2012

Lexical organization in Mandarin-speaking children: insights from the semantic fluency task

Su-Mei Chen
University of Iowa

LEXICAL ORGANIZATION IN MANDARIN-SPEAKING CHILDREN: INSIGHTS FROM THE SEMANTIC FLUENCY TASK

by

Su-Mei Chen

A thesis submitted in partial fulfillment of the
requirements for the Master of Arts degree
in Speech Pathology and Audiology
in the Graduate College of
The University of Iowa

December 2012

Thesis Supervisor: Professor Karla McGregor

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

____

MASTER'S THESIS

____

This is to certify that the Master's thesis of

Su-Mei Chen

has been approved by the Examining Committee for the thesis requirement for the Master of Arts degree in Speech Pathology and Audiology at the December 2012 graduation.

Thesis Committee: ________________________________
Karla McGregor, Thesis Supervisor

____
Melissa Duff

____
Jean K. Gordon

____
Amanda Owen Van Horne

____
J. Bruce Tomblin

To Hsusheng

ACKNOWLEDGEMENTS

Before the completion of this project it was hard to imagine that it is even more difficult to write the acknowledgements than the thesis itself—so many people have been helping me out in different ways. This research project would not have been possible without the support of lots of people and organizations.

I would like to express my deepest gratitude to my academic advisor, Dr. Karla McGregor, who is not only an amazing researcher but a caring, patient, talented, passionate mentor. Whenever I need help of any kind, she is always patient and supportive, trying to work something out—even before I ask for help. I am really lucky to have her as my advisor.

Huge thanks also go to the members of my committee, Dr. Melissa Duff, Dr. Jean K. Gordon, Dr. Amanda Owen Van Horne, and Dr. J. Bruce Tomblin, who have provided me with many insightful comments and suggestions for my study plan, this project, and future studies.

I also thank Dr. Jean Gordon and the members of her lab, Language in Aging & Aphasia Lab, who gave me a lot of useful feedback on analyzing the semantic fluency data when I was completing research hours at the lab in spring 2011.

I want to thank all the participating children and their parents, who made this thesis possible. Thanks to the teachers and administrators in the participating kindergartens and elementary schools in Taipei City and Taichung City, Taiwan for believing in the value of research and helping me on participant recruitment.

Thanks to Yi-Wen Su for coding part of my data for reliability analysis. For coding semantic fluency data in this study it is essential but difficult to think like a child. Without her input, I would have missed some interesting insights into children’s perspective on categorization.

A special thank-you to Dr. Rick Arenas, who provided assistance on calculating the proximity scores for the AddTree analysis. The program he wrote had saved me enormous time in carrying out the analysis.

Thanks to my previous and current lab mates of the Word Learning Lab, especially Nichole Eden, Dr. Timothy Arbisi-Kelm, and Dr. Gwyneth Rost, who had been giving me helpful comments on this project. Very special thanks go to Nichole Eden for always standing by me throughout these years.

Thanks to my fellow students at the Department of Communication Sciences and Disorders, particularly Joanna C. Lee and Shanju Lin, who had been giving me suggestions on my research and helping me out even before I came to the University of Iowa.

Thanks to my classmates in the course Scientific Writing and the instructor Dr. Karla McGregor for their feedback on my writing during spring 2012: Bryan Brown, Benjamin Kirby, Shanju Lin, Rachel Scheperle, Amanda Silberer, and Sojeong Yoon.

Special thanks go to the friends in a study group: Elizabeth Chang, Ying-Ling Jao, Yu-Hsin Lin, and Yu-Yu Tien, who often encouraged me when I was writing and provided a number of “incentives” for attendance.

Profound thanks go to my family for their constant support and love. Thanks to my mom, Bi-Jiau Chiu (邱碧嬌), for always believing in me, listening to me, being supportive unconditionally.

ABSTRACT

Our purpose was to explore developmental changes in the organization and access to the mental lexicon between the ages of three-, five-, and seven years. Six-hundred and seventy three Mandarin-speaking participants listed all exemplars of animals and foods that came to mind within two one-minute intervals. Compared to younger participants, the older children demonstrated more correct responses and fewer errors, suggesting that they have greater knowledge of category-relevant vocabulary. They produced more subcategories, many of which involved embedding and overlapping, which suggests they have more sophisticated lexical-semantic organization. Also, they produced fewer and less closely spaced repetitions, suggesting they could more effectively monitor retrieval responses. We conclude that between the ages of three to seven, children expand and refine the organization of their mental lexicons. Improved monitoring may reflect growth in executive functioning.

Key words: *semantic fluency*; *taxonomic relation*; *contextual relation*; *categorization*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Take a moment to list the five animals that first come to your mind.

Let's suppose you named *dog*, *cat*, *tiger*, *lion*, and *elephant.* In this example, *dog* and *cat* are pets; *cat*, *tiger*, and *lion* are felines; *tiger*, *lion*, and *elephant* are African animals. This example illustrates that when trying to efficiently retrieve words from a semantic category, people tend to make use of semantic relations between words. It is likely that consecutive words on your list were related as well. One semantic relation used to organize word-to-word relationships is taxonomic. Taxonomies in natural kinds are organized according to ontological similarity. For instance, when asked to list animals, people often list *cat* and *lion* together because they are both felines. Another way to organize the lexicon is the contextual relation. Entities can be grouped together because they may occur in the same environment or context. When thinking of *turtle*, some people may come up with *rabbit*. Instead of being similar kinds of animals, *turtle* and *rabbit* are related because they appear in a well-known story, *The Tortoise and the Hare*. In this study, we used a word retrieval task, the semantic fluency task, to investigate how words link to one another in children's mental lexicon and how children's lexical organization develops.

What is a semantic fluency task?

In a semantic fluency task, participants are asked to provide as many words as they can think of in a semantic category (e.g., *animal*, *food*, and *supermarket*). This task has been widely used to explore children's lexical semantic knowledge, lexical organization, and word retrieval (e.g. Hurks et al., 2010; Koren, Kofman, & Berger, 2005). The more words in the category that exist in their mental lexicon, the larger their total number of responses. Additionally, responses often reveal how the lexicon is organized. Through activation of neighboring words and through strategic searching, the responses are often well organized: consecutive words are semantically related to one another (Sauzéon, Lestage, Raboutet, N'Kaoua, & Claverie, 2004; Troyer, Moscovitch, & Winocur, 1997).

What cognitive abilities are involved?

Two strategies maximize the efficiency of responding in a semantic fluency task: clustering and switching (Fagundo et al., 2008; Hurks et al., 2010; Lanting, Haugrud, & Crossley, 2009; Raboutet et al., 2010; Troyer et al., 1997; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). Clustering is producing words in the same subcategory (e.g., *cat, tiger, lion*) whereas switching is moving from one category to another (e.g., switching from *cat, lion* to *butterfly, beetle*). The most efficient way to approach a semantic fluency task may be to cluster words on the basis of semantic subcategories, and to switch to another when the subcategory is exhausted (Troyer et al., 1997). This hypothesis is supported by the evidence that the intervals between words selected within a cluster are shorter than those between clusters (Koren et al., 2005).

Clustering has been used to assess a word store or richness of word knowledge (e.g., Nash & Snowling, 2008). Some researchers have explicitly related cluster size[1] to the size of the lexicon (Hurks et al., 2010; Sauzéon et al., 2004) or functioning of the semantic memory network (Tröster et al., 1995). Although they agree that clustering involves richness of word knowledge, researchers have proposed different approaches to the measurement of clusters. Some studies counted single words as clusters (Troyer et al., 1997) whereas others argued that single words should be excluded because producing single words does not involve clustering (Kavé, Kigel, & Kochva, 2008; Koren et al., 2005). Another concern is that including single words may underestimate the cluster size when the participant produces many single words. As for identifying clusters, researchers have been using different methods. Troyer and colleagues (1997) have proposed potential semantic subcategories for three types of categorization: living environment (Africa, Australia, Arctic/Far North, farm, North America, water), human use (beast of burden, fur, pets), and zoological categories (bird, bovine, canine, deer, feline, fish, insect, insectivores, primate, rabbit, reptile/amphibian,

[1] The measurement, cluster size, is derived from the average number of items per cluster.

rodent, weasel). Although they argued that the determination was "derived from the actual patterns of words generated by participants during test performance, rather than on an a priori organizational scheme" (Troyer et al., 1997, p. 140) the semantic categories they listed on the basis of healthy adult data have been widely used not only for data from the clinical population (e.g., Raoux et al., 2008) but also for child data (e.g., Hurks et al., 2010; Kavé et al., 2008). Using a pre-existing semantic structure or a structure derived from a different population to code semantic fluency data can be problematic because children's semantic organization changes over time and can be different from adults' (Bowerman, 1978, 2005). Another approach to examining clustering is to generate the structure through the similarity between two items, derived from the distance of two items and frequency of co-occurrence. The rationale is that if two items, like *tiger* and *lion*, co-occur frequently and always occur consecutively they must be similar in some way. In contrast, in some cases two items like *frog* and *elephant* may co-occur next to each other very rarely, and for the most cases there are some items in between. Pairs like this would yield very low similarity score. Crowe and Prescot (2003) have used the AddTree program, proposed by Sattath & Tversky (1977) and further refined by Corter (1998), to explore the semantic networks for the category "animals" in 155 children at first grade, third grade, and fifth grade (6-, 8-, 10-year-old) in the UK. According to Crowe and Prescot's algorithm, the proximity scores were determined by distance and frequency of co-occurrence. They found that older children have more and larger clusters within an addtree. Also, older children produced more nonmamals than younger children.

In contrast, switching is related to executive function (Koren et al., 2005) and searching efficiency (Tröster et al., 1995). To switch from one category to another, participants have to inhibit the information from the previous category and activate items from the next category. In addition, another measure, the switches per category sampled, is believed to indicate searching efficiency (e.g., Tröster et al., 1995). The underlying assumption is that efficiency is best achieved if shifts occur only when necessary. That is, exemplars in one category are uttered in a single cluster (e.g., *apple, banana, pear, grape* | *rice, potato*) instead of separate ones (*apple* | *rice* | *banana, pear, potato, grape*).

Most studies of children's semantic fluency have included children who are seven years old or older. Total correct words and switches grow with age whereas mean cluster size reaches the ceiling before age 8 (Kavé, 2006; Koren et al., 2005; Sauzéon et al., 2004). Also, older children have shorter intervals than younger children within and between clusters (Koren et al. 2005).

Most studies of semantic fluency have used the indices of total number of correct responses, number of errors, number of repetitions, mean cluster size, and number of switches However, none of these measures can capture the distinctions between types of semantic relations (e.g., contextual vs. taxonomic). Previous studies using word association tasks have found that younger children tended to use contextual relations whereas older children preferred taxonomic relations perhaps as a result of schooling or reading acquisition (Cronin, 2002). Additionally, the hierarchy of categorization and the emerging complexity of a semantic network have been neglected because most studies have followed Troyer and colleagues' (1997) coding procedure, only identifying the largest clusters. Thus, although vocabulary knowledge increases across the life span (Verhaeghen, 2003), the mean cluster size reaches the ceiling before the third grade (Kavé, 2006; Koren et al., 2005; Sauzéon et al., 2004) because this measure fails to capture developmental changes in the complexity of taxonomy. Although some researchers labeled their method as hierarchical semantic network exploitation (e.g., Sauzéon et al., 2004), this method oversimplifies semantic hierarchies in that only one layer of subcategory was identified. A more refined coding system is needed because examining the internal structure of clustering can provide a nuanced picture about the development of lexical semantic organization.

Methodological innovations

The measures and coding systems in previous studies fail to fully capture the development and reorganization of the mental lexicon. Accordingly, we developed alternative methods to examine the developmental tendencies of strategic responses to the semantic fluency task in addition to the widely-used measures of switching and clustering.

A *post-hoc* approach to identifying clusters

Instead of using pre-existing semantic hierarchies proposed by previous studies (e.g.,Raoux et al., 2008 used Troyer et al.'s structure), we argue for the need to use native speakers' cultural and language knowledge as well as the frequency of grouping in children's responses to infer whether words are in a cluster. That is, we did not impose adults' category structures on the children's data; instead, a *post-hoc* approach to identifying subcategories was used. For instance, *triceratops* and *rhinoceros* do not belong to the same category in the zoological system, but given the fact that they have similar shapes and that they often co-occurred in young children's responses, the responses were coded as a cluster. One way to allow semantic clusters to emerge from the children's data is to use the program AddTree. The computer program, AddTree, proposed by Sattath & Tversky (1977) and further developed by Corters (1998), can generate trees on the basis of proximity. For the semantic fluency data, proximity for each word pair like *elephant* and *zebra,* can be calculated from the distance of each pair that occurs frequently in children's responses. In the current study, we applied the objective method AddTree and supplemented it with other hand-coded analyses, which allow us to use the whole data set, as described next.

Overlapping clusters: Richer knowledge or inability to cluster?

Previous investigators have chosen to ignore overlapping clusters (e.g., *cat* as in *lion, tiger, cat, dog* can be a part of the first cluster, *felines,* as well as the second, *pets*) (e.g., Troyer et al., 1997). By doing so, they likely missed insight into the children's strategic searches. Overlapping clusters may be evidence of strategic searching. Using overlapping clusters can make searching more efficient because all intervals between the words in the overlapping clusters are within a cluster and thus shorter.

Overlapping clusters may reveal participants' richer lexical knowledge. In this example, *finch, bat, mouse*, the first two items can be a cluster of *animals that can fly* and the last two can be *small mammals*. To produce overlapping clusters like this, one have to know not only *bats fly* but *bats are mammals*. In this sense, overlapping clusters can be evidence of richer lexical knowledge or a more sophisticated semantic network, and thus older children should be more likely to produce overlapping clusters.

On the other hand, overlapping clusters can be the outcome of an inability to continue clustering or of an inability to inhibit irrelevant information. A child may produce responses like *tiger*, *lion*, *cat*, and *dog*. To keep naming felines after producing *cat*, the child has to inhibit the irrelevant information *cats are pets*. Given the finding that younger children are less able to inhibit irrelevant information (Diamond & Taylor, 1996) we may predict that younger children produce more overlapping clusters. However, this account does not fit the data presented in previous studies that have shown younger children switch less than older children. To determine whether producing overlapping clusters is better considered an efficient or inefficient strategy, we examined its correlation to the total number of words produced and its relationship to age.

Relations other than coordinate relations

Semantic levels

To explore the organization of responses, we also examined the semantic levels of the items. This allowed us to explore strategies such as moving from specific to general (e.g., *pike, fish*) or moving from general to specific (e.g., *bird*, *robin*). Nash and Snowling (2008) are among the few researchers to explore such patterns in responses to a semantic fluency task. They found that children with Down syndrome tended to provide the superordinate label before listing exemplars (e.g., *bird*, *robin*, *duck*) whereas typically developing children were less likely to do so. They argued that this may be because children with Down syndrome need to verbalize the category label to facilitate retrieval.

Phonological clusters: Task-discrepant?

Two types of clustering have been found in a verbal fluency task: task-consistent clustering (i.e., semantic clusters in a semantic fluency task and phonological clusters in a phonological fluency task) and task-discrepant clusters (i.e., phonological clusters in a semantic fluency task and semantic clusters in a phonological fluency task). Some researchers have analyzed task-consistent clustering only, arguing that task-discrepant clusters rarely occur (Troyer et al. 1997). On the other hand, task-discrepant clusters have also been seen as evidence of more advanced strategic use (Abwender et al. 2001). For instance, Koren and colleagues (2005) found fifth-graders use more task-discrepant clusters than third-graders, and they linked the use of task-discrepant clusters to greater cognitive flexibility and more effortful searching.

In Mandarin, phonological clusters during a semantic fluency task might not necessarily be task-discrepant. As an analytic language Mandarin has a different morphological system from languages like English. In Mandarin, words in a semantic category tend to share the same word root (i.e., a morpheme, or a character, mostly CV or CV-nasal/glide structure) that indicates the semantic category such as *dog*, with other characters indicating other information such as appearance or habitat for a subcategory of animals. Therefore it might be easy for Mandarin-speaking children to understand and take advantage of the morphological regularity because the word root or morpheme can also stand alone by itself, denoting the superordinate or the unmarked category. For instance, whereas English has *bison*, *buffalo*, *cow*, *musk ox*, *yak*, *dairy cattle*, *bull,* and *calf,* all these words for bovines in Mandarin end with the word *niu2* ‘cow’. Another example the word for zebra, *ban1ma3*, ‘striped horse’ shares the same word root with the word for horse *ma3* as they are both equine. In this sense, a large proportion of phonological clusters can be semantic clusters at the same time. Therefore, instead of only coding for phonological clustering, we also identified whether the items shared the same word root and whether they were in the same semantic category. This scoring system can help us to tease apart the development of morphological awareness and strategic use of phonological relationships.

Embedded clusters

Whereas most studies have only coded the largest clusters, we argue that examining sub-clusters within a large cluster sheds light on children's lexical organization. The internal structure in a large cluster may become more complex with increasing age although the mean cluster size of the largest clusters and even the composition of items may remain the same. Suppose a 3-year-old produces *donkey*, *cow, chicken*, *horse,* and *duck*, whereas a 7-year-old produces *cow, horse*, *donkey*, *duck,* and *chicken*. Although both examples yield exactly the same cluster size, the 7-year-old seems to appreciate the ontological similarity between *horse* and *donkey* as well as *duck* and *chicken* whereas the 3-year-old's response only reveals a one-layer category of farm animals. Ignoring the embedded clusters may impede our understanding of developmental changes in the semantic network.

Contextual relations vs. taxonomic relations

We have identified contextual relations like prey-predator pairs (e.g., *cat*, *fish*), or animals appearing in a story (e.g., *The Tortoise and the Hare*) whereas most studies only identify taxonomic relationships among items as clusters. It has been well documented that children's preference for taxonomic relations and contextual relations changes with age (e.g., Blaye & Jacques, 2009; Dunham & Dunham, 1995). This influential body of research originated from studies that reported *syntagmatic-paradigmatic shift* occurs between the ages of five and eight (e.g., Anderson & Beh, 1968; Nelson, 1977). In word association tasks, younger children tend to respond with a word that falls into a different syntactic category from the syntactic class of the stimulus and that would co-occur in a syntactic structure with the stimuli (i.e., syntagmatic responses, e.g., *eat* > *apple*). In contrast, older children are more likely to respond with a word that falls in the same syntactic class (i.e., paradigmatic responses, e.g., *walk* > *run*). One hypothesis is that developmental changes in the salience of conceptual relationships underlie this shift. In tasks as varied as free recall (Siaw, 1984), picture-matching, or object-grouping (e.g., Blaye & Jacques, 2009; Dunham & Dunham, 1995), younger children tend to prefer contextual relationships, that is, they notice that two words or objects related to the same theme are bound temporally, functionally, or spatially (e.g. syntagmatic responses such as *eat* > *apple*). In contrast, in these same tasks, older children and adults tend to prefer taxonomic relationships, that is, words or objects belonging to the same category (e.g. paradigmatic responses like *dog* > *cat*).

It should be noted that the contextual relations defined in this study differ from those defined by studies using the word association task or the picture-matching task because of different task demands. Just as paradigmatic responses can encompass both taxonomic relations (*apple* > *pear*) and contextual relations (*pencil* > *paper*), the taxonomic relation as defined in object-grouping or picture-matching studies is broader than the one considered in the current study because the semantic fluency task is constrained in that correct responses come from a natural taxonomic category (e.g., *animal*). For instance, some studies used a target picture *dog* and asked participants to choose from two potential matches: a *snail* (taxonomic) and a *doghouse* (contextual) (e.g., Blaye & Jacques, 2009). Although the choice of *snail* is designed for the taxonomic relation because *dog* and *snail* are both animals, *dog* and *snail* are never coded as a cluster in responses to a semantic fluency task. In addition to the pictures, the instruction like "which one goes with it?" may contribute to a bias toward contextual relations. Therefore, given that the tasks impose different demands, results from the semantic fluency task may yield different results.

The contribution of formal schooling to the change in preferences from thematic to taxonomic as observed in the picture-matching task is thought to be causally important (Cronin, 2002). This current study compared 3-year-olds, 5-year-olds, and 7-year-olds to test whether there is a shift in preference of semantic relations within two semantic networks:

ANIMAL and FOOD. If schooling turns children's attention to taxonomic relations, we would expect 7-year-olds to use more taxonomic relations than younger children. Nevertheless, another competing hypothesis is that contextual relations like the predator-prey relation (e.g., *aardvark-ant*) may be likely to be used by older children because the relations require more knowledge than just experiencing those items occurring in the same environment (e.g., *fish* and *frog* as WATER ANIMALS) or the same story (e.g., *turtle* and *rabbit*).

Purpose

Our purpose was to investigate developmental changes in lexical organization within two categories, ANIMAL and FOOD, from the age of three years to seven years in Mandarin-speaking children through responses to the semantic fluency task. We proposed coding procedures that capture the emerging complexity of semantic networks and preference for semantic relations. We hypothesized that, because older children have richer semantic vocabulary knowledge, their responses might be more sophisticated and show more strategic use of semantic relations.

METHODS

Participants

Participants included Mandarin-speaking children in three age groups: 230 three-year-olds, 212 five-year-olds, and 241 seven-year-olds, as shown in Table 1. This age range was selected because of the research gap and because it can allow us to explore the schooling effect. We recruited more than 200 children for each age group for the purpose of generating addtrees that can reflect the general tendency of children's lexical organization**.** The children were recruited from kindergartens and elementary schools in Taichung City and Taipei City in Taiwan. All the children were Mandarin speakers with normal language development and without reported cognitive or perceptual impairments. One 3-year-old and one 7-year-old were excluded from analysis because they learned Mandarin as a second language. The recruitment and experimental protocols, approved by Institutional Review Board at the University of Iowa, complied with the requirements for the protection of human subjects.

Procedure

The semantic fluency tasks were administered in the children's schools. The experimenter administered the task to each child individually. In the practice trial, the children were asked to produce as many words as they could think of in the category of BODY PARTS in thirty seconds. If the children failed to provide any words, the experimenter would prompt them to respond by pointing to a specific body part or giving examples from the category of BODY PARTS. After the practice trial, the children were told to produce as many words as they could think of in the two categories, ANIMAL and FOOD, within two one-minute intervals. The order of the two categories was counterbalanced across participants. For the two categories ANIMAL and FOOD, the instruction did not refer to any subcategories (e.g., *water animals* for ANIMAL) and did not give any examples (e.g., *dog* for ANIMAL). The children's responses were recorded, transcribed, and coded for the following analysis.

Analysis

Correct responses

Total correct words

The semantic fluency for each category was defined as the number of total correct words. Total correct words were calculated by excluding errors and repetitions for each participant. For the following analysis on correct responses, although we took repetitions and errors into consideration concerning whether there was a switch between categories[2], repetitions and errors were not counted for scores of all the following measures for correct responses[3].

Number of switches

A switch was defined as a shift between clusters or single words. For each participant, the number of switches for each category was calculated. One-way ANOVA was used to compare the number of switches for the three age groups.

---

[2] For instance, the response of *tiger, cat | frog, butterfly | dog, cat, rabbit, dog* would yield three clusters (felines, metamorphosis, pets), each of which contains two words. The repetitions *cat* and *dog* are still in the same cluster with other pet animals *dog* and *rabbit* but do not contribute to the score of cluster size. In another example *cat, tiger |lizard python| cat| frog turtle,* there are three clusters, each of which contains two words. The repetition *cat* was not counted for the scoring of mean cluster size or switches but it serves to mark a switch and to prevent us from counting *lizard, python, frog, turtle* as a cluster.

[3] Troyer and colleagues (1997) included repetitions and even errors for their coding of clustering as they argued that repetitions reflect the psychological functioning process. This argument might apply to our data in that repetitions often occur in the middle of a cluster because it becomes difficult to inhibit the repeated item. However, Troyer and colleagues studied adult data where repetitions rarely occur and are likely activated by clustered words (e.g., in *cat, dog, rabbit| tiger, lion, cat*, the repetition *cat* is activated by the feline cluster) whereas part of our data came from very young children who repeated a lot—even within the same cluster (e.g., in *tiger, lion, elephant, lion, tiger, elephant* all the items including repetitions belong to the African-animal cluster). Including the repetitions will inflate the cluster size too much.

Number of clusters

A cluster was defined as consecutive words that are in the same subcategories or relate to one another in any way (e.g., taxonomic and contextual relations). For each participant, the number of clusters was calculated. One-way ANOVA was used to compare the mean of number of clusters for the three age groups.

Mean cluster size

Mean cluster size was derived from the number of words per cluster on average for an individual. Cluster size is usually counted starting from the second word in a cluster (i.e., the number of words in a cluster minus one). For instance a cluster of *cat, lion, tiger* would yield a cluster size of two.

Two approaches to the mean cluster size were used. The first approach includes single words, which yield scores of zero for cluster size. The mean cluster size was then calculated by the sum of cluster size divided by the number of clusters and single words. For instance, *milk, juice| hotdog | oil | ginger | white rice, brown rice, rice ball, sushi* would yield a mean cluster size of 0.8:

$$\frac{(1+0+0+0+3)}{5} = 0.8$$

The second approach did not include single words as clusters. The mean cluster size for the second approach would be 2:

$$\frac{(1+3)}{2} = 2$$

The averages of mean cluster size for each age group were compared.

Number of overlapping items

Some items can belong to each of two or more consecutive clusters. To capture the interaction between executive function and semantic knowledge, we calculated a score for overlapping clusters.

For instance, in the example, *leopard, tiger, cat, dog, wolf, fish, frog*, there are four clusters: felines (*leopard, tiger, cat*), pets (*cat, dog*), canines (*dog wolf*), and water animals (*fish, frog*). The item *cat* occurs in the two clusters, felines and pets, whereas *dog* occurs in the two clusters, pets and canines. The overlapping item score would be two (*cat* and *dog*). The score was calculated for each child and one-way ANOVA was used to analyze group differences. Also, *Chi*-square tests were used to determine whether there was an age effect on the number of children who produced overlapping clusters.

<u>Number of embedded items</u>

To explore the development in the complexity of the semantic networks we analyzed the embedded clusters, defined as smaller clusters occurring within a larger cluster. For each child, an embedded score was calculated.

This example, *cow, <u>chicken, duck</u>, <u>donkey, horse</u> | fish, frog| cat | zebra,* consists of two clusters (five farm animals: *cow, chicken, duck, donkey, hourse* and two water animals: *fish, frog*) and two single words (*cat* and *zebra*). Within the cluster of farm animal, there are two smaller embedded clusters, consisting of four words: birds (*chicken, duck*) and equines (*donkey, horse*). While seven words are in clusters, four words are in smaller clusters, yielding the embedded score of four. The score was calculated for each child and one-way ANOVA was used to analyze group differences. Also, *Chi*-square tests were used to analyze if there was an age effect on the number of children who produced embedded clusters.

<u>Phonological clusters</u>

Phonological clusters were coded independently of semantic clusters. The following criteria were used to identify phonological clusters: (1) consecutive words that share the same onset, (2) consecutive words that rhyme, and (3) consecutive words that share more than 50% of phonemes. The number of phonological clusters and the number of children who produced one or more phonological clusters was calculated for each age group. Also each phonological cluster was coded as whether items shared the same morpheme or not and occurred within a semantic category or not.

Evaluation of the indices

We recruited more than 200 children for each age group because a large sample size is required to compute proximity scores for pairs of words to generate addtrees. However, the large sample size may not be necessary for other measures we developed. To evaluate the possibility of applying the developed measures to clinical assessment in future studies, we calculated the minimum of sample size for power of 80% on the basis of current findings.

Distribution of semantic relations as the basis of clustering

Each identified cluster was further coded for its semantic relations. The example *chicken wing, chicken nugget| red bean, soy milk |French fries, hamburger*, would be coded as three clusters: chicken, bean, fast food. To give the child maximum credit, semantic relations were coded at the finest level. For example, *beef, chicken, pork* would be coded as meat while *chicken wing*, *chicken nugget* would be coded as chicken (see Appendix A and Appendix B for more example). As shown in Appendix A and Appendix B, many items belong to more than one type of subcategory (e.g., *milk* as in beverage and breakfast). Specific semantic relations were only coded for clusters, not for single items, and thus determined by co-occurrence of the items.

Taxonomic relations vs. contextual relations

The identified clusters were also coded for type of semantic relation as either taxonomic relation or contextual relation depending on how the subcategory was organized (see Appendix A and Appendix B). The taxonomic relation was coded when the relation was based on ontological similarity among the referents (e.g., *cat* and *lion* as felines) whereas the contextual relation was coded when the relation was based on the fact that the referents often occur in the same context or can be grouped functionally (e.g., *milk* and *bread* as breakfast).

The proportion of taxonomic relations vs. contextual relations to all clusters was calculated for each group and *Chi*-square tests were used to see if there was significant difference between age groups.

AddTree Analysis

The pairs that occurred in children's responses were used to calculate the proximity. The proximity was determined by the frequency of the co-occurrence of the two items and the distance between them on the basis of the algorithms proposed by Crowe & Prescott (2003). Only the items that were produced by at least 15% of children in a given age group were selected for the AddTree analysis. The AddTree analyses were carried out for each age group and each category using the computer program AddTree proposed by Sattath and Tversky and further developed by Corter (1998).

Repetitions and errors

Repetition occurrence and repetition distance

In addition to exact repetitions, some responses were identified as repetitions when they referred to essentially the same type as a previous word. The occurrences of repetitions for all groups were calculated.

For both categories, children often used modifiers (e.g., *brown* in *brown bear; strawberry* vs. *strawberry cake*). When the modifier and the head noun as a word can refer to a difference species/kind of food or a narrower/broader category, it was counted as a different correct item instead of a repetition. For instance, *spider monkey* is a specific kind of *monkey* and thus would be counted as a different response than *monkey*. In contrast, *baby bear* was considered a repetition of *bear* because *baby bear* does not refer to a different species. The same is true for *zebra-shaped chocolate* or *strawberry cake* in the FOOD category. When it was not obvious whether or not the word indicates a different type of species, the coder referred to the WordNet (Miller, 1995) and checked if there was an item in the system. If so, the word would be coded as a different word (e.g., *wild horse* vs. *horse; green apple* vs. *apple*). If not, the word would be coded as a repetition (e.g., *wild dog* is a repetition of *dog*).

To explore the development of executive function we developed an additional measure, repetition distance. The repetition distance was defined as the number of words between a given item and its repetition. In the sequence of *tiger, zebra, dog, cat, tiger, leopard*, the word *tiger* is repeated, occurring in the first and the fifth position. The repetition distance for this case is then three because there are three words in between. To prevent repeating, the participants had to keep in mind what they had already said at the same time they were listing names. One-way ANOVA was used to examine whether there was a significant difference in the repetition distance between the three age groups.

Errors

For the FOOD category, twelve types of errors were identified as shown Table 2a. For the ANIMAL category, ten types of errors were identified as shown in Table 2b.

The different error types were identified to reflect different degrees of relatedness to the words in the categories ANIMAL and FOOD. For each of the two categories, we compared the distribution of error types between age groups as well as number of errors.

Reliability analysis

Reliability analyses were carried out to make sure the *post hoc* coding system is replicable. The coding was done by the author, a graduate student from Taiwan who is a native speaker of Mandarin. Another graduate student who was blind to the hypothesis as well as blind to the age groups and from the same cultural/language environment was trained to identify and code the clusters for both categories. The coder coded at least 20% of data randomly extracted from each age group. The agreement scores were calculated from each item. The average agreement was 93% for ANIMAL and 97% for FOOD category.

RESULTS

Table 3a summarized participants who produced at least one response for the FOOD category. For the FOOD task, thirty-seven 3-year-olds, three 5-year-olds, and one 7-year-old didn't produce responses. Also, seven 3-year-olds produced only errors. As for the ANIMAL task, 19 three-year-olds did not provide any responses and four 3-year-olds produced only errors. Participants who produced only errors were included for our analysis, receiving a semantic fluency score of zero, while participants who produced no responses were not included.

Correct responses

Frequency

Table 4a summarized the most frequent responses for the FOOD category in each age group. The composition of the most frequent words was similar across age groups. Among the FOOD responses, *rice*, *apple*, *veggie*, *banana*, and *meat* were the most frequent for all age groups. For ANIMAL responses, African animals, such as *giraffe*, *elephant*, *tiger*, *lion*, and *monkey* were produced by the most children. This result suggested that the typicality of these words in the food category emerge early and remain stable with age. Interestingly, although findings from productive vocabulary indicated that children utter the word *dog/cat* very early[4], the responses of *dog*, did not reach 50% until age 7.

The results of the correct words, number of switches, number of clusters, mean cluster size (including or excluding single words), overlapping items, embedded items, and phonological clusters are shown in Table 5. The three age groups were significantly different in all the measures, *p*s < .0001.

---

[4] Results from Communicative Development Inventory (Mandarin version) showed that both items *dog* and *cat* reach the age of acquisition by the age of 19 months: more than 50% of 17-month-olds produced the items and more than 75% of 19-month-olds produced the items in spontaneous speech (Chen, 2008).

Total correct words

For both FOOD and ANIMAL task, the number of correct words increased with age. ANOVA indicated significant differences, FOOD: $F_{(2,638)} = 202.12$, $p <.0001$, $\eta^2=.39$; ANIMAL: $F_{(2,660)} =313.72$, $p <.0001$, $\eta^2=.48$. The Bonferroni test showed that all three age groups differed significantly from each other, *p*s < .0001.

We have found that many measures such as switches may co-vary with number of correct words. To exclude the possibility that the number of correct words works as a confounding factor, we calculated ratios by dividing each index by the number of total correct words for each participant. Only participants who produced at least one correct response were included in the analysis as the denominator, the number of total correct words, had to be larger than zero. The number of participants was shown in Table 6. The ratios were transformed to arcsin values for ANOVA. Table 7 summarized the results from the ANOVA of switches, number of clusters, mean cluster size, overlapping clusters, embedded clusters, and phonological clusters.

Number of switches

One-way ANOVA showed significant difference(s) in the switch ratios between age groups, *FOOD:* $F_{(2,631)} = 20.10$, $p < .0001$, $\eta^2=.06$; *ANIMAL:* $F_{(2,649)} = 49.82$, $p < .0001$, $\eta^2=.13$ as shown in Table 7. The Bonferroni test showed that for FOOD, all three age groups were significantly different, *p*s <.0001. As for ANIMAL, the difference between age 3 and age 5 was significant as well as age 3 and age 7, $p < .0001$, whereas the difference between age 5 and age 7 was not significant.

Number of clusters

One-way ANOVA showed significant difference(s) in the cluster ratios between age groups for FOOD, $F_{(2,631)} = 8.73$, $p < .0005$, $\eta^2=.03$ but no significant difference for ANIMAL was found. The Bonferroni tests showed that for FOOD only the difference between age 3 and age 5 as well as between age 3 and age 7 was significant, *p*s <.0001.

Mean cluster size

The ratio of mean cluster size to number of correct words is the only one ratio that decreases with age, indicating that the mean cluster size did not increase as either a cause or result of increasing correct words. As shown in Table 5, one-way ANOVA showed significant difference(s) in mean cluster size between age groups, including single words: FOOD: $F_{(2,638)} = 9.05$, $p < .0005$, η^2=.03, ANIMAL: $F_{(2,660)} = 5.98$, $p < .0005$, η^2=.01; excluding single words: FOOD:$F_{(2,638)} = 15.35$, $p < .0001$, η^2=.05, ANIMAL: $F_{(2,660)} = 17.64$, $p < .0001$, η^2=.05. The Bonferroni test showed different patterns of development for the two categories. When single words were included the difference between age 3 and age 7 was significant, FOOD: $p < .0001$, ANIMAL: $p <.005$ for. However, for the FOOD category, the differences within two years of age were only marginally significant, age 3 vs. age 5: $p = .18$, age 5 vs. age 7: p=.06. In contrast, for the ANIMAL category, the difference between age 3 and age 5 was marginally significant, p =.05, whereas no significant difference between age 5 and age 7 was found, $p = 1$.

Nevertheless, for the cases of the mean cluster size that excluded single words, results for the FOOD category showed that all three age groups differ from one another, age 3 vs. age 5: $p < .05$; age 5 vs. age 7: $p < .01$. In contrast, results for the ANIMAL category showed that only the difference between age 3 and age 5 was significant, p<.0001.

This difference in the results of the two measures indicated the mean cluster size that excludes single words may be more sensitive to the development within this age range. Also, mean cluster size for the ANIMAL category may develop earlier than FOOD category.

Items in overlapping clusters

To investigate whether using overlapping clusters is an efficient strategy to produce more words or evidence of failure to inhibit irrelevant information, we calculated the correlation between number of items in overlapping clusters and number of correct words.

Results showed that number of overlapping items significantly correlated with number of correct words, FOOD: $r = .26$, $p < .0001$, ANIMAL: $r = .51$, $p<.0001$.

As shown in Table 7, one-way ANOVA showed significant difference(s) of the ratio of overlapping items to correct words in the three age groups, FOOD: $F_{(2,631)} = 7.46$, $p < .001$, $\eta^2=.02$; ANIMAL $F_{(2,649)} = 31.43$, $p < .0005$, $\eta^2=.09$. Bonferroni tests showed that for FOOD, age 3 was significantly different from age 5 and from age 7, *p*s <.0001, whereas the difference between age 5 and age 7 was not significant. As for ANIMAL, Bonferroni tests showed all three age groups were significantly different, 3- vs. 5-year-olds: $p < .05$, 5- vs.7-year-olds: $p <.0001$. This result indicated that this measure may be more sensitive to the development of the category ANIMAL than the category FOOD.

In addition, as shown in Table 8, the proportion of children who produced at least one overlapping cluster increased significantly with age, FOOD: $\chi_{(2)}^2=22.60$, $p < .0001$, ANIMAL: $\chi_{(2)}^2=92.91$, $p < .0001$. Also, in terms of types, older children produced a larger variety of overlapping items (7-year-olds: 28 types; ANIMAL: 55 types) than younger children (3-year-olds FOOD: 4 types; ANMAL: 19 types). Furthermore, for the FOOD category, 35 types of words accounted for the 115 items occurring in overlapping clusters. Twenty-nine types out of 35 types were related to one cluster contextually and related to another taxonomically (e.g., *chicken nugget* as in fast food and meat; *milk* as in beverage and breakfast; *rice* as in meal and main food) whereas only six related to both clusters in the same type of relation (e.g., *tomato* as in vegetable and fruit clusters; *braised pork rice* as in meat and rice). Similar patterns were found for the ANIMAL task. Most items in overlapping clusters were related to one cluster contextually and related to another contextually (e.g., *cat* as in pets and felines).

Items in embedded clusters

The ratio of the number of items in embedded clusters to number of correct words increased with age, FOOD: $F_{(2,631)} =43.06$, $p < .0001$, $\eta^2=.12$; ANIMAL $F_{(2,649)} = 15.51$, *p*

$< .0001$, η^2=.10. The Bonferroni tests showed that for FOOD all three age groups were significantly different, 3- vs. 5-year-olds: $p < .05$, 5- vs.7-year-olds and 3- vs.7-year-olds: $p < .0001$. As for ANIMAL, the difference between age 3 and age 5 was significant, $p < .0005$ as well as between age 3 and age 7, $p < .0001$ whereas no significant difference was found between age 5 and age 7.

Also, as shown in Table 9, the proportion of children who produced at least one overlapping item increased with age, FOOD: χ^2=92.91, $p < .0001$, ANIMAL: χ^2=74.91, $p < .0001$.

Phonological clusters

As shown in Table 7, for both categories the ratio of number of items in phonological clusters to total correct words increased with age, FOOD: $F_{(2,631)}$ =46.09, $p < .0001$, η^2=.13; ANIMAL: $F_{(2,649)}$ = 45.77, $p < .0001$, η^2=.12. The Bonferroni tests showed that for both tasks all three age groups differed significantly from one another, *p*s < .0001.

As shown in Table 10, the number of children who produced at least one phonological cluster increased with age, FOOD: χ^2=109.53, $p < .0001$, ANIMAL: χ^2=129.09, $p < .0001$. The proportion increased from 12.30% at age 3 to 65% at age 7 for ANIMAL, from 4.15% to 49.37% for FOOD.

In addition, as shown in Table 11, the phonological relations among the phonological clusters predominantly shared the same word roots: 85.71% for ANIMAL, 92.59% for FOOD. Seven-year-olds started to produce more phonological clusters with other phonological relations such as rhyming or sharing the same onsets, FOOD: χ^2=6.17, $p < .05$, ANIMAL: χ^2=16.11, $p < .0001$.

The phonological similarity that children used reliably indicates semantic categories; that is, their sound-based similarity reflects a shared root that indicates a common semantic category. As shown in Table 12, 95.14% of ANIMAL phonological clusters and 76.21% of FOOD were this type. For the FOOD category, no significant difference across age groups was found, $\chi^2 = 2.86$, $p = .24$. For the ANIMAL category, the proportion of phonological clusters sharing root that indicates a semantic category decreased with age, $\chi^2 = 23.24$, $p < .0001$, although the raw number increased with age.

Evaluation of the indices

To explore the potential use of the indices in clinical studies where sample sizes are likely to be smaller than in the current study, we determined the minimum sample size required to replicate each significant age-group to age-group comparison. We set power (1-β) at 80% and *p* (α) at .05. Table 13 summarized the results from the raw number whereas Table 14 summarized the results on the basis of transformed ratios of the measures to the number of total correct words. We found that different measures were sensitive to development during different age ranges. For the ANIMAL task as shown in Table 13b, the measure, number of embedded items, requires a smaller sample size to detect the difference between age 3 and age 5 than the measure overlapping items, and the opposite is true for age 5 and age 7. On the other hand, for the FOOD task, the pattern was reversed. Also, for both categories, among the ratio scores as shown in Table 14 , the measure of items in phonological clusters requires the smallest sample size to detect the difference between age 3 and age 7. This measure is also sensitive to development within two years of age (i.e., age 3 vs. age 5 and age 5 vs. age 7).

Taxonomic relation vs. contextual relation

The *Chi*-square test showed that for the FOOD task, there was no significant difference in the proportion of contextual relation vs. taxonomic relation across age groups (see Table 15a). In all age groups, children used more taxonomic relations to cluster words than contextual relations: taxonomic relations accounted for 70% of clustering. In contrast, as shown in Table 15b, for the ANIMAL task, children used more contextual relations than taxonomic relations to cluster animal items. Also, children from different age groups have different preference for contextual relations or taxonomic relations, $\chi_4^2 = 20.55$, $p < .0001$.

Five-year-olds used more taxonomic relations than 3-year-olds whereas the proportion stayed similar from age 5 to age 7.

Distribution of semantic relations as the basis of clustering

Food

We have identified the specific semantic relations that were used as the bases of clustering. Table 16a showed the distribution of contextual relations for the FOOD category. The use of snack/dessert and fast food as basis of clustering increased after age 5 whereas the use of meal decreased in terms of both occurrence and proportion. For the taxonomic relations shown in Table 16b, whereas 3-year-olds mainly used larger categories to cluster responses, such as fruit, vegetable, main food, 5- and 7-year-olds were shifting to smaller categories, such as root and rice, suggesting an emergent complexity within semantic networks. Also, idiosyncratic clusters were found in 3-year-olds (e.g., *carrot, tomato, chili pepper, ketchup, strawberry* as a cluster of red food). In contrast, strategies to cluster words such as using the protein category emerged in 7-year-olds, perhaps as a result of schooling.

Animal

Table 17a showed the distribution of contextual relations for the ANIMAL task. African and zoo animals combined for 79% of contextual relations for 3-year-olds. With the increase of age, the semantic relations used to cluster words became more diverse. Some contextual relations such as zodiac animal and prey-predator, which require more advanced semantic knowledge, increased with age.

Similarly, as shown in Table 17b, the taxonomic relations used were more diverse in older children, as a result of richer semantic knowledge. Idiosyncratic clusters, such as *dinosaur* and *kangaroo*, were also found in 3- and 5-year-olds. It should be noted that although we coded the smallest semantic subcategory for a cluster to give credit to children who were able to make finer semantic distinction, only the largest cluster was coded for this analysis. The use of finer semantic categories might be underestimated.

AddTree Analysis

The addtrees generated from FOOD and ANIMAL items were shown in Figure 1 and Figure 2. For the FOOD category shown in Figure 1, the cluster that emerged the earliest is <u>meal</u> (i.e., *rice* and *veggie/side)*. At age 5, fruit clusters started to be generated: *apple* and *banana* were not in a cluster as shown in the addtree for age 3, but they were grouped at age 5 and also occurred within a larger cluster with *grape*. Furthermore, vegetable clusters emerged at age 7. Addtrees for age 5 and age 7 also showed a closer relationship between vegetable clusters and fruit clusters.

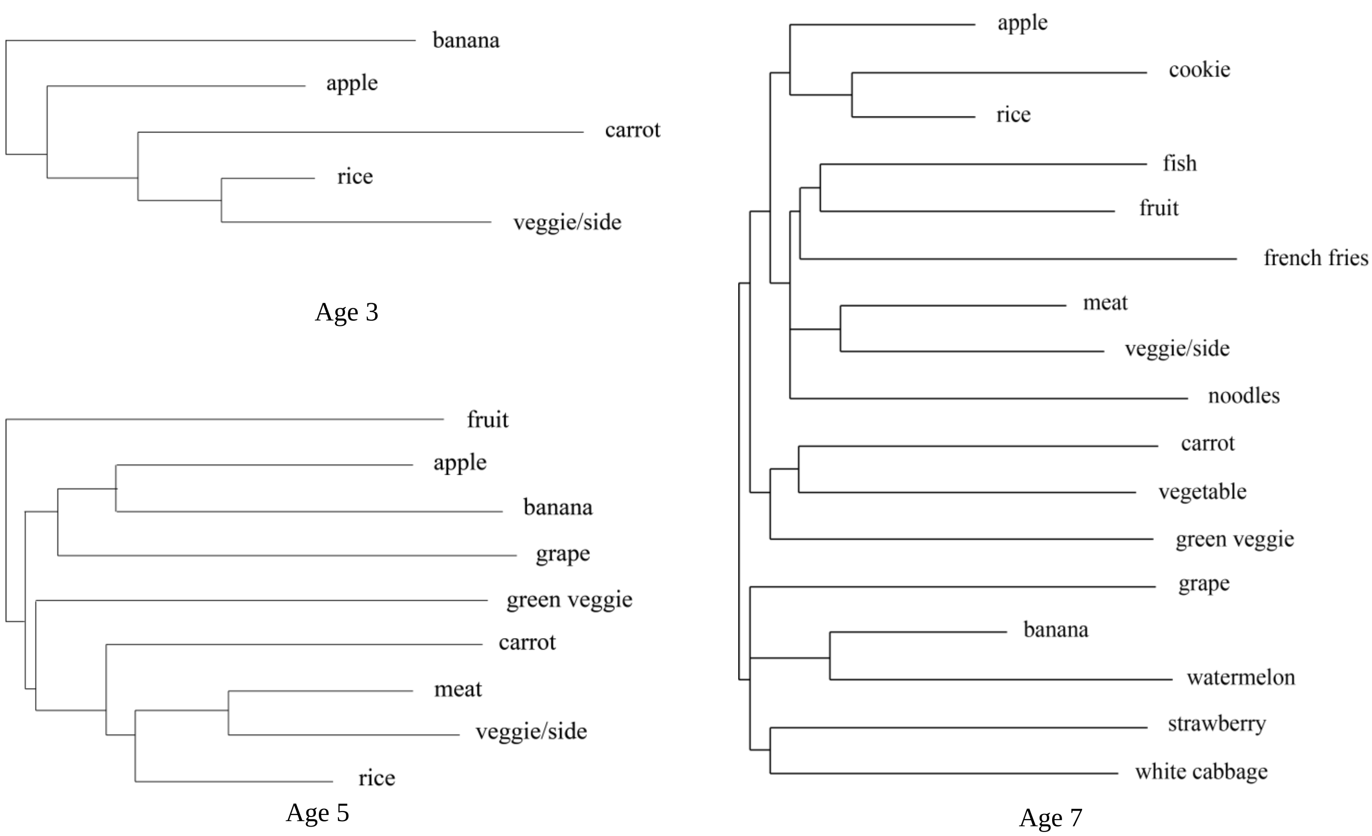

Figure 1. AddTree analyses of inter-item proximity for FOOD items

Figure 2 showed the addtrees for ANIMAL items. At age 3, only African animals reach the criteria that can be entered to the program. At age 5, the items were also predominantly African animals or zoo animals, with some pet animals, *cat* and *dog*. Interestingly *cat* was clustered together with *leopard*, which is both phonologically related and taxonomically related to *cat*[5]. In addition to African/zoo animals, 7-year-olds started to produce farm-animal clusters (*cattle*, *sheep/goat*, *pig*). In addition, *snake* is close to the farm-animal cluster, likely as a result of closer distance in the sequence of zodiac animals (i.e., 12 animals used to designate years in the lunar calendar, see Appendix B).

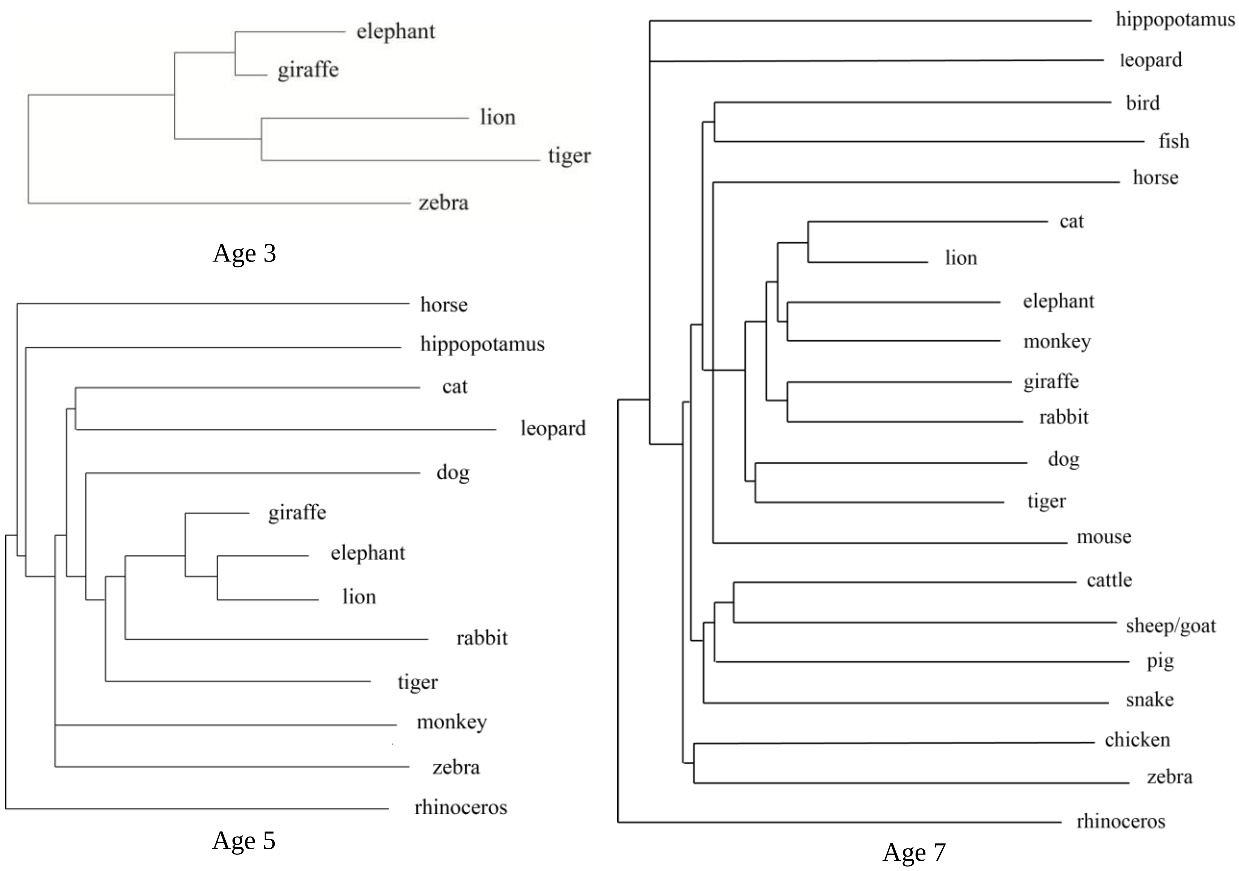

Figure 2. AddTree analyses of inter-item proximity for ANIMAL items

*The word *yang2*, 'sheep-goat,' is a general item that can refer to both sheep and goat.

[5] The two Mandarin words rhyme: *bao4* 'leopard' *mao1*'cat.'

Repetitions and errors

Repetition analysis

As shown in Table 18a, from age 3 to age 7, the proportion of repetitions to the total responses from the FOOD category dropped significantly from 9.04% to 2.63%. The same is true for the ANIMAL category as shown in Table 18b: the proportion dropped from 13.42% to 4.03%, $\chi_{(2)}^{2}$=140.28, $p < .0001$. In addition, one-way ANOVA showed that the three age groups significant differ in repetition distance, $F_{(2, 638)}$=16.77, p <.0001 for FOOD; $F_{(2, 660)}$= =41.68, p <.0001 for ANIMAL. The repetition distance increased from 1.76 words in 3-year-olds to 4.43 in 7-year-olds. Bonferroni tests showed that for both categories, the repetition distance was significantly longer in 7-year-olds than 5-year-olds, who produced significantly longer repetition distance than 3-year-olds (*p*s < .01).

Error analysis

As shown in Table 19, from age 3 to age 5, the proportion of errors dramatically dropped from 13.6% to 1.07% for FOOD, from 8.58% to 0.87% for ANIMAL but the difference between age 5 and age 7 was subtle, suggesting the accuracy rate for the two categories reached the ceiling at age 5. Nevertheless, the distribution of types of errors as shown in Table 20 revealed that the categorization is shifting throughout the age range. For 3-year-olds, the most frequent error types were unspecified items, animals and not related words. Unspecified items, such as *food*, can be seen as the participant’s repeating the experimenter’s words. For the FOOD task, Among the 15 three-year-olds who produced 35 occurrences of the error type of animals, 14 three-year-olds, who were responsible for 34 occurrences, had been asked to produce items from the ANIMAL category prior to the FOOD category task. Similarly, for the ANIMAL task, six of eight 3-year-olds who were responsible for error type, foods, had been asked to produce the FOOD category prior to the ANIMAL task, suggesting a carry-over effect. Also, 3-year-olds produced much more “not related words” and “tableware” than older children. For 5-year-olds, the errors were more related to food, such as specific brands, meals, and idiosyncratic expressions. For 7-year-olds, most of the errors were words referring to something that can be consumed (e.g., specific brand, supplements and medicine) or to a general concept of food (i.e., meal, e.g., *lunch*). Similar pattern was observed in the ANIMAL task: with the increase of age, the errors were more related to animals.

DISCUSSION

Our primary purpose was to investigate developmental changes in semantic networks using the semantic fluency task. We also used the semantic fluency task as a window into the development of executive function. Finally we sought to develop appropriate measures that can capture those developmental changes.

Development in complexity of the semantic network

We found that from age 3 to age 7 the semantic fluency (total correct words) and number of switches grow with age. However, the mean cluster size was less sensitive to the developmental changes, especially when we included single words as clusters. This discrepancy may be due to the strength of the relation between what is developing and what is measured. The semantic fluency task works as an indicator of development because it captures many aspects of language that develop with age, e.g., executive function, vocabulary size, and so on. Specifically, we replicated the finding of increasing switches with age, which has been argued as reflecting the development of executive function throughout childhood (e.g., Kavé et al., 2008; Raboutet et al., 2010). In contrast, the mean cluster size including the single words cannot fully capture the development of semantic networks of children at the ages of three years to seven years—who were even younger than participants in previous studies—because what is developing is more related to organization rather than to size. Previous studies have linked the measure of cluster size to the richness of semantic knowledge. Studies on children ranging from age 7 to adolescence found that the cluster size reaches a ceiling before age 8 (Kavé et al., 2008; Koren et al., 2005) or even reduces after age 7 (Sauzéon et al., 2004). However, evidence from previous studies supports that the richness of semantic knowledge grows in early ages and keeps growing throughout the life span (Verhaeghen, 2003). If the measure indeed captures the richness of semantic knowledge, it should not reach the ceiling before age 8 and we should have found differences between age 3 and age 5 as well as age 5 and age 7.

The failure of the mean cluster size measure to capture developmental changes motivated our analyses of overlapping clusters, embedded clusters, phonological clusters, error analysis, and semantic relations. Our results revealed that children's semantic networks become more sophisticated and the categories are refined with the increase of age. The increasing use of overlapping clusters indicates richer semantic knowledge because in order to produce overlapping clusters, children have to know at least two features of the item that belongs to two consecutive clusters. The item that belongs to two overlapping clusters can serve as a link between two categories. Another finding is that children's categorization reorganizes with development, as indicated by the results on the semantic relations as bases of clustering. Three-year-olds' categories tend to be larger than older children's. Three-year-olds may be more likely to rely on large categories such as fruit and main food. With development, children shift to use smaller categories as bases of grouping, such as rice and root. The evidence from the embedded clusters also supported the developmental trend. Older children are more likely to use smaller categories embedded in a larger category. Also growth of some cluster types such as fast food and snack might be attributable to experience. Another finding from the ANIMAL category is that typicality of words for a category changes with age. Three-year-olds mostly produce African animals but rarely produce *dog* or *cat* in a semantic fluency task, even though most children were found to produce *dog* and *cat* in spontaneous speech before 19-month of age (Chen, 2008). With development, the responses of *dog* and *cat* to the ANIMAL task increased. This result can also serve as evidence of how children at different ages link words to a category or superordinates.

Results of the AddTree analyses also showed a development of connections within the semantic network. Some items that occur frequently for each age group connected to one another in different ways. For instance, *apple* and *banana* were among the most frequent responses for each age group but were not linked until age 5. Some taxonomic relations emerge from age 3 to age 5. For instance, *lion* was clustered with *cat* in the addtree of 5-year-olds. Also, trees generated from 7-year-olds' responses were more complicated and revealed a greater variety of semantic relations than trees generated from younger children's. For the ANIMAL category, at age 3 and age 5, the responses were organized mainly by the contextual relation: African animals. At age 7, the interpretation of some clusters is less obvious because the addtrees of 7-year-olds did not fit well with the zoological structure. This might be because children started to use a wider variety of relations, such as zodiac animals, which can explain well why *snake* was close to other farm animals (*cattle*, *sheep-goat*) in the addtree. This result also parallels findings by Crowe & Prescott (2003) who found that some clusters, like *fish, fox, rabbit, hamster, snake*, in the addtree of 10-year-olds were more difficult to interpret whereas clusters in 8-year-olds seemed to make more sense (*fish, dolphin, shark*). Another possible account is that 7-year-olds produced a much wider variety of items. Only the ones that were common in the age group can enter the AddTree program for analysis. Therefore, the addtree generated from 7-year-olds' responses might not reflect the complete network in the mental lexicon.

Three-year-olds' clustering tended to be more idiosyncratic. For example, one 3-year-old used "color" to group exemplars for FOOD and some 3- and 5-year-olds used "shape" to cluster animals (e.g., *dinosaur* and *kangaroo*), which may not fit well with adults' categorizations. Other idiosyncratic relations might have occurred but would be difficult to identify unless we could think like a child. With increase of age, children converge in their strategies of clustering perhaps as a result of schooling. For instance, for the task of FOOD 7-year-olds started to use "protein" as a strategy for clustering, which encompasses a large variety of food and requires knowledge of nutrition. Seven-year-olds also tended to produce the superordinate terms, such as *grain-root category* and *egg-bean-fish-meat category*, before they started to name the exemplars in the subcategories. For the task of ANIMAL 7-year-olds started to cluster zodiac animals.

Data from Mandarin-speaking children can also be informative for the strategies of phonological clustering given that Mandarin nouns usually end with a word root that indicates a certain semantic category. Our results showed that few 3-year-olds used the morpho-phonological relationships to cluster words. However, some 3-year-olds seemed aware of the morphological regularity: they produced responses like *lion-animal*, *tiger-animal*, though the expressions are not conventional at all. With increase of age, many more children adopt the morpho-phonological relationships to cluster words. In 3- and 5-year-olds, most phonological relationships were actually the same word roots, indicating the subcategories, whereas many 7-year-olds used other phonological relations such as rhyming or sharing onsets to group words.

Our results concerning semantic relations differ from previous studies. We did not find a shift of preference from contextual relations to taxonomic relations as children get older. Instead, we found different preferences within two categories: children tend to use taxonomic relations for FOOD while they use more contextual relations for ANIMAL. That is, the way that words are linked to one another can be specific to a semantic domain. This difference might reflect variations in how or where semantic categories are learned. Animal names might be learned from books organized by context (e.g., books on farm animals, jungle animals etc.) or from zoos organized by habitats, whereas foods might be learned from eating a variety of foods within a single context (e.g., at home). Another possibility is that the semantic taxonomies for foods are better cued by the morphology of Mandarin than the taxonomies for animals. Specifically, in children's responses there are many more shared roots for foods within a given taxonomy (e.g., meat, rice, flour product) than for animals (e.g., bovine but not feline or canine). This also reflects different morphological tendencies at different levels for the two semantic domains. For the FOOD category, shared word roots are more likely superordinates (e.g., *meat*) while in the ANIMAL category, shared word roots tend to be at the basic level (e.g., *dog, cat, leopard, cow*). Given that young children tend to produce basic-level words in the semantic fluency task, morpho-phonological relations facilitate taxonomic relations for foods (e.g., *ji1rou4* 'chicken-meat,' *zhu1rou4* 'pig-meat,' *niu2rou4* 'cow-meat') rather than animals (e.g., basic-level words *mau1* 'cat' and *shi1zi* 'lion' vs. subordinates *zhe2er3mau1* 'fold-ear-cat' and *po1si1mau* 'Persian-cat'). In addition, one of contextual relations used for animals is habitats, which is sometimes reflected by a shared root like *sea, river, water, arctic, mountain*. The contextual relations that are used to organize foods (e.g., breakfast) are rarely reflected by the morpho-phonological relations. This morphological tendency can also account for the result of different patterns of phonological clusters with shared word roots in FOOD and ANIMAL as shown in Table 11; and the result of different patterns of use of word roots that indicate semantic relations as shown in Table 12. We found that for FOOD as children get older, children tend to produce more word roots in terms of both raw number and proportion, and those word roots are more likely to indicate semantic relations; for ANIMAL, the proportion of word roots in phonological clusters decreases with age and the word roots are less likely to indicate semantic relations. These differences might be partially because morpho-phonological relations are more reliable in FOOD than in ANIMAL.

In summary, the lexical semantic network becomes more complex between ages three to seven. The changes involve increased overlapping clusters, embedded categories, variety of links between words, convergent clustering strategies, and morpho-phonological awareness. These developments are important as a rich semantic lexicon has been linked to success in reading and academic success (Clarke, Snowling, Truelove, & Hulme, 2010; Cunningham & Stanovich, 1997; Pearson, Hiebert, & Kamil, 2007).

Developmental changes in executive function

The semantic fluency task also taps children's executive function because it requires online retrieval and monitoring responses. Our results indicated the growth of executive function across the age range from three years to seven years. As indicated in previous studies (e.g., Kavé et al. 2008 and Troyer et al. 1997), switches might be tentatively related to executive function as participants need to inhibit previous category and initiate a new search for a new subcategory. Our results of switches for children from the age of three years to seven years were consistent with the findings of previous studies on older children: older children switched more than younger children (Kavé et al., 2008). Also, the decreasing repetition occurrence and increasing repetition distance as shown in this study reflected the development of executive function. To prevent repeating responses, the children had to keep in mind what they just said and refrain from producing those words again while they were producing responses. Our results revealed that the repetition occurrence among older children was much fewer than among younger children, suggesting older children performed better at monitoring. In addition, older children produced repetitions at much longer intervals between the repeated word and the repetition than younger children, suggesting older children would fail only when they had to hold the information much longer. Also for 7-year-olds, a repetition usually occurred in another cluster (e.g., *cat*, *tiger* | *fish* | *elephant*, *zabra*, *tiger*), where inhibition of uttered words became more challenging because activation by related words. Additionally, when 7-year-olds repeated words, they were aware of the fact—they often asked "*did I say it?*" after producing a repetition. Errors due to carry-over effect (i.e., producing animals for the FOOD task or vice versa) were predominantly produced by 3-year-olds, indicating that 3-year-olds had difficulty inhibiting irrelevant information that was activated by a previous task.

In summary, executive function increases between ages three and seven as evident in fewer repetitions, longer repetition distance, more awareness of repetition, and fewer task-discrepant errors due to the carry-over effect. These changes are important in that good executive function is highly predictive of academic and professional success (Blair & Diamond, 2008).

Semantic fluency task: A window into the complexity of semantic networks

Vocabulary tests like PPVT have been widely used as indicators of children's semantic vocabulary knowledge. Nevertheless, we have demonstrated that semantic vocabulary knowledge is not just about size. Although the semantic fluency task measures how many words one can list in one minute, the index is not just a function of how many words are in one's mental lexicon. Rather, the internal structure in the response reflects how words activate one another in a sophisticated way.

Although mean cluster size and number of switches in a semantic fluency task have been seen as indicators of semantic knowledge and executive function respectively in studies on clinical populations (e.g., Murphy, Rich, & Troyer, 2006), these measures may not reflect the complexity of semantic networks. We have developed some alternatives to measure, analyze, and describe the emerging complexity in lexical organization across childhood. Specifically, our approach to coding overlapping clusters and embedded clusters has never been reported in the literature. Our results showed that children's lexical organization becomes more sophisticated and their semantic categories become more refined with age. Also, a larger number of phonological clusters were found in older children's responses, which might be a result of growing morpho-phonological awareness as well as more organized strategic use. Mandarin presents an excellent opportunity for determining this growth as many category members share a morphological root. Future studies are desirable to replicate the above findings, provide cross-linguistic comparisons and examine the applications and implications of the new measures for clinical assessment. For clinical studies in particular, number of correct words, number of switches, and number of clusters might be ideal in that differences can be obtained with relatively small samples.

TABLES

Table 1. Mean age and number of participants in age groups

Age	*n*	Female	Male	Mean age	Std. Deviation (months)
3	230	110	120	3;6	3.49
5	212	105	107	5;5	3.02
7	241	120	121	7;7	3.55

Table 2. Error types

a. FOOD task

Error types	Examples
(1) Specific brands	*Pepsi*
(2) Idiosyncratic expressions	*pizza rice for baked rice with cheese*
(3) Meals	*Breakfast, lunch, buffet, hot pot*
(4) Places to eat	*McDonald's, night market*
(5) Supplements and medicine	*Vitamins, pills*
(6) Plants	*Grass, flower, tree*
(7) Tableware	*Spoon, chopstick, bowl*
(8) Unspecified categories	*Things to eat, food, dog's food,*
(9) Body parts	*Skin, brain*
(10) Animals	*Zebra, dog*
(11) Unintelligible	-
(12) Not related words	*Chair, book*

b. ANIMAL task

Error types	Examples
(1) Proper names	*Winnie-the-pooh*
(2) Imaginary animals	*Mermaid, monster, Elmo*
(3) Parts of animal	*Tail, spot, head*
(4) Occupation, kinship terms	*Grandpa, sister, king*
(5) Unspecified categories	*Animal, water animal*
(6) Food	*Seafood, tomato*
(7) Plants	*Grass*
(8) Vehicle	*Train, airplane,*
(9) Unintelligible	-
(10) Others	*robot wooden horse, zoo*

Table 3. Participants who produced at least one response

a. FOOD category

Age group	*n*	(Number of participating children)	%	Mean age	Std. deviation (months)
3	193	230	83.91	3;8	3.30
5	209	212	98.58	5;5	3.00
7	239	241	99.17	7;7	3.55

b. ANIMAL task

Age group	*n*	(Number of participating children)	%	Mean age	Std. deviation (months)
3	211	230	91.74	3;7	3.40
5	212	212	100.00	5;5	3.02
7	240	241	99.59	7;7	3.55

Table 4. The most frequent responses (repetitions/errors excluded)

a. FOOD task

Age 3	*n*=193	%	Age 5	*n*=209	%	Age 7	*n*=239	%
rice	95	49.22	rice	111	53.11	apple	92	38.49
apple	48	24.87	meat	83	39.71	rice	89	37.24
veggie/side	46	23.83	veggie/side	76	36.36	meat	85	35.56
noodles	45	23.32	carrot	66	31.58	banana	76	31.80
carrot	34	17.62	apple	53	25.36	veggie/side	72	30.13
fruit	25	12.95	green veggie	46	22.01	fish	69	28.87
banana	24	12.44	white cabbage	44	21.05	fruit	67	28.03
watermelon	24	12.44	banana	41	19.62	cookie	63	26.36
meat	23	11.92	watermelon	36	17.22	white cabbage	63	26.36
cookie	18	9.33	grape	33	15.79	vegetable	58	24.27
green veggie	17	8.81	fruit	31	14.83	green veggie	56	23.43
fish	15	7.77	noodles	30	14.35	noodles	52	21.76
guava	15	7.77	broccoli	28	13.40	carrot	51	21.34
bread	13	6.74	tomato	27	12.92	strawberry	45	18.83
corn	13	6.74	cookie	23	11.00	watermelon	45	18.83
egg	13	6.74	strawberry	23	11.00	candy	44	18.41
tangerine	13	6.74	soup	21	10.05	French fries	43	17.99
candy	12	6.22	egg	20	9.57	grape	41	17.15
grape	10	5.18	vegetable	20	9.57	egg	40	16.74
dried mushroom	9	4.66	candy	19	9.09	burger	39	16.32

Table 4. Continued

b. ANIMAL task (repetitions/errors excluded)

Age 3	*n*=211	%	Age 5	*n*=212	%	Age 7	*n*=240	%
elephant	107	50.71	giraffe	145	68.40	lion	174	72.50
giraffe	98	46.45	elephant	136	64.15	tiger	141	58.75
lion	79	37.44	lion	134	63.21	elephant	139	57.92
tiger	68	32.23	tiger	103	48.58	monkey	133	55.42
zebra	34	16.11	rabbit	79	37.26	giraffe	131	54.58
monkey	31	14.69	dog	64	30.19	dog	125	52.08
rabbit	31	14.69	monkey	60	28.30	cat	123	51.25
hippopotamus	25	11.85	zebra	56	26.42	rabbit	121	50.42
cattle	21	9.95	cat	54	25.47	mouse	80	33.33
dinosaur	19	9.00	hippopotamus	45	21.23	snake	74	30.83
frog	19	9.00	horse	39	18.40	bird	72	30.00
horse	17	8.06	frog	38	17.92	cattle	67	27.92
crocodile	16	7.58	cattle	32	15.09	horse	66	27.50
dog	14	6.64	leopard	32	15.09	fish	55	22.92
fish	14	6.64	rhinoceros	31	14.62	pig	55	22.92
leopard	14	6.64	snake	31	14.62	zebra	55	22.92
bird	13	6.16	koala	26	12.26	hippopotamus	51	21.25
cat	13	6.16	fish	24	11.32	sheep/goat	50	20.83
turtle	13	6.16	mouse	22	10.38	chicken	63	26.25
panda	12	5.69	panda	22	10.38	leopard	42	17.50

Table 5. Mean of total correct words, switches, mean cluster size and comparison between age groups

a. FOOD task

	Age 3 (*n*=193)	Age 5 (*n*=209)	Age 7 (*n*=239)	Total (*n*=641)	*F*	*df*	*p*	η^2
Total correct words	4.09 ± 2.50	7.44 ± 2.88	10.54 ± 4.15	7.59 ± 4.23	202.12	2, 638	<.0001	.39
Switches	1.50 ± 1.60	3.14 ± 1.94	4.05 ± 2.20	2.99 ± 2.21	92.43	2, 638	<.0001	.22
Number of clusters	1.04 ± .87	2.01 ± 1.14	2.77 ± 1.44	2.00 ± 1.39	141.77	2,638	<.0001	.26
Mean cluster size (including single words)	1.83 ± 1.16	2.10 ± 1.15	2.36 ± 1.28	2.12 ± 1.22	10.33	2,638	<.0001	.03
Mean cluster size (excluding single words)	2.60 ± 1.07	2.90 ± 1.18	3.25 ± 1.36	2.94 ± 1.25	15.35	2, 638	<.0001	.05
Overlapping items	.05 ± .22	.18 ± .42	.24 ± .49	.16 ± .41	11.78	2, 638	<.0001	.04
Embedded items	.06 ± .35	.28 ± .82	1.39 ± 2.26	.63 ± 1.58	52.274	2,638	<.0001	.14
Items in phonological clusters	.11 ± .59	.75 ± 1.45	1.77 ± 2.32	.94 ± 1.81	70.62	2,638	<.0001	.15

b. ANIMAL task

	Age 3 (*n*=211)	Age 5 (*n*=212)	Age 7 (*n*=240)	Total (*n*=663)	*F*	*df*	*p*	η^2
Total correct words	4.05 ± 2.39	8.20 ± 3.13	12.00 ± 4.19	8.25 ± 4.69	313.72	2,660	<.0001	.49
Switches	1.00 ± 1.35	2.75 ± 1.96	4.05 ± 2.23	2.66 ± 2.28	144.98	2,660	<.0001	.31
Number of clusters	1.21 ± .90	2.29 ± 1.40	3.48 ± 1.65	2.38 ± 1.66	154.49	2,660	<.0001	.32
Mean cluster size (including single words)	1.45 ± 1.29	1.77 ± 1.56	1.89 ± 1.31	1.71 ± 1.40	5.98	2,660	<.0005	.01
Mean cluster size (excluding single words)	1.88 ± 1.15	2.50 ± 1.49	2.56 ± 1.35	2.33 ± 1.37	17.64	2,660	<.0001	.05
Overlapping items	.18 ± .60	.41 ± .81	.87 ± 1.06	.50 ± .91	38.74	2,660	<.0001	.11
Embedded items	.61 ± 1.22	1.67 ± 2.08	2.35 ± 2.40	1.58 ± 2.04	46.52	2,660	<.0001	.12
Items in phonological clusters	.29 ± .82	1.13 ± 1.66	2.23 ± 2.25	1.28 ± 1.90	70.62	2,660	<.0001	.18

Table 6. Participants who produced at least one correct response

a. FOOD task

Age group	*n*	(Number of participating children)	%	Mean age	Std. deviation (months)
3	186	230	80.87	3;8	3.26
5	209	212	98.58	5;5	3.00
7	239	241	99.17	7;7	3.55

b. ANIMAL task

Age group	*n*	(Number of participating children)	%	Mean age	Std. deviation (months)
3	200	230	86.96	3;8	3.36
5	212	212	100.00	5;5	3.02
7	240	241	99.59	7;7	3.55

Table 7. Mean of ratios of measures to the number of correct words and comparison between age groups

a. FOOD task

		Age 3 (*n*=186)	Age 5 (*n*=209)	Age 7 (*n*=239)	Total (*n*=634)	*F*	*df*	*p*	η^2
Switches	(raw)	.31 ± .26	.42 ± .20	.38 ± .17	.37 ± .21				
	(corrected)	.50 ± .38	.68 ± .25	.64 ± .23	.61 ± .29	20.10	2,631	<.0001	.06
Number of clusters	(raw)	.25 ± .18	.27 ± .12	.26 ± .09	.26 ± .13				
	(corrected)	.46 ± .28	.53 ± .16	.53 ± .11	.51 ± .19	8.73	2,631	<.0005	.03
Mean cluster size (including single words)	(raw)	.23 ± .24	.17 ± .19	.16 ± .17	.18 ± .20				
	(corrected)	.43 ± .33	.39 ± .24	.38 ± .21	.40 ± .26	1.93	2,631	.14	.01
Mean cluster size (excluding single words)	(raw)	.31 ± .24	.27 ± .19	.24 ± .18	.27 ± .20				
	(corrected)	52 ± .34	.52 ± .23	.50 ± .20	.51 ± .25	0.49	2,631	.61	.00
Overlapping items	(raw)	.01 ± .05	.02 ± .06	.02 ± .05	.02 ± .05				
	(corrected)	.02 ± .10	.06 ± .14	.07 ± .14	.05 ± .13	7.46	2,631	<.001	.02
Embedded items	(raw)	.01 ± .07	.04 ± .10	.10 ± .16	.05 ± .13				
	(corrected)	.02 ± .12	.07 ± .19	.21 ± .28	.11 ± .23	43.06	2,631	<.0001	.12
Items in phonological clusters	(raw)	.02 ± .11	.08 ± .16	.14 ± .18	.09 ± .16				
	(corrected)	.03 ± .18	.16 ± .27	.28 ± .30	.17 ± .28	46.09	2,631	<.0001	.13

Table 7. Continued

b. ANIMAL task

		Age 3 (*n*=200)	Age 5 (*n*=212)	Age 7 (*n*=240)	Total (*n*=652)	*F*	*df*	*p*	η^2
Switches	(raw)	.20 ± .22	.31 ± .19	.33 ± .15	.29 ± .20				
	(corrected)	.35 ± .35	.55 ± .27	.60 ± .19	.50 ± .29	49.82	2,649	<.0001	.13
Number of clusters	(raw)	.30 ± .15	.27 ± .10	.29 ± .10	.29 ± .12				
	(corrected)	.55 ± .21	.54 ± .13	.56 ± .12	.55 ± .16	1.08	2,649	.34	.00
Mean cluster size (including single words)	(raw)	.38 ± .28	.25 ± .24	.18 ± .16	.26 ± .24				
	(corrected)	.62 ± .34	.50 ± .28	.41 ± .19	.50 ± .29	32.18	2,649	<.0001	.09
Mean cluster size (excluding single words)	(raw)	.44 ± .25	.35 ± .24	.24 ± .17	.33 ± .23				
	(corrected)	.70 ± .31	.61 ± .27	.49 ± .20	.60 ± .27	35.30	2,649	<.0001	.10
Overlapping items	(raw)	.03 ± .08	.04 ± .07	.07 ± .08	.05 ± .08				
	(corrected)	.06 ± .16	.10 ± .18	.19 ± .19	.12 ± .18	31.43	2,649	<.0001	.09
Embedded items	(raw)	.12 ± .23	.20 ± .24	.20 ± .20	.18 ± .23				
	(corrected)	.20 ± .36	.34 ± .37	.37 ± .37	.31 ± .36	15.51	2,649	<.0001	.10
Items in phonological clusters	(raw)	.06 ± .16	.12 ± .16	.18 ± .17	12 ± .17				
	(corrected)	.10 ± .27	.24 ± .29	.36 ± .29	.24 ± .30	45.77	2,649	<.0001	.12

Table 8. Number of children producing overlapping cluster(s)*age cross-tabulation

a. FOOD task

	Age			
	3	5	7	Total
No. of children who produced one or more overlapping items	10	34	51	95
% within Age	5.18	16.27	21.34	14.82
No. of children who produced no overlapping items	183	175	188	546
% within Age	94.82	83.73	78.66	85.18
Total	193	209	239	641

b. ANIMAL task

	Age			
	3	5	7	Total
No. of children who produced one or more overlapping items	24	57	127	208
% within Age	11.37	26.89	52.92	31.37
No. of children who produced no overlapping items	187	155	113	455
% within Age	88.63	73.11	47.08	68.63
Total	211	212	240	663

Table 9. Number of children producing embedded clusters*age cross-tabulation

a. FOOD task

	Age			
	3	5	7	Total
No. of children who produced one or more embedded items	6	25	90	121
% within Age	3.11	11.96	37.66	18.88
No. of children who produced no embedded items	187	184	149	520
% within Age	96.89	88.04	62.34	81.12
Total	193	209	239	641

b. ANIMAL task

	Age			
	3	5	7	Total
No. of children who produced one or more embedded items	52	109	156	317
% within Age	24.64	51.42	65.00	47.81
No. of children who produced no embedded items	159	103	84	346
% within Age	75.36	48.58	35.00	52.19
Total	211	212	240	663

Table 10. Number of children producing phonological cluster(s) *age cross-tabulation

a. FOOD task

	Age			
	3	5	7	Total
No. of children who produced one or more phonological items	8	53	118	179
% within Age	4.15	25.36	49.37	27.93
No. of children who produced no phonological items	185	156	121	462
% within Age	95.85	74.64	50.63	72.07
Total	193	209	239	641

b. ANIMAL task

	Age			
	3	5	7	Total
No. of children who produced one or more phonological items	26	89	156	271
% within Age	12.32	41.98	65.00	40.87
No. of children who produced no phonological items	185	123	84	392
% within Age	87.68	58.02	35.00	59.13
Total	211	212	240	663

Table 11. Phonological clusters that share the same word root*age cross-tabulation

a. FOOD task

	Age			Total
	3	5	7	
No. of phonological clusters where words share the same word root	8	65	177	250
% within Age	80.00	98.48	91.24	92.59
No. of phonological clusters where words do not share the same word root	2	1	17	20
% within Age	20.00	1.52	8.76	7.41
Total	10	66	194	270

b. ANIMAL task

	Age			Total
	3	5	7	
No. of phonological clusters where words share the same word root	24	99	167	290
% within Age	85.71	84.62	66.27	85.71
No. of phonological clusters where words do not share the same word root	4	18	85	4
% within Age	14.29	15.38	33.73	14.29
Total	28	117	252	397

Table 12. Number of clusters sharing roots that were semantically*age cross-tabulation

a. FOOD task

	Age			
	3	5	7	Total
No. of phonological clusters sharing roots in a semantic cluster	7	64	164	235
% within Age	87.50	98.46	94.25	95.14
No. of phonological clusters sharing roots not in a semantic cluster	1	1	10	12
% within Age	12.50	1.54	5.75	4.86
Total clusters that share the same word root	8	65	174	247

b. ANIMAL task

	Age			
	3	5	7	Total
No. of phonological clusters sharing roots in a semantic cluster	22	89	110	221
% within Age	91.67	89.90	65.87	76.21
No. of phonological clusters sharing roots not in a semantic cluster	2	10	57	69
% within Age	8.33	10.10	34.13	23.79
Total clusters that share the same word root	24	99	167	290

Table 13. Minimum sample size per group for power of 80% , α=.05, two tailed, for counts of total correct words, switches, mean cluster size, overlapping items, embedded items, and items in phonological clusters

a. FOOD task

	Comparison		
	Age 3 vs.5	Age 5 vs. 7	Age 3 vs. 7
Total correct words	16	22	6
Switches	19	-	10
Number of clusters	18	49	9
Mean cluster size (including single words)	-	-	73
Mean cluster size (excluding single words)	75	133	25
Overlapping items	111	-	65
Embedded items	109	41	25
Items in phonological clusters	48	58	18

b. ANIMAL task

	Comparison		
	Age 3 vs.5	Age 5 vs. 7	Age 3 vs. 7
Total correct words	9	16	5
Switches	18	41	8
Number of clusters	23	27	7
Mean cluster size (including single words)	-	-	200
Mean cluster size (excluding single words)	63	-	52
Overlapping items	160	67	26
Embedded items	46	156	19
Items in phonological clusters	39	52	13

Table 14. Minimum sample size per group for power of 80%, α=.05, two tailed, for ratios of switches, mean cluster size, overlapping items, embedded items, and items in phonological clusters to correct words

a. FOOD task

	Comparison		
	Age 3 vs.5	Age 5 vs. 7	Age 3 vs. 7
Switches	55	-	82
Number of clusters	160	-	160
Mean cluster size (including single words)	-	-	-
Mean cluster size (excluding single words)	-	-	-
Overlapping items	156	-	105
Embedded items	141	53	23
Items in phonological clusters	56	90	18

b. ANIMAL task

	Comparison		
	Age 3 vs.5	Age 5 vs. 7	Age 3 vs. 7
Switches	39	-	22
Number of clusters	-	-	-
Mean cluster size (including single words)	99	130	29
Mean cluster size (excluding single words)	201	58	27
Overlapping items	212	74	30
Embedded items	109	-	57
Items in phonological clusters	69	92	20

Table 15. Semantic relation*age cross-tabulation

a. FOOD task

	Age			
	3	5	7	Total
Contextual	52	130	204	407
%	28.00	32.40	31.60	31.40
Taxonomic	134	271	442	890
%	72.00	67.60	68.40	68.60
Total	186	401	646	1297
%	100	100	100	100

b. ANIMAL task

	Age			
	3	5	7	Total
Contextual	190	334	573	1097
%	76.61	69.01	69.04	70.23
Taxonomic	49	120	238	407
%	19.76	24.79	28.67	26.06
Contextual& Taxonomic	9	30	19	58
%	3.63	6.20	2.29	3.71
Total	248	484	830	1562
%	100	100	100	100

Table 16. Distribution of clusters in the FOOD task

a. Clustering based on contextual relations

	Age							
	3		5		7		Total	
	n	%	*n*	%	*n*	%	*n*	%
Meal	35	67.31	83	63.85	79	38.73	197	51.04
Snack/dessert	11	21.15	21	16.15	63	30.88	95	24.61
Fast food	3	5.77	12	9.23	40	19.61	55	14.25
Bread/butter	1	1.92	2	1.54	6	2.94	9	2.33
Breakfast	1	1.92	5	3.85	9	4.41	15	3.89
Soup with ~,~	1	1.92	4	3.08	4	1.96	9	2.33
Curry	0	0.00	3	2.31	3	1.47	6	1.55
Total	52	100	130	100	204	100	386	100

Table 16. Continued

b. Clustering based on taxonomic relations

	Age						Total	
	3		5		7			
	n	%	*n*	%	*n*	%	*n*	%
Fruit	59	44.03	83	30.63	123	27.83	265	31.29
Main food	29	21.64	20	7.38	44	9.95	93	10.98
Vegetable	26	19.40	92	33.95	108	24.43	226	26.68
Meat	5	3.73	39	14.39	70	15.84	114	13.46
Flour	4	2.99	5	1.85	15	3.39	24	2.83
Seafood	3	2.24	4	1.48	2	0.45	9	1.06
Liquid	2	1.49	4	1.48	19	4.30	25	2.95
Protein	1	0.75	0	0.00	14	3.17	15	1.77
Rice	1	0.75	9	3.32	12	2.71	22	2.60
Bean	1	0.75	1	0.37	4	0.90	6	0.71
Noodles	1	0.75	1	0.37	1	0.23	3	0.35
Corn	1	0.75	3	1.11	0	0.00	4	0.47
Red	1	0.75	0	0.00	0	0.00	1	0.12
Egg	0	0.00	2	0.74	4	0.90	6	0.71
Radish	0	0.00	2	0.74	1	0.23	3	0.35
Seasoning	0	0.00	2	0.74	0	0.00	2	0.24
Fish	0	0.00	1	0.37	2	0.45	3	0.35
Sugar	0	0.00	1	0.37	1	0.23	2	0.24
Milk	0	0.00	1	0.37	0	0.00	1	0.12
Tomato	0	0.00	1	0.37	0	0.00	1	0.12
Chicken	0	0.00	0	0.00	5	1.13	5	0.59
Root	0	0.00	0	0.00	4	0.90	4	0.47
Ice	0	0.00	0	0.00	2	0.45	2	0.24
Kelp	0	0.00	0	0.00	2	0.45	2	0.24
[Meat]ball	0	0.00	0	0.00	1	0.23	1	0.12
Chocolate	0	0.00	0	0.00	1	0.23	1	0.12
Dumpling	0	0.00	0	0.00	1	0.23	1	0.12
Fried	0	0.00	0	0.00	1	0.23	1	0.12
Oil	0	0.00	0	0.00	1	0.23	1	0.12
Potato	0	0.00	0	0.00	1	0.23	1	0.12
Sandwich	0	0.00	0	0.00	1	0.23	1	0.12
Total	134	100	271	100	442	100	847	100

Table 17. Distribution of clusters in the ANIMAL task

a. Clustering based on contextual relations

	Age							
	3		5		7		Total	
	n	%	*n*	%	*n*	%	*n*	%
African	84	44.21	105	31.44	68	11.87	257	23.43
zoo	66	34.74	105	31.44	197	34.38	368	33.55
Water	18	9.47	26	7.78	58	10.12	102	9.30
Pet	15	7.89	52	15.57	114	19.90	181	16.50
Prey-predator	3	1.58	14	4.19	44	7.68	55	5.01
Farm	3	1.58	10	2.99	25	4.36	38	3.46
Australian	1	0.53	2	0.60	0	0.00	3	0.27
Zodiac	0	0.00	6	1.80	40	6.98	46	4.19
Story	0	0.00	7	2.10	7	1.22	14	1.28
Scary	0	0.00	1	0.30	0	0.00	1	0.09
House	0	0.00	1	0.30	6	1.05	7	0.64
Evict	0	0.00	1	0.30	12	2.09	13	1.19
Carry	0	0.00	1	0.30	1	0.17	2	0.18
Arctic	0	0.00	3	0.90	1	0.17	4	0.36
Total	190	100	334	100	573	100	1097	100

Table 17. Continued

b. Clustering based on taxonomic relations

	Age							
	3		5		7		Total	
	n	%	*n*	%	*n*	%	*n*	%
Predator	14	28.57	13	10.83	29	12.18	56	13.76
Bird	5	10.20	9	7.50	51	21.43	65	15.97
Shape	4	8.16	8	6.67	11	4.62	23	5.65
Furry, soft	4	8.16	4	3.33	4	1.68	12	2.95
Insect	3	6.12	19	15.83	27	11.34	49	12.04
Feline	3	6.12	12	10.00	14	5.88	29	7.13
Reptile	3	6.12	4	3.33	9	3.78	16	3.93
Rodent	2	4.08	4	3.33	11	4.62	17	4.18
Fly	2	4.08	4	3.33	8	3.36	14	3.44
Cold-blooded	2	4.08	3	2.50	0	0.00	5	1.23
Bear	1	2.04	8	6.67	8	3.36	17	4.18
Small mammal	1	2.04	7	5.83	8	3.36	16	3.93
Equine	1	2.04	4	3.33	6	2.52	11	2.70
Young-adult	1	2.04	3	2.50	2	0.84	6	1.47
Dinosaur	1	2.04	1	0.83	1	0.42	3	0.74
Metamorphosis	1	2.04	0	0.00	4	1.68	5	1.23
Elephant	1	2.04	0	0.00	1	0.42	2	0.49
Crawl	0	0.00	4	3.33	2	0.84	6	1.47
Canine	0	0.00	2	1.67	4	1.68	6	1.47
Worm/Bug	0	0.00	1	0.83	4	1.68	5	1.23
Worm	0	0.00	1	0.83	3	1.26	4	0.98
Swine	0	0.00	1	0.83	2	0.84	3	0.74
Size	0	0.00	1	0.83	1	0.42	2	0.49
Frog	0	0.00	1	0.83	1	0.42	2	0.49
Deer	0	0.00	1	0.83	1	0.42	2	0.49
Cow	0	0.00	1	0.83	1	0.42	2	0.49
Turtle	0	0.00	1	0.83	0	0.00	1	0.25
Shell	0	0.00	1	0.83	0	0.00	1	0.25
Lizard	0	0.00	1	0.83	0	0.00	1	0.25
Fox	0	0.00	1	0.83	0	0.00	1	0.25
Fish	0	0.00	0	0.00	7	2.94	7	1.72
Primate	0	0.00	0	0.00	6	2.52	6	1.47
Snake	0	0.00	0	0.00	2	0.84	2	0.49

Table 17b. Continued

Cattle	0	0.00	0	0.00	2	0.84	2	0.49
Behavior	0	0.00	0	0.00	2	0.84	2	0.49
Beetle	0	0.00	0	0.00	2	0.84	2	0.49
Young	0	0.00	0	0.00	1	0.42	1	0.25
Sheep	0	0.00	0	0.00	1	0.42	1	0.25
Rhino	0	0.00	0	0.00	1	0.42	1	0.25
Eagle	0	0.00	0	0.00	1	0.42	1	0.25
	49	100	120	100	238	100	407	100

c. Clustering based on both contextual and taxonomic relations

	Age							
	3		5		7		Total	
	n	%	*n*	%	*n*	%	*n*	%
African feline	4	44.44	14	46.67	12	63.16	30	51.72
Grassland; grass eater	2	22.22	7	23.33	2	10.53	11	18.97
Farm mammal	2	22.22	7	23.33	2	10.53	11	18.97
Farm bird	1	11.11	0	0.00	2	10.53	3	5.17
Water bird	0	0.00	0	0.00	1	5.26	1	1.72
Water mammal	0	0.00	1	3.33	0	0.00	1	1.72
African; shape	0	0.00	1	3.33	0	0.00	1	1.72
Total	9	100	30	100	19	100	58	100

Table 18. Number of repetitions and repetition distance

a. FOOD task

Age group	Total responses	Repetition	%	Repetition distance Mean *(range)*	Repetition distance Std. deviation
3	1007	91	9.04	1.76 *(0-7)*	1.76
5	1632	66	4.04	2.98 *(0-12)*	2.60
7	2626	69	2.63	4.43 *(0-17)*	4.12
Total	5265	226	15.71	2.88	3.03

b. ANIMAL task

Age group	Total responses	Repetition	%	Repetition distance Mean *(range)*	Repetition distance Std. deviation
3	1095	147	13.42	2.17 *(0-8)*	1.71
5	1835	79	4.31	3.56 *(0-12)*	2.68
7	3025	122	4.03	5.72 *(0-28)*	4.55
Total	5955	348	5.84	3.73	3.53

Table 19. The number and proportion of errors

a. FOOD task

Age	*n*	%	Total responses
3	137	13.60	1007
5	20	1.23	1632
7	28	1.07	2626
Total	185	3.51	5265

b. ANIMAL task

Age	*n*	%	Total responses
3	94	8.58	1095
5	16	0.87	1835
7	24	0.79	3025
Total	134	2.25	5955

Table 20. The distribution of error types

a. FOOD task

Age 3	*n*	%	Age 5	*n*	%	Age 7	*n*	%
Unspecified categories	39	28.47	Unspecified categories	4	20.00	Meal	11	39.29
Animals	35	25.55	Specific brand	3	15.00	Specific brand	5	17.86
Not related	30	21.90	Unintelligible	3	15.00	Plant	5	17.86
Tableware	8	5.84	Meal	3	15.00	Supplements and medicine	3	10.71
Place to eat	5	3.65	Idiosyncratic expressions	2	10.00	Unspecified categories	2	7.14
Unintelligible	5	3.65	Plant	2	10.00	Body part	1	3.57
Meal	4	2.92	Place to eat	1	5.00	Animals	1	3.57
Plant	4	2.92	Not related	1	5.00	Total	28	100
Specific brand	3	2.19	Body part	1	5.00			
Supplements and medicine	2	1.46	Total	20	100			
Idiosyncratic expression	1	.73						
Body part	1	.73						
Total	137	100						

Table 20. Continued

b. ANIMAL task

Age 3	*n*	%	Age 5	*n*	%	Age 7	*n*	%
Unspecified categories	26	27.66	Imaginary animals	4	25.00	Unspecified categories	6	25.00
Unintelligible	12	12.77	Unintelligible	4	25.00	Proper names	4	16.67
Food	10	10.64	Proper names	3	18.75	Unintelligible	4	16.67
Vehicles	10	10.64	Others	2	12.50	Imaginary animals	3	12.50
Proper names	8	8.51	Occupation, kinship terms	1	6.25	Plants	2	8.33
Occupation, kinship terms	8	8.51	Unspecified categories	1	6.25	Others	2	8.33
Parts of animal	6	6.38	Plants	1	6.25	Parts of animal	1	4.17
Others	6	6.38	Total	16	10	Occupation, kinship terms	1	4.17
Imaginary animals	5	5.32				Food	1	4.17
Plants	3	3.19				Total	24	100
Total	94	100						

REFERENCES

Anderson, S. W., & Beh, W. (1968). The reorganization of verbal memory in childhood. *Journal of Verbal Learning and Verbal Behavior, 7*(6), 1049-1053.

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*(3), 899-911.

Blaye, A., & Jacques, S. (2009). Categorical flexibility in preschoolers: contributions of conceptual knowledge and executive control. *Developmental Science, 12*(6), 863-873.

Bowerman, M. (1978). Systematizing semantic knowledge: Changes over time in the child's organization of word meaning. *Child Development, 49*(4), 977-987.

Bowerman, M. (2005). Why can't you "open" a nut or "break" a cooked noodle? Learning covert object categories in action word meanings. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 209-243). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Chen, S. (2008). *Nature or nurture? Imageability, frequency, and age of acquisition of early Mandarin vocabulary*. Paper presented at the Second Conference on Language, Discourse and Cognition.

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties. *Psychological Science, 21*(8), 1106-1116.

Corter, J. E. (1998). An efficient metric combinatorial algorithm for fitting additive trees. *Multivariate Behav Res, 33*(2), 249-271.

Cronin, V. S. (2002). The syntagmatic-paradigmatic shift and reading development. *Journal of Child Language, 29*(01), 189-204.

Crowe, S. J., & Prescott, T. J. (2003). Continuity and change in the development of category structure: Insights from the semantic fluency task. *International Journal of Behavioral Development, 27*(5), 467-479.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934-945.

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to “Do as I say, not as I do”. *Developmental Psychobiology, 29*(4), 315-334.

Dunham, P., & Dunham, F. (1995). Developmental antecedents of taxonomic and thematic strategies at 3 years of age. *Developmental Psychology, 31*(3), 483-493.

Fagundo, A. B., López, S., Romero, M., Guarch, J., Marcos, T., & Salamero, M. (2008). Clustering and switching in semantic fluency: Predictors of the development of Alzheimer's disease. *International Journal of Geriatric Psychiatry, 23*(10), 1007-1013.

Hurks, P. P. M., Schrans, D., Meijs, C., Wassenberg, R., Feron, F. J. M., & Jolles, J. (2010). Developmental changes in semantic verbal fluency: Analyses of word productivity as a function of time, clustering, and switching. *Child Neuropsychology, 16*(4), 366-387.

Kavé, G. (2006). The development of naming and word fluency: Evidence from hebrew-speaking children between ages 8 and 17. *Developmental Neuropsychology, 29*(3), 493-508.

Kavé, G., Kigel, S., & Kochva, R. (2008). Switching and clustering in verbal fluency tasks throughout childhood. *Journal of Clinical and Experimental Neuropsychology, 30*(3), 349-359.

Koren, R., Kofman, O., & Berger, A. (2005). Analysis of word clustering in verbal fluency of school-aged children. *Archives of Clinical Neuropsychology, 20*(8), 1087-1104.

Lanting, S., Haugrud, N., & Crossley, M. (2009). The effect of age and sex on clustering and switching during speeded verbal fluency tasks. *Journal of the International Neuropsychological Society, 15*(2), 196-204.

Miller, G. A. (1995). WordNet: A Lexical Database for English. . *Communications of the ACM, 38*(11), 39-41.

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the International Neuropsychological Society, 12*(4), 570-574.

Nash, H. M., & Snowling, M. J. (2008). Semantic and phonological fluency in children with Down syndrome: Atypical organization of language or less efficient retrieval strategies? *Cognitive Neuropsychology, 25*(5), 690-703.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin, 84*(1), 93-116.

Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly, 42*(2), 282-296.

Raboutet, C., Sauzéon, H., Corsini, M.-M., Rodrigues, J., Langevin, S., & N'Kaoua, B. (2010). Performance on a semantic verbal fluency task across time: Dissociation between clustering, switching, and categorical exploitation processes. *Journal of Clinical and Experimental Neuropsychology, 32*(3), 268-280.

Raoux, N., Amieva, H. l. n., Le Goff, M. l., Auriacombe, S., Carcaillon, L., Letenneur, L., et al. (2008). Clustering and switching processes in semantic verbal fluency in the course of Alzheimer's disease subjects: Results from the PAQUID longitudinal study. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 44*(9), 1188-1196.

Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika, 42*(3), 319-345.

Sauzéon, 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.

Siaw, S. N. (1984). Developmental and population comparisons of taxonomic and thematic organization in free recall. *Journal of Educational Psychology, 76*(5), 755-765.

Tröster, A. I., Warmflash, V., Osorio, I., Paolo, A. M., Alexander, L. J., & Barr, W. B. (1995). The roles of semantic networks and search efficiency in verbal fluency performance in intractable temporal lobe epilepsy. Epilepsy Research, 21(1), 19-26.

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.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. Neuropsychologia, 36(6), 499-504.

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. Psychology and Aging, 18(2), 332-339.

APPENDIX A. SUBCATEGORIES IN THE FOOD CATEGORY

Table A1. Subcategories organized by contextual relations (FOOD)

Categories	Examples
Meal	"rice, veggie, meat, fish"
Breakfast	bread, milk, cereal, Taiwanese omelet, radish cake
Fast food	French fries, hamburger, cola
Snack/dessert	cake, cookie, candy
Soup with ~, ~	mushroom soup, mushroom
Curry	curry chicken, rice
Butter and Bread	butter, bread/ toast, jam

Note: The examples are not exhaustive.

Table A2. Subcategories organized by taxonomic relations (FOOD)

Categories	Examples			
Liquid	beverage	milk	milk, goat milk, yogurt drink	
		tea	green tea, black tea, milk tea	
		other	coke	
	oil			
Protein (Label Name: Egg-Bean-Fish-Meat Category)	meat seafood	meat	chicken	drumstick, wing
			pork	rib, pork leg,
				sausage, hotdog
			beef	beef noodle, steak
			snake	
		seafood	fish	salmon, codfish, shark meat
			crab, shrimp, clam, oyster	
	egg	egg roll, omelet, tea-flavored egg, steamed egg, boiled egg, egg white, yolk		
	bean	Red bean, soy bean, soy milk		
Main food	rice	rice roll, sushi, brown rice, white rice		
	flour	noodles	Stir-fried noodles,	
		bread, bun cookie bread, noodles, flour		
	root	potato, taro, sweet potato/yam, carrots		
		radish		Radish, radish cake
Fruit	orange, tangerine			
	strawberry, blueberry			
	watermelon, muskmelon			
	juice			
Vegetable	white cabbage, celery cabbage, tomato, corn, potato			
	needle mushroom, dried mushroom			
	corn	popcorn, corn, corn soup		
	celery	American celery, wild celery		
Seasoning	sugar	Caramel, honey		
	salt, garlic, soy sauce, salad dressing			
Red food	carrot, tomato, chili pepper, ketchup, strawberry			

Note: The examples are not exhaustive.

APPENDIX B. SUBCATEGORIES IN THE ANIMAL CATEGORY

Table B1. Subcategories organized by contextual relations (ANIMAL)

Categories	Examples	
African animal	monkey, gorilla, lion, tiger, ostrich, hippopotamus, rhinoceros, giraffe, crocodile	
Arctic animal	polar bear, polar fox, penguin	
Farm animal	buffalo, cattle, dairy cattle, horse, donkey, pig, rabbit, sheep, goat goose, swan, chicken, duck Fish, pig	
Prey-predator	cat, bird, rat; snake, squirrel; tiger, deer;	
Zoo animals	African (monkey, elephant, lion, leopard, hippo, rhino, gorilla), bear (bear, koala, panda) penguin, snake deer, horse	
House	pet	Turtle, bird, cat, chameleon, chick, dog, fish, guinea pig, lizard, mouse, parrot, Pekingese pig, poodle, rabbit,
	to be evicted	Ant, cockroach, gecko, mosquito, mouse, snake
	human	
Story	turtle, rabbit; mouse, elephant	
Water animal	whale, frog,	
Zodiac animals	rat, ox, tiger, rabbit, dragon, snake, horse, goat/sheep, monkey, chicken, dog, pig	

Note: The examples are not exhaustive.

Table B2. Subcategories organized by taxonomic relations (ANIMAL)

<table>
<tr><th>Categories</th><th colspan="3">Examples</th></tr>
<tr><td>Feline</td><td colspan="3">cat, Indian leopard, lion, tiger, leopard, Siberia tiger</td></tr>
<tr><td>Canine</td><td colspan="3">dog, fox, wolf</td></tr>
<tr><td>Bear</td><td colspan="3">Formosan black bear, polar bear, koala, panda</td></tr>
<tr><td>Bird</td><td colspan="3">chicken, duck, finch, woodpecker, eagle, peacock, sparrow, dove, owl, sea gull, cuckoo, sparrow, bald turtledove, egret, swan, robin, vulture</td></tr>
<tr><td>Cow</td><td colspan="3">cattle, dairy cattle</td></tr>
<tr><td>Dinosaur</td><td colspan="3">pterosaur, stegosaur, Tyrannosaurus rex, Allosaurus</td></tr>
<tr><td rowspan="5">Worm/Bug</td><td rowspan="2">insect</td><td>beetle</td><td>rhinoceros beetle, scarab, stag beetle</td></tr>
<tr><td colspan="2">dragonfly, praying mantis, bee, cicada, cockroach, cricket, damselfly, firefly, fly, grasshopper, ladybug</td></tr>
<tr><td colspan="3">ant, bee, beetle, bug, butterfly, caterpillar</td></tr>
<tr><td colspan="3">centipede, chilopod, earthworm,</td></tr>
<tr><td colspan="3">spider</td></tr>
<tr><td rowspan="2">Small mammal</td><td colspan="3">rabbit, dog, cat</td></tr>
<tr><td>rodent</td><td colspan="2">squirrel, mouse</td></tr>
<tr><td>Furry, tender</td><td colspan="3">deer, koala, rabbit, panda, pig, sheep, goat</td></tr>
<tr><td>Fish</td><td colspan="3">cannibalistic fish, clown anemonefish, great white shark, shark, skate, whale, soft-shelled turtle</td></tr>
<tr><td>Equine</td><td colspan="3">donkey, horse, unicorn, zebra</td></tr>
<tr><td>Eagle</td><td colspan="3">bald eagle, god eagle</td></tr>
<tr><td>Elephant</td><td colspan="3">elephant, mammoth, sea elephant</td></tr>
<tr><td>Deer</td><td colspan="3">Formosan sika deer, giraffe, moose</td></tr>
<tr><td>Predator</td><td colspan="3">tiger, dinosaur, Allosaurus, bald eagle, bear, wolf dog, cheetah, crocodile, eagle</td></tr>
<tr><td>Herbivore</td><td colspan="3">camel, cattle, dairy cattle, deer, elephant, giraffe, goat, horse, moose, rhinoceros, Tibet antelope, zebra</td></tr>
<tr><td rowspan="4">Cold-blooded</td><td rowspan="2">reptile</td><td>lizard</td><td>Agama, Komodo Dragon</td></tr>
<tr><td>snake</td><td>Python, rattlesnake</td></tr>
<tr><td></td><td colspan="2">chameleon, crocodile</td></tr>
<tr><td>amphibian</td><td colspan="2">frog, Od red frog, toad</td></tr>
<tr><td>Animals that go through metamorphosis</td><td colspan="3">frog, butterfly, silk worm, caterpillar</td></tr>
</table>

Table B2. Continued

Young-adult	tadpole-frog; caterpillar-butterfly
Shape	fish, snake; rhino, hippo; rhino, Triceratops; dinosaur/kangaroo; anteater-porcupine
Fly	bird, butterfly, bat, albatross, cicada, eagle, dragonfly, fly, mosquito, Chinese bulbul, parrot, polatouche
Crawl	snake, snail, turtle
Behavior	human, parrot

Note: The examples are not exhaustive.

Table B3. Subcategories organized by contextual and taxonomic relations (ANIMAL)

Categories	Examples
African feline	lion, tiger, leopard
Farm mammal	cattle, sheep/goat, pig
Farm bird	goose, swan, chicken, duck
Water mammal	whale, dolphin
Water bird	duck, goose, penguin, egret
African shape	Hippocampus, rhinoceros
Grassland; herbivore	antelope, leopard, deer, horse, goat/sheep

Note: The examples are not exhaustive.