Auditory Serial Position Effects in Story Retelling for Non-Brain-Injured Participants and Persons With Aphasia

Martin B. Brodsky
VA Pittsburgh Health Care System, Pittsburgh, PA
University of Pittsburgh, Pittsburgh, PA
and
Medical University of South Carolina, Charleston, SC

Malcolm R. McNeil
Patrick J. Doyle
Tepanata R. D. Fossett
VA Pittsburgh Health Care System
and
University of Pittsburgh

Neil H. Timm
University of Pittsburgh

Grace H. Park
National Institute on Deafness and Other Communication Disorders, Bethesda, MD

Using story retelling as an index of language ability, it is difficult to disambiguate comprehension and memory deficits. Collecting data on the serial position effect (SPE), however, illuminates the memory component. This study examined the SPE of the percentage of information units (%IU) produced in the connected speech samples of adults with aphasia and age-matched, non-brain-injured (NBI) participants. The NBI participants produced significantly more direct and alternate IUs than participants with aphasia. Significant age and gender differences were found in subsamples of the NBI controls, with younger and female participants generating significantly more direct IUs than male and older NBI participants. Alternate IU productions did not generate an SPE from any group. There was a significant linear increase from the initial (primacy) to the final (recency) portion of the recalled alternate IUs for both the NBI group and the group of participants with aphasia.

Results provide evidence that individuals with aphasia recall discourse length information using similar memory functions as the nonimpaired population, though at a reduced level of efficiency or quantity. A quadratic model is suggested for the recall of information directly recalled from discourse-length language material.

KEY WORDS: serial position effect, memory, aphasia, story retelling, discourse

Aging brings a predisposition to injury and disease that can affect cognitive and language functions. Among the most common diseases/disorders are dementia, head injury, and stroke. Investigators have reported decreased performance, even without injury or disease, in the comprehension of spoken (e.g., Burke & Laver, 1990; Wingfield & Stine-Morrow, 2000) and written language (e.g., Adams, Smith, Nyquist, & Perlmutter, 1997). Memory for language (see Balota, Dolan, & Duchek, 2000, and Anderson & Craik, 2000, for reviews) has also been shown to be associated with increased age. The serial position effect (SPE) stands in bold relief among the many metrics used for quantifying and investigating the underlying nature of memory impairment.

The SPE is the phenomenon whereby information occurring early and late in a series is remembered better than information presented in the middle of the sequence. According to Crowder and Greene (2000), the SPE has its roots in serial learning theory dating back to the late 1800s. Groeger (1997) described the pattern that typically appears from

1092-4388/03/4605-1124

SPEs as a U-shaped curve in which there is an increased likelihood to recall the first items (primacy effect) and the last items (recency effect). This effect has been demonstrated in non-brain-injured (NBI) participants for various recall activities including the presentation of lists of words (Capitani, Della Sala, Logie, & Spinnler, 1992; Carlesimo, Sabbadini, Fadda, & Caltagirone, 1997; Glanzer & Cunitz, 1966; Korsnes & Magnussen, 1996; Murdock, 1962), faces (Bruyer & Vanberten, 1998), nonverbal patterns (Korsnes & Gilinsky, 1993; Korsnes & Magnussen, 1996), questions probing information from paragraphs or dialogue (Newhouse & Holen, 1975; Roberts, 1966; Yasuda, Nakamura, & Beckman, 2000), and free recall of text or discourse-length material (Freebody & Anderson, 1986; Roberts, 1966). Whereas data are abundant from memorized lists, only a few studies have addressed the serial position effects of recall from connected discourse, either in NBI populations (Meyer & McConkie, 1973; Newhouse & Holen, 1975; Roberts, 1966) or in those with disorders (Hall & Bornstein, 1991).

In one connected speech study, Newhouse and Holen (1975) presented a 20-min audiotaped lecture to 25 graduate students. Participants were then presented with three sets of 10 randomized multiple-choice questions. Questions were derived from evenly spaced material throughout the lecture and selected from the beginning, middle, and end of the presentation. An SPE was demonstrated as initial and final sections of answers being recalled significantly more accurately than the middle section.

Roberts (1966) orally presented two subtests from the Stanford-Binet Intelligence Scale (Memory for Stories: *The Wet Fall* and Repeating Thoughts of Passages: *Value of Life*) to two different groups. Both groups' recall scores were based on seven ideas from the four sentences in the paragraph and both subtests yielded a reverse SPE—the material in the middle of the passages was recalled significantly more often than the material in the beginning and end of the passages. Roberts suggested that the difficulty of the items at the beginning and end of the passages and the level of meaningfulness between items were the probable causes for the unusual pattern.

Similar to Roberts (1966), Meyer and McConkie (1973) completed a study in which three groups of undergraduate students enrolled in an introductory psychology course listened to two paragraphs chosen from *Scientific American* magazine. The number of ideas recalled from their written account of each of the passages was examined. One of the two passages yielded no SPE, whereas the expected U-shaped curve was found for the other. Thus, the presence of the SPE was inconsistent with NBI participants in this study. They concluded that logical structure (i.e., ideas related to other ideas) in passages was a necessary component to show an SPE.

The former studies examined participants without impairments. To date, the only study to have addressed SPEs for narrative paragraph recall in a disordered population was conducted using participants with mild closed-head injury and an education and age-matched group of NBI controls (Hall & Bornstein, 1991). Each participant was administered Story A from the Wechsler Memory Scale–Revised (Wechsler, 1987) for immediate oral recall. The 25 items in this subtest were divided into thirds representing the primacy, middle, and recency portions. Whereas both groups demonstrated a significant primacy and recency effect, significantly fewer items were recalled by the persons with aphasia.

It has been found that the nature of the stimulus material has an effect on the elicitation of an SPE with discourse-length stimuli in both NBI and brain-injured populations. This phenomenon, therefore, has the potential to inform the contributions of memory in connected speech samples where the target discourse is known and to further the understanding of such targets, such as the Story Retell Procedure (SRP) described by Doyle, McNeil, and colleagues (Doyle et al., 1998, 2000; Hula, McNeil, Doyle, Rubinsky, & Fossett, 2003; McNeil, Doyle, Fossett, Park, & Goda, 2001; McNeil, Doyle, Park, Fossett, & Brodsky, 2002). The SRP is a standardized procedure during which individuals provide an immediate oral retell for each of 12 audio-recorded stories originally derived from the Discourse Comprehension Test (Brookshire & Nicholas, 1993). The number of possible information units (IUs) is predetermined for each story, and information transfer is measured by the percentage of IUs produced relative to those in the stimulus story (McNeil et al., 2001). An IU is defined as a word or word string that is intelligible, informative, and provides accurate content relevant to the stimulus story. They are coded as specific words reproduced from a story (direct IUs) or as their legitimate synonyms (alternate IUs). Alternate IUs are potentially interesting because they are believed not to be retrieved simply from phonological code, and they reflect a deeper level of lexical–semantic processing. The distinction between direct and alternate IUs holds a possibility of illuminating differential mechanisms for word retrieval failure.

Performance on the SRP is limited by the amount of information the reteller comprehends and remembers, the formulation and production requirements of this task, and the language limitations of the reteller. One way to determine and describe the memory limitations of the reteller is to investigate the presence of an SPE. Whereas the presence of an SPE would suggest the combination of short-term (recency effect) and long-term (primacy effect) memory processes, the absence of this effect would imply that different memory processes or strategies are required and/or operative during this task.

Decreased performance on linguistic memory tasks is well documented in individuals with aphasia (Burgio & Basso, 1997; McNeil, 1988; Ostergaard & Meudell, 1984), and an SPE for word lists in persons with aphasia (Ostergaard & Meudell, 1984) has been found. However, no known study has investigated the SPE in a connected language story retell task with this population.

In order to further define the comprehension, memory, and production limitations of persons with aphasia presented with connected language material, and to further explore the processing demands of the SRP, an SPE was investigated for %IU production in NBI controls and in individuals with aphasia. Given that the SPE has been demonstrated over a wide array of memory tasks and in both NBI controls and persons with brain injury, we predicted that an SPE would be evident for both groups. We hypothesized that the amount of information produced and the slope of the SPE would be reduced in the group of participants with aphasia relative to the NBI control group and, based on data from Hall and Bornstein (1991) and McNeil (1988), that there should be a difference in performance characteristics for recall between direct (those stated in the stimulus story) and alternate (legitimate synonym) IUs over the course of the story. As such, we expected that no SPE would be demonstrated for alternate IUs because they are not directly input into the working memory system, but are acceptable semantic derivatives of the direct IUs.

Secondary hypotheses and analyses were also formed. Based on the sizeable literature supporting age and gender differences in verbal memory, these differences were investigated for NBI controls. We predicted that young NBI controls would produce significantly more direct IUs than older NBI controls (Baddeley, 1999; Hess & Arnould, 1986; Schugens, Daum, Spindler, & Birbaumer, 1997; Stine, 1990) and that NBI females would produce significantly more direct IUs than NBI males (Maitland, Intrieri, Schaie, & Willis, 2000; Vacanti, Hamm, Cammeron, & Peterson, 1977). Because of the relatively small sample of participants with aphasia, no age or gender predictions were made for this investigation. Gender and age differences were not expected for alternate IUs for NBI controls.

Method

Participants

Thirty-one adults volunteered as controls. Using questionnaires, we screened these participants for positive histories of neurological impairment. Participants reporting experiences with speech and/or cognitive impairments due to stroke, head injury, or neurological disease were excluded. NBI controls passed the Hearing Handicap Inventory for Adults (Newman, Weinstein, Jacobson, & Hug, 1990; Ventry & Weinstein, 1982) and showed no more than a two item difference in retell performance from immediate to delayed retellings on the story retelling subtests of the Arizona Battery for Communication Disorders of Dementia (Bayles & Tomoeda, 1993), performance that is consistent with normal immediate and short-term memory skills. No controls were excluded because they failed to meet these two screening criteria.

Fifteen native English-speaking adults with mild to moderate aphasia, as defined by McNeil and Pratt (2001) and as measured by the Porch Index of Communicative Ability (Porch, 1981; M = 79th percentile, range = 53rd to 97th percentile), volunteered for this study. All participants with aphasia passed a pure-tone audiometric screening at 35 dB HL at 500, 1000, 2000, and 4000 Hz. Table 1 contains further descriptive information on all participants with aphasia.

Instrumentation

The 12 SRP stories were presented randomly to each participant for this investigation. Participants were seated comfortably in front of a computer with a 15-in. monitor for the presentation of the stories. A unidirectional microphone connected to the computer was used

Table 1. Participant descriptions.

Participant	Age	PICA VRB percentile	PICA AUD percentile	PICA OA percentile	RTT (5-item) percentile
1	62	78	73	92	73
2	67	63	72	59	19
3	47	54	64	65	4
4	51	60	99	87	3
5	69	86	99	85	77
6	56	89	99	87	95
7	74	97	99	94	96
8	55	71	72	75	63
9	67	76	72	80	94
10	57	75	99	86	58
11	65	78	69	86	54
12	71	37	54	43	5
13	52	91	99	87	80
14	74	70	99	78	66
15	74	54	54	63	21
M	62.73	71.93	81.53	77.80	53.87
SD	9.14	16.19	17.85	14.30	34.50

Note. PICA = Porch Index of Communicative Ability (Porch, 1981), percentile compared to adults with left hemisphere damage; VRB = verbal; AUD = auditory; OA = overall; RTT = Revised Token Test (Arvedson, McNeil, & West, 1986; McNeil & Prescott, 1978; Park, McNeil, & Tompkins, 2000), percentile scores for adults with left-hemisphere damage.

to record the participants' retellings. The stories were presented at 170 syllables per minute by a male speaker at approximately 70–75 dB SPL via computer speakers placed on either side of the monitor in a quiet environment. This level was measured by a sound-level meter placed at the participant's ear and calibrated for each participant before the task.

Procedure

Prior to the presentation of the story, participants were instructed via a digital recording as to the nature of the task and the need to retell the story in their own words following its completion. During the story, six full-screen, black and white illustrations were presented in temporal association with the plot of the story to all participants. At the completion of the narration, instructions to retell the story in their own words, via a digital recording, were again presented to every participant and immediately followed by a computer screen showing a composite of all 6 pictures shown as smaller versions of the original pictures in two rows of three pictures each. This method of concurrent picture presentation has been examined in previous research and has shown no statistically significant effect with or without the pictures during story presentation and story recall (Doyle et al., 1998). However, a trend of increased production of IUs from memory was noted with pictures present during story presentation and story recall.

Direct and alternate IUs were identified for each story (M = 152, SD = 14.5, range = 111–162). Trained scorers orthographically transcribed the retellings from the recordings of each story from each participant. McNeil et al. (2001) reported percentage agreement averages for interjudge reliability of 96% for coding both NBI and aphasic participant retells. Point-to-point reliability was calculated as number of agreements divided by the number agreements + disagreements for each of the four raters for each measure. In a second reliability study, Hula et al. (2003) categorized IUs from each participant's retelling as either a direct IU or alternate IU. They reported intraclass correlation coefficients of .993, .979, and .885 for NBI participants, and .995, .986, and .944 for participants with aphasia, for total, direct, and alternate %IUs/minute, respectively. Point-to-point reliability averaged 91% (range = 85%–95%) for both NBI and aphasic participant retells.

Analysis

For statistical purposes, a transformation was made from the original number of IUs in each story to an equivalent scale across all 12 stories. This was accomplished by first listing the IUs in the order in which they occurred in each story, then numbering them serially, creating a position number. Once each story's IUs were numbered, the IUs' position numbers were converted to a decimal based on the total number of IUs from the first story (see Equation 1). In effect, the decimal derived from this conversion became the new rank order for each IU relative to all 12 stories. After all 12 stories' IUs were converted using Equation 1, all 1,819 IUs derived from the stimulus stories were sorted based on their newly calculated position, ultimately interlacing the IUs from all 12 stories. Once the IUs were interlaced, a 20-point moving average across all 12 stories was computed based on the results of the autoregressive iterative moving average (ARIMA) procedure in SAS/ETS (SAS Institute Inc., 1999).

$$\frac{\left[\text{Total number of IUs in Story 1 (111)}\right]}{\left[\text{Total number of IUs for the story being converted}\right]}\left(\text{IU position in the story}\right) = \text{A decimal replacing an IU's original position in a story with the position relative to the IUs for } \textit{all 12 stories} \quad (1)$$

IUs were believed to be interdependent within sentences and within stories due to their lexical relationship from one to the next. Simply, the predicate of a sentence is dependent on the subject of the same sentence, and subsequent sentences are dependent on the previous sentences to maintain the plot of the story. As such, high and positive correlations between IUs were predicted. The ARIMA procedure was used in an attempt to remove these autocorrelations. The ARIMA procedure also allowed for the assignment of a moving average to the data whereby the function for data presentation could be smoothed. To reduce the amount of dependence in the data while maximizing the smoothing for data presentation, experiments used the autocorrelation from the ARIMA procedure showing the least correlation between data points. An ordinary least squares model for regression could not be used because the model assumes independent data, and our data clearly violate this assumption. Instead, the autoregression (AUTOREG) procedure in SAS/ETS (SAS Institute, 1999) was used. Both linear and quadratic autoregression models were attempted for best fit in each of the participant groups and for both direct and alternate IUs within each of the six participant groups (e.g., aphasic, controls, young controls, elderly controls, male controls, female controls). The presence (or absence) of a serial position effect was determined by the model that best fit the data.

The system of linear regression equations (SYSLIN) procedure in SAS/ETS (SAS Institute, 1999) using Bonferroni corrections was used to test differences in both direct and alternate IU models. All descriptive and inferential statistical analyses were completed using SPSS (SPSS Inc., 2001). Statistical significance was set at $\alpha \leq .01$ for normality and overall significance was set at $\alpha \leq .05$.

NBI participants were separated into a young group and an elderly group by the median age of the control participant pool. NBI controls were also divided into groups by gender. A chi-square analysis was used to determine whether there was a significant difference in the composition of this subdivision of the control group by gender. Paired t tests were used to determine the presence of significant differences in age and education between groups. A total of five tests between models were completed to determine significant differences—NBI controls versus participants with aphasia, young controls versus participants with aphasia, elderly controls versus participants with aphasia, young controls versus elderly controls independent of gender, and male controls versus female controls independent of age.

Results

Figures 1 through 5 depict the 20-point moving average for the participant groups with their respective recall of direct IUs and alternate IUs. The vertical bars in each of these graphs represent the average frequency of recalled IUs (direct and alternate, respectively) across all 12 stories after the IUs were interlaced using Equation 1. Direct and alternate IUs were interlaced separately. Consequently, these two sets of IUs are depicted separately in the graphs. Regression lines for each of these 20-point moving averages are also contained in each figure. The U-shaped curves represent quadratic regression equations for the IUs presented in the same color. Straight lines represent linear regression equations for the IUs presented in the same color. Results for each comparison are as follows.

NBI Control Participants Versus Participants With Aphasia

The 31 NBI control participants consisted of 15 males and 16 females ranging in age from 22 to 80 years (M = 43.7 years, SD = 17.2 years). The 15 participants with aphasia, 11 male and 4 female, ranged in age from 47 to 74 years (M = 62.7 years, SD = 9.1).

Figure 1. Direct and alternate information units (IU)s with corresponding regression lines for all 31 non-brain-injured (NBI) controls and all 15 participants with aphasia.

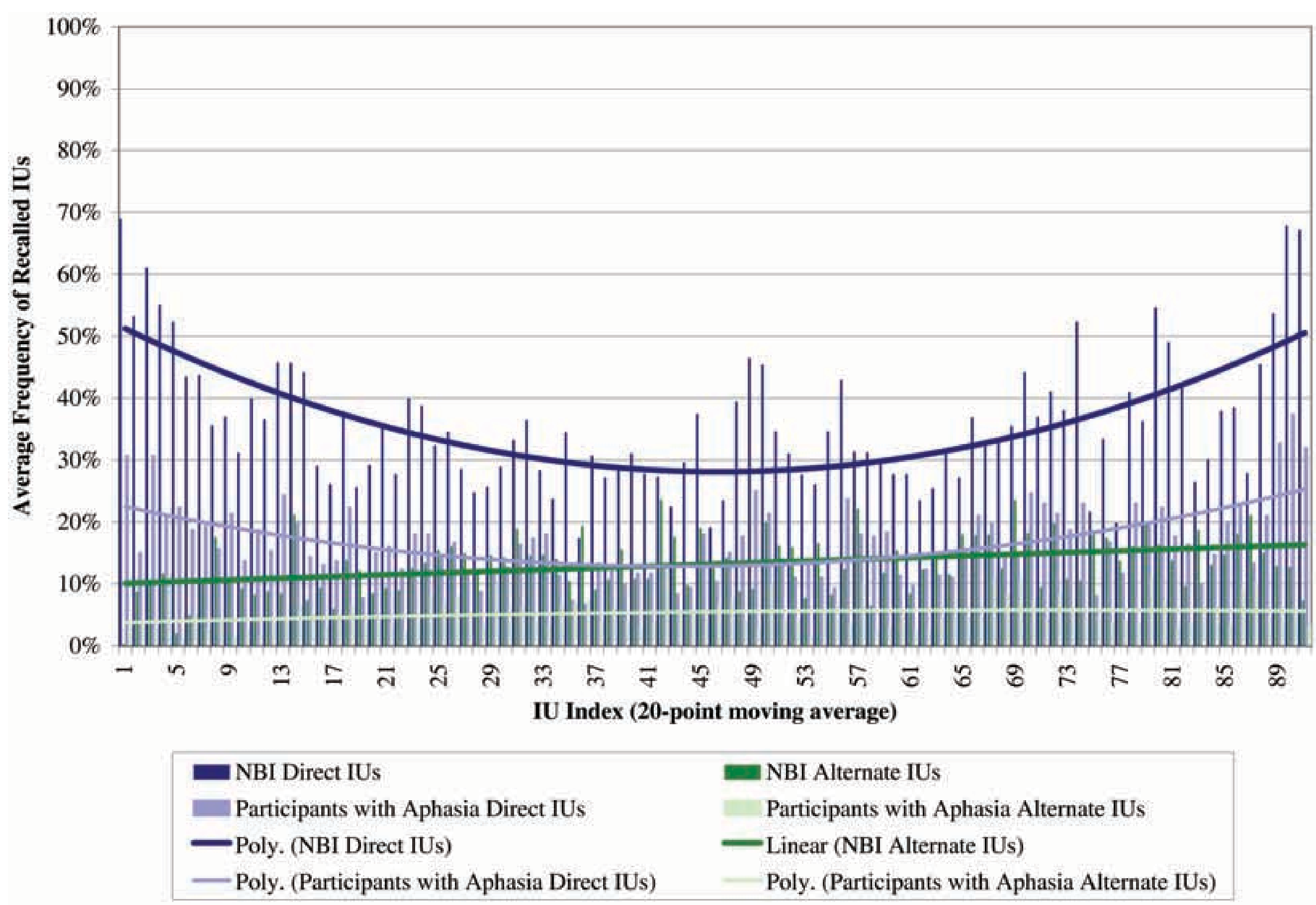
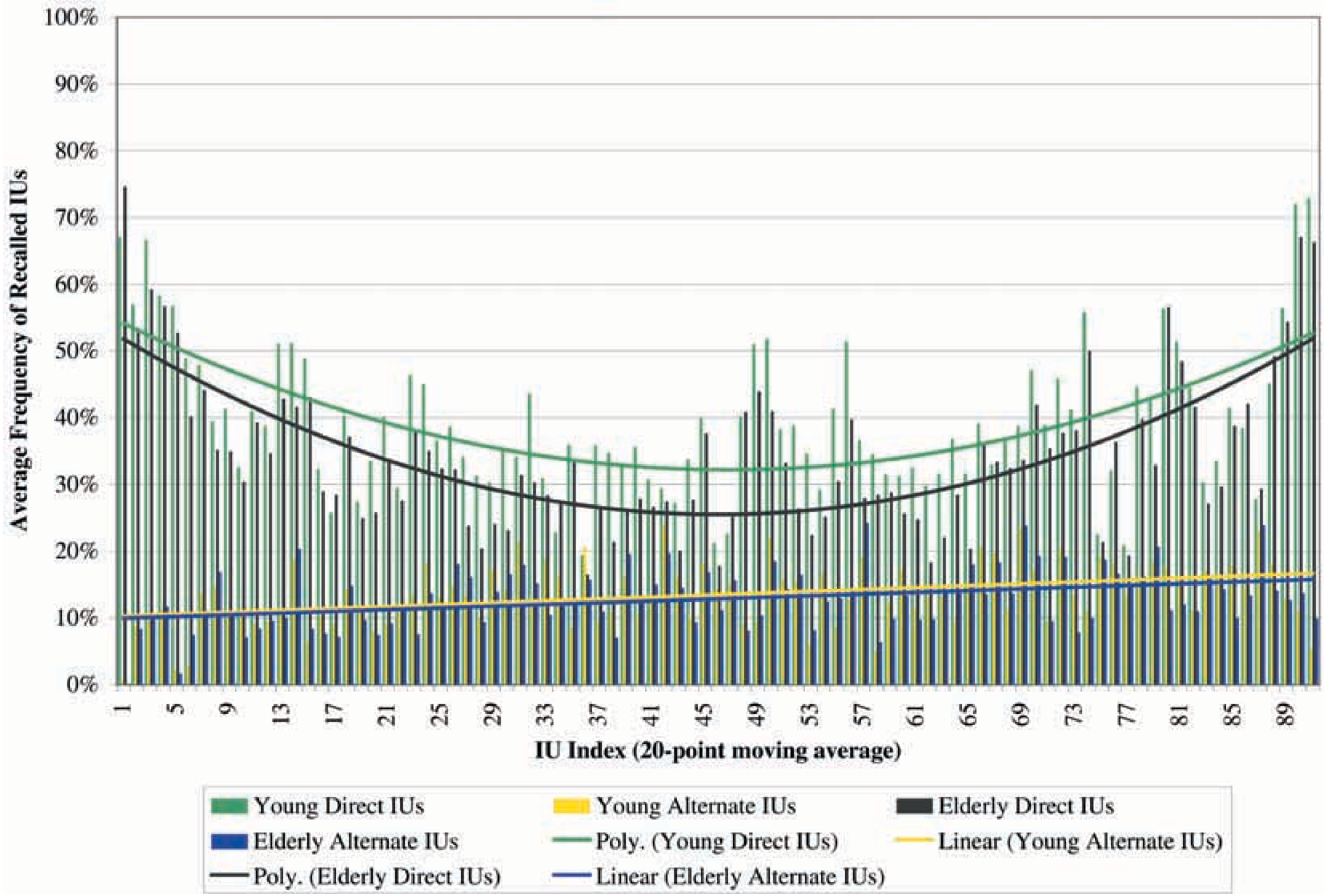


Figure 2. Direct and alternate IUs with corresponding regression lines for all 16 young and 15 elderly NBI controls.

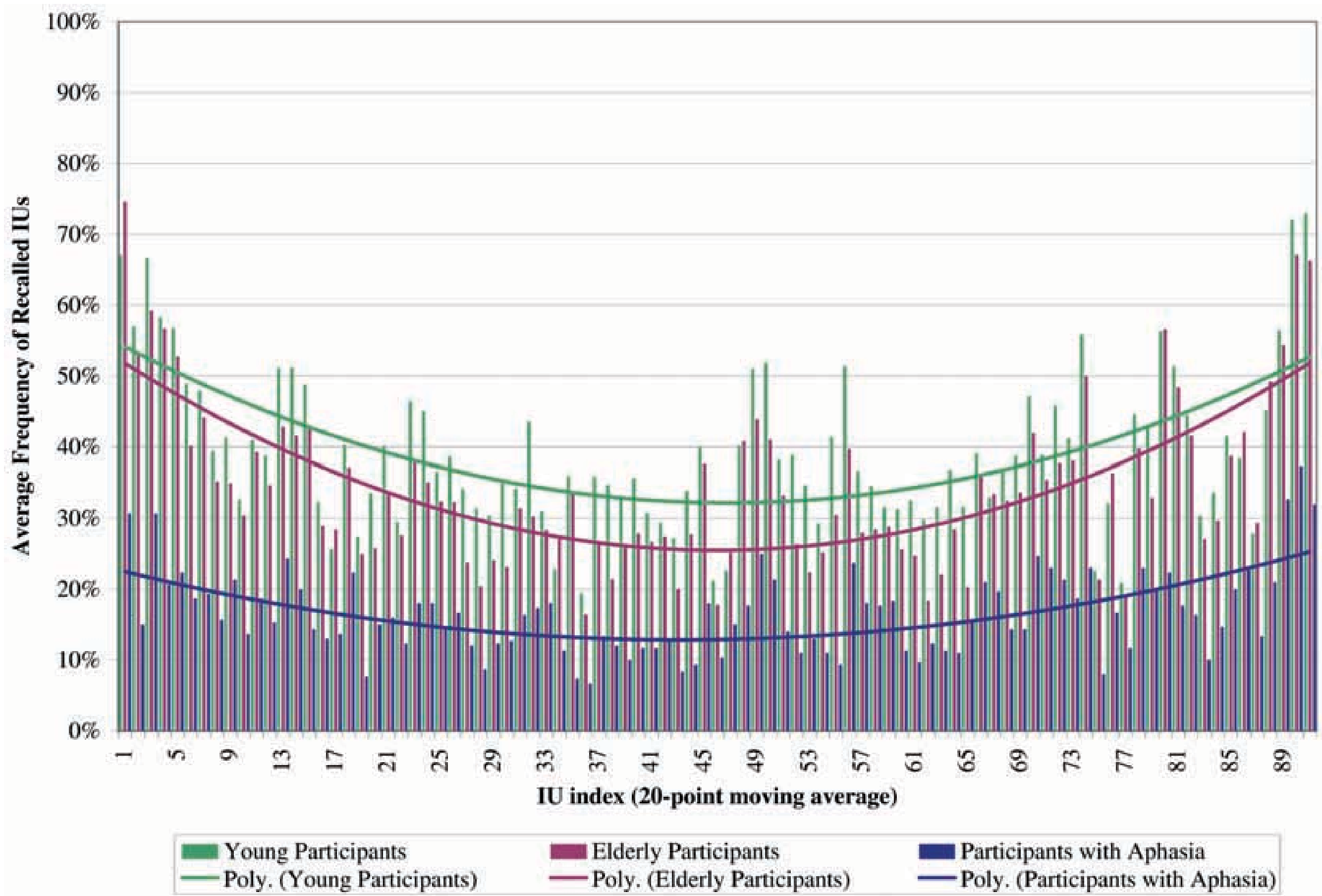


The results of the analyses performed on the direct IUs for both NBI controls and participants with aphasia showed high levels of correlation between IUs. The ARIMA procedure results showed the autocorrelations between direct IUs for the NBI participants to decrease from .55 between any 2 IUs to .04 when separated by 20 IUs. In order to utilize the lowest correlation coefficient between IUs, to produce the smoothest function curve and make all data sets equivalent, 20 IUs were chosen as the criterion to define the moving average for all four sets of data (direct and alternate IUs for NBI controls and direct and alternate IUs for participants with aphasia). The moving average was computed by considering the IUs from all 12 stories (N = 1,819) and dividing them serially, according to the newly computed position order (see Equation 1) by every 20 IUs, thus reducing these four data sets to 91 data points. Therefore, each point in the newly created IU index represents the 20 IUs, produced in their serial order, from all 12 stories.

The ARIMA procedure reduced the autocorrelation in each series, though it did not completely eliminate the dependence within the data sets. The direct IUs for both groups of participants showed stationary time series and an autocorrelation close to 0 with a 20-point lag (correlations computed at 20 consecutive IU points; R^2 = .04 for control participants and .12 for participants with aphasia). Significant autocorrelations [AR(1) terms] of −.40 for control participants and −.25 for participants with aphasia remained after the data set was transformed to the 91-point IU index for the direct IU series (the primacy end of the index is defined as the first 30 IUs, and the last 30 IUs comprise the recency end). Hence, the resulting series were analyzed using regression models with autocorrelated errors. Both linear and quadratic autoregression models were fit to each series, though the quadratic series still included significant negative AR(1) terms outlined above. We examined higher order models that did fit the data marginally better, however, in this case, the additional variance accounted for by a cubic term was small and the associated standard error of the coefficient was large. Consequently, we limited the expressions to a quadratic model.

A quadratic autoregressive regression line fit to the direct IUs from control participants had an R^2 value of .50, and the function for the participants with aphasia resulted in an R^2 of .37. These functions depicted SPEs

Figure 3. Direct IUs with corresponding regression lines for young and elderly NBI controls and participants with aphasia.



for both groups' production of direct IUs. The alternate IUs for control participants and participants with aphasia had linear regression lines (no SPE) with small values for R^2 (.21 and .06, respectively) and a significant ($p < .05$) slope. The NBI controls' slope changed from an average production of alternate IUs at the primacy end of 10.7% to 15.2% at the recency end, whereas the participants with aphasia changed from 4.6% to 5.9%, respectively. The histograms and their respective regression lines for the four data sets of all participants are seen in Figure 1.

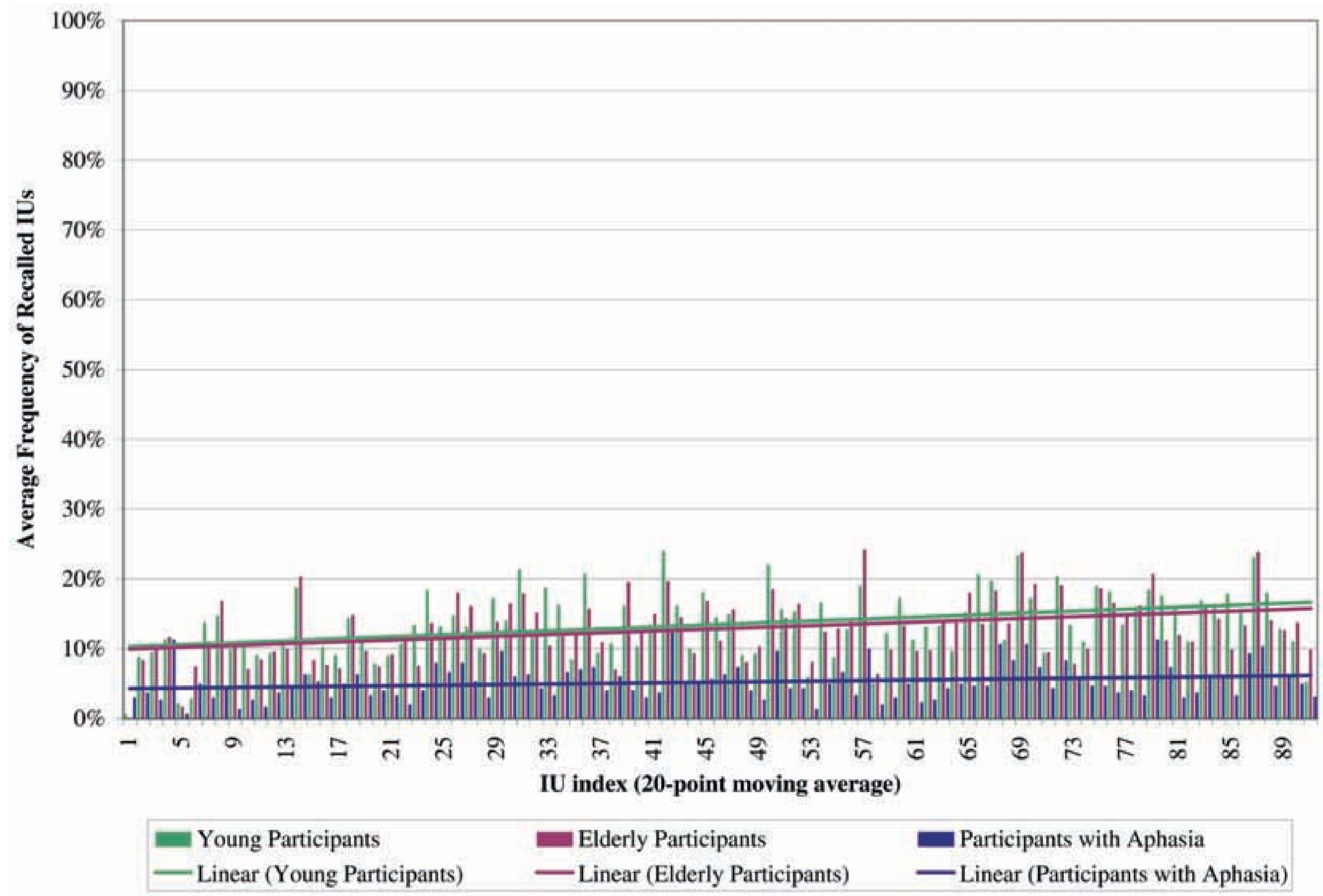
NBI controls produced a combined average (direct + alternate IUs) of 52% (*SD* = 5.23%, range = 42%–62%) of the total possible IUs in all stories. Specifically, 36% of all possible IUs were recalled as direct IUs (*SD* = 10.77%; range = 17%–69%) and 13% were recalled as alternate IUs (*SD* = 4.62%; range = <1%–24%) for the 12 stories.[1] Participants with aphasia produced a combined average (direct + alternate IUs) of 22% (*SD* = 3.56%, range = 15%–26%) of the total possible IUs in all stories. An average of 17% (*SD* = 6.02%; range = 7%–37%) were recalled as direct IUs and 5% (*SD* = 2.52%; range = <1%–12%) were recalled as alternate IUs across all stories. The direct IU and alternate IU mathematical models for control participants and the participants with aphasia were tested for differences using protected *t* tests with the Bonferroni correction. Significant differences, $F(3, 176) = 404.50$, $p < .05$, were found between the two participant groups for both direct IUs and alternate IUs.

Young NBI Participants Versus Elderly NBI Participants

The 31 participants in the NBI participant pool were divided by their median age (40 years) into two groups: young NBI controls and elderly NBI controls. The younger group, consisting of 16 members, had an average age of 29.69 years (*SD* = 4.95 years, range = 23–40 years). The older group had 15 members with an average age of 58.60 years (*SD* = 11.93 years, range = 42–80 years). There was a statistically significant difference

[1] Direct IUs and alternate IUs do not add up to 52% due to rounding errors.

Figure 4. Alternate IUs with corresponding regression lines for young and elderly NBI controls and participants with aphasia.



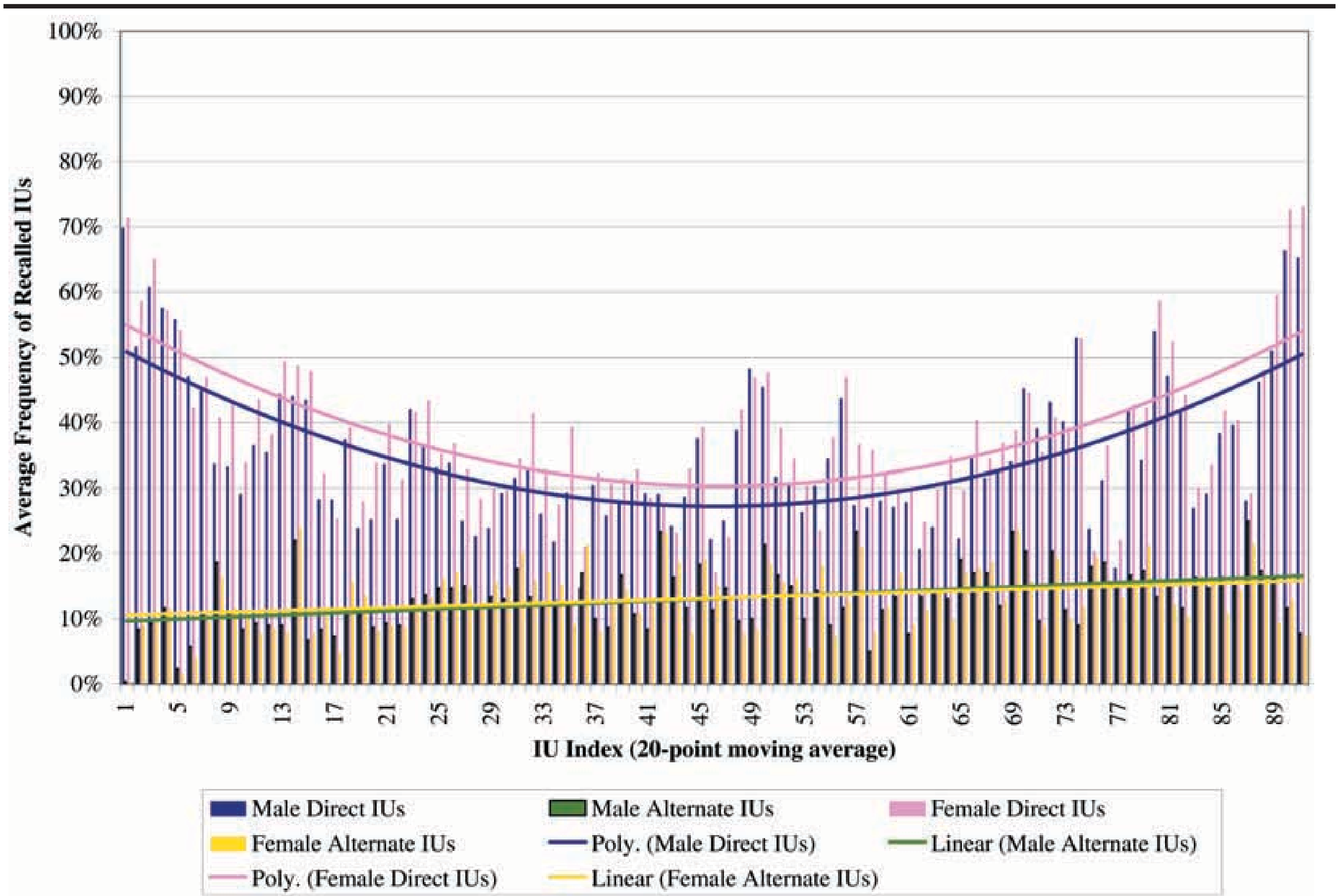
in age, $t(14) = -7.616$, $p < .05$, between the two groups. Additionally, the younger participants had an average of 16.69 years of education (*SD* = 3.05 years) and the elderly participants had an average of 15.27 years (*SD* = 3.56 years). The difference in education between the two groups was small and nonsignificant, $t(14) = 1.058$, $p > .05$.

The direct IUs for both the young and elderly control participant groups showed the same stationary time series and a low autocorrelation close to zero with a 20-point lag (correlations computed at 20 consecutive IU points; $R^2 = .02$ for each). A significant AR(1) term of –.42 for the young controls and –.36 for the elderly controls remained for the direct IU series.

A quadratic autoregressive regression line was fit to the direct IUs, though the quadratic series still included a significant negative AR(1) term. The young control participants had an R^2 value of .45, and the function for the elderly control participants resulted in an R^2 of .55. These functions depicted SPEs for both groups' production of direct IUs. Alternate IUs for both young and elderly control participants continued to show linear trends (no SPE) with small values for R^2 (.20 and .19, respectively) and significant ($p < .05$) slopes. Young NBI participants' production of alternate IUs averaged 10.9% at the primacy end, and the recency end averaged 15.5%. The elderly participants produced averages of 10.5% and 14.6%, respectively. The histograms and their respective regression lines for these four data sets are presented in Figure 2.

Young control participants produced an average of 39% of the total possible direct IUs (*SD* = 11.09%; range = 19%–73%) and 13% of alternate IUs (33% of the direct IUs; *SD* = 4.80%; range = <1%–24%) for each story. Elderly control participants produced 35% of direct IUs (*SD* = 11.53%; range = 16%–75%) and 13% of alternate IUs (37% of the direct IUs; *SD* = 4.53%; range = 0%–24%) for each story. A significant, $F(3, 176) = 49.56$, $p < .05$, difference was found between the young and the elderly control participants for direct IU production. No significant difference, $F(2, 178) = 0.5047$, $p > .05$, was found between the young and elderly control participants for alternate IU production. Compared to the individuals with aphasia, young controls, $F(3, 176) = 408.36$, $p < .05$, and elderly controls, $F(3, 176) = 333.10$, $p < .05$, produced

Figure 5. Direct IUs with corresponding regression lines for female and male NBI controls.



significantly greater numbers of direct IUs. Likewise, significant differences were found between the individuals with aphasia and both young, $F(2, 178) = 122.91, p < .05$, and elderly control participants, $F(2, 178) = 113.42, p < .05$, for alternate IU production. The histograms of these data with their respective regression lines are shown in Figures 3 and 4.

NBI Male Participants Versus NBI Female Participants

The 31 participants in the NBI participant pool were subdivided into two groups by gender: 15 male participants and 16 female participants. There were no significant differences in age, $t(14) = 0.617, p > .05$, or education $t(14) = 0.514, p > .05$, between the two groups.

The direct IUs for both the male and female control groups showed the same stationary time series and a low autocorrelation close to zero with a 20-point lag (correlations computed at 20 consecutive IU points; $R^2 = .03$ and .02, respectively). A significant AR(1) term of –.42 for the male participants and –.38 for the female participants remained for the direct IU series.

A quadratic autoregressive regression line was fit to the direct IUs, though the quadratic series still included a significant negative AR(1) term. The male participants had an R^2 value of .51, and the function for the female participants resulted in an R^2 of .50. These functions depicted SPEs for both groups' production of direct IUs. Alternate IUs for both male and female NBI participants continued to show linear trends (no SPE) with small values for R^2 (.22 and .17, respectively) and significant ($p < .05$) slopes. Male participants' production of alternate IUs averaged 10.5% at the primacy end, and the recency end averaged 15.5%. The female participants produced averages of 10.9% and 14.7%, respectively. The histograms and their respective regression lines for the four data sets are presented comparatively in Figure 5.

Male NBI participants produced an average of 35% of the total possible direct IUs (SD = 11%; range = 15%–70%) and 13% alternate IUs (37% of the direct IUs; SD = 5%; range = <1%–25%) for each story. Female NBI participants produced 39% direct (SD = 11%; range = 17%–73%) and 13% alternate IUs (33% of the direct IUs; SD = 5%; range = <1%–24%) for each story. There was a significant difference, $F(3, 176) = 26.00, p < .05$, between the direct IU mathematical models by gender, with females producing more direct IUs. No significant difference, $F(2, 178) = 0.2768$, $p > .05$, was present in the alternate IU models.

Discussion

Over a century of research on the serial position curve has shown a typically robust effect across the temporal recall of various materials. In the current experiment, direct IUs were recalled expressly from stated information in the stimulus story (phonetically and semantically identical with the target) substantially more frequently than alternate IUs for both controls and individuals with aphasia. As expected, young controls recalled significantly more direct IUs than the elderly controls. Further, female controls recalled significantly more direct IUs than their male counterparts. All of these groups demonstrated an SPE for the free verbal recall of direct IUs from stories, but no SPE was evident for the alternate IU productions. This effect raises the question as to why primacy and recency effects would be shown for direct, but not alternate IUs.

Several hypotheses provide a possible explanation for such a finding. One difference in the linguistic material accessed is that those lexical items recalled directly from statements made in the story could be evoked with more independence from semantic or associated levels of processing. That is, direct IUs may have been retrieved from their phonology alone. Conversely, alternate lexical items must be retrieved from the semantic store, whether or not phonological effects on these synonyms influence them. If the hypothesis that only phonological words show a SPE were accurate, then those alternates that are phonologically similar to the target might also show an SPE. Contacting semantic memory more deeply, or alternatively, failing to retain the phonological form of the word in working memory, may eliminate the serial position effect and create increased dependency on semantic memory or word order syntactic dependencies.

This semantic/syntactic relationship can be shown statistically. The number of IUs generated from both the NBI control participants (and within each of the subgroups) and the participants with aphasia showed a high degree of autocorrelation. That is, each IU was highly correlated with those surrounding it. In fact, until IUs were separated by approximately 20 consecutive IUs, there was a significant autocorrelation shown among them. A moving average of 20 IUs reduced the correlation between IUs; however, there still remained a small but significant autocorrelation with Lag 1 (IUs correlated with 1 IU on either side), thus suggesting that there is a great deal of predictability among adjacent items recalled.

This finding is consistent with those of Meyer and McConkie (1973), who found a clustering of ideas in their study of undergraduates' auditory story recall of two articles from *Scientific American*. When a specific idea was recalled, the one immediately above it in the hierarchical structure had a high probability of being recalled as well. They concluded that the first ideas served as cues for ones that followed in the story. Further, Meyer and McConkie proposed that the position of the ideas in their logical structure could account for most of the variance attributable to the serial effects, despite the fact that they did not show SPEs for either of their experimental passages.

At the outset, one might expect that an SPE should not be shown for alternate IUs, despite the fact that their order of recall is serially produced from directly input stimulus material. It is hypothesized that because alternate IUs are the result of a failure of phonological recall and require the maintenance of semantic association memory, they might not approximate the same pattern of memory recall as the direct IUs. Additionally, because it is necessary to make semantic associations in order to generate alternate IUs, other memory computational devices and strategies (e.g., chaining, chunking) may be enacted that could disrupt the longer-term memory advantage evidenced by the primacy effect. If this account were accurate, lexical recall of story information accessed phonologically would show the SPE, whereas lexical information accessed semantically would not. An unexpected finding was the rising slope (i.e. recency effect) for the recall of alternate IUs.

One explanation for this rise in slope may be that as the story progresses, there is a movement from the phonologic dependency seen in the primacy end to a merging of the phonologic and semantic buffers. The recency effect may just be the consolidation between these two buffers—the phonologic buffer still is used to recall information from the story, and the semantic buffer is used for simultaneous comprehension of the story due to added information and subsequent processing with time. While the activation of phonologic information may dominate as the mechanism used to recall information in the SRP, the increased production of alternate IUs later in the retell may be the result of a deeper level of processing leading to a greater understanding of the material's content and synthesis of the phonologic input.

Hypothesized and found in this study, corroborated by Hall and Bornstein (1991) in individuals with traumatic brain injury and generalized by McNeil (1988), persons with aphasia have reduced memory function. In the current study, this was evident in the reduced production of both direct and alternate IUs relative to those produced by the NBI control group, with participants with aphasia producing approximately half of those produced by the NBI control group. Thus, in this study, persons with aphasia demonstrated both phonologic and semantic memory deficits; however, the role of other nonlinguistic factors, such as the demands of the task and the role of attention in this task, may have contributed to performance results. Additionally, it is important to remember that the analyses in this study examined correctly produced language or actual produced language and were not analyses of error productions. Failure to produce a direct IU could be attributed to any number of factors, from failure to conceptually encode the direct IU to a lack of access to the phonological form. The nature of the task in this investigation disallows the identification of a specific memory process as responsible for the decreased performance of persons with aphasia relative to NBI participants. There was no concurrently administered assessment of memorial functions in this investigation that would allow a more precise interpretation of these findings. Only additional research will clarify this issue.

Overall, the NBI participants recalled about one half of the possible IUs presented in the stimulus stories. Approximately two thirds of all IUs recalled were phonologically and semantically identical lexical items (i.e., direct IUs), and one third were only semantically derived lexical items (i.e., alternate IUs). Likewise, persons with aphasia produced about half as many total IUs as the control participants and about two thirds were direct IUs, approximately one third of which were alternate IUs. Despite the difference in the number of IUs produced, the results from both participant groups in this study, as well as those of Hall and Bornstein (1991), show similar SPEs. These findings illustrate the roles of both short- and long-term auditory memory in the SRP for both NBI individuals and individuals with aphasia, and further define the processing demands of this task. That is, there appears to be a substantive auditory memory demand in the SRP that affects the amount of information recalled. The number of initially comprehended and/or recalled information units is significantly limited in persons with aphasia relative to matched controls. Furthermore, results from this study suggest that the mechanisms responsible for preserving primacy and recency of information are maintained in persons with aphasia, despite a significantly reduced overall quantity of information comprehended and/or recalled, as evidenced by the obtained SPE for direct IUs.

This study contributes to the literature on the SPE in two ways. First, it adds to the limited number of studies (Hall & Bornstein, 1991; Newhouse & Holen, 1975; Roberts, 1966) addressing the SPE in story-level connected speech recall. Second, this study further defines the memory-production mechanisms and deficits in persons with aphasia. From this, we know that word retrieval is more efficient/effective at the beginning and end of the story. To our knowledge, the study by Hall and Bornstein is the only one that addresses the differences between NBI participants and those with a cognitive–linguistic disorder. The current study also addresses deficits in memory capacity that may be attributable to age and gender. As previously stated, Hall and Bornstein found significant differences between their two participant groups, despite controlling for age. Though the present study showed significant differences in recall of total %IUs between the young and elderly control participants, significant differences between both of these groups and participants with aphasia were present.

One minor statistical concern posed by this study is the remarkably high correlations between and among IUs. Despite efforts employing the ARIMA procedure to remove the dependency between IUs, we were unable to eradicate all of the correlations between IUs. When the initial ARIMA procedure began, we found correlations between IUs to be 40 deep (20 on either side of the original information unit). After the procedure was completed and the 20-point moving average was implemented, there still remained a lag of 1, meaning that there was a 1-IU correlation on either side of the information unit in question. Discourse length material probably prohibits removing all correlations between IUs due to its very nature. The IUs, in the format of subject–verb–object for most of the sentences in the Discourse Comprehension Test (Brookshire & Nicholas, 1993), by definition are correlated, and any statistical procedure would be put to great challenges to remove the correlation. Despite its inability to remove the significant correlations between IUs, the ARIMA procedure provides an interpretable statistical context. This correlation value (R^2) aside, the data points as shown in the graphs (Figures 1–5) show a very clear SPE.

Rubin and Wenzel (1996) reviewed over 200 studies directed toward the application of mathematical models to plot the function of memory retention and recall. Unfortunately, no single model or mathematical function ideally described the relationship between the two. Rubin and Wenzel further stated that there seemed to be an unending number of functions describing memory retention through mathematical and number theory, most with limited success. Even with their extensive review, Rubin and Wenzel could only narrow the field by one function from the five most frequently considered mathematical functions to explain retention. They suggest that the four functions best able to describe retention are the power, logarithmic, exponential, and hyperbolic functions, removing the linear function as a model.

The regression model that provided the best fit to the direct IU data in this study was a quadratic one, fitting nicely into the exponential model suggested by Rubin and Wenzel (1996). Murdock (1962) posited an idealized serial position curve and described such a curve as one that has a "rather steep (possibly exponential) primacy effect, an S-shaped recency effect, and a horizontal asymptote extending between the primacy and recency effect" (p. 486). Murdock's model is based not only on his data, but also on the data of several other researchers. Our attempt to fit this model to the NBI participants' data in this study yielded an R^2 value that accounted for only 37% of the variance. As discussed in the results, the quadratic model accounted for 50% of the variance—an increase of 13% over that of Murdock. Whereas the independent variable in Murdock's and others' studies was free recall of word lists, the independent variable in this study was the recall of story length information with a story-level discourse structure. Our two studies suggest that a different mathematical model exists for story-length recall and that the idealized serial position curve suggested by Murdock does not seem appropriately placed for this paradigm. Replication of this study with story-level recall will better test this assertion.

The model proposed by Murdock (1962) is logarithmic in nature, whereas the present study's model is quadratic. Given these analyses, we believe that the quadratic model is best used to explain the direct IU data in this study because it accounts for more of the variance than the other models tested. Likewise, the outright rejection of the linear model by Rubin and Wenzel (1996) is inconsistent with the alternate IU data in this study. Larger recency effects relative to the primacy effects have been noted many times in free recall and serial recall tasks, though in the present investigation no such differences were noted. With limited serial position effect research in story retelling and discourse recall, this finding is difficult to explain. Further examinations in word patterns/priming, word frequencies, and semantic relationships might explain this pattern. Additionally, normative data and investigations that extend beyond the disorders of aphasia and head injury are needed to understand more fully the specificity, validity, and clinical use of the SRP and the serial position data. This includes the overall %IU score, the relative amount of direct versus alternate %IUs, and the SPE for both direct and alternate IUs.

The possible patterns that are available for the direct-to-alternate IU ratio and the various possible patterns available over time allow for the speculation of differential performance patterns based on the nature of the specific memory deficit. Although recent studies (Brown, Preece, & Hulme, 2000; Vousden, Brown, & Harley, 2000) have examined the interaction of memory and phonological encoding (i.e., phoneme ordering) with respect to the serial position effect, to our knowledge there are no models of serial position effect that examine the role of memory in a free recall task that do not control for the time course of lemma and lexeme access. Thus, based on the findings from the current study, we would predict a larger proportion of direct to alternate IUs with a normal SPE for the direct IUs for persons with primarily semantic access deficits. For persons with a primarily semantic working memory capacity deficit, we would predict a reduced overall percentage of IUs relative to NBI participants and a larger direct-to-alternate IU ratio and a recency SPE. For persons with a primarily phonological access deficit, we would predict a normal SPE and a smaller direct-to-alternate ratio relative to the expected 2/3:1/3 ratio. For persons with a phonological working memory deficit, we would predict both a reduced direct-to-alternate ratio and an alternate IU recency effect. Finally, for persons with a general verbal working memory capacity deficit, we would predict the expected direct-to-alternate IU ratio and a recency SPE (see Table 2 for a summary of these predictions). Future research will address these differential diagnostic issues. Additional research should focus on a broader understanding of memory for discourse length material and its application to various clinical populations for a better understanding of the disorder and a more complete understanding of the patient.

Table 2. Summary of predicted IU and SPE outcomes for populations with memory and language deficits based on data from the present study.

Deficits	Reduced %IUs/ normal	Higher ratio of direct-to-alternate IUs	Lower ratio of direct-to-alternate IUs	Normal SPE	Recency SPE
Semantic work memory (storage)	X	X			
Semantic accessing		X		X	
Phonological working memory (storage)			X		X
Phonological access			X	X	
General verbal working memory	X				⸮X

Note. IU = information unit; SPE = serial position effect.

Acknowledgment

This study was supported by a grant from the Veterans Administration Department of Rehabilitation Research and Development Project B2265-R, awarded to Malcolm R. McNeil and Patrick J. Doyle.

References

Adams, C., Smith, M. C., Nyquist, L., & Perlmutter, M. (1997). Adult age-group differences in recall for the literal and interpretive meanings of narrative text. *Journal of Gerontology, 52B*(4), 187–195.

Anderson, N. D., & Craik, F. I. M. (2000). Memory in the aging brain. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 411–425). Oxford, England: Oxford University Press.

Arvdeson, J. C., McNeil, M. R., & West, T. L. (1986). Prediction of Revised Token Test overall, subtest, and linguistic unit scores by two shortened versions. *Clinical Aphasiology, 16*, 57–63.

Baddeley, A. D. (1999). Memory and ageing. In A. D. Baddeley (Ed.), *Essentials of human memory* (pp. 251–273). Hove, East Sussex, England: Psychology Press.

Balota, D. A., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 395–409). Oxford, England: Oxford University Press.

Bayles, K. A., & Tomoeda, C. K. (1993). *Arizona Battery for Communication Disorders of Dementia*. Tucson, AZ: Canyonlands Publishing.

Brookshire, R. H., & Nicholas, L. E. (1993). *Discourse Comprehension Test*. Tucson, AZ: Communication Skill Builders.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). An oscillator-based model of memory for serial order. *Psychological Review, 107,* 127–181.

Bruyer, R., & Vanberten, M. (1998). Short-term memory for faces: Ageing and the serial position effect. *Perceptual and Motor Skills, 87*, 323–327.

Burgio, F., & Basso, A. (1997). Memory and aphasia. *Neuropsychologia, 35*, 759–766.

Burke, D. M., & Laver, G. D. (1990). Aging and word retrieval: Selective age deficits in language. In E. A. Lovelace (Ed.), *Aging and cognition: Mental processes, self awareness and interventions* (pp. 281–300). Amsterdam: Elsevier.

Capitani, E., Della Sala, S., Logie, R. H., & Spinnler, H. (1992). Recency, primacy, and memory: Reappraising and standardizing the serial position curve. *Cortex, 24*, 315–342.

Carlesimo, G. A., Sabbadini, M., Fadda, L., & Caltagirone, C. (1997). Word-list forgetting in young and elderly participants: Evidence for age-related decline in transferring information from transitory to permanent memory condition. *Cortex, 33*, 155–166.

Crowder, R. G., & Greene, R. L. (2000). Serial learning: Cognition and behavior. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 125–135). Oxford, England: Oxford University Press.

Doyle, P. J., McNeil, M. R., Park, G., Goda, A., Rubenstein, E., Spencer, K., Carroll, B., Lustig, A., & Szwarc, L. (2000). Linguistic validation of four parallel forms of a story retelling procedure. *Aphasiology, 14*, 537–549.

Doyle, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology, 12*, 561–574.

Freebody, P., & Anderson, R. C. (1986). Serial position and rated importance in the recall of text. *Discourse Processes, 9*, 31–36.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351–360.

Groeger, J. A. (1997). *Memory and remembering: Everyday memory in context*. New York: Addison Wesley Longman.

Hall, S., & Bornstein, R. A. (1991). Serial-position effects in paragraph recall following mild closed-head injury. *Perceptual and Motor Skills, 72*, 1295–1298.

Hess, T. M., & Arnould, D. (1986). Adult age differences in memory for explicit and implicit sentence information. *Journal of Gerontology, 41*(2), 191–194.

Hula, W. D., McNeil, M. R., Doyle, P. J, Rubinsky, H. R., & Fossett, T. R. D. (2003). The inter-rater reliability of the story retell procedure. *Aphasiology, 17,* 523–528.

Korsnes, M. S., & Gilinsky, A. S. (1993). Aging and serial list picture memory. *Perceptual and Motor Skills, 76*, 1011–1014.

Korsnes, M. S., & Magnussen, S. (1996). Age comparisons of serial position effects in short-term memory. *Acta Psychologica, 94*, 133–143.

Maitland, S. B., Intrieri, R. C., Schaie, K. W., & Willis, S. L. (2000). Gender differences and changes in cognitive abilities across the adult life span. *Aging, Neuropsychology, and Cognition, 7*(1), 32–53.

McNeil, M. R. (1988). Aphasia in the adult. In N. J. Lass, L. V. McReynolds, J. Northern, & D. E. Yoder (Eds.), *Handbook of speech, language, and hearing* (pp. 738–786). Philadelphia: W. B. Saunders.

McNeil, M. R., Doyle, P. J., Fossett, T. R. D., Park, G. H., & Goda, A. J. (2001). Reliability and concurrent validity of the Percent Information Unit (%IU) scoring metric for the RAPP Story Retelling Procedure. *Aphasiology, 15*, 991–1006.

McNeil, M. R., Doyle, P. J., Park, G. H., Fossett, T. R. D., & Brodsky, M. B. (2002). Increasing the sensitivity of the Story Retell Procedure for the discrimination of normal elderly participants from person with aphasia. *Aphasiology, 16*, 815–822.

McNeil, M. R., & Pratt, S. R. (2001). A standard definition of aphasia: Toward a general theory of aphasia. *Aphasiology, 15*, 901–911.

McNeil, M. R., & Prescott, T. (1978). *The Revised Token Test*. Austin, TX: Pro-Ed.

Meyer, B. J. F., & McConkie, G. W. (1973). What is recalled after hearing a passage? *Journal of Educational Psychology, 65*(1), 109–117.

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*, 482–488.

Newhouse, R. C., & Holen, M. C. (1975). Individual difference in serial position effects with connected discourse. *Psychological Reports, 37*(3, pt. 1), 824.

Newman, C. W., Weinstein, B. E., Jacobson, G. P., & Hug, G. A. (1990). The Hearing Handicap Inventory for Adults: Psychometric adequacy and audiometric correlates. *Ear and Hearing, 11*, 430–433.

Ostergaard, A. L., & Meudell, P. R. (1984). Immediate memory span, recognition memory for subspan series or words, and serial position effects in recognition memory for supraspan series of verbal and nonverbal items in Broca's and Wernicke's aphasia. *Brain and Language, 22*, 1–13.

Park, G. H., McNeil, M. R., & Tompkins, C. A. (2000). Reliability of the Five-Item Revised Token Test for individuals with aphasia. *Aphasiology, 14*, 527–535.

Porch, B. A. (1981). *Porch Index of Communicative Ability*. Austin, TX: Pro-Ed.

Roberts, D. M. (1966). Serial position effects in two Stanford-Binet subtests. *Psychology: A Journal of Human Behavior, 3*(3), 2–4.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103*, 734–760.

SAS Institute Inc. (1999). *SAS/ETS user's guide Version 8*. Cary, NC: Author.

Schugens, M. M., Daum, I., Spindler, M., & Birbaumer, N. (1997). Differential effects of aging on explicit and implicit memory. *Aging, Neuropsychology, and Cognition, 4*(1), 33–44.

SPSS Inc. (2001, November 15). SPSS for Windows (Release 11.0.1) [Computer software]. Chicago: Author.

Stine, E. A. L. (1990). The way reading and listening work: A tutorial review of discourse processing and aging. In E. A. Lovelace (Ed.), *Aging and cognition: Mental processes, self awareness and interventions* (pp. 301–327). North-Holland: Elsevier Science.

Vacanti, J. M., Hamm, N. H., Cammeron, S. A., & Peterson, C. (1977). Sex differences in encoding strategies for active and passive words. *Psychological Reports, 40*(3, pt. 2), 1043–1048.

Ventry, I. M., & Weinstein, B. E. (1982). The Hearing Handicap Inventory for the Elderly: A new tool. *Ear and Hearing, 3*(3), 128–134.

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology, 41*, 101–175.

Wechsler, D. (1987). *Manual for the Wechsler Memory Scale—Revised*. San Antonio, TX: The Psychological Corporation.

Wingfield, A., & Stine-Morrow, E. A. L. (2000). Language and speech. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 359–416). Mahwah, NJ: Erlbaum.

Yasuda, K., Nakamura, T., & Beckman, B. (2000). Comprehension and storage of four serially presented radio news stories by mild aphasic participants. *Brain and Language, 75*, 399–415.

Received September 30, 2002

Accepted March 25, 2003

DOI: 10.1044/1092-4388(2003/088)

Contact author: Malcolm R. McNeil, PhD, Department of Communication Science and Disorders, 4033 Forbes Tower, University of Pittsburgh, Pittsburgh, PA 15260. E-mail: mcneil+@pitt.edu