Semantic neighborhood effects in verbal fluency.

Treena Blake
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/3424

This online database contains the full-text of PhD dissertations and Masters’ theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000 ext. 3208.
NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.
Successful performance on verbal fluency tasks is dependent upon systematic search and retrieval from the lexicon and the semantic organization of memory stores. Previous research has established that analyzing patterns of word retrieval on verbal fluency tasks can provide useful information regarding the underlying semantic component but few studies have examined the contribution of semantics in the letter fluency task. The purpose of this study was to explore the contribution of semantics in verbal fluency performance of healthy young participants under a language-based view of semantics. Thirty-eight participants were asked to generate as many words as possible according to letter, initial phoneme, and semantic category cues. The data were analyzed using a novel computational model of semantics based on the co-occurrence of words in large samples of text. Semantic relatedness, as quantified by the average semantic weight of the words generated, was not correlated with word production but did vary as a function of time spent on the task. Overall, the results indicate that this is a promising approach to studying the contribution of semantics to lexical retrieval in a widely used neuropsychological test.
# TABLE OF CONTENTS

ABSTRACT iii

LIST OF TABLES v

LIST OF FIGURES vi

INTRODUCTION 1

- Influence of semantics in verbal fluency performance 3
- Models of semantics 9
- The present study 14

METHOD 18

- Participants 18
- Procedure 18
- Statistical analyses 20

RESULTS 21

- Number of words generated 21
- Initial phoneme performance and orthographic influences 25
- Semantic variables 25
- Category exemplars 32

DISCUSSION 35

- Directions for future research 43
- Conclusions 44

REFERENCES 45

VITA AUCTORIS 52
LIST OF TABLES

Table 1. Mean number of correct responses for each cue in the fluency tasks. 23
Table 2. Correlation coefficients for word production for three minute trials. 28
Table 3. Correlation coefficients for word production for the first minute of each trial. 29
Table 4. Top ten category exemplars according to CATSCAN. 32
Table 5. Distribution of correlations between sequence of category exemplars in CATSCAN and participants’ word lists. 34
LIST OF FIGURES

Figure 1: Word production as a function of time across fluency tasks. 24

Figure 2: Mean semantic weight as a function of time across fluency tasks. 31
Semantic Neighborhood Effects in Verbal Fluency

Introduction

Verbal fluency tasks are widely used by researchers and clinicians to examine cognitive functioning. These tasks require individuals to generate as many words as possible according to a given criterion within an allotted time period, usually one minute. In letter fluency, subjects are asked to orally produce words that begin with a specified letter, such as $c,f,$ or $l$ (Benton, Hamsher, & Sivan, 1994). Less frequently, subjects are asked to produce words that begin with specific phonemes, such as $/sh/$ or $/ma/$ (Sauzeon, Lestage, Raboutet, N'Kaoua, & Claverie, 2004). In category fluency, subjects are asked to produce exemplars from a specified semantic category, such as *animals* or *fruits and vegetables* (Goodglass & Kaplan, 1983; Newcombe, 1969). As reviewed subsequently, the underlying cognitive components of fluency performance are thought to generally reflect executive processes and the organization of semantic memory stores. The present study addresses the semantic component of fluency tasks. Following a review of literature pertaining to the influence of semantics on fluency performance, two broad classes of models of semantics are introduced. The goal of this study was to assess the contribution of semantic factors in category, initial phoneme, and letter fluency under a language-based view of semantics.

Successful performance on both letter and category fluency tasks is dependent upon systematic search and retrieval from the lexicon and semantic memory, the organized body of knowledge about the world, which consists of knowledge about facts, concepts, and words along with their meanings and associations (Tulving, 1986). As...
such, fluency tasks impose substantial demands on language as well as executive functions, including attention, effortful self-initiation, self-monitoring to prevent repetition and error, and cognitive flexibility (Monsch et al., 1994; Rosen & Engle, 1997; Ruff, Light, Parker, & Levin, 1997). Letter fluency is traditionally considered more dependent on executive functions than is category fluency because generating words on the basis of orthographic criteria is thought to be more abstract or novel and requires that individuals suppress the habit of using words according to their meaning (Perret, 1974). In contrast, category fluency is primarily based on the meaning and semantic associations of the words produced, with performance essentially reflecting the integrity of semantic memory (Hodges, Salmon, & Butters, 1992; Martin & Fedio, 1983).

The dissociation between letter and category fluency has been extended to hypotheses regarding the neural correlates underlying performance on the tasks. That is, based on a wealth of evidence that the neural substrates of executive processes lie in the frontal cortex and that temporal structures are particularly responsible for semantic memory, it has been suggested that letter fluency is primarily mediated by the frontal lobes whereas category fluency is more dependent upon temporal lobe functioning (e.g., Martin, Wiggs, Lalonde, & Mack, 1994; Moscovitch, 1994). Although there is some evidence that letter and category fluency are differentially affected by frontal and temporal lesions (Milner, 1964; Newcombe, 1969; Perret, 1974; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998) and secondary interference tasks (Martin et al., 1994; Moscovitch, 1994), empirical support for the localizing specificity of verbal fluency deficits is inconsistent (Baldo & Shimamura, 1998; Henry & Crawford, 2004; Joanette & Goulet, 1986).
Neurologically intact individuals generally produce more items on category fluency than on letter fluency tasks (Lezak, Howieson, & Loring, 2004; Tombaugh, Kozak, & Rees, 1999). This difference in performance has often been explained in terms of the degree of structure and effort required in the search and retrieval processes involved in each task. Nelson and McEvoy (1979) argued that semantic categories are smaller and better defined than letter categories and this small, organized structure should facilitate retrieval. Moreover, searching for exemplars of a category relies on well-established search strategies consistent with the inherent organization of semantic knowledge (Martin et al., 1994). Such reasoning is congruent with the assumption of many models of semantic memory in which concepts are organized in a hierarchic manner within taxonomic categories (e.g., Collins & Loftus, 1975). Accordingly, the search for category exemplars is said to depend, in large part, on the relatively passive spread of activation between closely related items in the lexicon. In contrast, generating words that begin with the same first letter requires a more effortful, strategic search because of the greater number of possible candidates and little inherent organization to guide retrieval (Martin et al., 1994).

Influence of Semantics in Verbal Fluency Performance

Although verbal fluency performance is conventionally measured by the total number of words generated, there have been attempts to characterize qualitative aspects of the data, such as the number and type of errors produced, productivity as a function of time, characteristics of the actual words generated, and patterns among the words (e.g., Chan, Butters, Paulsen, Salmon, Swenson, & Maloney, 1993; Crowe, 1998; Ober, Dronkers, Koss, Delis, & Friedland, 1986; Troyer, Moscovitch, & Winocur, 1997).
Researchers have used qualitative aspects of performance to explore the precise nature of deficient performance in patient populations and to study the structure of semantic knowledge. For example, Ober and colleagues (Ober et al., 1986) examined verbal fluency performance in patients with Alzheimer’s disease (AD) and found that even mildly demented patients gave a lower proportion of correct responses and a higher proportion of errors (perseverations and noncategory responses) than did normal elderly controls. The patients’ rate of production also tended to asymptote more quickly, relative to controls. Moreover, the authors reported that the dominance and frequency structure of the word lists did not differ between patients and controls. Specifically, despite differences in the rate of production, the mean category dominance (i.e., typicality) of the responses produced by AD patients did not differ from that of controls on the category fluency task, and both controls and demented patients produced high-frequency (i.e., very common) words early on during the letter fluency task and low-frequency words later in the task. When asked to name items found in a supermarket, patients with moderate to severe AD produced fewer responses than did their mildly-impaired counterparts and also accessed fewer subcategories (e.g., fruits or meats) with fewer responses per subcategory. Based on these findings, the authors argued that AD patients’ decreased fluency is not merely due to a slowing in the rate at which they searched memory but that AD patients are less organized in their retrieval. The findings of Ober et al. (1986) not only support a progressive breakdown in semantic memory function in AD but also provide evidence that qualitative aspects of performance on tests of verbal fluency are sensitive measures of this deficit.
Another finding that has emerged from examining qualitative output on verbal fluency tasks is that words tend to be produced in clusters. For example, it has long been known that word generation tends to occur in brief spurts over time, and words produced within these temporal clusters are often semantically related (Bousefield & Sedgewick, 1944; Gruenewald & Lockhead, 1980). A proposed model to account for the temporal pattern of responding on category fluency tests involves two processes: 1) a search for semantic subcategories and 2) an output mechanism to produce as many words as possible within the subcategories once they are identified (Gruenewald & Lockhead, 1980, Wixted & Rohrer, 1994). Troyer et al. (1997) operationalized these processes as clustering (the production of words within semantic or phonemic subcategories) and switching (the ability to shift efficiently to a new cluster when a subcategory is exhausted). According to Troyer et al.'s (1997) methodology, phonemic clusters on letter fluency are defined as successively generated words that begin with the same first two letters, differ only by a vowel sound, rhyme, or are homonyms. On the category fluency task (animal naming), semantic clusters are defined as successively generated words belonging to the same semantic subcategory, including those related to living environment (e.g., farm animals, African animals), human use (e.g., pets, fur), and various zoological categories (e.g., primates, birds, insects). Of note, the semantic subcategories were not based on an a priori organizational scheme but were derived from the actual patterns of words generated by participants during test performance. The investigators argued that the production of semantic clusters on letter fluency tasks and phonemic clusters on category tasks is infrequent and therefore, such measures are not calculated. Switching is defined as the number of transitions between clusters for both the
letter and category fluency tests. Characterized as such, Troyer and colleagues (1997) contend that clustering and switching are necessary for optimal fluency performance and together determine the total number of words generated. Furthermore, the definitions imply that temporal lobe-mediated processes such as verbal memory and lexical or semantic storage are reflected in clustering, whereas frontal lobe-mediated processes such as strategic searching, cognitive flexibility, and set-shifting are reflected in switching scores (Troyer et al., 1997). Accordingly, in healthy adults, both clustering and switching scores were significantly correlated with the total number of words generated on the category fluency task, and switching was more strongly related to word production on the letter fluency task. When this approach was applied to AD populations, patients produced smaller clusters than did normal controls on both letter and category tasks, and the contribution of clustering to word generation was reduced (Beatty, Testa, English, & Winn, 1997; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998). These findings were interpreted as consistent with impoverished or disorganized semantic memory, including lexical-phonemic stores, in Alzheimer’s disease.

While enthusiasm for Troyer et al.'s (1997) methodology has been demonstrated through its application to a number of neuropsychological disorders (e.g., Rich, Bylsmam, Troyer, & Brandt, 1999; Robert et al., 1998; Troster et al., 1998), these studies have not always yielded expected results and significant practical and theoretical limitations have been documented (see Abwender, Swan, Bowerman, & Connolly, 2001 for a review). In terms of semantic organization, clustering scores suggest that individuals' responses on the category task are influenced by semantic associations but provide little information regarding the strength of these associations. Moreover, the
definition of semantic clustering is limited to superordinate relationships or category membership and may not capture other ways in which words can be semantically related. Finally, this approach does not allow for the exploration of associative semantic effects on letter fluency tasks.

An approach to analyzing patterns of responding on verbal fluency tasks that more directly examines the organization of semantic knowledge involves the use of multidimensional statistical procedures. Such methods allow words to be plotted as points on a two- or three-dimensional map, where the distances between points are assumed to reflect the psychological proximity between the respective items in an internal semantic network (Chan et al., 1993). Chan and colleagues (Chan et al., 1993) compared the semantic network of AD patients to that of healthy controls using multidimensional scaling and additive tree clustering analyses applied to 12 commonly generated exemplars on an animal naming task. Consistent with the notion of automatic spreading activation within interconnected semantic networks, the authors argued that when individuals are asked to generate names of animals, animals that are closely associated (e.g., dog, cat) tend to cluster together, whereas animals not closely associated (e.g., dog, rhinoceros) tend not to follow each other in subjects' sequential responses. This clustering measure was considered an index of the proximity of the nodes in the semantic network of animals. The investigators then labeled each dimension on the resulting two-dimensional map with regard to the positioning of the exemplars; the dimensions were deemed to reflect domesticity and size. Relative to those of normal controls, AD patients’ cognitive maps contained extended distances between animal names, suggesting that the associations between related items (e.g., dog, cat) are weaker. In addition, AD patients
produced abnormal associations and uninterpretable clusters. These findings were reported as further evidence of a breakdown in the structure of semantic knowledge in AD.

Schwartz, Baldo, Graves, and Brugger (2003) used similar multidimensional statistical techniques to investigate the influence of semantic factors on word retrieval in the verbal fluency performance of healthy young participants. In this study, participants were given a two-choice task for both letter and category fluency (generate words starting with the letter “A” or “F” for the letter fluency task and generate names of “animals” or “fruits” for the category task) so that runs of successive words corresponding to either of the two possible cues could be easily identified. Differences in individual word lists were translated into distances on bidimensional maps using correspondence analysis followed by hierarchical clustering to delineate clusters on the maps. Similar to the finding of Chan et al. (1993) the resultant dimensions on the category task distinguished between domestic and exotic items and size, although the size dimension appeared less salient. Clustering analysis distinguished four main clusters: prototypical items from both semantic categories, wild animals, farm animals, and exotic fruits. Importantly, semantic influences were also evidenced in the letter fluency task such that participants’ responses were organized along an animate-inanimate dichotomy (i.e., airplane, floor, versus family, fish, ant). Participants’ word lists were also analyzed for successive runs of phonemically or semantically related words. In the letter fluency task, a phonemic run was defined as successive words beginning with the same letter, whereas in the category fluency task, semantic runs were defined as successive words from the same category (i.e., animals or fruits). Mean run length was greater in the category than the letter
fluency task. Finally, analysis of inter-item times revealed strong sequential priming effects in both category and letter tasks. In accordance with a connectionist perspective, the results were taken to suggest that the links within semantic categories like animals and fruits are tighter than phonemic links, resulting in a quicker spreading activation between semantic nodes of a given category than between words beginning with the same letter. Overall, the authors concluded that their results provided the first report that semantic facilitation is pervasive in word retrieval processes, in both letter and category fluency.

The results of studies employing multidimensional statistics are notable for the information they provide regarding the underlying semantic organization of responses generated in fluency tasks. However, this approach only allows words to be represented in semantic space along two dimensions (domesticity and size in this case) and thus, may fail to capture important aspects of the data since words can be related in many ways.

Models of Semantics

The approaches reviewed above clearly demonstrate that analyzing the sequential pattern of responses on verbal fluency tasks contributes to the understanding of the influence of semantics in verbal fluency performance. The findings of Troyer et al. (1997), Chan et al. (1993), and Schwartz et al. (2003) are consistent with the well-established observation that semantic associations facilitate the generation of related concepts. In this manner, the retrieval of words is primed or activated by preceding semantically related words, particularly on the category fluency task (Gruenewald & Lockhead, 1980). In the aforementioned studies, it is implied that semantic relationships are based on perceptual features or on taxonomic category membership. That is, cat and
Dog are said to occur in close proximity within individuals' word lists (and by extension, in their semantic networks) because these concepts share a number of features (e.g., four legs, tail, fur, etc.) or belong to the same category (e.g., domestic pets). In this sense, concepts that share many features, such as CAT and DOG, would be considered close semantic neighbors. The spread of activation between semantic neighbors is dependent upon the number of shared features (i.e., as features for CAT become activated, several features of its neighbor DOG also become activated) because the more features two concepts have in common, the more connections that exist between the two nodes via these features. Such an explanation is consistent with a feature- or object-based model of semantics (Buchanan, Westbury, & Burgess, 2001).

An alternative viewpoint is that words are meaningfully related on the basis of association or co-occurrence in language use (Buchanan et al., 2001). Whereas object-based models are based on properties shared by the concepts underlying words, associative relationships are based on the properties of the words themselves, as well as on the frequency with which words co-occur in spoken or written language. Language-based models regard the words cat and dog as close semantic neighbors, not because they share features, but because they appear in similar contexts in large samples of written text or through frequent spatiotemporal co-occurrence in language use (e.g., Landauer & Dumais, 1997; Lund & Burgess, 1996; Lund, Burgess, & Audet, 1996; Nelson, McEvoy, & Schreiber, 1998). As such, association or language-based models of semantics allow concepts to be considered semantic neighbors despite the absence of shared features. For example, the concepts DOG and HOUSE do not share features and would not be considered semantically related under an object-based view of semantics. However, in
terms of a language-based viewpoint, DOG and HOUSE would be considered semantic neighbors on the basis of one's experience with these concepts in similar contexts (Locker, Simpson, & Yates, 2003).

Support for language-based models of semantic organization stems largely from semantic priming experiments, which consistently demonstrate that processing of a target word is facilitated when it is preceded by a related word, compared with a neutral or unrelated word. Because semantic priming effects are observed on tasks that do not require participants to access a word’s meaning (e.g., lexical decision tasks), it has been suggested that semantic priming reflects the underlying organization of concepts in semantic memory (Meyer & Schvaneveldt, 1971). Semantic priming has reliably been found to occur among category members and word pairs that overlap in terms of features, but also for word pairs that do not share features such as knife-bread (e.g., Moss, Ostrin, Tyler, & Marslen-Wilson, 1995). This finding does not accord with an object-based view of semantics since activation would not be expected to spread from KNIFE to BREAD because they do not share features. Furthermore, in a recent meta-analysis, Hutchinson (2003) concluded that word pairs that share both feature-based and associative relationships show larger priming effects than nonassociated pairs, an observation termed an associative boost. Glosser, Friedman, Grugan, Lee, and Grossman (1998) also assessed priming for associatively related words in a sample of patients with AD. Whereas healthy elderly controls showed the effects of priming for words that were associatively related and also for words that were related by virtue of category membership, AD patients evidenced facilitatory priming of lexical associations but impaired priming effects for category member primes. These results suggested that
lexical associations are relatively preserved in the context of disrupted semantic knowledge in AD and further demonstrate the relevance of associational relationships in the structure of semantic knowledge. As Buchanan et al. (2001) note, it is likely that both object- and language-based relationships play an important role in the structure of the semantic system, although the extent and precise nature of these roles is yet to be fully understood.

An example of a language-based conceptualization of how meaning becomes represented in semantic memory is presented in the hyperspace analogue to language (HAL) model developed by Lund and Burgess (1996). Briefly, this computational model produces a high-dimensional semantic space by using a global co-occurrence learning algorithm to track lexical co-occurrences across a large sample of written text. HAL encodes the contexts in which words are found as weighted co-occurrences and then sums these co-occurrence values, resulting in a high-dimensional co-occurrence matrix containing information from the entire learning history of the model. Each word in the corpus is then represented by a vector and these vectors constitute points in high-dimensional semantic space such that words with similar vectors occur relatively close together. Of importance, overall meaning of a word can be extracted by examining the surrounding context in which the word appears and thus, words in similar regions of the high-dimensional space are considered semantically related (see Burgess, 1998). This is true for words that co-occur in similar contexts and for words that seldom co-occur locally (e.g., road and street rarely occur in the same sentence, but they appear in similar contexts and therefore have very similar vector representations). Results from a number
of simulations provide support for HAL as a viable representational model of semantic
memory (see Lund & Burgess, 1996).

One of the primary advantages of language-based models such as HAL is that they are able to provide quantitative representations of a word's semantic characteristics. For example, the semantic neighborhood of a word can be defined as some set of words that are close to it in the semantic space constructed by HAL. In a recent study, Buchanan et al. (2001) employed the HAL model to investigate the effects of semantic neighborhood size on visual word recognition. They defined the semantic distance of a word as the mean distance between that word and its 10 closest neighbors in high-dimensional semantic space. Therefore, within a specified radius in semantic space, a word with a high semantic distance will, on average, have fewer semantic neighbors, or a smaller semantic neighborhood, than will a word with a low semantic distance. Similarly, words with low semantic distance values will have large semantic neighborhoods. Using this measure of semantic distance together with regression techniques, Buchanan et al. (2001) examined whether semantic neighborhood size impacted performance on lexical decision and naming tasks. They reported that semantic distance predicted lexical decision latencies and this effect was not due to the values being similar to nonsemantic measures such as frequency and orthographic neighborhood size. Moreover, the positive partial correlation obtained indicated that as semantic distance decreased, so did response latency. Thus, larger semantic neighborhoods (as defined by low semantic distance) assisted in lexical decision. Semantic distance also correlated with reaction times even when a more traditional semantic variable (imageability) was included in the analysis, suggesting that semantic distance captured unique, semantically relevant information.
about words. On the basis of these results and those of a series of factorial experiments that followed, Buchanan and colleagues concluded that semantic neighborhood, as defined by semantic distance, plays an important role in the processing of words in lexical decision and this effect is facilitatory.

The intriguing findings of Buchanan et al. (2001) imply that measures of semantic distance derived from language-based models hold promise as a means for investigating neighborhood effects on tasks assumed to rely on semantic activation. Indeed, the effects of semantic distance have been investigated in a variety of paradigms including false memories (Buchanan, Brown, Cabeza, & Maitson, 1999), semantic categorization tasks (Siakaluk, Buchanan, & Westbury, 2003), and recognition of ambiguous words (Locker, Simpson, & Yates, 2003). It remains open to question whether these variables contribute to performance on other tasks that have been used to investigate semantic structure, including verbal fluency.

**The Present Study**

The purpose of the present study was to explore the contribution of semantic factors, as quantified by a computational model of semantics, in letter, phoneme, and category fluency by analyzing patterns of word retrieval in healthy young participants. The current investigation used semantic variables derived from a recently developed language-based model of semantics, CATSCAN (Computational Analysis of Text: Semantic Co-occurrence Association Norms; Durda, Casey, Buchanan, & Caron, in press). Akin to the HAL model (Lund & Burgess, 1996), CATSCAN is a computational model of semantic associations based on global co-occurrence of words in text (see Durda et al., in press, for a comparison of this method to existing computational models).
CATSCAN constructed a high-dimensional semantic space from a lexical co-occurrence matrix that was formed by analyzing a word corpus of approximately 267 million words obtained from works of literature, the British National Corpus, as well as many government, advertising, and technical documents collected from the internet. Similar to HAL, each word in the corpus is represented by a vector made up of the other words with which it co-occurs. Since words that are closely related tend to co-occur with the same other words, these vectors are assumed to represent a word's meaning. The CATSCAN method calculates a semantic weight for each word by averaging the 10 highest lexical co-occurrence values, essentially representing the 10 closest neighbors in high-dimensional semantic space. Note that semantic weight corresponds to the notion of semantic distance as defined in Buchanan et al. (2001). However, words with a high semantic distance according to the definition of Buchanan et al. (2001) will have a low semantic weight, and words with a low semantic distance will have a high semantic weight. To avoid confusion arising from the inverse relationship, these measures can be described in relation to semantic neighborhoods: A large semantic neighborhood is associated with a high semantic weight value, but a low semantic distance value. This word can be said to have a dense semantic neighborhood in that the word is relatively close to its 10 closest neighbors in semantic space. Conversely, a word with a low semantic weight value, therefore having a high semantic distance value, would have a sparse neighborhood as it is relatively distant from its 10 closest neighbors in high-dimensional semantic space. The decision to use the ten highest co-occurrence values in determining semantic weight follows from Buchanan et al.'s (2001) assertion that this value provides a reasonable indication of the likelihood that a selected cutoff point would
allow the inclusion of many neighbors for a "close" neighborhood and few neighbors for a "distant" neighborhood.

A primary goal of the present study was to determine whether semantic neighborhoods, as defined by semantic weight, play a role in verbal fluency performance. Semantic neighborhood effects have been found to predict word recognition performance (Buchanan et al., 1999; Buchanan et al., 2001; Locker, Simpson, & Yates, 2003; Siakaluk et al., 2003) and it may well be that similar effects can be obtained on other tasks assumed to rely on semantic activation. Based on previous research, it was expected that the influence of semantic weight on verbal fluency would be facilitatory in nature. That is, it was assumed that the retrieval process is sensitive to the distance between words in semantic space such that having many close neighbors will facilitate lexical access. This is in line with spreading activation theory which indicates that activation within semantic networks spreads automatically between close nodes, and decays with increasing distance between nodes (Collins & Loftus, 1975). Therefore, higher semantic weights within individuals' word lists were expected to be associated with greater word production on the task. In view of the fact that Schwartz et al. (2003) demonstrated semantic facilitation on both category and letter tasks, it was predicted that semantic neighborhood effects would be found in both letter and category fluency tasks.

Another aim of the current investigation was to explore how semantic characteristics vary over the course of verbal fluency tasks. In addition to assessing whether individuals tend to retrieve words that reside in large or small semantic neighborhoods, the present study explored the distribution of semantic weights of words within individuals' lists. For instance, individuals may begin generating words that are
closely connected and readily activated (i.e., words with high semantic weights) and then generate words with fewer neighbors (i.e., words with low semantic weights) later in the task. Participants may also alter their search strategy based on task demands such that patterns on letter fluency differ from those found on category fluency.

Previous studies have established that analyzing patterns of word retrieval on verbal fluency tasks can provide useful information regarding the semantic component underlying these tasks. The present study expanded upon previous findings by invoking a novel approach to analyzing output on letter and category fluency tasks. Whereas previous studies have examined semantic associations on verbal fluency tasks from an object-based perspective, the current investigation sought to determine whether language-based relationships also contribute to performance. The use of a computational model such as CATSCAN to compute distance between words in high-dimensional space has several advantages over previous studies that defined semantic proximity as closeness on a two-dimensional cognitive map (e.g., Chan et al., 1993; Schwartz et al., 2003) or by coding responses based on category membership (Troyer et al. 1997). Firstly, this computational method does not rely on human judgments. Second, this model allows words to be related in many ways, not just on the basis of superordinate category membership, or along a small number of dimensions. Finally, words that do not share features can be semantically related in CATSCAN. This aspect may be of particular importance to investigating the influence of semantics on letter fluency tasks, which have been under-examined to date.
Method

Participants

Forty-four undergraduate psychology students from the University of Windsor participated in this study for course credit. All participants spoke English as their first language. Data from six participants with a known history of psychiatric or neurological disorder, learning disability, attention deficit hyperactivity disorder, or history of substance abuse was excluded from the analyses. The remaining 38 participants (29 females, 9 males) had an average age of 24 years (range = 18 – 48 years). The protocol was approved by the ethics committee of the university and written informed consent was obtained from each student.

Procedure

Participants were tested individually in a quiet room. Demographic information including age, gender, years of education, as well as information regarding exclusionary criteria was obtained prior to administration of the fluency tasks. Each participant performed three trials of category fluency, letter fluency, and initial phoneme fluency. The initial phoneme task was intended to limit orthographic constraints on the task. The order of the letter and initial phoneme tasks was counterbalanced across participants and the category task was always administered last. The order of the trials within each type of task was counterbalanced among participants.

For letter fluency, participants were instructed to name as many words as possible that begin with the letters A, L, and W, excluding proper names and variations on previously produced words, according to the procedures outlined by Spreen and Strauss (1998). For initial phoneme fluency, participants were asked to name as many words as
possible that begin with the sounds /f/, /k/, and /dʒ/, excluding proper names and variations on previously produced words. These cues were selected because possible responses can vary in terms of the initial letter (i.e., for the cue /f/, participants could generate words that begin with the letter “f” or “ph”), thereby allowing the possible influence of orthographic constraints to be assessed. Half the participants received the phonemic cue followed by an example word starting with a specified letter (e.g., /f/ as in fox, /k/ as in key, /dʒ/ as in jar) and the other half received an example corresponding to the other possible spelling (e.g., /f/ as in phone, /k/ as in cake, /dʒ/ as in gem).

Immediately following each trial of the letter and initial phoneme tasks, any words with multiple spellings (e.g., for, fore, four) were reviewed with the participant to clarify the intended response.

For category fluency, participants were instructed to name as many animals, fruits, and tools as possible (adapted from Spreen & Strauss, 1998). Participants were asked to avoid different subtypes of the same fruit (e.g., Macintosh apples, Golden Delicious apples, etc.). Participants were informed that tools may include any hand-held instrument with a specific function in order to expand the range of possible responses from typical exemplars (e.g., hammer, screwdriver) to include such items as fork, spoon, and comb.

Each trial lasted three minutes rather than the standard one-minute time period, in order to exhaust participants’ lexicons. To prevent fatigue, participants were allowed brief breaks between trials, as needed. All responses were tape-recorded and simultaneously transcribed by the examiner.
Statistical Analyses

Due to the exploratory nature of the study, the approach to data analysis emphasized descriptive and correlational methods. The results were explored in the following manner: (a) the total number of words generated by participants in each of the fluency tasks was compared; (b) responses on the initial phoneme task were further analyzed to explore any effect of the prompt given; (c) CATSCAN-derived semantic measures were examined for each fluency task; and (d) responses on the category fluency tasks were compared to the most strongly related exemplars of a category, as generated by the CATSCAN method.
Results

Number of Words Generated

The total number of words produced during each trial, excluding repetitions and errors (e.g., proper names, non-category responses) was calculated for each participant. Prior to statistical analyses, potential outliers were identified using boxplots and z-values (any z-value > ± 2.5 was removed from the data set). Where applicable, Mauchly’s Test was examined to determine whether the assumption of sphericity was met. Sphericity is the assumption of compound symmetry of the covariance matrix, which is required for the application of the repeated measures analysis of variance procedure (Stevens, 2002). In all instances, the assumption of sphericity was met.

The data for trials within the letter, initial phoneme, and category conditions were collapsed in order to compare overall word production across tasks. A repeated measures ANOVA indicated that the number of words generated by participants differed significantly across tasks, $F(2, 56) = 30.16, p < .0001$. A priori contrasts were set up based on previous research indicating that participants generate more words on category tasks than on letter tasks. These analyses revealed that participants generated significantly more words overall in the category fluency task ($M = 74.34, SD = 13.92$) than in the letter fluency ($M = 63.48, SD = 16.78$), or initial phoneme task ($M = 51.38, SD = 18.21$), $F(1, 28) = 33.75, p < .0001$. Furthermore, participants generated significantly more words overall in the letter fluency task than in the initial phoneme task, $F(1, 28) = 23.62, p < .0001$.

Subsequent analyses were performed on the separate cues in each fluency task. On letter fluency, participants produced a comparable number of words beginning with
the letters A, L, and W, $F(2,68) = 2.97$. A significant difference among cues was found on the initial phoneme task, $F(2, 74) = 50.43, p < .0001$. Post hoc comparisons using multiple $t$ tests with a Bonferroni correction of $p < .0167$ showed that participants produced significantly fewer words beginning with the sound /dʒ/ than beginning with the sounds /f/ ($t(37) = 8.91, p < .0001$), and /k/ ($t(37) = 7.81, p < .0001$). In contrast, the number of words that participants produced beginning with the sounds /f/ and /k/ did not differ, $t(37) = 1.66$. A significant difference among the cues on the category fluency task was also found, $F(2, 62) = 68.18, p < .0001$. Post hoc comparisons indicated that participants generated significantly more animals than fruits ($t(31) = 8.12$, $p < .0001$), and tools ($t(36) = 11.59, p < .0001$). In turn, participants named significantly more fruits than tools, $t(31) = 2.70, p < .0167$. The mean number of words generated for each cue in the three fluency tasks is presented in Table 1. Additionally, word production for the fluency tasks was graphed over time and can be seen in Figure 1.
Table 1

*Mean Number of Correct Responses for Each Cue in the Fluency Tasks*

<table>
<thead>
<tr>
<th>Fluency Task</th>
<th>Cue</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter</td>
<td>A</td>
<td>21.03</td>
<td>6.42</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>22.94</td>
<td>8.68</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>21.83</td>
<td>5.39</td>
</tr>
<tr>
<td>Initial Phoneme</td>
<td>/f/</td>
<td>21.55</td>
<td>8.58</td>
</tr>
<tr>
<td></td>
<td>/k/</td>
<td>19.95</td>
<td>8.19</td>
</tr>
<tr>
<td></td>
<td>/dʒ/</td>
<td>11.95</td>
<td>4.99</td>
</tr>
<tr>
<td>Category</td>
<td>animals</td>
<td>33.69</td>
<td>7.68</td>
</tr>
<tr>
<td></td>
<td>fruits</td>
<td>22.25</td>
<td>3.98</td>
</tr>
<tr>
<td></td>
<td>tools</td>
<td>19.16</td>
<td>6.95</td>
</tr>
</tbody>
</table>
Figure 1. Word production as a function of time across fluency tasks.
Initial Phoneme Performance and Orthographic Influences

When asked to generate words on the basis of initial phoneme, participants produced a disproportionate number of words spelled with one of the two possible initial letters, irrespective of the prompt that was administered: 94% of the /f/ words produced began with the letter “f”, 83% of the /dʒ/ words produced began with the letter “j”, and 82% of the /k/ words produced began with the letter “c”. To determine whether the prompt given during the instructions had an effect on the proportion of words that begin with either letter, independent samples t tests were conducted. The prompt administered during task instructions significantly influenced the composition of participants’ word lists in two of the initial phoneme tasks. The average proportion of words that begin with the letters “ph” was greater for participants who were asked to generate words beginning with the sound /f/ “as in phone” than for participants who were instructed to generate words beginning with the sound /f/ “as in fox”, t(36) = 2.31, p < .05. Similarly, the average proportion of words that begin with the letter “g” was greater for participants who were asked to generate words beginning with the sound /dʒ/ “as in gem” than for participants who were instructed to generate words beginning with the sound /dʒ/ “as in jar”, t(36) = 2.91, p < .05. The proportion of words produced that began with the letter “k” did not differ between participants who were instructed to give words that begin with the sound /k/ “as in key”, and those who were given the prompt /k/ “as in cake.”

Semantic Variables

The semantic weight of every correct response in each participant’s word lists, as computed by the CATSCAN method, was determined. The category tools was excluded from further analyses due to the number of two word responses given to this cue (e.g.,
can opener, measuring tape, hair brush). At the present time, the CATSCAN database contains information only for single words. Further, semantic weights for certain words that appeared in individuals' lists were not available. For instance, the CATSCAN database does not contain information for the words as, at, are, any; consequently, these words were omitted from the analyses for the letter-A fluency task. Generally, few words in each list were omitted due to missing semantic weight values (mean across cues = 1.6, range = 0 - 7). For consistency, all responses in the category fluency task were converted to singular form prior to deriving semantic weights.

To determine whether semantic weight was related to word production, bivariate correlations between the average semantic weight of responses within each participant's word list and the total number of words generated were calculated. Due to the variability in the total words generated across different cues in the fluency task conditions, separate correlations were calculated for each cue. Prior to these analyses, the association between semantic weight and word production was depicted graphically via scatterplots and outliers were identified and removed. As can be seen in Table 2, a significant association between the total words generated and the semantic weight of the words was obtained only for the letter-A fluency task, $r(37) = .393, p < .05$. In this case, as the average semantic weight of words on an individual's list increased, so did the length of the list. In order to investigate whether the association between semantic weight and word production was attenuated by the extended time period used in the present study, these analyses were repeated for the first minute of each trial. As can be seen in Table 3, a consistent pattern of association between average semantic weight and word production did not emerge when the more conventional time period was used. Word production
during the first minute of the trial and semantic weight were not significantly correlated for any of the letter or initial phoneme cues. However, for category fluency, the number of animals generated during the first minute of the trial was inversely related to the semantic weight of the words contained in the list, $r(37) = -.343, p < .05$. On average, participants who generated more animal names tended to provide animal names with lower semantic weights during the first minute of the task. A similar trend was observed for the fruit category, although this correlation failed to reach statistical significance, $r(32) = -.323, p = .07$.

Because the organization of the lexicon is based on features besides semantic information, each word was also coded for frequency and orthographic neighborhood size. The relationship between these variables and word production is shown in Table 2. Word frequency was significantly correlated with word production on category fluency such that lists containing less typical animal and fruit names overall tended to be longer in length ($r(36) = -.590, p < .001$ for animals, $r(31) = -.486, p < .001$ for fruits).

Orthographic neighborhood size refers to the number of words that differ from a given word by only a single letter. As shown by the correlations in Table 2, orthographic neighborhood size was positively correlated with the total number of words produced with A and W as the initial letter, and inversely related to number of animals named. When these analyses were applied to data from the first minute of each fluency trial, only the correlation between frequency and productivity on the animal task remained statistically significant. In addition, word frequency was significantly correlated with productivity during the first minute of the letter-A ($r(31) = .333, p < .05$), and letter-W ($r(33) = -.374, p < .05$) tasks, as shown in Table 3.
Table 2

Correlation Coefficients for Word Production for Three Minute Trials

<table>
<thead>
<tr>
<th>Words generated</th>
<th>(n)</th>
<th>SW</th>
<th>Logfreq</th>
<th>ON</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Letter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>37</td>
<td>.393*</td>
<td>.315</td>
<td>.343*</td>
</tr>
<tr>
<td>L</td>
<td>35</td>
<td>.116</td>
<td>-.246</td>
<td>.248</td>
</tr>
<tr>
<td>W</td>
<td>38</td>
<td>---</td>
<td>-.230</td>
<td>.359*</td>
</tr>
<tr>
<td><strong>Initial Phoneme</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/f/</td>
<td>35</td>
<td>.050</td>
<td>-.251</td>
<td>.084</td>
</tr>
<tr>
<td>/dʒ/</td>
<td>38</td>
<td>-.153</td>
<td>-.052</td>
<td>-.055</td>
</tr>
<tr>
<td>/k/</td>
<td>35</td>
<td>.064</td>
<td>-.187</td>
<td>-.004</td>
</tr>
<tr>
<td><strong>Category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>animals</td>
<td>36</td>
<td>-.309</td>
<td>-.590**</td>
<td>-.355*</td>
</tr>
<tr>
<td>fruits</td>
<td>31</td>
<td>.050</td>
<td>-.486**</td>
<td>-.240</td>
</tr>
</tbody>
</table>

*p < .05  **p < .001

Note. SW = average semantic weight of words in individuals’ lists, Logfreq = average log frequency of words in individuals’ lists, ON = average number of orthographic neighbors for words in individuals’ lists.

The scatterplot indicated that the relationship between average semantic weight and total words beginning with the letter W was not well described by a linear function such that calculation of a Pearson correlation coefficient was inappropriate.
Table 3

*Correlation Coefficients for Word Production for the First Minute of Each Trial*

<table>
<thead>
<tr>
<th>Words generated in first minute of trial</th>
<th>Variable</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SW</td>
<td>Logfreq</td>
<td>ON</td>
</tr>
<tr>
<td>Letter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>31</td>
<td>-.280</td>
<td>.333*</td>
<td>-.122</td>
</tr>
<tr>
<td>L</td>
<td>35</td>
<td>.331</td>
<td>-.296</td>
<td>.174</td>
</tr>
<tr>
<td>W</td>
<td>33</td>
<td>.133</td>
<td>-.374*</td>
<td>-.031</td>
</tr>
<tr>
<td>Initial Phoneme</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/b/</td>
<td>35</td>
<td>-.205</td>
<td>-.129</td>
<td>-.127</td>
</tr>
<tr>
<td>/d3/</td>
<td>34</td>
<td>.148</td>
<td>.178</td>
<td>.078</td>
</tr>
<tr>
<td>/k/</td>
<td>37</td>
<td>.094</td>
<td>.127</td>
<td>.114</td>
</tr>
<tr>
<td>Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>animals</td>
<td>37</td>
<td>-.343*</td>
<td>-.471**</td>
<td>-.196</td>
</tr>
<tr>
<td>fruits</td>
<td>32</td>
<td>-.323</td>
<td>-.263</td>
<td>.081</td>
</tr>
</tbody>
</table>

*p < .05   **p < .001

*Note. SW = average semantic weight of words in individuals’ lists, Logfreq = average log frequency of words in individuals’ lists, ON = average number of orthographic neighbors for words in individuals’ lists.*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
To explore whether semantic weight varies as a function of time spent on the task, the average semantic weight of words produced during each minute of the task was analyzed in a series of 3 x 3 (Time interval x Cue) repeated measures ANOVAs. The Greenhouse-Geisser correction was applied when the assumption of sphericity was not met (Stevens, 2002). In the letter fluency condition, there was a main effect of time, $F(2, 72) = 6.905, p < .01$ (Greenhouse-Geisser), a main effect of cue, $F(2, 72) = 6.219, p < 01$ (Greenhouse-Geisser), and a significant interaction, $F(4,144) = 3.977, p < .05$ (Greenhouse-Geisser). The interaction indicated that the average semantic weight of the words generated was not the same for the three letter cues across the time intervals (see Figure 2). An analysis of the simple main effects revealed that semantic weight did not vary across time intervals for the letters L and W; however, the average semantic weight of responses beginning with A decreased significantly over the first two minutes of the task.

In the Initial Phoneme condition, there was only a significant main effect of cue, $F(2,74) = 3.895, p < .05$ (Greenhouse-Geisser). As illustrated in Figure 2, combined across the three time intervals, the average semantic weight of words beginning with the sound /dʒ/ was less than the average semantic weight of words beginning with the sounds /ʃ/ or /kl/.

Finally, in the Category condition, a significant main effect of time was found, $F(2, 64) = 4.337, p < .05$ (Greenhouse-Geisser) indicating that the average semantic weight of animal names and fruit names decreased over the course of the task (see Figure 2).
Figure 2. Mean semantic weight as a function of time across fluency tasks.
Category Exemplars

Durda et al. (in press) reported the ten most strongly related category exemplars for each category cue, as determined by the vectors computed by the CATSCAN method (see Table 4).

Table 4

Top Ten Category Exemplars According to CATSCAN

<table>
<thead>
<tr>
<th>Animal</th>
<th>Fruit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>Apple</td>
</tr>
<tr>
<td>Dog</td>
<td>Orange</td>
</tr>
<tr>
<td>Horse</td>
<td>Grapes</td>
</tr>
<tr>
<td>Cat</td>
<td>Peach</td>
</tr>
<tr>
<td>Cattle</td>
<td>Melon</td>
</tr>
<tr>
<td>Bird</td>
<td>Pear</td>
</tr>
<tr>
<td>Fish</td>
<td>Strawberries</td>
</tr>
<tr>
<td>Sheep</td>
<td>Fig</td>
</tr>
<tr>
<td>Bull</td>
<td>Kiwi</td>
</tr>
<tr>
<td>Fox</td>
<td>Banana</td>
</tr>
</tbody>
</table>

On average, participants named about 6 of the animals that CATSCAN identified as the top exemplars ($M = 5.6 \text{ words}, SD = 1.79, \text{ range} = 2 -10), and 58\% of participants generated at least half of the animal exemplars. Participants named more fruits that CATSCAN identified as the top exemplars ($M = 7.64, SD = 1.25, \text{ range} = 4-10), and 97\% of participants generated at least half of the fruit exemplars.
The order of the exemplars in the above lists is said to reflect the strength of association to the category superordinate (Durda et al., in press). To determine whether the order that these exemplars appeared in individuals’ word lists was similar to the order specified by CATSCAN, the sequence of these exemplars within each individual’s word lists was identified and Spearman’s rho correlations were computed. The distribution of these correlations on the animal and fruit naming tasks is shown in Table 5. As can be seen, the sequence of exemplars in participants’ lists was at least moderately correlated with the sequence provided by CATSCAN in 63% of cases for animal naming and 52% of cases for the fruit naming task.
Table 5

*Distribution of Correlations Between Sequence of Category Exemplars in CATSCAN and Participants' Word Lists*

<table>
<thead>
<tr>
<th>Rho Score Range</th>
<th>Animal Naming (n=38)</th>
<th>Fruit Naming (n=33)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>.9-.999</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>.8-.899</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>.7-.799</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>.6-.699</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>.5-.599</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>.4-.499</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>.3-.399</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>.2-.299</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>.1-.199</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>0 -.099</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>&lt;0</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

*Note.* Dashed line indicates cutoff for a moderate correlation.
Discussion

Quantitative analyses revealed that participants generated more words overall in the category fluency task than in the letter fluency task, consistent with previous findings for normal individuals (e.g., Crowe, 1998; Lezak et al., 2004; Monsch et al., 1994; Tombaugh et al., 1999; Schwartz et al., 2003). This pattern of results supports the conception that letter fluency is a more difficult task, at least for the letters and categories studied. As noted by Azuma et al. (1997), semantic categories follow hierarchical memory organization and have natural subcategories that letter categories lack, thereby facilitating search and retrieval. In contrast, search and retrieval guided by letter cues imposes greater demands on working memory and cannot rely on the inherent organization of semantic memory (Azuma et al., 1997; Martin et al., 1994). Thus, the letter task is more effortful and participants produce fewer correct responses. The observed disparity between letter and initial phoneme fluency performance may also reflect differences in the relative difficulty of the tasks. Unfortunately, few studies have compared performance on letter and initial phoneme fluency tasks as the latter is typically used in place of a letter fluency task in circumstances in which orthographic knowledge may confound results, such as with young children (e.g., Sauzeon et al., 2004) or illiterate populations (e.g., Ratcliff et al., 1998). Like letter categories, initial phoneme categories consist of large sets of possible candidates with little inherent organization to guide retrieval. Therefore, it is not surprising that participants produced fewer words on the initial phoneme fluency task than on the category fluency task.

The finding that participants also produced fewer correct responses in the initial phoneme condition than in the letter condition may indicate that generating words to a
phonemic rule is a more difficult task. Indeed, the search strategy employed in a letter fluency task could be considered less novel and more consistent with the organization of the world. For instance, research in the field of metacognition has demonstrated that college students frequently use first-letter mnemonics in an attempt to organize information and facilitate later retrieval, despite empirical evidence that this technique is not particularly effective (West, 1995). Furthermore, the familiarity of alphabetical systems of organization (e.g., telephone directories, library systems) may enhance the ability of individuals to retrieve words on the basis of orthography. In addition, the demographic characteristics of the sample suggests that the majority of participants in this study likely learned to read through whole language instruction rather than by phonics, supporting the notion that retrieving words based on sounds would be relatively unfamiliar. Any hypotheses regarding underlying differences in the generation of words to initial letter versus initial phoneme cues require empirical validation.

An alternative explanation arises upon examination of word production to the separate cues in the initial phoneme task. Participants generated fewer words beginning with the sound /dʒ/ than to the other phonemic cues, and this difference may account for the disparity in overall word production on the letter and phoneme tasks. In fact, when the average number of words generated in each fluency task was compared without the /dʒ/ trial, performance on the letter and phoneme tasks did not differ (a priori contrast, $F(1,31) = 1.363$). Thus, considerable variability in performance for individual phoneme cues contributed to the observed difference in word production between letter and phoneme fluency tasks. Further research is clearly needed to replicate this finding and explore potential reasons that it might be more difficult to generate words in response to
this particular criterion. Nonetheless, the present results suggest that conclusions about cognitive abilities derived from fluency performance can be highly influenced by the individual cues used in the tasks. Furthermore, these results indicate that caution should be exercised in clinical settings wherein performance on phoneme fluency tasks may be interpreted in the context of normative expectations on letter fluency tasks.

Variability within the category fluency condition was also demonstrated, as participants generated more animal names than types of fruits, followed by names of tools. This pattern of responding is consistent with previous findings (Azuma, 2004; Azuma et al., 1997) and may reflect, in part, differences in category size. The category animal has more exemplars than the category fruit and larger semantic categories tend to have a greater degree of organizational structure in memory (Azuma et al., 1997). That is, large semantic categories such as animals have a number of subcategories which can be used to facilitate or organize retrieval (i.e., are more amenable to clustering). Taken together, the present results support the work of Azuma and colleagues (Azuma, 2004; Azuma et al., 1997) that calls attention to the need for researchers to consider variations in the nature of the categories used in fluency tests before drawing conclusions about the relative difficulty of, or effect of disease on, verbal fluency tasks.

When participants were instructed to generate words to a phonemic rule, a markedly disproportionate number of their responses began with one of the two possible initial letters, irrespective of the prompt that was provided (i.e., 94% of the /f/ words produced began with the letter “f”, 83% of the /dʒ/ words produced began with the letter “j”, and 82% of the /k/ words produced began with the letter “c”). This finding may indicate that participants translated the initial phoneme cue to an associated grapheme.
and then generated responses on the basis of this criterion. If this were the case, similar search and retrieval strategies were used in the initial phoneme and letter fluency tasks.

Although orthographic influences were pervasive on this task, as reflected in the proportion of responses beginning with one of the possible letters, individuals' lists contained some words beginning with the alternate spelling. This finding indicates that participants were not altogether insensitive to phonological representations while performing the task. Moreover, the prompt that was administered during the instructions significantly influenced the composition of participants' word lists on two of the phoneme trials. While the overall proportion of responses beginning with the letter “ph” and “g” was small, participants that were given the corresponding examples (i.e., “phone” and “gem”), generated lists with a significantly greater proportion of words beginning with these letters.

In addition to the effect of task instructions, differences in the orthographic composition of the word lists was likely influenced by the number of possible candidates in the lexicon. Accordingly, the number of three- to ten-letter words beginning with each of the possible letter combinations was acquired from the MRC Psycholinguistic Database. This analysis revealed a pattern congruent with that obtained in the fluency task data: far more “f” versus “ph” words, “j” versus permissible “g” words, and permissible “c” versus “k” words exist in the English language. However, this explanation cannot fully account for the results because there were hundreds of possible candidates even in the smallest category, and therefore, participants did not merely deplete the available word store during the task. Undoubtedly, individuals’ lexicons do not contain all the entries found in a dictionary and word frequency plays an important
role in the retrieval process. For example, of the 245 words in the online dictionary beginning with the letters “ph”, only four have written word frequencies greater than 50 per million (phase, philosophy, phone, physical) which likely contributed to the relative difficulty of generating responses to this cue. Further research is needed to consistently establish the relative difficulty of the cues used in this task and to compare performance on initial phoneme and letter fluency tasks in normal individuals. Subsequent studies could then compare the strategies used by illiterate subjects versus skilled readers to further elucidate the influence of orthography on initial phoneme versions of the fluency task.

In keeping with the primary goal of exploring the contribution of semantic factors in verbal fluency, the semantic weight of every correct response was obtained from the CATSCAN database and averaged across each word list. This value represents a global measure of the semantic relatedness of the words generated to each cue. As follows, lists with high average semantic weights contain words drawn from dense semantic neighborhoods, whereas lists with lower average semantic weights contain words drawn from sparser neighborhoods. Consistent with spreading activation theory, higher semantic weights within individuals’ word lists were expected to be associated with increased word production. The present results do not support this hypothesis. A consistent relationship between average semantic weight and productivity across fluency tasks was not found. In fact, a significant correlation was found only for words generated to the letter-A cue. As the semantic weights of the “A” words increased so did the length of the list. However, this finding did not extend to the other letter cues or fluency conditions.
Therefore, few conclusions can be drawn about the relationship between productivity and overall semantic weight of responses.

Notably, the results also did not provide evidence of a relationship between frequency and word production in the letter or initial phoneme tasks, despite previous findings indicating moderate to high correlations on letter fluency tasks in healthy young adults (Crowe, 1998). In contrast, significant correlations between word production and frequency were found in the category fluency task such that lists containing more low-frequency words overall, tended to be longer. Although the inverse relationship between frequency and productivity might be considered a curious finding, given that frequency is thought to facilitate retrieval, it indicates that participants who continue to search their lexicon and retrieve low-frequency words after more common words have been exhausted generally produce more words. This was well-illustrated by the participant who generated the most animal names during the task -- not only did she provide the common responses cat, dog, horse, which are high-frequency words, but she also included the low-frequency words jellyfish, centipede, and lemur that most participants did not provide. The observed association between word production and orthographic neighborhood size on the letter fluency task is consistent with previous research showing the frequent production of phonemic clusters (i.e., words that begin with the same first two letters or differ only by a vowel sound) on letter fluency tasks (Gruenwald & Lockhead, 1980; Troster et al., 1998; Troyer et al., 1997).

Restricting the analyses to the first minute of the task in line with a more conventional administration of verbal fluency also failed to elucidate a relationship between semantic weight and word production. One possible explanation for the lack of
an association in these results is that the approach to analysis was overly insensitive. That is, the semantic weights of individual responses vary greatly throughout a word list, and calculating the overall average of these semantic variables may essentially dilute any relationship that does exist. Furthermore, word generation varied over the course of the task, with production decreasing as time on the task increased (see Figure 1). Perhaps the time intervals examined (first minute and entire three-minute trials) were too broad to capture an association between productivity and semantic weight. Indeed, a study designed to assess the contribution of frequency to fluency performance (Crowe, 1998) used 15-second time intervals and found that word frequency decreased over time but the differences between the intervals were small and they asymptote quickly. Thus, it appears that a more fine-grained analysis of the output on fluency tasks is required and this idea is more fully explored later.

Alternatively, the lack of an association between overall semantic weight and performance on fluency tasks in the present study may suggest that other factors that influence performance outweigh the constraints imposed by the structure of the semantic system. For instance, individual differences in vocabulary size and working memory have been shown to correlate with performance on fluency tasks (Ruff et al., 1997; Rosen & Engle, 1997). It is important to bear in mind that while the present study aimed to address semantic factors underlying performance on fluency tasks, executive processes also play an important role. In addition, language-based models of semantics have been investigated using tasks that generally involve automatic processes such as lexical decision, whereas the present study represents an attempt to extend the application to tasks that involve strategic, controlled processing as well. Further research is needed to
determine whether language-based models can be fruitfully applied to tasks that involve strategic processing.

The present results also indicate that the average semantic weight of generated words varies as a function of time on the category fluency task. Specifically, the average semantic weight of words decreased as time on the task increased. As such, during the early portion of the task, words were retrieved from dense semantic neighborhoods and over time, the semantic neighborhoods became sparser. Within the context of decreased productivity over time, this finding may suggest that the search and retrieval process becomes more effortful with time due, in part, to semantic neighborhood effects. That is, exemplars might be produced on the basis of the spread of activation between close semantic neighbors early on in the task while greater effort may be required to retrieve words as they become more dispersed in semantic space. The pattern of semantic relatedness as a function of time spent on the task in healthy young adults can be compared to the patterns obtained by patients with AD or other patient populations. Any differences might reflect disruptions to the semantic system of these patients.

A comparison between the number of CATSCAN-derived exemplars generated by participants in the current study and those produced by patients who experience semantic deficits might yield intriguing results. Normal individuals in this study, on average, named a sizeable proportion of the fruits and animals that were derived on the basis of co-occurrence in language. Moreover, the majority of participants tended to generate these exemplars in a similar sequence which essentially reflects the dominance (strength of the association between a word and its superordinate category) of the words. These findings provide further support that the CATSCAN method captures semantic
information and validates the use of this technique as a basis for a model of semantics. Finally, the approach used in this analysis represents a promising method for scoring the sequencing of exemplars on category fluency tasks and provides a normative basis for comparison of these scores.

*Directions for Future Research*

The current study is the first attempt to analyze verbal fluency data using a quantifiable, language-based view of semantics. As such, it provides a framework for a wealth of future research aimed at replicating and expanding the findings. Although the semantic weight of words generated on the fluency tasks was not correlated with overall word production in the present study, this may have been due to the global approach used to analyze the data. Previous studies have focused on the analysis of successively generated words and a similar, more refined approach may be warranted in the present circumstances. Rather than using semantic weight values as a measure of semantic relatedness, the distance between the actual word vectors in CATSCAN-defined semantic space could be used. In order to delineate clusters within letter and initial phoneme fluency responses, a cut-off for semantic relatedness (i.e., the distance between vectors) could be established on the basis of rather well-defined clusters on the animal naming task. The clusters obtained in this manner could then be compared to those resulting from other approaches found in the literature such as that of Troyer and colleagues (Troyer et al., 1997).

The present results also highlight the need to establish the relative difficulty of initial phoneme and letter fluency tasks for normal individuals. Moreover, possible explanations for the differences in word generation to the different cues used in this study
should be explored. For example, it may be that individuals produce fewer words beginning with the letter “g” than “j” in response to the /dʒ/ cue because retrieving permissible words that begin with the letter “g” involves actively inhibiting words with the phoneme /g/ such as gate, get, give in addition to the other cognitive demands of the task.

Conclusions

In conclusion, this study was designed to assess the contribution of semantics in verbal fluency performance under a language-based view of semantics. Word retrieval in healthy young adults was analyzed using a novel computational model of semantics to quantify the semantic neighborhoods of responses. Results suggest that a global measure of the semantic weight of responses does not predict performance on fluency tasks but that this measure of semantic relatedness does vary as a function of time. In general, the current results provide preliminary evidence that this is a promising approach to studying the contribution of semantics to lexical retrieval in a widely used neuropsychological test. However, further research is clearly needed to fully understand the role of semantics in verbal fluency performance.
References


*Journal of Experimental Psychology: Human Learning and Memory, 6*, 225-240.


VITA AUCTORIS

NAME: Treena Blake

PLACE OF BIRTH: Pincher Creek, Alberta

YEAR OF BIRTH: 1977

EDUCATION: Kitscoty Senior High School, Kitscoty, Alberta 1992-1995

University of Alberta, Edmonton, Alberta 1995-1999 B.Sc. (Hon.)