

Word Free Recall Search Scales Linearly With Number of Items Recalled

Eugen Tarnow, Ph.D.

18-11 Radburn Road

Fair Lawn, NJ 07410

etarnow@avabiz.com (e-mail)

Abstract:

I find that the total search time in word free recall, averaged over item position, increases linearly with the number of items recalled. Thus the word free recall search algorithm scales the same as the low-error recognition of integers (Sternberg, 1966). The result suggests that both simple integer recognition and the more complex word free recall use the same search algorithm. The proportionality constant of 2-4 seconds per item (a hundred times larger than for integer recognition) is a power function of the proportion not remembered and seems to be the same function for word free recall in young and old subjects, high and low presentation rates and delayed and immediate free recall. The linear scaling of the search algorithm is different from what is commonly assumed to be the word free recall search algorithm, search by random sampling. The linearity of the word free recall extends down to 3 items which presents a challenge to the prevalent working memory theory in which 3-5 items are proposed to be stored in a separate high-availability store.

Keywords: Free recall; short term memory; memory search, Sternberg

In short term memory research most data elude a common theoretical understanding. This is true for data that is extremely simple like the Sternberg (1966) beautiful straight lines from integer recognition, or data that is rather complex with recency and primacy effects such as the Murdock (1962) word free recall experiments. Fifty years after their publication the original papers are citation classics, both types of experiments have been redone many times with variations, but there is not an accepted theory of short term memory that describes either one to the satisfaction of the full research community.

The two experiments probe different parameters of short term memory. Sternberg (1966) designed his experiment to be extremely simple and heavily controlled: the items presented were single digits. The error rate was designed to be small (1.4% errors). The subject processing was limited - in less than a second the subjects had to answer only whether a displayed integer was in the presented list. The resulting data formed straight lines. On the contrary, the earlier Murdock (1962) was designed to show how memory decays in experiment designed to be fast compared to its predecessors. The items presented were complex - words, and the error rate was very high, sometimes at 70% (Tarnow, 2011). In addition the failure to recall some items may be larger than 90% (Murdock, 1962). The Murdock free recall experiment also gave a lot of freedom to the subjects, they had a full minute to recall the words any way they wanted. The resulting data formed strange u-shaped curves.

Hintzman (2011) argued that the freedom of free recall has lead to its "downfall" and wrote that only "few researchers ... dare to ask their subjects about the strategies they used" and if they do they will find that there are significant individual differences. But even though the two experiments are so different, there is at least one bridge between the two types of data: The Sternberg (1966) experiments do show a recency effect under some conditions (see, for example, Duncan & Murdock, 2000 and references therein). I will now show in addition that the word free recall search algorithm scales the same as does the search in Sternberg (1966), both scale linearly.

Thus in the word free recall data of Murdock & Okada, 1970 and Kahana, Zaromb & Wingfield, 2001 (Fig. 1) I find that there is a linear relationship between the total free recall search time, defined as the sum of all response times, including the first. The linear regression accounts for 97-99% of the variance in the word free recall response times while in the Sternberg data the linear regression accounts for 99.4% of variance in the recognition response time. The slope of 2-4 seconds, which is the additional time needed to respond if there is an additional item that can be recalled, is a hundred

times larger than the 38 msec per item in Sternberg (1966). The slope varies between different experimental subjects and delays: it is smaller for slow compared to fast presentation rates, it is smaller for young compared to old subjects and also smaller if the free recall starts immediately than if the recall is delayed.

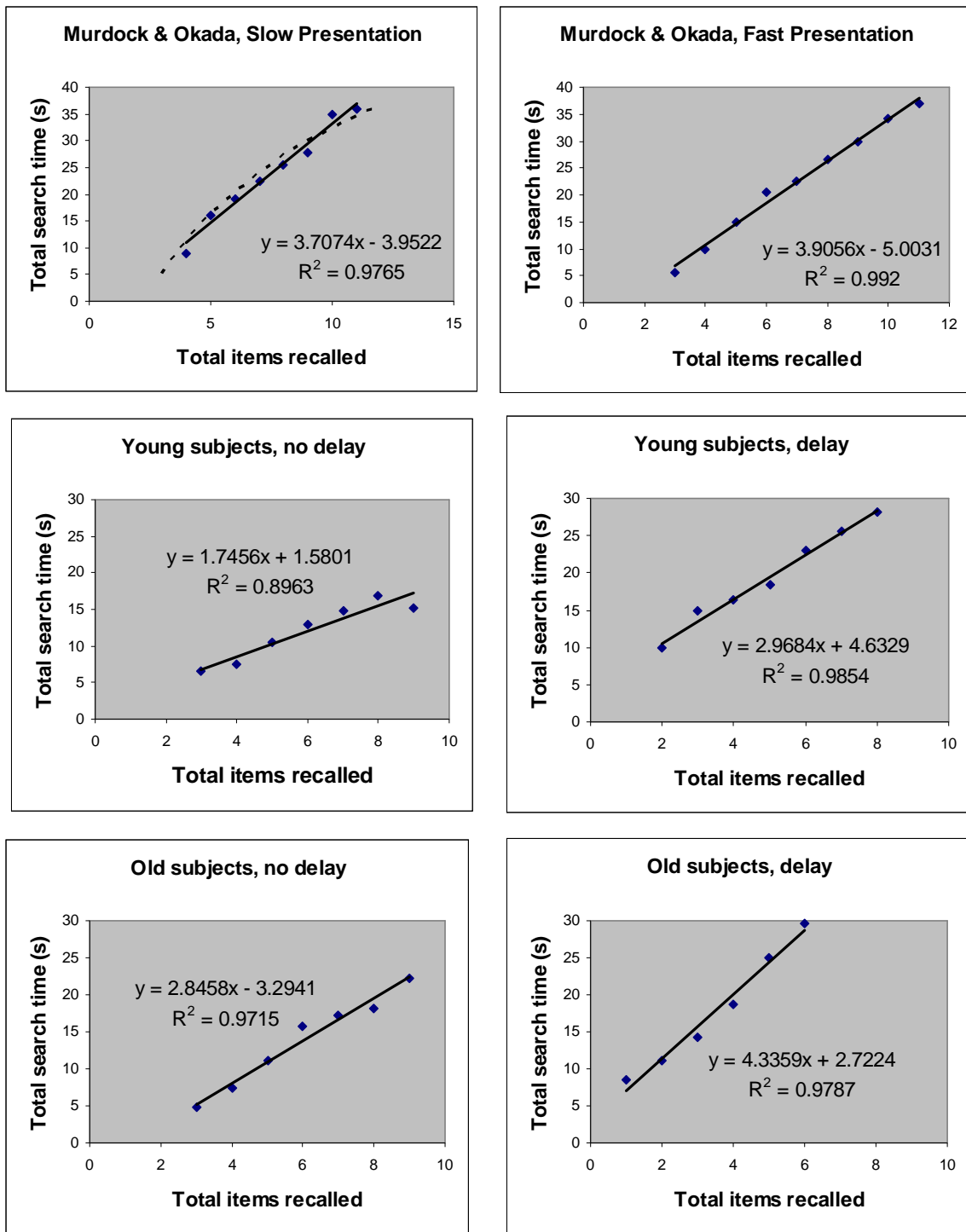


Fig1. The sum of all RTs, i.e. the total search time, is a linear function of the total number of items in memory for word free recall. The top graph corresponds to the immediate free recall data in the Mudock & Okada (1970) and the four graphs below to the data in Kahana, Zaromb and Wingfield (2001). The top graph also has a solid line indicating the expected behavior of a search by random sampling. The middle left graph corresponds to immediate free recall by young subjects,

the middle right graph corresponds to delayed free recall by young subjects, the lower left graph corresponds to immediate free recall by old subjects and the remaining graph in the lower right hand corner corresponds to delayed free recall by old subjects. Note that the slope increases from young to old and from immediate to delayed free recall in the Kahana, Zaromb and Wingfield (2001) data.

In Fig 2 is shown that the slope is closely related to the overall number of words not recalled. A high slope is correlated with many words not recalled. This relationship, that words with low probability of recall take a long time to recall, is similar to what activation theory would have it (see, for example, Tarnow, 2008): a word that is not in a highly activated state has a low probability of being recalled and takes time to reactivate. If the presented word items formed a sentence presumably the number of words not recalled would be close to 0% and the delay per item would be similarly close to zero (simply reciting the sentence), suggesting a power law (the line in Fig. 2) if not a proportionality. The exact relationship cannot be inferred from the rather small range of overall recall, experiments in which the overall number of words not recalled is closer to 0-25% would be better able to pin down the relationship.

These linear relationships in Fig. 1 give some food for thought. First they are surprisingly good considering Hintzman's (2011) pessimistic point about different individuals using different search algorithms. They are also surprisingly good considering that individual response times vary widely from half a second to ten seconds or even longer. That these relationships occur both for extremely simple recognition and for word free recall, that the corresponding search algorithms both scale linearly with the number of items, suggests that both processes just might use the same search algorithm. Second, the linear scaling of the word free recall search shows that previous assumptions of a random sampling algorithm (see, for example, Rohrer & Wixted, 1994) is incorrect. If the latter was correct, the bowing of the dotted curve in the top graph of Figure 1 would apply and an additional item would only take $\text{const}/(N+1)$ to find (Johnson, Johnson & Mark, 1951). The actual algorithm subjects use scales worse than random sampling. Third, the data may contradict the prevalent working memory theory (see, for example, Cowan, Rouder, Blume & Saults, 2012) in which 3-5 items are thought to be stored in a separate store from subsequent items: the corresponding discontinuity in the slope at 3-5 items is not there, in fact, the slope is the same for 3 items as for 9 items (see also Laming, 1999, Tarnow, 2010 and Tarnow, 2011, who noticed other discrepancies with the working memory model in the same free recall data).

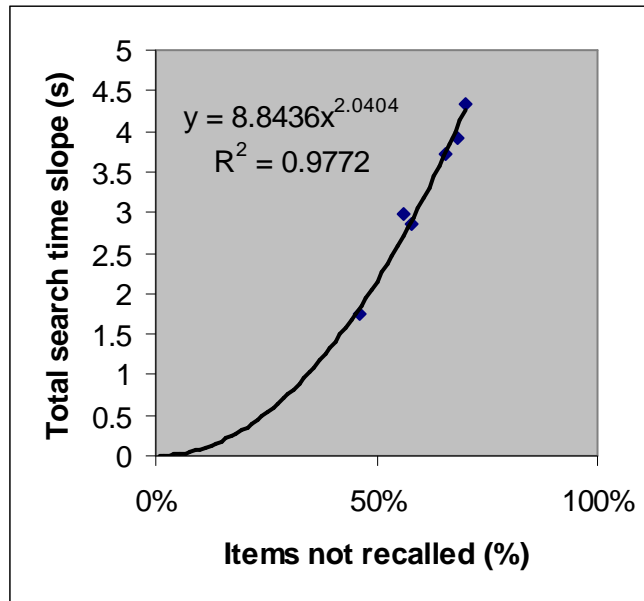


Fig 2. Average time to find an additional item as a function of the total percent words not recalled. The more likely the complete word list is to be recalled the shorter the additional search time. The power law fit suggests that the longest search time for words that are very hard to recall is 8.8 seconds per word. To find out the true form of the relationship between the slope and overall recall, one would need additional data for experiments in which the overall number of words not recalled is 0-25%. The Sternberg (1966) data would fall close to the origin of the graph.

APPENDIX I: The recall data

The Murdock (1962) data can be downloaded from the Computational Memory Lab at University of Pennsylvania (<http://memory.psych.upenn.edu/DataArchive>). Lists of words were read to groups of subjects to the accompaniment of a metronome (to fix the word presentation rate) after which the subjects wrote down as many of the words as they could remember in any order for a period of 1.5 minutes. There were two presentation rates: 1 or 2 seconds and there were five word list lengths – 10, 15, 20 for the two second presentation rate and 20, 30 and 40 for the one second presentation rate. The six sets of data that resulted are typically labeled M-N in which M is the number of list words and N is the number of seconds between word presentations (1 or 2).

In Kahana et al (2002) two groups of 25 subjects, one older (74 mean age) and one younger (19 mean age), studied lists of ten word words from a computer screen. Immediately following the list presentation participants were given a 16-s arithmetic distractor task and only after this delay were the subjects asked to recall the words. This data can also be downloaded from the Computational Memory Lab at University of Pennsylvania (<http://memory.psych.upenn.edu/DataArchive>).

For both sets of data, trials with errors were discarded in order to keep the number of items to be remembered or having been recalled well defined. Data points with fewer than 10 measurements were not included in the graphs to minimize noise.

References

- Cowan, Nelson; Rouder, Jeffrey N.; Blume, Christopher L.; Saults, J. Scott 2012 "Models of verbal working memory capacity: What does it take to make them work?" *Psychological Review*, Vol 119(3), Jul 2012, 480-499
- Matthew Duncan, Bennet Murdock (2000) "Recognition and Recall with Precuing and Postcuing", *Journal of Memory and Language* Volume 42, Issue 3, April 2000, Pages 301–313.
- Hintzman D, (2011), *Research Strategy in the Study of Memory: Fads, Fallacies, and the Search for the "Coordinates of Truth"*, *Perspectives on Psychological Science* 6(3) 253–271.
- Johnson DM, Johnson RC Mark AL (1951). A mathematical analysis of verbal fluency. *J General Psychology* 44, 121-128.
- Kahana JM, Zaromb F, Wingfield A (2001) Age dissociates recency and lag-recency effects in free recall *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28(3) 530-540. Kahana M, Associative retrieval processes in free recall, *Memory & Cognition* 1996,24 (1), 103-109
- Laming D (1999) Testing the Idea of Distinct Storage Mechanisms in Memory, *International Journal of Psychology* 34, 5-6, 419-426.
- Murdock Jr., Bennet B. "The immediate retention of unrelated words." *Journal of Experimental Psychology*, Vol 60(4), Oct 1960, 222-234
- Murdock B (1962) The serial position effect of free recall *Journal of Experimental Psychology* 64(5) 482-488.
- Rohrer D, Wixted J (1994) An analysis of latency and interresponse time in free recall, *Memory & Cognition* 22(5):511-524.
- Tarnow E (2010) There is no capacity limited buffer in the Murdock (1962) free recall data. *Cognitive Neurodynamics* December 2010, Volume 4, Issue 4, pp 395-397

Tarnow E (2011) The free recall search process introduces errors in short term memory but apparently not in long term memory. See <http://cogprints.org/7337/>