

Running head: AGE VS. ALZHEIMER'S: A COMPUTATIONAL MODEL

Age vs. Alzheimer's: A Computational Model of Changes in Representation

Patrick Conley and Curt Burgess

University of California, Riverside

Guila Glosser

Department of Neurology, University of Pennsylvania

Patrick Conley  
Department of Psychology  
University of California, Riverside  
Riverside, CA 92521  
Phone: (909) 787-2102  
E-mail: [patrick@cassandra.ucr.edu](mailto:patrick@cassandra.ucr.edu)

### Abstract

Previous research has demonstrated that the language of older adults leads to denser representations in a high dimensional model of memory than does the language of younger adults (Conley & Burgess, in press), and thus that density in the model (HAL, or the Hyperspace Analogue to Language) may constitute an useful metric in comparing memory for younger and older adults. This paper extends the previous research by examining the role of density in language generated by adults with Alzheimer's and comparing the results with age-matched normal controls. We found that, just as older adults have denser representations in semantic space than do younger adults, adults with Alzheimer's have still denser representations than normal older adults. These results support the hypothesis that greater density, normally associated in the model with good semantic depth, may in fact reach a "saturation point" and affect retrieval in older adults and especially adults with Alzheimer's.

### Age vs. Dementia: A Computational Model of Alzheimer's-Related Memory Effects

The effects of aging and dementia on semantic memory have been the subject of a large body of literature in recent years. It is well established that the normal aging process causes some deterioration in the memory process. Age has been shown to cause deficits in recall and recognition memory (Burke & Light, 1981). However, the nature and ultimate cause of these deficits is not clear, as demonstrated by cases in which aging does not affect memory processes. Implicit memory, for instance, has been shown in several studies to be resistant to aging (Light & Singh, 1987; Hashtroudi, Chrosniak, & Schwartz, 1991; Schacter, Cooper, & Valdiserri, 1992).

Similarly, the effect of Alzheimer's Disease (AD) on semantic memory is difficult to determine; clearly, deficits are present and detectable, but are these deficits due to actual semantic degradation, or are processing requirements overwhelming the capabilities of the AD patients (Nebes, 1989)? This sets up the traditional "processing vs. representation" argument also prevalent in other domains of cognition. Most existing models of AD, as well as normal aging, are processing based. However, our memory model, called HAL, (Hyperspace Analogue to Language; Burgess & Lund, 1997; Lund & Burgess, 1996) is a representational model. This allows us to examine this problem from another direction, and in this paper we attempt to determine to potential role of representation in memory deficits associated with AD and normal aging.

#### **The HAL Model**

The HAL model of memory has been used to investigate semantic priming and many other processes in memory and cognition (Burgess & Lund, 1999). The HAL model of memory produces semantic representations for words from a 320 million word corpus text, derived from the Usenet. HAL computes weighted co-occurrence information by parsing a ten-word window

over the corpus of text, and stores this co-occurrence information in a  $70,000^2$  dimensional matrix. By combining row and column information (preceding and following co-occurrence values), a vector for each of the 70,000 most frequent words in the text corpus is generated. This vector thus contains the coordinates for plotting the word's location in high-dimensional space. The contextual similarity between words is inversely related to their distance in this semantic space; closer words are more semantically related.

Previous research (Conley & Burgess, in press) has demonstrated that the language produced by older adults leads to *denser* representations within the semantic space generated by the HAL model than is the case for the language of younger adults. Density, in these studies, is operationalized as the mean distance between any word and the closest ten neighbors to that word in semantic space. Normally, increasing the size of the text corpus available to the model will result in an increase in density in semantic space, but in the cases previously mentioned, text size was held constant across the corpora being compared. These differences in density have been replicated across both Usenet text and interview transcript data. Other factors -- such as the type/token word differences between the older and younger corpora, topic breadth, and possible oversensitivity of the density metric to any changes between different text corpora -- have been controlled for or experimentally examined. Therefore, we believe that the HAL model is detecting actual representational differences between the language produced by younger and older adults, and that the variation in mean density between matrices generated from each group is not merely an artifact of word frequency differences or topic breadth.

The next question we address is whether differences in representation will be detected between matrices when factors other than age are considered. Specifically, we wished to

determine whether the HAL model, and specifically, the density metric, would differentiate between adults with AD and normal older adult controls.

### **Method and Results**

Two corpora were generated, both based on interview text provided by Glosser (from Glosser & Deser, 1990, 1992). Each experimental group was asked the same interview questions. The first text corpus (AD corpus) consisted of 19,000 words and was generated from the interview texts of 23 patients (mean age 64 years) meeting research diagnostic criteria for probable Alzheimer's disease (McKhann et al., 1984). The second text corpus (normal corpus) also contained 19,000 words, and was generated by 41 normal older adult controls (mean age 55 years). As the second text corpus was truncated to match the length of the first, only 15 of the older adults' interview texts were included in the second text corpus. The two corpora shared 877 words.

Each corpus was then used to generate a high-dimensional matrix. To obtain a measure of density in semantic space for each matrix, the mean distance of the ten closest neighbors to each word in the matrix was computed. Calculating a mean distance for every word in the matrix insures that any density changes detected are systematic (i.e., not due to sampling error). The obtained mean distances, therefore, are a measure of density in semantic space, and these distances were then compared across the matrices.

The matrix generated from the AD corpus had denser representations ( $M = 437$  RCUs, or Riverside Context Units, an distance measure scaled to resemble human RTs) than the normal corpus ( $M = 452$  RCUs),  $t(876) = 3.657, p < .001$ . Thus, we see that the language generated by the adults with AD is leading to denser representations within the model than is the case with language produced by normal older adults, even though the factors that normally influence

density in the semantic space (amount of text, topic breadth, number of unique words) have all been controlled for. This finding extends previous research (Conley & Burgess, in press), in which we detected a similar density difference between younger and older adults, with older adults' language producing denser representations in the HAL model.

### **Discussion**

In this study, we detected a density difference between the memory matrices generated from language spoken by adults with AD and language spoken by normal older adults. This finding is similar to an earlier study in which the language of older adults led to denser representations in high-dimensional space than the same amount of text from younger adults. Thus, density in the HAL model appears to change as a function of population, with younger adults producing the sparsest representations, older adults generating denser representations, and adults with AD leading to yet denser representations in semantic space. We previously postulated density (Burgess & Conley, 1998) as being a source of the difficulty commonly experienced in the retrieval of proper names. The results of the current study support a cautious extension of this hypothesis: perhaps "density saturation" is one source of the memory difficulties commonly found in aging, and exacerbated still further in Alzheimer's disease.

As an example, disproportionate difficulty with proper names is frequently reported by older adults (Cohen & Burke, 1993). In the HAL model, names occupy their own area of semantic space, and name neighborhoods have been found to be denser than the neighborhoods of frequency-matched common nouns. Perhaps names offer disproportionate difficulty because they are already denser than common nouns, and thus more susceptible to this saturation effect. Errors of retrieval in common nouns would usually result in a word that is semantically related to the desired word, which the retriever could then use to recover to the desired word. While this

process would result in the desired word, it would also require more time and processing power than just accessing the correct word during the original retrieval attempt. Thus, density, which usually provides the model with better representations and greater semantic depth, could, in time, cause the system to decrease in efficiency if the density was increased beyond a level of diminishing returns, or if the individual's ability to retrieve information efficiently was compromised in some fashion, as might be the case in normal aging and especially in Alzheimer's.

HAL is more of a learning and representational model than a cognitive processing model, yet when differences in mental processing are operationalized as the language produced by groups with different memory characteristics, we hypothesize that the model is "borrowing" the processing of information provided by each of the groups studied -- when they are simply trying to answer the interview questions in an appropriate manner -- and that it is this difference in quality of processing between different groups of adults that is leading to the variation in representations between matrices. Most theories of AD and normal aging necessarily focus on processing issues (see Nebes, 1989, for a review), yet we argue that the memory differences brought on by both normal aging and abnormal dementia might be reflected in representation as well as process. These results demonstrate how representational issues might affect processing: density in HAL would normally suggest greater semantic depth, and thus a richer representation; words cluster closely in the semantic space, signifying that the model has a very good semantic "understanding" of words located in that space. However, if the processing (or retrieval) has been adversely affected by other factors, such as AD or normal age-related deficits, then the very factors that lead to richer representation may lead to increased difficulties in processing, as the

large number of close semantic competitors will require processing resources that might not be available to resolve the confusion and retrieve the desired word.

## References

- Burgess, C., & Conley, P. (1998) Developing semantic representations for proper names. *Proceedings of the Cognitive Science Society* (p.185-190). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Burgess, C., & Lund, K. (1997). Modeling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*. 12, 177-210.
- Burgess, C. & Lund, K. (1999) The dynamics of meaning in memory. In Dietrich & Markman (Eds.), *Cognitive Dynamics: Conceptual Change in Humans and Machines*.
- Burke, D. M. & Light, L. L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin*, 90, 513-546.
- Cohen, G., & Burke, D. M., (1993). Memory for proper names: A review. *Memory*, 1, 249-263.
- Conley, P., & Burgess, C. (In press). Age effects in a computational model of memory. *Brain and Cognition*.
- Glosser, G. & Deser, T. (1990). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language*, 40, 67-88.
- Glosser, G. & Deser, T. (1992). A comparison of changes in macrolinguistic and microlinguistic aspects of discourse production in normal aging. *Journal of Gerontology*, 47, 266-272.
- Hashtroudi, S., Chrosniak, L. D., & Schwartz, B. L. (1991). Effects of aging on priming and skill learning. *Psychology & Aging*, 6, 605-615.
- Light, L. L., & Singh, A. (1987). Implicit and explicit memory in young and older adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 13, 531-541.

Lund, K. & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrences. *Behavior Research Methods, Instruments and Computers*, 28, 203-208.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlen, E.M. (1984). Clinical diagnosis of Alzheimer's disease. *Neurology*, 34, 939-944.

Nebes, R. D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin*, 106, 377-394.

Schacter, D. L., Cooper, L.A., & Valdiserri, M. (1992). Implicit and explicit memory for novel visual objects in older and younger adults. *Psychology & Aging*, 7, 299-308.