BRAIN AND COGNITION **32**, 91–348 (1996)
ARTICLE NO. 0062–0067

TENNET VII: Theoretical and Experimental Neuropsychology

August 14–16, 1996

Pavillon Judith-Jasmin, Salle Marie-Gérin-Lajoie, Université du Québec à Montréal, Montréal, Québec, Canada

W-3. Overcrowding in Semantic Neighborhoods: Modeling Deep Dyslexia

LORI BUCHANAN

The Hospital for Sick Children

CURT BURGESS AND KEVIN LUND

University of California

Deep dyslexic (DD) readers produce semantic errors during word naming and are impaired at nonword naming. Previous models of DD have explained this co-occurrence of deficits by postulating damage to both lexical and nonlexical pathways in the reading system. Buchanan, Hildebrandt and MacKinnon (1994) offered an alternative explanation that resulted in the prediction that words with several semantic and phonological neighbors would be read with less success by DDs than would words with few neighbors. This paper briefly describes a test of the semantic neighborhood hypothesis using HAL, a computational model of semantic space developed by Lund and Burgess (in press).

pgs 111 - 114

Patients with deep dyslexia (DD) are formerly literate adults who cannot read nonwords such as ''frip'' aloud and who also have severe difficulties in word naming, such as the production of semantic paralexias. The pathway responsible for nonword naming is assumed to be independent of the whole-word pathway containing lexical and semantic information. Thus, the production of semantic errors in patients who are also unable to read nonwords presents a challenge for the dual route model. Morton and Patterson (1980) explained the co-occurrence of semantic errors with the inability to read nonwords by postulating multiple damage to the system: DD resulted from damage in the semantic system and an impaired phonological pathway (also supported by computational approaches to deep dyslexia, Plaut & Shallice, 1993).

However, recent evidence suggesting that deep dyslexics do use the phonological pathway (Buchanan, Hildebrandt, & MacKinnon, 1994; submitted) led Buchanan et al. (submitted) to propose that a selection impairment in the phonological output lexicon is responsible for the reading deficits in DD. From this theory it follows that increasing the number of active lexical representations should decrease the likelihood that a patient can read the word correctly. Buchanan et al. therefore predicted that both semantic and phonological neighborhood sizes would be negatively correlated with the reading performance of deep dyslexics. While this was true for phonological neighborhood, semantic neighborhood size was not correlated with performance.

The metric for semantic neighborhood density in that study was obtained using word association norms from 218 normal subjects. Paradoxically, these associative relationships did not account for variability in the patients' naming performance. Part of the answer to this dilemma may reside in the distinction between associative and semantic relationships. Word association norms used in the development of stimuli for tests of semantic memory confound the associational and the semantic (or categorical) aspects of word meaning. Lund, Burgess, and Atchley (1995) have provided empirical and computational data that suggests that associative and semantic components of memory retrieval can be dissociated. Thus, semantic errors characteristic of DD may be more semantic than associative in nature.

In this paper we describe a test of the semantic neighborhood hypothesis utilizing the Hyperspace Analogue to Language (HAL) model of semantic memory (Burgess & Lund, in press; Lund & Burgess, in press; Lund, Burgess, & Atchley, 1995). The HAL model uses a high-dimensional semantic space constructed from a co-occurrence matrix formed by analyzing a 160 million word corpus. The analysis procedure uses a ''window,'' representing a span of words, that is passed over the corpus. Words within the window are recorded as co-occurring with a strength inversely proportional to the number of other words separating them within the window. At each one-word increment of window movement along the corpus, the co-occurrence values are summed in a 490 million cell data matrix that represents the raw co-occurrence information for 70,000 lexical items. Each lexical item has a corresponding row and a column of co-occurrence information. In HAL each word is represented by a 140,000 unit vector, however, only 200 or so of the most variant vector elements are required to replicate a broad range of semantic experiments.

These 200 elements in each word's vector representation serve as coordinates in high-dimensional semantic space. Similar words tend to be clustered in the same space, thus making it straightforward to define semantic neighborhoods of particular words given some distance parameter.

Semantic density can be defined as the standard deviation of N items in the semantic neighborhood. In this study, N = 50, so the density measure used in the analysis represents the standard deviation of 50 items in each semantic neighborhood.

Method

Patient descriptions. Three patients (GZ, PB and JC) fit the DD profile; they produce semantic errors, are impaired in nonword reading, and are sensitive to the lexical status and imageability of words (Buchanan et al., submitted).

Stimulus set and procedure. The stimulus set consisted of 300 three-to-seven-letter words with varied orthographic neighborhoods, phonological neighborhoods, and semantic neighborhoods. Stimuli were printed on individual cards, and patients read aloud each item.

The goal in this study was to determine whether the semantic density (SDen) of the semantic neighborhood generated by HAL can account for the semantic errors. Density measures (the standard deviations of the neighborhoods) for each of the words in the list was correlated with patient performance. Errors were coded as "0" and correct responses as "1"; a positive correlation indicates a facilitory relationship.

Results

The Buchanan et al. theory predicts that semantic neighborhood size will be negatively correlated with performance since it was assumed to increase the competition for selection during response production. In contrast, a less dense semantic neighborhood should produce fewer competitors in the phonological output lexicon so that SDen should be positively correlated with performance.

SDen accounts for some variability in the patients' errors (see Table 1). While only JC produced data that resulted in a reliable effect, the correlations for the HAL measure were greater than for the associative measure in every case. The combined effect was calculated by adding the log probabilities of the individual correlations (Guilford, 1956) and thus provides something

TABLE 1
Semantic Effects

	JC	GZ	PB
Assoc. neighborhood (Buchanan et al.) Combined $\chi^2 = 10.36$	$r = -.10$	$r = -.04$	$r = -.02$
Semantic density (From HAL) combined $\chi^2 = 20.59$*	.23*	.07	.04

*$p < .05$.

akin to a sign test. This resulted in a reliable effect of SDen but not of neighborhood.

The dissociation between SDen and neighborhood provides further support for the idea that the semantic errors caused by deep dyslexia are in fact essentially semantic in nature and not simple associative mistakes.

Furthermore, we conclude that semantic density does play a role in the reading performance of deep dyslexics, and, as suggested by Buchanan et al., words with high density neighborhoods produce more errors than words with low density neighborhoods.

References

Buchanan, L., Hildebrandt, N., & MacKinnon, G. E. 1994. Phonological processing by a deep dyslexic patient: A Rowse is implicitly a rose. *The Journal of Neurolinguistics,* **8.**

Buchanan, L., Hildebrandt, N., & MacKinnon, G. E. Submitted. Selection impairments in deep dyslexia: Neighborhood confusion during phonological output.

Burgess, C., & Lund, K. In press. Modeling cerebral asymmetries of semantic memory using high-dimensional semantic space. In Beeman, M., & Chiarello, C. (Eds.), *Getting it right: The cognitive neuroscience of right hemisphere language comprehension.* Hillsdale, NJ: Erlbaum.

Guildford, J. P. 1956. *Fundamental statistics in psychology and education.* New York: McGraw–Hill.

Lund, K., & Burgess, C. In press. Producing high-dimensional semantic spaces from lexical co-occurrence. Behavior Research Methods, Instrumentation, and Computers.

Lund, K., Burgess, C., & Atchley, R. A. 1995. Semantic and associative priming in high-dimensional semantic space. *Proceedings of the Cognitive Science Society* (pp. 660–665). Hillsdale, NJ: Erlbaum.

Morton, J., & Patterson, K. E. 1980. A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. Patterson and J. C. Marshall (Eds.), Deep dyslexia (pp. 91–118). London: Routledge & Kegan Paul.

Plaut, D., & Shallice, T. 1993. Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology,* **10,** 377–500.