Lions And Tigers And Bears: Are They Merely Mammals Or Really Scary?

Nora Isacoff & Karin Stromswold

In this paper we argue that lexical meaning is represented by multiple levels, each of which has previously garnered support from psychologists, linguists, and philosophers. We spell out two of these levels: that which is activated when a person is explicitly asked to group items and that which is activated automatically when hearing a word. To investigate the explicit level, we asked monolingual English-speaking adults (n = 71) to sort 20 animals into groups and to name these groups. To investigate the automatic level, we gave a verbal fluency (i.e. naming animals) task to a different group of monolingual English-speaking adults (n = 72). We analyze the verbal fluency data by searching for clusters of animals sharing semantic features. We introduce a novel method of evaluating semantic clustering that eliminates some of the problems inherent to previous methods. We find that adults do semantically cluster during verbal fluency tasks. Additionally, we suggest that the semantic features that are important for explicit categorization are those that might be learned in a classroom (e.g. mammal) while those that are important for automatic access are more personally relevant (e.g. scary). This work has implications for the nature of lexical representation, lexical access, and the word-concept relationship.

Lexicon, Lexical Access, Meaning, Clustering, Verbal Fluency, Categorization

1.1 INTRODUCTION

If it walks like a duck and quacks like a duck, is it a duck? Adults reason that an animal can lack all of the characteristic features of a duck and still be a duck (Carey, 1985). Previous research has demonstrated that concept-specific features are not necessary for people's judgments of category inclusion, but rather, that domain-general features such as parenthood and DNA (for natural kinds) and functionality (for artifacts) govern these judgments (Keil, 1986, Gelman & Markman, 1987). In particular, researchers (e.g. Gelman, 2003, Carey, 1985, Strevens, 2000) have argued that people believe natural kinds (including, for adults, specific animals) have a "causal essence" (Hirschfeld, 1996), empirically discoverable but beyond intuitive grasp. This claim meshes nicely with the intuition that most, if not all of the facts people believe about a kind—e.g. that ducks live on the Earth, or that they are not made of jello—do not seem integral to conceptual constitution. Additionally, Malt, Slobin, & Gennari (2003) have demonstrated that linguistic categories are not necessarily isomorphic to conceptual categories, and Genome & Lombrozo (2012) have argued that neither description nor causal information can fully account for judgments about concept reference. Given the vagueness of "causal essence" and the fact that we cannot equate world knowledge or conceptual content with meaning, how are words mentally represented, how are they accessed during lexical retrieval, and what aspects of meaning are tapped during automatic and explicit verbal tasks? In this paper we argue that although features are neither necessary nor sufficient for category inclusion—and although such features do not necessarily constitute conceptual meaning—lexical meaning is at least partially organized around *semantic* features, which aid in efficient and accurate lexical retrieval. Additionally, we find that explicit grouping tends to reflect the knowledge that one might learn in a classroom, while automatic lexical retrieval is more likely to reflect intuitive features.

The precise structural relationship between a word's linguistic content and the concept it denotes has been a central point of debate in linguistics, philosophy, and psychology. Jackendoff (1985) posits a direct link between grammatical and conceptual information, while Katz and Fodor (1963) contend that an additional layer of semantic meaning straddles grammatical and conceptual meaning. Taking a third approach**,** some psychologists (e.g. Collins & Quillian, 1969, Landauer & Dumais, 1997) have operationalized a word's meaning as its relationship to semantically related words linked with them in a network. These approaches can be reconciled if lexical meaning is multitiered. Under a multitiered model, the elements of lexical meaning that are accessed during explicit sorting are represented as concept-like lexical meaning (a la Jackendoff), while the elements of lexical meaning that are more automatically accessed are represented in an intermediate layer (a la Katz & Fodor. Each of these layers is instantiated as semantic webs in the tradition of Collins and Quillian. Finally, the nonverbal element of conceptual meaning, which is accessed most automatically and is distinct from verbal meaning, emerges in the research of Carey, Keil, and Gelman & Markman. Work on embodied concepts (e.g. Gallese & Lakoff, 2005), as an alternative framework, might most closely capture this nonverbal level of concept meaning.

Although the explicitness and cognitive economy of networks makes them appealing, one problem with the claim that meaning is reducible to links in a network is that some relationships are more integral to our intuitive[1] sense of word meaning. "Barking" is a prototypical but non-necessary feature of DOG; "having a liver" is necessary but not prototypical; and neither seem to approach the meaning of the word. At the very least, it seems necessary to allow for different distances in the links between words (as suggested by Collins & Loftus, 1975), but some researchers have argued that such a model is not falsifiable (Johnson-Laird, Herrmann & Chaffin, 1984). The distance between words might be dependent on their co-occurrence frequency (Griffiths, Steyvers, & Tenenbaum, 2007), but, although co-occurrence might be one important element of word meaning, it does not seem to capture meaning fully. For example, although DOG and FISH more frequently co-occur, DOG and WOLF are intuitively closer in meaning, and both FISH and WOLF prime DOG (Ferrand & New, 2003).

A natural way to divide the literature on lexical meaning is to distinguish between the meanings that participants give when explicitly asked to group items or give their intuitions about essential characteristics and those that emerge during automatic tasks.

[1] See Discussion section for a consideration of the role of intuition in investigating mental content.

To investigate how people determine membership to a lexical category when they are explicitly asked to do so, some researchers (e.g. Guastavino, 2007) have used sorting tasks, in which they give participants a list of words, or cards containing words, and ask them to sort the words into groups by features of their choosing (see our Explicit Grouping task below). Other researchers (e.g. Rosch, 1973, Schmitz & Wentura, 2012) have used semantic categorization tasks, in which participants are asked whether an item is a member of a category given by the researcher. Finally, to investigate the neurological correlates of explicit lexical meaning, researchers have employed several cognitive neuroscience techniques during categorization tasks including PET (Sergent, Zuck, Levesque, & MacDonald, 1992), MEG (Low, et al., 2003), ERP (Mari-Beffa, et al., 2005), and fMRI (Mahon & Caramazza, 2010).

Similarly, psychologists have investigated automatic lexical access (and by extension, the organization of the lexicon) via a variety of techniques including priming (e.g. Forster, Mohan, & Hector, 2003); lexical decision tasks (e.g. Meyer & Schvanevelt, 1971); analyses of retrieval failures (e.g. Brown & McNeil's (1966) work on "tip of the tongue" and Fromkin's (1980) work on slips of the tongue); and neuroimaging techniques including PET (e.g. Frith, Friston, Liddle & Frackowiak, 1991), ERP (e.g. Federmeier, McLennan, de Ochoa & Kutas, 2002); MEG (e.g. Amunts, Weiss, Mohlberg, et al., 2004), fMRI (e.g. Gauthier, Duyme, Zanca, & Capron, 2009), and NIRS (e.g. Takahashi, Takikawa, Kawagoe, et al., 2011).

These studies suggest that there are multiple routes to lexical access, including semantic (e.g. Frenck-Mestre & Bueno, 1999, Troyer, 2000), but also associative (e.g. Buchanan, Brown, Cabeza, & Maitson, 1999), phonological (e.g. Slowiaczek &

Hamburger, 1992, Yee & Sedivy, 2006), orthographic (Chereau, Gaskell, & Dumay, 2007) and by frequency (e.g. Segui, Frauenfelder, & Morton, 1982). A plausible model of lexical meaning must capture the fact that normally people can access words rapidly and effortlessly, which suggests that words and the concepts they denote are tagged systematically for efficient retrieval.

A technique that capitalizes on the efficiency of lexical retrieval is the verbal fluency task, in which participants are given a short period to name members of a category (e.g. they are asked to name animals or foods), and the order in which items are named is taken to reflect lexical organization. For example, if three quarters of the jungle animals that participants name are consecutive, but the brown animals that participants name are dispersed throughout their lists, this would suggest that, in the lexicon, animals are tagged by location but not by color features.

Verbal fluency data are generally analyzed in one of two ways. In the first, researchers examine participants' lists (e.g. lists of animals or foods) and search for consecutive responses that intuitively share some feature. For example, Troyer, Moscovitch, & Winocur (1997) evaluated individual participants' lists of animals and, post hoc, identified clusters of what they argued were "obvious" animal subcategories based on biological type, location, domesticity, and other semantic subcategories. In identifying clusters, Troyer et al. gave "participants the benefit of the doubt regarding their use of clusters" (pp. 140). Subsequently, da Silva, Petersson, Faisca, Ingvar, & Reis (2004) employed Troyer et al.'s method to evaluate verbal fluency of both animals and supermarket items and found that literate and non-literate populations produced similar clusters, both quantitatively and qualitatively. In another study using a similar method, participants named supermarket items, and Sauzeon, Lestage, Raboutet, N'Kaoua, & Claverie (2004) identified semantic clusters by sorting the items into one of ten predetermined categories including fruits, meats, and desserts. Troyer et al.'s criteria has also been used in neuropsychological assessments that analyze clustering and switching in verbal fluency (e.g. Strauss, Sherman, & Spreen, 2006).

Although studies that rely on researchers' intuitions have been invaluable in providing the groundwork for using verbal fluency tasks to investigate lexical structure, one concern with this method is that researchers may over-identify features by observing a feature that is not used in lexical retrieval and does not reflect lexical structure. For example, if three quarters of the animals that participants name are mammals, researchers are likely to observe several consecutive mammals and declare the existence of a mammal cluster, even though statistically speaking, a participant is likely to name several consecutive mammals by chance alone. Conversely, researchers might under-identify features: ignoring an unintuitive but psychologically important feature. For example, if researchers do not entertain the possibility that SCARY is a feature, they may fail to notice if people name scary animals consecutively, above chance.

An additional concern with the implementation of this technique is that some researchers do not tag items with multiple features. For example, in their study of supermarket items, Sauzeon et al. could have tagged "milk" as both DAIRY and DRINK—two of their features—but they only tagged "milk" as the former. Limiting an item to a single category does not necessarily yield the most psychologically plausible model.

The biggest concern with researchers identifying features based on their intuitions is that if intuition were sufficient for uncovering lexical meaning, there would be no reason to conduct an experiment.

In the second general approach to analyzing verbal fluency, researchers use clustering algorithms to analyze verbal fluency data. Employing the information theoretic paradigm initially adopted in psychology to investigate memory (e.g. Tulving, 1962), researchers have analyzed verbal fluency data via a variety of techniques including a next-to similarity matrix (Rubin & Olson, 1980), latent semantic analysis (e.g. Landauer, Foltz, & Laham, 1998), correspondence analysis and hierarchical clustering (Schwartz, Baldo, Graves, & Brugger, 2003), dynamical models such as the random inheritance model (Borge-Holthoefer & Arenas, 2009), and network theory (Goni, Arrondo, Sepucre, Martincorena, et al., 2010). Although each technique is computationally distinct, they are similar in that they compute co-occurrence frequencies for items in verbal fluency lists, generating a multidimensional map of clusters.

A major advantage of clustering algorithms is that they detect patterns without projecting preexisting notions of what features--if any--people use to retrieve lexical items. The disadvantage, given our research interest, is that the output is merely descriptive. Once the model outputs clusters of items, the researcher must label the clusters (or at least conjecture why people tend to name some items together). For researchers concerned exclusively with modeling, lack of explicit features may not be a disadvantage at all. However, for researchers such as ourselves who seek an explanation for the underlying structure of the lexicon, this methodology is not ideal. Researchers who do label the clusters created by these computational models often provide labels that are not intuitively compelling. For example, Goni and colleagues tag both “brown bear” and “starfish” with their BEAR AND POLAR feature (which includes any bear *or* polar animal) and assign UNCLASSIFIABLE as the feature linking items for which they could not decipher a common thread. In sum, although verbal fluency tasks present a promising avenue for exploring lexical structure, there are limitations to current methods for uncovering clusters from verbal fluency data.

Previous research suggests some overlap between the order in which participants name items and measures of more explicit lexical meaning. For example, Henley (1969) demonstrated that the proximity of animals named in verbal fluency tasks was highly correlated with both the similarity ranking that participants gave pairs of animals and also with which triads of animals participants chose as most similar when given a larger set of animals. Similarly, Rosch, Simpson, & Miller (1976) found a correlation between the order in which participants named items and other participants’ prototypicality ratings of those items. However, verbal fluency data is generally not analyzed in conjunction with data from explicit semantic tasks.

There are two primary goals of this paper. The first is to present a new technique for extracting semantic clusters from verbal fluency data that reduces some of the problems with currently existing techniques. The second is to elucidate the differences between the semantic features people use when they are explicitly asked to group and those they use during automatic lexical retrieval.

2.1 METHODS

Participants

One hundred forty-three monolingual English speaking adults participated in this study. All participants were Rutgers University undergraduates who received course credit for participating. None of the participants had ever been diagnosed with a language disorder. Half of the participants (*n* = 72) performed a verbal fluency task in which they rapidly named animals, and the other half (*n* = 71) performed a free sort of the animals most frequently named by the Verbal Fluency participants.

Instructions

Verbal Fluency The Verbal Fluency participants were given a sheet of lined paper with the following instructions: *You will have 60 seconds. When I say go, I want you to write down as many animals as you can.*

Explicit Grouping The Explicit Grouping participants grouped the 20 animals most frequently named by the Verbal Fluency participants (*bear, bird, cat, cow, deer, dog, elephant, fish, giraffe, hamster, horse, lion, lizard, monkey, mouse, pig, snake, squirrel, tiger,* and *zebra).*

Participants were instructed:

Please make up categories in which you could group these animals. Make a list of these categories, and write the appropriate animals next to each category, using only the animals on this list. You can write an animal in more than one category. THERE IS NO RIGHT OR WRONG ANSWER.

Half of the group saw the list of animals in one randomized order, and half of the group saw the list in the reverse order. Note that we gave no indication of what types of features to use, the number of features to use, or how many times to use a given feature.

3.1 EXPLICIT GROUPING RESULTS

The seventy-one Explicit Grouping participants cumulatively produced forty-five distinct features[2]. There was a great deal of overlap in how participants grouped animals. For example, the feature PET was produced by over half of the participants. Twelve features were produced by ten or more participants and fifteen features were produced by five or more participants. Twenty-two feature names were produced by more than one participant (henceforth, Explicit Grouping features).

Although we did not specify which types of features to use, all but three of the forty-five distinct explicit grouping features were (broadly-speaking) semantic. Only one participant used an orthographic feature (NUMBER OF [orthographic] VOWELS, with *mouse* being grouped under 3 VOWELS); one participant used a grammatical/phonological feature (animals that SOUND THE SAME SINGULAR OR PLURAL, e.g. *fish* and *deer*); and one participant used sound (DISTINCT SOUND). Although lexical retrieval failures such as speech error studies (e.g. Fromkin, 1980; Moller, Jansma, Rodriguez-Fornells, & Munte, 2007) and tip-of-the-tongue studies (e.g. Brown & McNeil, 1966; Schwartz & Metcalfe, 2011) suggest that there might be a phonological route to lexical retrieval, not one Explicit Grouping participant used a truly phonological feature (e.g. onsets, number of syllables, stress). Furthermore, no participant used the perceptual features color, shape, smell, or touch (e.g. soft) to group animals.

Although all Explicit Grouping features were semantic, the features otherwise varied widely. Roughly speaking, there were biological, IS-A features (mammal, reptile, bird, rodent, feline, ape, herbivore, carnivore, quadruped); location features (farm,

[2] We combined synonymous feature names, e.g. BIG and LARGE, QUADRUPED and 4 LEGGED

household pet, backyard, circus, Africa, water); behavior feature (flies); human use features (EATEN, RIDDEN); and descriptive features (wild, scary, disgusting, large). Table 1 shows the number of Explicit Grouping participants who grouped by each feature (e.g. 19 participants grouped by MAMMAL).

Feature	**# Participants Grouping by Each Feature (Out of 71)**
PET	36
WILD	21
LARGE	19
MAMMAL	19
FARM	18
AFRICA	15
REPTILE	14
SCARY	14
CARNIVORE	13
WATER	13
HERB	12
QUADRUPED	10
FLIES	7
EATEN	6
RODENT	5
FELINE	4
BIRD	3
APE	2
BACKYARD	2
CIRCUS	2
DISGUSTING	2
RIDDEN	2

Table 1. Number of Explicit Grouping participants who grouped by each feature

3.2 VERBAL FLUENCY RESULTS

Collectively, the Verbal Fluency group named 174 distinct animals, with participants naming an average of 18.1 animals (SE = .45, range 10-29 animals). Table 2 shows the number of distinct animals tagged with each of the Explicit Grouping featuers.

For example, 113 of the 174 animals were mammals. Statistically speaking, then, it is likely that participants would have named multiple mammals consecutively even if they named animals randomly. In contrast, only 9 of the 174 animals were from the ape family. Therefore, it is unlikely that participants would have named multiple apes consecutively if they are naming animals randomly.

Feature	**# Animals with Each Feature (Out of 174)**
CARNIVORE	118
MAMMAL	113
HERBIVORE	112
QUADRUPED	110
WILD	101
LARGE	78
SCARY	57
PET	36
BACKYARD	34
WATER	34
DISGUSTING	27
AFRICA	26
FLIES	26
BIRD	24
EATEN	24
FELINE	18
RODENT	15
FARM	14
REPTILE	12
RIDDEN	11
APE	9
CIRCUS	9

Table 2. Number of Verbal Fluency animals tagged with each Explicit Grouping feature

4.1 CLUSTERING ANALYSES FOR VERBAL FLUENCY

We transformed each of the 174 verbal fluency animals into a set of twenty-two binary values, corresponding to the twenty-two Explicit Grouping features (e.g. WHALE

= +MAMMAL, -PET, -FELINE, +WATER, etc.). We operationalized clustering as two or more consecutive positive instances of a single feature and calculated two indices of clustering: mean cluster run and maximum cluster length. (See example below.)

To test whether Verbal Fluency participants semantically clustered at above chance level, we randomized each participant's animal list and calculated mean cluster run and maximum cluster length for each of these randomized lists. We conducted paired *t*-tests comparing these randomized cluster indices with actual cluster indices.

4.2 Example of clustering analyses

Table 3A represents a toy example of a single participant's list of animals. In this example, the participant generated five clusters: two consecutive mammals, another six consecutive mammals, three consecutive pets, four consecutive felines, and four consecutive water animals. The participant's mean cluster run is 3.8 ((2 mammals + 6 mammals + 3 pets + 4 felines + 4 water)/5 clusters). The participant's maximum cluster length is 6, because there are 6 consecutive mammals.

Table 3B represents the randomized version of the participant's list of animals. In our randomized toy example, the mean cluster run is 3.33 ((5 mammals +3 mammals +2 water)/3 clusters), and the maximum cluster length is 5, because there are 5 consecutive mammals. If this were a real participant's data, we would use a paired t-test to compare the participant's actual cluster run (3.8) to the mean randomized cluster run (3.33) and the participant's actual cluster length (6) to the mean randomized cluster length (5).

	Mammal	Pet	Feline	Water	...
Whale	✓			✓	
Dolphin	✓			✓	
Shark				✓	
Fish		✓		✓	
Dog	✓	✓			
Cat	✓	✓	✓		
Lion	✓		✓		
Tiger	✓		✓		
Cheetah	✓		✓		
Elephant	✓				

A. Actual List

	Mammal	Pet	Feline	Water	...
Lion	✓				
Whale	✓				
Cat	✓				
Shark				✓	
Fish				✓	
Tiger	✓				
Elephant	✓				
Dolphin	✓				
Dog	✓				
Cheetah	✓				

B. Randomized List

Table 3. Comparison of clustering in toy example

4.3 Clustering Collapsed Across Features

Collapsing across features, mean cluster size was significantly greater in actual lists than in randomized lists (3.7 & 3.3, respectively, $t\ (71) = 5.7$, $p = .001$). See Figure 1.

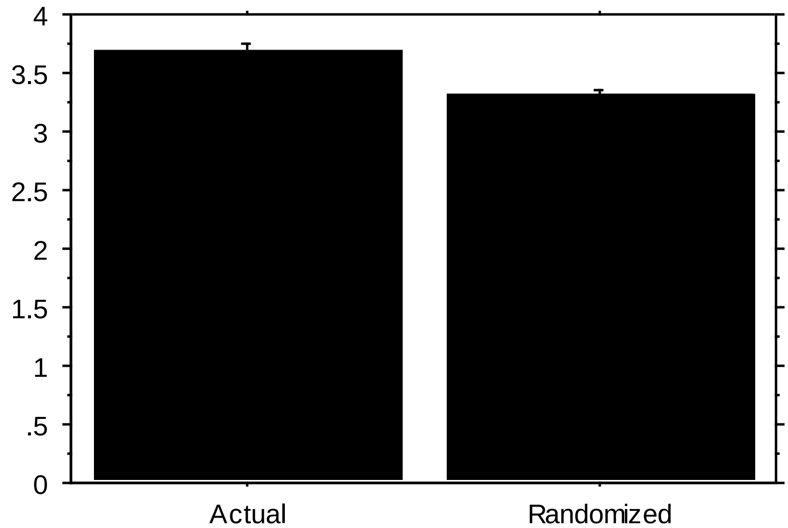

Fig. 1. Mean cluster size (error bars = SEM, p = .001)

Similarly, collapsing across features, maximum cluster length was significantly greater in actual lists than in randomized lists (10.8 & 9.8, respectively, t (71) = 3.5, p = .001). See Figure 2.

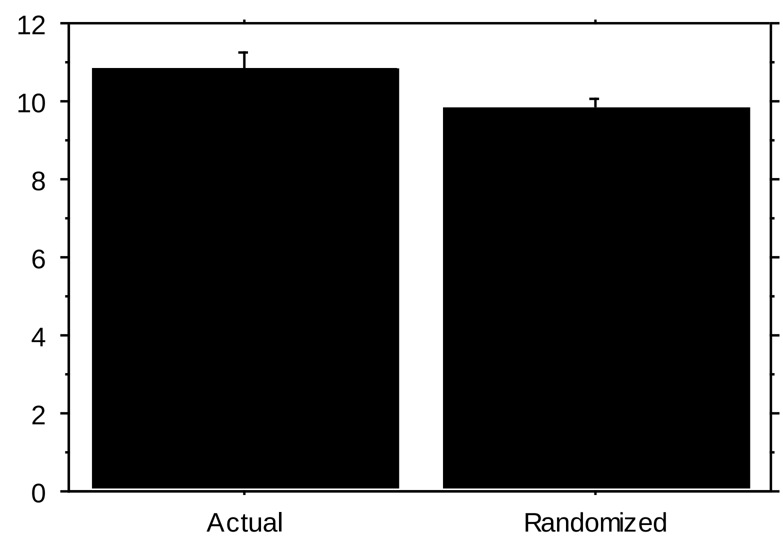

Fig. 2. Maximum cluster length (error bars = SEM, p = .001)

4.4 Clustering of Individual Semantic Features

The previous results demonstrate that adults use semantic features in accessing lexical items; however, they do not discern whether adults use some features more than others. To address this question, we used paired *t*-tests to compare the amount of clustering used for a given feature on actual versus randomized lists. We again compared clustering in 2 ways: mean of each participant's mean cluster run for each feature and mean of each participant's maximum cluster length for each feature. Given the number of *t*-tests, we set our $\alpha = .001$.

Mean cluster run: For twelve features, participants' actual mean cluster runs were significantly longer than randomized versions of these lists (*p*'s $< .001$, see Table 4). Two of these features (WILD, LARGE) played a large role in lexical access (Cohen's *d*'s $\geq .8$); seven features (PET, FELINE, AFRICA, RODENT, SCARY, FARM, EATEN) played a moderate role (Cohen's *d*'s = .5-.8); and three features (DISGUSTING, REPTILE, WATER) played a modest role (Cohen's *d*'s = .2-.5:).

Maximum cluster length : Analyzes of maximum cluster length yielded similar results, with one additional significant feature (CARNIVORE). Of these thirteen features, ten features (WILD, LARGE, PET, FELINE, AFRICA, RODENT, SCARY, FARM, EATEN, WATER) played a moderate role (Cohen's *d*'s = .5 - .8), and three features (DISGUSTING, REPTILE, CARNIVORE) played a modest role (Cohen's *d*'s = .2-.5).

	Mean Cluster Size		Maximum Longest Cluster	
FEATURE	**T statistic**	**Cohen's d**	**T statistic**	**Cohen's d**
WILD	5.53	0.90	6.44	0.77
LARGE	5.98	0.86	6.62	0.72
PET	7.18	0.79	6.75	0.77
FELINE	5.54	0.78	4.93	0.71
AFRICA	6.22	0.74	5.79	0.73
RODENT	4.46	0.61	5.04	0.50
SCARY	4.52	0.58	5.82	0.70

FARM	4.54	0.57	5.24	0.54
EATEN	4.09	0.50	5.37	0.57
DISGUSTING	4.28	0.49	3.76	0.31
REPTILE	3.40	0.49	3.67	0.36
WATER	3.84	0.47	4.62	0.53
CARNIVORE[3]	2.80	0.37	4.41	0.32

Table 3. Explicit Grouping features used in clustering. ($p < .001$)

4.5 Relationship between features explicitly named and features implicitly used

Finally, we investigated the degree of overlap between the explicit features used in the Explicit Grouping task and the implicit features used in the Verbal Fluency task. As shown in Figure 3, a regression analysis revealed a substantial but not perfect overlap between the number of Explicit Grouping participants who used a feature to group animals and that feature's effect size in the verbal fluency task, with about a third of the variance accounted for ($r = .56$, $p = .01$). Even if the outlier corresponding to PET is removed, more than half of the variance is still unaccounted for ($r = .69$, $p < .001$), suggesting that to some extent, different features are used for explicit and implicit lexical access.

[3] CARNIVORE was not significant for Mean Cluster Size.

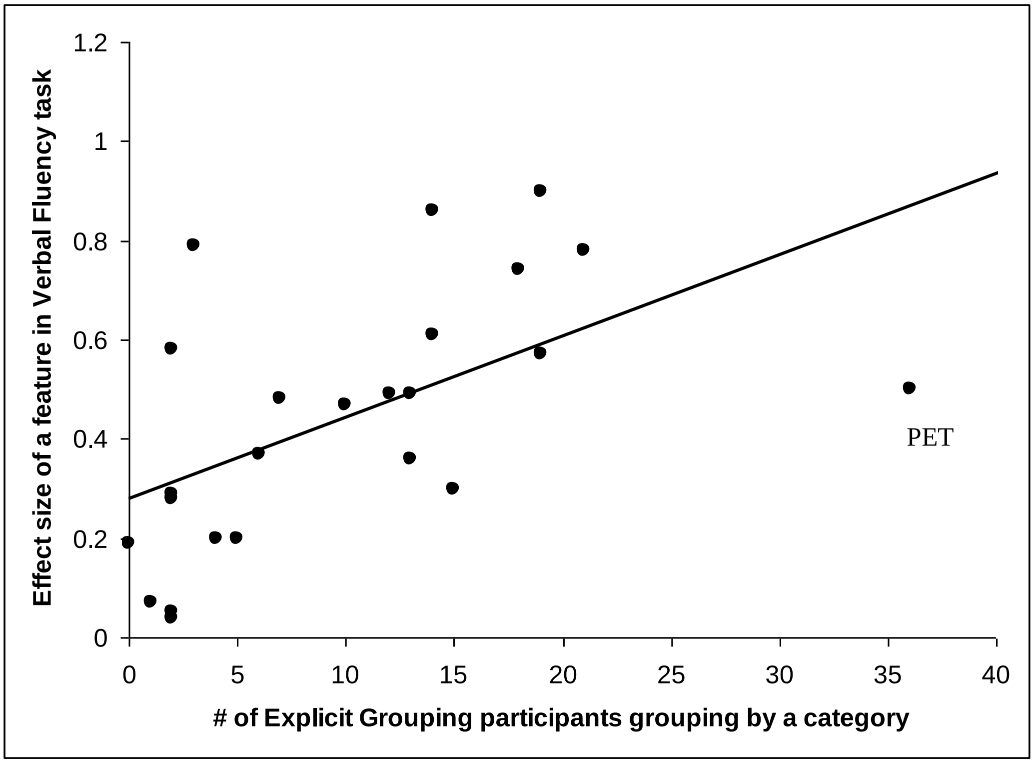

Fig 3. Overlap in importance of features in explicit and automatic tasks

5.1 DISCUSSION

Intuition is a widely-used tool for investigating mental content in linguistics, philosophy, and psychology (see Goldman, 2007, for a review), and we struggled with how much weight should be given to features being intuitively compelling. Since the fifteenth century when Descartes declared the mind fundamentally knowable by self-reflection, some philosophers have argued that intuition is valid window into cognition. Indeed, Kripke (1980) argued that intuition is ultimately the most conclusive evidence for investigating mental content. Similarly, many linguists (e.g. Chomsky, 1965) have relied heavily upon native speaker intuitions, and cognitive psychologists have frequently employed participant judgments in investigating the structure of concepts, categories, and the lexicon. After all, Rosch's prototypes would hold little weight if people did not concur that APPLE is a more typical fruit than OLIVE. Gelman's Essentialism would fall flat if people did not agree that natural kinds carry a causal essence. And so on.

On the other hand, psychological science prides itself on poking holes in folk psychology. If Descartes were correct that intuition is "indubitable," psychology would be superfluous. Consequently, epistemologists past and present have cautioned against an over-reliance on intuition (see, for example, papers in Depaul & Ramsey, 1998). Crucially, these admonitions are consistent in their criticism against intuition uncorroborated by empirical evidence, rather than against any appeal to intuition. In our approach, we sought to use empirical measures to account for intuitions. Given the enormity of the intuition problem, we do not pretend that our approach is perfect. Rather, we argue that, in terms of this problem, we have addressed some of the limitations of previous work.

In particular, our study makes a unique contribution to the investigation of semantic clustering in verbal fluency tasks. One problem we identified with some previous work is that researchers intuited the existence of semantic clusters and/or imposed intuitive feature labels on clusters without empirical support. As argued in the introduction, the problem with relying solely on intuitions in investigating semantic clusters or features is that doing so can lead to over-identifying or under-identifying clusters. Our results suggest that our concern is valid. In particular, studies using both of the methods of analyses described in the introduction (e.g. Troyer et al., 1997, Borge-Halthoefer, et al., 2009) suggested that people cluster using the features MAMMAL and BIRD. Our results suggest that this is not the case. Furthermore, we demonstrated that people do cluster using features previously overlooked by other researchers (e.g. WILD, LARGE).

Also worrisome is research that did not consider intuition at all. For example, as discussed in the introduction, Goni et. al's BEAR AND POLAR feature does not mesh with our intuitions about category structure, and the feature UNCLASSIFIABLE does not seem to mesh with other researchers' intuitions, either.

Our finding that adults semantically cluster has implications for the debate about what counts as a semantic primitive: whole concepts (e.g. DOG) or features (e.g. PET). If each animal type is a semantic primitive[4], then we would expect the order of naming to be random within a given set of animals. The fact that the order of naming was not random demonstrates that semantic features play a role in automatic lexical *retrieval*. The assumption in the lexical access literature is that discoveries about lexical access inform our understanding of lexical representation.

In our explicit grouping task, people overwhelmingly used semantic features to group animals. Although the features most frequently used to group animals tended to be the same features used in automatic lexical retrieval, there were some notable exceptions. While MAMMAL was the third most frequent feature used in the explicit grouping task (produced by over a quarter of participants) it was not important in the verbal fluency task. Conversely, only two Explicit Grouping participants grouped animals by the feature DISGUSTING, but this feature was important in the verbal fluency task. This contrast is consistent with our general finding that participants in the verbal fluency task appeared not to have relied on biological features (e.g. MAMMAL, BIRD) nor on purely

[4] Please note that we are *not* suggesting that our features are *universal, innate* primitives, in the sense of Wierzbicka's semantic primes (e.g. 1996) or Carey's core cognition representations (2009)

functional features (e.g. RIDDEN, CIRCUS) but rather on the most personally-relevant features (e.g. WILD, LARGE, PET).

Our feature set contains not only IS-A features (e.g. a DOG IS-A MAMMAL), but also descriptions (e.g. SCARY), locations (e.g. BACKYARD), things the animal does (e.g. FLIES) and things done to the animal (e.g. EATEN, RIDDEN). One benefit of limiting a model of word meaning to IS-A features is that such a model highlights a mechanism for economical deductive reasoning. For example, if ANIMAL is linked to MAMMAL which is linked to DOG, anything that is known about animals can be inferred to be true about dogs. The evidence for word meaning including a network of IS-A links is mixed. Collins and Quillian (1969) found that people are generally faster at verifying statements that require traversal of only one IS-A link (e.g. "A robin is a bird") than two IS-A LINKS (e.g. "A robin is an animal"). However, in a similar study, Rips, Shoben, and Smith (1973) found an exception: participants were faster to verify "a dog is an animal" than "a dog is a mammal." They argued that items whose prototypes are more similar are easier to verify as being of the same type, and that items with fewer intervening IS-A links typically have closer prototypes, but that their results follow from the fact that DOG is a more typical animal than mammal.

In keeping with our claim that multiple levels of meaning co-exist, it seems plausible that these IS-A links exist for explicit reasoning, but that when asked to verify sentences swiftly, people use probabilistic information as suggested by Rips and colleagues. Another way of viewing the IS-A/non IS-A feature distinction is that ontological features can only be instantiated as IS-A links. What an animal does and where it does it are not relevant for ontological status. It is notable that the Explicit

Grouping features (including those that were significant in the Verbal Fluency task) are composed of both ontological primitives and salient but—non-necessary—features. This finding is consistent with previous research showing that adults' similarity judgments are not constrained by their notions of ontological similarity. For example, when adults are presented with triplets of items containing a target item, a "taxonomic alternative" (i.e. biologically-related item), and a "script alternative" (i.e. thematically-related item) and are asked to use knowledge about the target item to make inferences about the alternatives, people can make inferences along both lines (Ross & Murphy, 1999)..

In contrast with some previous methods for analyzing verbal fluency data which only allow an item to be tagged with a single feature, the flexibility of this system permits us to tag animals with as many features as our Explicit Grouping participants saw fit. This is an advantage because there is no reason to believe, for example, that WHALE could only be a subordinate of MAMMAL or of WATER ANIMAL, but not of both. Additionally, previous research has suggested that the ability to switch between features in a verbal fluency task is a sign of normal cognitive functioning (e.g. Unsworth, Spillers, & Brewer, 2011). One aspect of the switching mechanism is particularly compatible with the multiple-feature assumption that underlies our method. It seems likely that participants sometimes transition between clusters via an item that shares one feature with a previous cluster and one feature with a subsequent cluster. For example, a participant might begin with a PET cluster (DOG, CAT, FISH) and then transition into a WATER animal cluster (FISH, WHALE, CRAB), with FISH fitting into both of these clusters. Use of multiple features captures this phenomenon.

Although use of feature lists was beneficial for highlighting semantic clusters, they are limited in that they do not capture the causal relations between features (Barsalou & Hale, 1993). Murphy and colleagues found that concepts composed of causally-related features (e.g. DRIVES IN JUNGLES with MADE IN AFRICA, as opposed to MADE IN THE ARCTIC) are easiest to learn (Murphy & Wisniewski, 1989, Murphy & Allopenna, 1994). However, Medin and colleagues (Medin, Wattenmaker, & Hampson, 1987) demonstrated that in explicit sorting tasks (using laboratory-constructed objects), participants only sort along one dimension at a time, even when the experiment is rigged to encourage people to sort along more than one dimension simultaneously (e.g. by including some fuzzy categories or by having experimenters highlight the causal relationships between features). Medin et al.'s research suggests that participants might truly be sorting along unidimensional lines, which our methodology would capture. Nonetheless, there are clear relationships between some of our features (e.g. BIRD and FLIES), and it is possible that feature relationships could be incorporated into a future model. One ramification of such a model might be a reduction in the number of significant features, in that it is possible that more than one of our current features could be subsumed under a single feature heading.

An additional limitation of our clustering method is that we used binary features, which do not capture their graded nature (see Rosch et al., 1976). For example, we might find that "dog" is a better example of a pet than "turtle" which is a better example of a pet than "lion." It could be worthwhile in future work to have participants rate animals on a non-binary scale for each feature and to incorporate these rankings into the search for clusters.

One notable finding of our study is that there was a great deal of overlap in the features by which participants grouped animals. This suggests that it is appropriate to generalize findings about lexical organization from one individual to a broader population, at least within the same language, culture, and educational background. Further research will be necessary to determine the generalizability of our findings across different populations. Some of the features produced and used by our participants are clearly culturally-specific (e.g. CIRCUS) while others (e.g. BIRD) might be more cross-cultural. (Of the culturally-specific features, some may constitute explicitly-taught, theoretical knowledge; others may be episodic or folk-biological. Hampton, 2010). One tentative hypothesis is that overall, the explicit grouping features (e.g. MAMMAL) tend to depend on explicitly taught features and therefore are more culturally specific. In contrast, the verbal fluency features (e.g. WILD, LARGE) are more generalizable.

Investigating the features used by speakers of different languages could provide insight into the effect of language on different levels of lexical structure. Linguistic relativity continues to be a hot-button issue. Pro-Whorfian scholars argue that lexical differences between languages cause speakers of these languages to think differently about objects (e.g. Lucy & Gaskins, 2001), time (e.g. Boroditsky, 2001), space (e.g. Casasanto, 2008), and color (e.g. Mo, Xu, Kay, & Tan, 2011). Anti-Whorfian scholars counter that in their own work, they have not found conceptual differences between speakers of different languages (e.g. Iwaski, Vinson, & Vigliocco, 2010); that lexical differences found in Pro-Whorfian studies do not extend to conceptual differences (e.g. Slobin, 1987, Vigliocco, Vinson, Paganelli, & Dworzynski, 2005); that lexical effects on spatial representations are dynamic rather than permanent (e.g. Landau, Desselegn, &

Goldberg, 2010) or are task-specific (e.g. Gennari, Sloman, Malt, & Fitch, 2002), that specific Pro-Whorfian studies are methodologically-flawed (e.g. January & Kako, 2007) or do not replicate (e.g. Chen, 2001); or that there are other explanations for cross-linguistic differences (e.g. Li, Abarbanell, Gletiman, & Papagragou, 2011). Perhaps an investigation into the relationships between a language's word-concept pairings and speakers' performance on both an Explicit Grouping task and a Verbal Fluency task could add new perspective. For example, if speakers of different languages are found to cluster by different features, this would support a Pro-Whorfian view; if speakers of different languages are found to cluster by the same features, this would support an anti-Whorfian view.

Especially challenging in any language-thought study is teasing apart language-driven differences from cultural differences. One prediction is that non-linguistic cultural and educational factors could constrain Explicit Grouping but that linguistic factors (e.g. phonology, and perhaps semantics) could play a role in Verbal Fluency. Bilinguals could be particularly interesting participants, in that if bilinguals are found to perform differently in different languages, this would support a Whorfian view, and vice versa.

Finally, a complete theory of meaning must reconcile stable lexical representation with flexible word use. Lakoff and Johnson (1980) observed that a waiter can use "the ham sandwich" to refer to the person who ordered it, and Barsalou (1983) further demonstrated that people can use words as ad hoc metaphors. In particular, animals have long served as metaphors for human characteristics, as evidenced by the Bible, Shakespeare, and modern American vernacular (Palmatier, 1995). Perhaps future research could investigate whether people more frequently employ explicit or implicit features in metaphor use. Additionally, the intractability of lexical meaning has been a central theme in the philosophy of language and related scholarship. Frege (1884) asserted that words only have meaning within the context of a proposition. Wittgenstein (1921) assented with his thought experiment demonstrating the undefinability of "game," and subsequently, linguists (e.g. Labov, 1973) and psychologists (e.g. Malt, 1994) followed suit by empirically demonstrating that people use words like "cup" and "water" in intractable ways. We suggest the problem of meaning becomes more tractable if multiple levels of meaning are considered. Exploring which types of features are important in different pragmatic contexts would provide further insight into the nature of lexical representation and access.

Acknowledgement

We thank Jacob Feldman and Reid Howard. This research was supported by NSF grants DGE-0549115 and BCS-0446850.

REFERENCES

Amunts, K., Weiss, P. H., Mohlberg, H., Pieperhoff, P., Eickhoff, S., Gurd, J. M., et al. (2004). Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space--the roles of brodmann areas 44 and 45. *NeuroImage, 22*(1), 42-56.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11*(3), 211-227.

Barsalou, L. W., & Hale, C. R. (1993). Components of conceptual representation: From feature lists to recursive frames. In I. Van Mechelen, J. Hampton, R. Michalski & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 97-144). San Diego, CA: Academic Press.

Borge-Holthoefer, J., & Arenas, A. (2009). Navigating word association norms to extract semantic information. *Proceedings of the 31st Annual Conference of the Cognitive Science Society, , 621.* pp. 2777-2782.

Boroditsky, L. (2001). Does language shape thought? mandarin and english speakers' conceptions of time. *Cognitive Psychology, 43*(1), 1-22.

Bousfield, W. A., & Puff, C. R. (1965). Determinants of the clustering of taxonomically and associatively related word pairs. *The Journal of General Psychology, 73*(2d Half), 211-221.

Brown, R., & McNeil, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325-337.

Buchanan, L., Brown, N. R., Cabeza, R., & Maitson, C. (1999). False memories and semantic lexicon arrangement. *Brain and Language, 68*(1-2), 172-177.

Cappelen, H., & Lepore, E. (2005). *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. Malden, MA.: Blackwell.

Carey, S. (1985). In Gleitman L., Carey S., Newport E. and Spelke E. (Eds.), *Conceptual change in childhood*. Cambridge, MA: The MIT Press.

Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.

Casasanto, D. (2008). Who's afraid of the big bag whorf? Cross-linguistic differences in temporal language and thought. *Language Learning, 58*(1), 63-79.

Chen, J. Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition, 104*(2), 427-436.

Chereau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming *Cognition, 102*(3), 341-360.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: M.I.T. Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-247.

da Silva, C. G., Petersson, K. M., Faisca, L., Ingvar, M., & Reis, A. (2004). The effects of literacy and education on the quantitative and qualitative aspects of semantic verbal fluency. *Journal of Clinical and Experimental Neuropsychology, 26*(2), 266-277.

DePaul, M. R., & Ramsey, W. (1997). *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry*. Lanham, Md.: Rowman & Littlefield.

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39*(2), 133-146.

Ferrand, L., & New, B. (2003). Semantic and associative priming in the mental lexicon. In P. Bonin (Ed.), *The mental lexicon* (pp. 25-43). New York: Nova Science Publishers.

Fodor, J. A. (1970). Three reasons for not deriving 'kill' from 'cause to die'. *Linguistic Inquiry, 1*, 429-438.

Frege, G. (1884/1974). *The foundations of arithmetic: A logico-mathematical enquiry into the concept of number* [Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl] (J. L. Austin Trans.). (2nd Revised Edition ed.). Evanston, Il.: Northwestern University Press.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. (1991). A PET study of word finding. *Neuropsychologia, 29*(12), 1137-1148.

Fromkin, V. (1980). *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. San Francisco: Academic Press.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22*(3), 455-479.

Gauthier, C. T., Duyme, M., Zanca, M., & Capron, C. (2009). Sex and performance level effects on brain activation during a verbal fluency task: A functional magnetic resonance imaging study. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 45*(2), 164-176.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*(6), 1532-1541.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition, 83*(1), 49-79.

Genome, J., & Lombrozo, T. (2012). Concept possession, experimental semantics, and hybrid theories of reference. *Philosophical Psychology, 25*, 1-26.

Goldman, A. (2007). Philosophical intuitions: Their target, their source, and their epistemic status. *Grazer Philosophische Studien, 74*, 1-26.

Goni, J., Arrondo, G., Sepulcre, J., Martincorena, I., Velez de Mendizabal, N., Corominas-Murtra, B., et al. (2011). The semantic organization of the animal category: Evidence from semantic verbal fluency and network theory. *Cognitive Processing, 12*(2), 183-196.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114*(2), 211-244.

Guastavino, C. (2007). Categorization of environmental sounds. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale, 61*(1), 54-63.

Hampton, J. (2010). Concepts in human adults. In D. Mareschal, P. C. Quinn & S. E. G. Lea (Eds.), *The making of human concepts* (14). Oxford: Oxford University Press.

Henley, N. M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8*, 176-184.

Hirschfeld, L. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: MIT Press.

Iwasaki, N., Vinson, D., & Vigliocco, G. (2010). Does the grammatical count/mass distinction affect semantic representations? Evidence from experiments in English and Japanese. *Language and Cognitive Processes, 25*(2), 189-223.

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

Johnson-Laird, P. N., Herrmann, D. J., & Chaffin, R. (1984). Only connections: A critique of semantic networks. *Psychol.Bull., 96*(2), 292-315.

Katz, J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39*, 170-210.

Keil, F. C. (1986). The acquisition of natural kind and artifact terms. In W. Demopoulos, & A. Marras (Eds.), *Language learning and concept acquisition* (pp. 133-153). Norwood, NJ: Ablex.

Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.

Labov, W. (1973). The boundaries of words and their meanings. In C. J. N. Bailey, & R. W. Shuy (Eds.), *New ways of analyzing variation in english* (pp. 340-373). Georgetown: Georgetown University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago.

Landau, B., Dessalegn, B., & Goldberg, A. M. (2010). Language and space: Momentary interactions. In P. Chilton, & V. Evans (Eds.), *Language, cognition and space: The state of the art and new directions*. London: Equinox Publishing.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211-240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes, 25*, 259-284.

Low, A., Bentin, S., Rockstroh, B., Silberman, Y., Gomolla, A., Cohen, R., et al. (2003). Semantic categorization in the human brain: Spatiotemporal dynamics revealed by magnetoencephalography. *Psychological Science, 14*(4), 367-372.

Lucy, J. A., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In S. C. Levinson, & M. Bowerman (Eds.), *Language acquisition and conceptual development* (pp. 257-283). Cambridge: Cambridge University Press.

Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain? *Trends in Cognitive Sciences, 15*(3), 97-103.

Malt, B. (1994). Water is not H20. *Cognitive Psychology, 27*, 41-70.

Malt, B., Sloman, S., & Gennari, S. (2003). Speaking versus thinking about objects and actions. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 81-112)

Mari-Beffa, P., Valdes, B., Cullen, D. J., Catena, A., & Houghton, G. (2005). ERP analyses of task effects on semantic processing from words. *Brain Research.Cognitive Brain Research, 23*(2-3), 293-305.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19*(2), 242-279.

Meyer, D. E., & Schvanevelt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operators. *Journal of Experimental Psychology, 90*(2), 227-234.

Mo, L., Xu, G., Kay, P., & Tan, L. H. Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences, , 108.* (34) pp. 14026-14030.

Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology.Learning, Memory, and Cognition, 20*(4), 904-919.

Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology.Learning, Memory, and Cognition, 15*(4), 572-586.

Palmatier, R. A. (1995). *Speaking of animals: A dictionary of animal metaphors*. Westport, CT.: Greenwood Press.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12*, 1-20.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance, 2*(491), 502.

Rubin, D. C., & Olson, M. J. (1980). Recall of semantic domains. *Memory & Cognition, 8*(4), 354-356.

Sauzeon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7-16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language*, *89*(1), 192-202.

Schmitz, M., & Wentura, D. (2012). Evaluative priming of naming and semantic categorization responses revisited: A mutual facilitation explanation. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Schwartz, S., Baldo, J., Graves, R. E., & Brugger, P. (2003). Pervasive influence of semantics in letter and category fluency: A multidimensional approach. *Brain and Language, 87*(3), 400-411.

Segui, J., Mehler, J., Frauenfelder, U., & Morton, J. (1982). The word frequency effect and lexical access. *Neuropsychologia, 20*(6), 615-627.

Sergent, J., Zuck, E., Levesque, M., & MacDonald, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. *Cerebral Cortex (New York, N.Y.: 1991), 2*(1), 68-80.

Slobin, D. (1987). Thinking for speaking. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society,* pp. 435-444.

Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology.Learning, Memory, and Cognition, 18*(6), 1239-1250.

Strauss, E., Sherman, E., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford: Oxford University Press.

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition, 74,* 149-175.

Takahashi, T., Takikawa, Y., Kawagoe, R., Shibuya, S., Iwano, T., & Kitazawa, S. (2011). Influence of skin blood flow on near-infrared spectroscopy signals measured on the forehead during a verbal fluency task. *NeuroImage, 57*(3), 991-1002.

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*(3), 370-378.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138-146.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69*(4), 344-354.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology (2006), 64*(3), 447-466.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: Implications for language learning and language use. *Journal of Experimental Psychology.General, 134*(4), 501-520.

Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford: Oxford University Press.

Wittgenstein, L. (1921/1922/2007). *Tractatus logico-philosophicus* [Logisch-Philosophische abhandlung, annalen der naturphilosophie] (C. K. Ogden Trans.). Bibliobazaar.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology.Learning, Memory, and Cognition, 32*(1), 1-14.