omission Error</td>
<td>[42]. H.M.: (describing one of the meanings of the ambiguous sentence, <em>The marine captain liked his new position.</em>) <em>He liked the new position because of being on a passenger line.</em> → “He liked the new position because of being, being [O] a passenger line.”</td>
</tr>
<tr>
<td></td>
<td>47 (data from MacKay, Burke et al., 1998)</td>
<td>Major Spoken Addition Error</td>
<td>[43]. H.M. (describing meanings of the ambiguous sentence, <em>He told me to go without any hesitation</em>): “Well I thought of a person being told something the straightforward *out [A] style that the other person was telling him without stopping him or stopping in [A] it or freely and uh, then I also think of a person [A] being told something by a person [A], uh, straight-forwardly, like the other, he told him right out, straight out, quick, ah, not quicker, but easy.”</td>
</tr>
<tr>
<td>Reading short sentences aloud</td>
<td>72 (data from MacKay &amp; James, 2001)</td>
<td>Major sentence reading error</td>
<td>[44]. H.M.: <strong>the boys who ate hot dogs got stomach aches</strong> → “the boys [O] ate hot dogs got stomach aches.”</td>
</tr>
<tr>
<td>Describing two meanings of ambiguous sentences</td>
<td>44 (data from MacKay, Burke, et al., 1998)</td>
<td>Major free association errors</td>
<td>[45]. H.M. (describing one of the meanings of the ambiguous sentence, <em>The marine captain liked his new position</em>): “that’s why he liked the position too, because he was above them and of all, most of all [FA].”</td>
</tr>
<tr>
<td>Describing two meanings of isolated</td>
<td>72.5 (data from MacKay et al., 2007)</td>
<td></td>
<td>[46]. H.M. (describing the two meanings of the ambiguous word <em>lots</em>): “And that could be many or more [FA].”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[47]. H.M. (describing the “draw lots” meaning of the</td>
</tr>
<tr>
<td>ambiguous words and phrases</td>
<td>I think it can refer to pieces of straw of various lengths that people use to choose between alternatives by chance. → “And uh well, uh it’d be uh, also it’d be that the- I think it’s, uh, [O] probably, straw [FA]… long and short ones [O]”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Using pre-specified words to create a single grammatical sentence that describes a picture</td>
<td>Minor word repetition and stuttering errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72 MacKay et al. (2008)</td>
<td>[48]. H.M.: “…on that bus, the scrawny bus [R] and have it drive it off… it [S] drives it [R] off.” [49]. H.M.: “…she wants to go on the bus…and it’s crowded… it’s crowded [R]… too crowded [R] to get on the bus.”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5

*Total Number of Errors by error type, with Percentages in parentheses for H.M. versus the Controls in Study 3 (see text for explanation)*

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Speaker(s)</th>
<th>H.M.</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substitution</td>
<td></td>
<td>81</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(51%)</td>
<td>(84%)</td>
</tr>
<tr>
<td>Omission</td>
<td></td>
<td>60</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(38%)</td>
<td>(10%)</td>
</tr>
<tr>
<td>Addition</td>
<td></td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11%)</td>
<td>(6%)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>158</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100%)</td>
<td>(100%)</td>
</tr>
</tbody>
</table>
Table 6

*Relative Frequency of Error Subtypes (in %) for H.M. versus the Controls in Study 3 (* indicates reliable differences; see text for explanation)*

<table>
<thead>
<tr>
<th>Error Subtypes</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H.M.</td>
</tr>
<tr>
<td><strong>Omission Error Subtypes</strong></td>
<td></td>
</tr>
<tr>
<td>Multi-word Omissions</td>
<td>53%</td>
</tr>
<tr>
<td>Single-word Omissions</td>
<td>47%</td>
</tr>
<tr>
<td><strong>Addition Error Subtypes</strong></td>
<td></td>
</tr>
<tr>
<td>Multi-word Additions</td>
<td>6%</td>
</tr>
<tr>
<td>Single-word Additions</td>
<td>94%</td>
</tr>
<tr>
<td><strong>Substitution Error Subtypes</strong></td>
<td></td>
</tr>
<tr>
<td>Multi-word Substitutions</td>
<td>24%</td>
</tr>
<tr>
<td>Single-word Substitutions</td>
<td>76%</td>
</tr>
<tr>
<td>Single-word Substitution Subtypes</td>
<td></td>
</tr>
<tr>
<td>Same syntactic category</td>
<td>51%</td>
</tr>
<tr>
<td>Different syntactic category</td>
<td>49%</td>
</tr>
<tr>
<td>Multi-word Substitution Subtypes</td>
<td></td>
</tr>
<tr>
<td>2-word substitutions</td>
<td>79%</td>
</tr>
<tr>
<td>3-word substitutions</td>
<td>14%</td>
</tr>
<tr>
<td>4-word substitutions</td>
<td>7%</td>
</tr>
</tbody>
</table>
Table 7. A sample of theory-linked inaccuracies in Skotko et al. (2005) (left panel), with readily verifiable corrections (right panel)

<table>
<thead>
<tr>
<th>Theory-linked Inaccuracies in Skotko et al.</th>
<th>Verifiable Corrections (see page numbers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skotko et al. (p. 397) claimed that MacKay, Burke, et al. (1998) “assert that [H.M.’s] previously reported memory shortfalls might be compounded or explained by language comprehension and production problems.”</td>
<td>Not only did MacKay, Burke, et al. (1998) not assert this, they explicitly denied this assertion: “We are not arguing that H.M.’s memory deficits are due to his production deficits or that his production deficits are due to his memory deficits” (p. 65).</td>
</tr>
<tr>
<td>Skotko et al. (p. 398) claimed that MacKay, Burke, et al. (1998) performed “a non-blind analysis” of the Corkin (1973) transcript.</td>
<td>Six judges who were “naïve” or “blind” to speaker identity (p. 53) performed this analysis in MacKay, Burke, et al., (1998), a point reiterated several times in MacKay et al., including in the abstract.</td>
</tr>
<tr>
<td>Skotko et al. (p. 398) described the analysis of H.M.’s meaning-definitions in James &amp; MacKay (2001) as ”unbaselined.”</td>
<td>This analysis did have a baseline; meaning-definitions of four controls similar to H.M. in age, education, employment history, and IQ (see James &amp; MacKay, 2001, p. 490).</td>
</tr>
<tr>
<td>According to Skotko et al. (p. 398), MacKay &amp; James (2001, p. 448) concluded from their reading experiments that ”H.M.’s errors in novel spoken discourse were so extensive as to render his output incoherent and incomprehensible.”</td>
<td>This passage in MacKay &amp; James (2001, p. 448) referred to H.M.’s incomprehensible and incoherent output in MacKay, Burke, et al. (1998) (reference deleted from the Skotko et al. quote). MacKay and James in fact concluded that H.M. made many fewer errors in reading (2001) than in spoken discourse (1998), i.e., the opposite of this Skotko et al. claim.</td>
</tr>
<tr>
<td>According to Skotko et al. (p. 398), Corkin (1984) reported that H.M.’s “language functions were essentially unimpaired.”</td>
<td>Corkin (1984, p. 254) reported that H.M.’s language functions were impaired, with deficits relative to norms on tests of semantic and symbolic fluency and language production.</td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1. Cartoon 3 (from Larson, 1993).

Figure 2. The number of single- versus multi-word omission and substitution errors for H.M. versus the controls in Study 3.
"Oh, I don't know. Billy's been having trouble in school, and Sally's always having some sort of crisis. I tell you, Edith, it's not easy raising the dead."
H.M.’s Language Production Errors

<table>
<thead>
<tr>
<th>OMISSIONS</th>
<th>SUBSTITUTIONS</th>
<th>ADDITIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omissions</td>
<td>Omissions</td>
<td>Substitutions</td>
</tr>
<tr>
<td>H.M.</td>
<td>Memory-Normal</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>32</td>
<td>43</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Number of Errors

This PDF was created using the Sonic PDF Creator.
To remove this watermark, please license this product at wwwinvestintech.com
Appendix V.
Sections 1-2 of H.M. describing Cartoon 1 and its caption: “No, Thursday’s out. How about never – is never good for you?” (see text for section definition).

Section 1 H.M.: Well… it’s a bye- about a person talking on the telephone, got a long cord on the extension. And l- looking at his notebook, besides. (EXP: Mhm.) And naturally, the window, behind him. (EXP: Good.) And other buildings over there, see. And d- there must be a street in between ... Because he's in his office. (EXP: Good.) (mumbles) And he's, um, these here are, these here are writin', and the note was up there. (EXP: Mhm.) And a pen. (EXP: Good.) And these right here, are just the edge of the uh, um, well half the uh, you could say, um, right in the middle of the wall.

EXP: OK. Now why is that supposed to be a funny cartoon?

Section 2 H.M.: And, this is, um, be- naturally the building’s behind him, and he's talking on the, to somebody, the gentleman who wrote him. (EXP: Mhm.) H.M.: And he’s not, looking out the window (burps), excuse me, (EXP: Mhm.) or anything like that, he’s just looking at- at his date book. (EXP: Yeah.) And he has a, a couple(?) down there, and two, and then of course the signature of the person who k- made the cartoon. (EXP: That's right.) Yeah, there.

EXP: Can you read that part down there?

H.M.: Yeah… uh, and so, just what he said, over the telephone. (EXP: OK.) To the person he's talking to ... And in, and- he’s making double correction.

EXP: He's doing what?

H.M.: He's making a double correction. Because, “it's never good for you,” means that, that never been good for the person he's talking- person he's talking to. (EXP: Mhm.) And, he has stated something, he stated it about, person who's out, and he's just statin' it, the other person always ma- said that, uh…

EXP: The other person always what?

H.M.: said it was never good. (EXP: OK.) And he’s just repeatin’ something.

(EXP: OK.)

EXP: OK. Is there anything else you want to say about that, or is that enough?

H.M.: And, the buildings in the background, some of them are just- s- constructed in a way, because they aren't finished in a way. (EXP: OK. OK.) I see three of them there. (EXP: Yep.) Just a- like the girders are up. (EXP: Right, just barely, barely-) And the other parts have got the windows on them.
Appendix W.

Sections 1-2 of H.M. describing Cartoon 2 and its caption: “The beatings will continue until morale improves.” (see text for Section definition).

Section 1 H.M.: It's um ... it's about this woman talking, to the ward there. And the uh, secretary is sitting down, writing. And then the woman is supposed to be listening to her, uh, listening to what the woman is saying. (EXP: OK.) And, uh, that picture they've got in the background there, it's, um, just a picture. (EXP: Mmhm.) But they're um, a business, in a way. In the area. Or in... maybe a distant view, because the mountain area in the back.

EXP: OK. What can you say about what she's saying here?

Section 2 H.M.: Well ... she's making a comment there that the beatings will continue ... Until... morality improves. And then- she said, "morale improves." (EXP: Good.) Morality, in a way. (EXP: OK.) Instead of "l" it's a "t". It should be a "t".

EXP: OK. Do you want to say anything more about that one? Or are you done with that one?

H.M.: And their window frame is slanted. (EXP: Yes.) Maybe slanted, but- and you don't know what is inside. Looking out a window in uh, place that they have it slanted that way, or is people supposed ... to be shaped f- in any- black stuff crawled under there. (EXP: That’s true.) On the wall. (EXP: That's true.) And that the- the guys are all out there and just at their table. They're uh, having the job prints out. And the secretary is out there and she's writing it down. She's using the pen now, pencil. (EXP: OK.) The pen. And this one is doing the talking. Um, you can't tell if she is the boss of the whole bunch, or, she just uh, what.
Appendix X.
Sections 1-2 of H.M. describing Cartoon 3 and its caption: “Oh, I don’t know. Billy’s been having trouble in school and Sally’s always having some sort of crisis. I tell you, Edith, it’s not easy raising the dead!” (see text for Section definition).

Section 1 H.M.: Well, this woman is talking to this woman. And uh, and this woman’s falling down the stairs. And, uh, I can’t tell just what- she crossed me more to make it her way, only her way. They’re in her way.
EXP: Mm, which one?
H.M.: This one over here.
EXP: Wants to make everything her way?
H.M.: Yeah, she's the one that’s talking. (EXP: OK, good.) To this woman.
Because the guy is partially down on the floor. And you can’t tell exactly what it is she's telling him um… picture or uh- what. Because she's afraid of falling off her chair. Both sides. Front, inside, of the chair. And uh, he's down on the floor. And she's fallin’ down the stairs. There, in the, in the- wonder about that- the railing they have there. For the stairway.
H.M.: And listen to a- what is brace, his braces here, or in between? Just that single. One. And he can't talk. No one will take him… possibly going against it, they all- stare- well, stairwell(?)-willing to give. (EXP: Right.) And, and that is a picture of a ghost. (EXP: Mmhm.) And that's a picture of a bat. (EXP: Mmhm.) And this one, is- is drawn wrong, because that woman sittin’ there, you see the top of the chair there. And not in- her hair also. And you can't tell just what her hair color is. (EXP: Mmhm.)
H.M.: And- it is, she uh, this uh, this of hers, you see dark on this side, and, what you can't tell oh, dots on this side. (EXP: Mmhm.) Continuation of it. (EXP: That’s true.)
And this one you naturally can’t see the color of her shoes. She’s just got white… on. And you can tell that that’s a… the base of the… chair that she just bl- the- blackens the whole way, and everything, because that was one of the rules. Then, uh, then just uh, then a shoe. Because a shoe would be narrower. That is really why there’s a rule.
EXP: OK. Can you read me the part on the bottom? Can you read me what it says?
Section 2 H.M.: Can’t read the name, it starts with an “O.”
EXP: Oh yeah, the- that one (the signature of the artist). But what about the, um, the typing down at the bottom?
H.M.: Oh. (slowly) I don’t know. Gus (pronounced ‘Guzz’) having- been having trouble in school and Sally having sh- always having some sort of crisis I tell Edith, it’s not… easy, the-raising the dead.
EXP: What do you make of that?
H.M.: And that’s um, mixed up in a way, because one person is talking, and … and one person is, is uh, saying what the- another person is. (EXP: Mmhm.) Causing a lot of trouble, each crisis and everything … Gus (pronounced “Guzz”) … Gus been having trouble in school. And that’s a very young kid. See, Billy. (EXP: Mmhm.) And in school … And this is a- an older man. (EXP: Mmhm.)
EXP: OK. Good. Do you want to say more things about that or do you think that’s enough?
H.M.: And um, that is uh, you can’t tell if that is uh, ghosts of what someone- some some, because it’s white over here, and the the- don’t have any eyes on it. (EXP: Mmhm.) And… it’s sorta whitish on her both arm- her wrist. And there are parts up the- white up there on her face.
Appendix Y.

BPC rules: Applications to H.M.’s utterances

This Appendix illustrates how we applied the five correction rules outlined in the introduction and argues that all five rules were necessary and sufficient to specify a BPC among competing descriptions of H.M.’s utterances in Studies 1-3. We first discuss examples where none of the rules for deciding between alternative descriptions were in conflict. Examples [2a-c] illustrate three alternative descriptions or ways to correct two problematic aspects of H.M.’s utterance in [2]: 1). H.M. should have said her rather than him in [2] because the woman ghost doing the talking is speaking to another woman ghost (see Figure 1); and 2). The cartoon caption indicates exactly what she’s telling the other woman ghost, rendering inaccurate H.M.’s “you can’t tell exactly what it is she's telling....” in [2]. Here Rule 2 (the best fit rule) alone sufficed to select [2c] as the BPC because the other alternatives, [2a] and [2b], added more words and included fewer words from H.M.’s original utterance (see [2]).

[2]. H.M.: “And you can't tell exactly what it is she's telling him um... picture or uh-what.” (segment in H.M.’s description of Figure 1)

[2a]. “You can't tell exactly what is going on, whether because of the picture or what.” (possible correction of [2]; added words are indicated in italics)

[2b]. “You can't tell exactly what it is she's saying about him, whether because of the picture or what.” (possible correction of [2]; added words are indicated in italics)

[2c]. “You can't tell exactly what it is she's telling her, whether because of the picture or what.” (possible correction of [2]; added words are indicated in italics)

However, BPC Rule 3 (coherence with immediately prior context) was essential for deciding between four contrasting interpretations of H.M.’s utterance in [3]. Under interpretation [3a], [3] only seems ungrammatical because prosody was not encoded in our transcript: Given the list-prosody hypothesized in [3a], H.M.’s utterance “then of course, the signature of the person who made the cartoon” is error-free because standard elision rules allow deletion of subject-verb repetitions in lists. However, we rejected [3a] for two reasons: 1). absence of list-prosody: After repeated listening to the relevant section of the tape, our phonologist reported a stress and intonation pattern inconsistent with list-elision; 2). To qualify for grammatical elision (see e.g., De Smedt & Kempen, 1987), an elided subject, when “re-inserted”, must designate the original referent. However, when re-inserted into [3] (see [3c]), he refers to the creator of the cartoon rather than its original referent (a cartoon character), which is impossible on logical grounds because a cartoon character cannot create itself. In short, [3a] violates Rule 3: maintain consistency with the prior utterance context (shown here in [3b]).

[3]. H.M.: “and then of course the signature of the person who k- made the cartoon. (EXP: That's right.) Yeah, there.” (segment in H.M.’s description of Cartoon 1)

[3a]. "and then, of course, [prosodic juncture indicating elision in a list] the signature of the person who made the cartoon." (hypothetical interpretation of [3]).

[3b]. “he’s just looking at- at his date book, and he has two dates down there.” (prior context for utterance [3])

[3c]. “he’s just looking at- at his date book, and he has two dates down there, and then of course he has the signature of the person who made the cartoon.” ([3a] with non-elision prosody because we have re-inserted H.M.’s hypothetical elision of he has in italics).

Turning to interpretations [3e]-[3f], Rule 2 (best fit) favored alternative [3e] over [3f], because [3e] adds fewer words to H.M.’s original utterance (see [3]), and Rule 5 also favored
[3e] (there’s) over [3d] (here’s or this’s) because of H.M.’s tendency to repeat words in other studies (see e.g., MacKay, James & Hadley, 2008): there is repeated in [1b] [“he’s just looking at- at his date book, and he has two dates down there, and then of course, there is the signature of the person who made the cartoon.”], whereas here (or this) is unrepeat ed in [3d]. [“he’s just looking at- at his date book, and he has two dates down there, and then of course, here is (or this’s) the signature of the person who made the cartoon.”]. Rules 2, 3 and 5 were therefore necessary to chose [e] as the BPC for [3] rather than [3a], [3d] or [3f].

[3d]. "and then, of course, here's (or this's) the signature of the person who made the cartoon." (possible correction of [3]; added words are in italics).
[3e]. "and then, of course, there's the signature of the person who made the cartoon." (possible correction of [3]; added words are in italics).
[3f]. "and then, of course, you can see the signature of the person who made the cartoon." (possible correction of [3]; added words are in italics).

We now turn to [4], a more complex example where Rule 5 was essential for resolving a conflict between Rule 2 and 3. Note the absence of a main verb and the indeterminate referent for the pronoun it in [4]. Based two types of context (the cartoon in Figure 1 and H.M.’s immediately prior utterance in [4a]), H.M.’s it in [4] must refer to one of two possible descriptors for the chairs seen through the transparent women ghosts in Figure 1: dots or darkness. However, neither referent appeared in the prior context (see [4a]), and Rule 3 (maintain consistency with the cartoon and the participant’s prior utterances) favored darkness as the singular referent for H.M.’s singular pronoun it, yielding the BPC in [4b] (with added words in italics), whereas Rule 2 (add as few words and include as many words as possible from what the participant said) favored dots, yielding the BPC in [4c] (which has fewer added words than [4b]). However, Rule 5 (maintain consistency with the nature of H.M.’s syntax, prosody and errors in other transcripts) resolved this conflict in favor of [4c] (consistent with H.M.’s many pronoun agreement errors in other studies, e.g., MacKay et al., 2008; MacKay, Stewart & Burke, 1998; and Schmolck, Stefanacci & Squire, 2000) rather than [4b] (which maintains number agreement with H.M.‘s it).

[4]. H.M.: “Continuation of it.” (free-standing segment describing Figure 1)
[4a]. H.M.: “And- it is, she uh, this uh, this of hers, you see dark on this side, and, what you can't tell oh, dots on this side” (context immediately prior to [4]; see Appendix X)
[4b]. “Here is a continuation of the darkness” (BPC for [4] under Rule 3)
[4c]. “Here is a continuation of the dots.” (BPC for [4] under Rule 2; added in italics).
Appendix Z
Relational Memory theory, Binding theory and H.M.’s parallel deficits and sparing in language comprehension and production, reading aloud, visual cognition, and memory.

Node structure binding theory, or binding theory for short (MacKay, Burke et al., 1998; MacKay & James, 2001, 2002; MacKay et al., 2007, 2008; and MacKay, Stewart, et al, 1998), readily explains H.M.’s parallel deficits and sparing in language comprehension and production, reading aloud, visual cognition, and memory (see especially MacKay et al., 2007), and relational memory theory (see e.g., Cohen & Eichenbaum, 1993; Eichenbaum & Cohen, 2001; Konkel, Warren, Duff, & Cohen, 2008; Shimamura, 2002; and Shimamura & Wickens, 2009) also explains aspects of these same parallels. Under relational memory theories, forming internal representations for familiar items such as “names, faces, or stimuli” engages non-hippocampal MTL areas, whereas forming arbitrary relations among constituent elements of events or scenes, including spatial relations, associative relations (e.g., the name associated with an object), and temporal relations (e.g., what follows what in an event sequence) engages hippocampal areas, so that relational binding suffers in amnesics with hippocampal damage (and perhaps also entorhinal, perirhinal, and parahippocampal damage; see Shimamura & Wickens, 2009).

Binding theory is more specific than relational memory theories in some respects and less specific in other respects. For example, binding theory specifies in detail the internal representations for semantic, phonological, and orthographic units and their neural instantiations in the cortex. Binding theory also specifies the structure of the connections between those units and the role of different types of MTL binding mechanisms in forming those connections (see MacKay, Stewart et al., 1998; MacKay & James, 2001; and MacKay et al., 2007). Binding theory also specifies what units and processes underlie both correct retrieval or use of an internal representation and incorrect retrieval, as in minor errors (see MacKay, 1987, pp. 49-61). Binding theory also specifies how aging and repeated retrieval affects those units and processes over the course of a lifetime (see James & MacKay, 2001; MacKay & James, 2002; and MacKay & Hadley, 2009). However, unlike relational memory theories, binding theory makes no attempt to specify what types of binding mechanisms occupy what brain loci, say, the hippocampus for episodic binding mechanisms to encode novel space-time-events versus the parahippocampal cortex for phonological binding mechanisms to encode never-previously-encountered syllables.

Mechanisms underlying item versus relational memory also differ in binding theory versus relational memory theories. Under relational memory theories, binding plays no role in encoding familiar items in episodic memory tasks, unlike binding theory, where binding is central to item encoding because what must be encoded is not the familiar item per se but the fact that it appeared in a just-presented list or task: Item recall therefore depends on forming a connection between internal representations for the item and its list- or task-context under binding theory (see e.g., Hadley & MacKay, 2006). Because new-learning deficits hinder the formation of such connections, binding theory (but not relational memory theories) readily explains why amnesics with hippocampus-specific damage exhibit reliable deficits relative to controls in item recall and recognition (see e.g., Konkel et al., 2008). Binding theory also explains why these same amnesics exhibit larger deficits in recalling relations than items (see Konkel et al.) because encoding the spatial relations examined in studies to date requires the formation of more new connections than encoding an item as occurring-in-the-task (see MacKay & James, 2009, for discussion of the number of new connections required to represent seemingly simple but never previously encountered objects in visual scenes).
This point highlights another difference between binding theory vs. relational memory theories. Unlike earlier relational memory theories (e.g., Lindsay & Norman, 1977, pp 383-411), current relational memory theories are under-specified: They neither define the concept of relations in general nor specify what theoretical (or neural) units encode specific spatial or non-spatial relations. Moreover, current definitions of relational binding fail to explain many of the deficits and sparing observed in amnesics. For example, contrary to the hypothesis that representing arbitrary relations always suffers in amnesics with hippocampal damage, H.M. exhibits deficits in processing arbitrary figure-ground relations involving unfamiliar but not familiar figures (see MacKay & James, 2009) in a task that is central to everyday perception (hidden figure detection; see Minsky, 2006, Thurstone, 1949), and hippocampal amnesics exhibit deficits in imagining new experiences such as “lying on a white sandy beach in a beautiful tropical bay,” where the relations between you, lying, beach and bay are not completely “arbitrary” (Hassabis, Kumaran, Vann & Maguire, 2007). Also challenging for relational memory but not binding theory, hippocampal amnesics can discriminate perceptually between familiar objects (insects) but not unfamiliar objects (blobs) (Barense et al., 2005; see also Barense, Gaffan & Graham, 2007; Lee, Barense & Graham, 2005; Lee et al., 2005).

Finally, relational memory theories fail to specify the processing procedures for encoding or retrieving relational information, and as a consequence, can neither predict nor explain the nature of the processing errors that occur during planning or encoding (major errors) and during retrieval (minor errors). Nor can current relational memory theories predict or explain why H.M.’s speech errors are usually anomalous and virtually never corrected whereas those of normal speakers are usually corrected, rarely anomalous, and virtually always corrected when anomalous (see Studies 1 and 2). Nor can current relational memory theories predict or explain why omissions are relatively more common but substitutions relatively less common for H.M. than controls; why single-word omissions are relatively less frequent and multi-word omissions relatively more frequent for H.M. than controls; and why single-word substitutions violate the syntactic class regularity relatively more often for H.M. than controls (see Study 3).

Moreover, even with added assumptions, relational memory theories fail as post hoc accounts of H.M.’s language production deficits. Under one such account (Shimamura, personal communication), H.M.’s speech was impaired in Studies 1 and 2 because maintaining the narrative conception of a cartoon and discussing it at the same time exceeded H.M.’s working memory capacity, unlike his everyday conversational speech. One problem with this relational-memory-capacity-limitation account is that double blind procedures indicate that H.M.’s conversational speech production (an ecologically valid task par excellence) is reliably impaired (except when producing clichés, responding “yes” or “no” to a question or repeating frequently repeated information; see MacKay, Burke et al., 1998). A second problem arises when considering the parallels between H.M.’s speech and reading aloud. To illustrate one such parallel, the three reliable idiosyncrasies that characterized H.M.’s speech errors in Studies 1 and 2 (non-correction, anomaly and omissions) also characterized his errors when reading isolated words aloud, as in [66a-c] from MacKay and James (2002) and [67a-c] from MacKay & Hadley (2009), where the stimulus word appears to the left, and H.M.’s reading error (in quotes) follows an arrow signifying “was misread as.” Parsimony therefore demands the same explanation for H.M.’s parallel idiosyncrasies (non-correction, anomaly and omissions) in speech and reading. However, explaining these parallels under the relational-memory-capacity-limitation account requires the absurd assumption that reading an isolated word aloud somehow exceeds H.M.’s working memory capacity, unlike producing a lengthy sentence from memory, e.g., “she can’t do
ANYthing, however, she has to do EVERYthing” in [1a]. By contrast, the damage to H.M.’s binding mechanisms in binding theory (with no new or added assumptions, absurd or otherwise) suffices to explain the many parallels between H.M.’s deficits and sparing in speech production, reading aloud, sentence comprehension, visual cognition, and episodic and procedural memory (without assuming that his other deficits depend on or stem from his memory deficits).

[66a]. *adumbrate* → “embryate” (uncorrected non-word that omits three speech sounds)
[66b]. *ellipsis* → “glipis” (uncorrected non-word that omits two speech sounds)
[66c]. *papyrus* → “papus” (uncorrected non-word that omits two speech sounds)
[67a]. *abolitionist* → “abolis” (uncorrected non-word that omits five speech sounds)
[67b]. *infinity* → “fis•tee” (uncorrected non-word that omits four speech sounds)
[67c]. *pretzel* → “PRET•zee” (uncorrected non-word that omits two speech sounds)