Russian Language Journal

Volume 65 | Issue 1 Article 2

2015

The Effect of Teaching Vocabulary in Semantic Groups: A Study in the Russian Language Classroom

Kate White

Russian Language Journal
Vol. 65, 2015
ARTICLES

The Effect of Teaching Vocabulary in Semantic Groups: A Study in the Russian Language Classroom

KATE WHITE

A long-standing assumption in the field of second language acquisition research is that learning new vocabulary items in semantic groupings has a positive effect on acquisition and retention (Finkbeiner and Nicol 2003). This assumption is common among researchers and instructors of second languages, as it seems to fit intuitively with the most popular current communicative approaches to teaching. However, researchers have begun to question this assumption, as it has not been supported by empirical evidence (Altarriba and Mathis 1997; Finkbeiner and Nicol 2003; Papathanasiou 2009). Previous research is not conclusive on the topic due to differences in methodology and design. In this study this issue is explored in more detail and with a language not previously investigated: Russian.

Definitions

For the purposes of this study and for the previous research that is reviewed below, *word learning* refers to initial learning of novel vocabulary items by second-language (L2) learners. Also, for this study, semantic and mixed groups were defined as follows: Semantic groups were defined according to the standards used in the studies of Finkbeiner and Nicol (2003), Tinkham (1997), and Hoshino (2010). Finkbeiner and Nicol do not articulate a definition for a semantic group, but provide the examples of animals, kitchen utensils, furniture, family members, body parts, items in a classroom, and places in the community. Thus, their definition of a semantic group seems to be the following: words that are related to a single context of a similar syntactic type. Tinkham seems to agree, as he defines a semantic cluster as a set "of semantically and syntactically similar words" which fall under "a common superordinate or covering concept… and are consequently gathered together as a result" of these shared characteristics (1997, 138–39). Tinkham relates these

clusters to the idea of semantic fields and provides the examples of colors, fruits, and professions. Hoshino (2010) used picture dictionaries to determine both semantic and thematic groupings[1]. The definition used here and proposed for semantic groupings is a set of words of the same part of speech that refer to the same segment of reality; this set consists of hyponyms[2] of a larger category, or hypernym. For example, whale, shark, crab, and jellyfish are all hyponyms of sea creature, their hypernym (or category/segment of reality). whale, shark, crab, and jellyfish are all hyponyms of sea creature, their hypernym (or category/segment of reality).

In this study, mixed groups were defined according to the groupings used by Tinkham (1997), Finkbeiner and Nicol (2003), and Papathanasiou (2009). Papathanasiou (2009) defines these groups as those containing words that are not semantically related. Tinkham defines mixed groups as linguistically unrelated sets, in which words "of the same form-class… do not directly descend from a common superordinate concept" (143). Tinkham provides an example of such a set: "cigar, wolf, lace, stone, chain, fuel, paint, funeral, recipe, market, uncle, ice" (151). Finkbeiner and Nicol created unrelated groups by taking an item from each of their semantic sets that were unrelated to each other (animals, kitchen utensils, body parts, and furniture). For this study, items in the mixed group were chosen from words that are not semantically related and that could not be considered hyponyms.

Review of Previous Research

The body of literature that is devoted to asking why instructors may not question this method of using semantically-related groupings in vocabulary presentation is small. In their review of researchers' and instructors' reasons for adopting this approach, Finkbeiner and Nicol (2003) cite assertions that grouping words by meaning provides precision for learners; in other words, having words presented in a semantic group helps learners to define the boundaries between the related words with specificity. Learning words that have been grouped semantically may also help to reinforce the overall meaning and helps learners notice fine-grained meaning distinctions (Finkbeiner and Nicol 2003; Papathanasiou 2009). To a language instructor, this practice seems common, as many current language textbooks use this method for vocabulary instruction[3]

(Tinkham 1997). While these reasons seem sound, a question remains: Is this method supported by empirical research?

Where is the evidence?

In their study, Altarriba and Mathis (1997) asked to what extent bilinguals are sensitive to semantic information in a translation decision task. The participants were English monolinguals (ML) and English-speaking second-language (L2) learners of Spanish (or bilinguals, BL). In the experiment, the authors introduced both groups of participants to new words on a computer with translation equivalents. The training tests required the participants to produce English translations of Spanish words, to fill in blanks in sentences with Spanish words, and to write Spanish words that best fit a given definition. The training phase was production-based and the post-test involved recognition in a translation decision post-test.

The authors found that the BLs at low and high levels of proficiency responded more slowly in the semantically-related condition, or when the translation prompt was semantically similar to the target word. Even the MLs were sensitive to the semantic information. Their results show that language learners and bilinguals are sensitive to the meaning, and not just the form, of a word from even initial stages of learning (Altarriba and Mathis 1997). Also, the authors showed that semantically-related items interfered with each other in recognition. This is supported by the findings of Isurin and McDonald (2001), who found that recalling words from a list was more difficult when the words were translation equivalents for a second list presented afterwards. In other words, the semantic similarity of the items on the second list interfered with retention and recall in memorization of the first list (Isurin and McDonald 2001).

Tinkham (1997) explored the effect of semantic and thematic clustering on word learning in learners of English as a second language. In his study, Tinkham separates groupings into two areas – semantic and thematic. Semantic groupings usually include lexical items that are one part of speech, as in nouns or verbs (i.e. peach, pear, plum, etc.). Thematic groupings may include a mixture and be more loosely related, such as lexical items relating to one specific place, schema, situation, or idea (i.e. pond, slimy, frog, lily pads, etc.). Tinkham used artificial words that he

created based on English phonotactics and his participants were first-language (L1) English speakers.

Tinkham found that new L2 vocabulary was learned with more difficulty when words were arranged in semantic groups. It is unclear what the effect of thematic groupings was, as it was detrimental for some but beneficial for others; however, it seemed to be beneficial rather than detrimental in more instances. The students also reported that learning the words in semantic groups was more difficult. In a replication study, Waring (1997) found the same effect with Japanese non-words. Tinkham points out some issues in applying the conclusions of the study, including the lack of a late post-test and the issue of lack of context for the words. Tinkham asserts that his results contradict the general assumption that semantic grouping is better for word learning, but concludes that more research is needed to determine why this is true (Tinkham 1997).

In a similar study, Finkbeiner and Nicol (2003) created 32 novel words based on English phonotactics (Ex. *birk*, 'cat'; *gorp*, 'cow'). These novel words fell into four semantic categories: animals, kitchen utensils, furniture, and body parts. The study took place over a two-day period and the participants were L1 speakers of English. The tasks included training and post-tests requiring production (translation) and recognition (word-decision). The words were presented aurally and images corresponding to the words were shown on a computer screen (Finkbeiner and Nicol 2003, 373). The word form was also given visually. The results of the post-tests showed that semantic groupings caused slower processing of new words. They tested word learning through measuring reaction time only.

Papathanasiou (2009) performed a study similar to that of Finkbeiner and Nicol (2003), but she used real English words with Greek learners of English as a foreign language. She administered the test to two different age groups, and the teaching and testing took place in the classroom. While her design was somewhat confounding due to conflation of age and proficiency, she found that the adults performed better with unrelated vocabulary. Children showed no significant difference in their performance for each group of vocabulary. She asserted that while semantic groupings may be more useful for instructors when planning classroom activities, there is no evidence that presentation of words in these groupings is beneficial (Papathanasiou 2009). The author

suggests that in previous studies, it has not been clear what would happen with a real language in a classroom setting.

Each of these studies has shown that semantic groupings are not significantly helpful to L2 vocabulary learning. Finkbeiner and Nicol (2003) assert that the source of this assumption about semantic groups lies in memory studies. Proponents of using semantic groupings often cite memory studies that require the participants to memorize lists of words (Finkbeiner and Nicol 2003; Erten and Tekin 2008). These studies have caused researchers to conclude that semantic groupings facilitate word learning, as the results of the studies show that participants who memorized semantically-grouped words recalled more items in the post-test (Bousfield 1953; Cofer 1966; Cohen 1963; Lewis 1971; Tulving and Pearlstone 1966; Tulving and Psotka 1971).

However, there are two issues with this approach. First, these studies used monolingual speakers. Most researchers in second language acquisition accept that monolinguals and bilinguals cannot be compared in this way (Cook 2002; Grosjean 1998) because of differences in their cognitive structure. Even low-proficiency L2 learners exhibit cognitive differences from monolinguals, such as semantic sensitivity (Altarriba and Mathis 1997) and processing speed (Kroll and Stewart 1994). The second issue is that monolinguals in the above studies were not learning new words; rather, they were memorizing lists of words they already knew. There is little evidence to support the idea that a meaningful comparison can be made between novel-word learning by bilinguals and word memorization by monolinguals of already-known words (Barcroft, 2002), especially as these are separate cognitive processes (Finkbeiner and Nicol 2003, 371). Overall, as shown in the studies summarized above, there is little empirical evidence to support the assumption that presenting novel vocabulary words in semantically-related groups is beneficial for learning.

Processing Depth

Though the studies above have shown that semantic groups may not facilitate word learning, there are some remaining questions. First, none of those studies used late post-tests, and as a result, the issue of processing depth might have been overlooked. When learning new words, attention is not infinite, and one type of information dominating the resources

available will detract from other aspects. For example, attention to meaning and form are in direct opposition, as any attention given to form will diminish resources available for attention to meaning, and vice versa (Barcroft 2002; Barcroft 2004; Lee and VanPatten 2003). In order to fully acquire a new word, processing resources must be devoted to encoding the word form in memory, activating appropriate semantic information (including collocational, syntactic, and other information), and creating a connection between the form and the meaning (Barcroft 2012). Also, as stated in the TOPRA, or "type of processing-resource allocation", model, an overload of one type of information (e.g., semantic elaboration) will result in a lack of resources available for other aspects of word learning (Barcroft 2004).

In terms of word groupings, the results of some previous studies show that it takes longer for participants to learn labels for new words when these words are grouped semantically, as discussed above (Finkbeiner and Nicol 2003; Higa 1963; Kintsch and Kintsch 1969; Kroll and Stewart 1994; Nation 2000; Tinkham 1993; Tinkham 1997; Underwood, Ekstrand, and Keppel 1965; Waring 1997). As Craik and Lockhart (1972) argue, presenting words in semantic groupings may cause the participants to use deeper processing because of the need to distinguish each item's semantic area and because of increased semantic elaboration (Craik and Lockhart 1972; Craik and Tulving 1975; Schneider, Healy, and Bourne 2002). Semantic elaboration is the process of increasing focus on the semantic value of a word (Craik and Lockhart 1972). They argue that this causes the words to be learned more fully, and that this fuller learning will be evident in slower learning.

However, semantic elaboration may also result in inhibition in novel-word learning (Barcroft 2004). It is possible that it simply takes participants longer to encode the words because grouping them semantically makes it more difficult for learners to process them, or to separate semantic distinctions, quickly. It may also result in too much focus on semantic information in the input to the exclusion of structural information, inhibiting the learner's acquisition of the novel word (Barcroft 2004). In 2002, Barcroft found that increasing the amount of semantic processing by requiring more elaborate manipulation of information (on multiple levels, including semantic, syntactic, and lexical) can inhibit a learner's ability to encode the formal properties of a new

word, supporting the TOPRA model. Learners have limited ability to use various processing resources at once (Barcroft 2002, 353), and there are limited available processing resources for language (de Groot 2011; Robinson 2003; Schmitt 2008). Robinson 2003; Schmitt 2008).

In other words, the deeper processing involved in learning semantically-grouped words may be used for distinguishing semantic relations between the words rather than contributing to their better retention. This means that the processor is busy with the semantics and does not have enough space left to encode the novel word form, as predicted in the TOPRA model (Barcroft 2004; Erten and Tekin 2008). Conversely, the distinctiveness of unrelated words could allow better retention due to the deeper processing, as less processing is taken up by the need to distinguish semantic areas. If deeper processing is used for analyzing the differences between semantically-related items, the words could be more difficult to retrieve later due to lack of encoding (Barcroft 2004; Finkbeiner and Nicol 2003).

Overall, there is a consensus that learners have limited processing resources for language learning, and that these are depleted in various ways by attending to portions of input (de Groot 2011). Between these limitations, there are many questions that must be answered regarding how instruction can be effective for vocabulary acquisition when learner cognition is taken into account. As these and other researchers assert, the question remains: Does slower learning mean deeper processing and better retention?

Proficiency Level

As mentioned above, Altarriba and Mathis (1997) explored the availability of semantic information in word learning. They also introduced an important variable that may relate to ease of word learning: proficiency level. From the part of their experiment regarding semantic groupings, the authors suggested that conceptual information plays a role in L2 word learning (Altarriba and Mathis 1997, 558). They concluded that new words in an L2 are connected very early with their corresponding concepts, as well as with translation equivalents in the L1. Both levels of participants (beginners and higher-proficiency L2 learners) showed semantic interference in the post-tests, though the amount of semantic interference diminished as proficiency increased (Altarriba and Mathis

1997, 559). The important question remains: What is the exact effect of semantic groupings for participants at different proficiency levels?

To summarize, in previous studies authors have shown that the idea that semantic groupings facilitate word learning is misguided (Altarriba and Mathis 1997; Erten and Tekin 2008; Finkbeiner and Nicol 2003; Papathanasiou 2009). While results from previous studies may be contradictory due to issues with methodological design, what is important is that such a basic assumption regarding language instruction should be supported empirically. Previous research in this area has a few weaknesses that the current study attempts to address. The first weakness is the confusion in previous studies' post-tests; it is difficult to understand what conclusions can be made when there is no consistency in testing approaches and some studies lack late post-tests, which could be used to investigate processing depth (Finkbeiner and Nicol 2003; Papathanasiou 2009). Second, few researchers have investigated the possible effects of proficiency levels (Altarriba and Mathis 1997). Third, as Papathanasiou (2009) asserts, few researchers have used a real language or conducted this research in the classroom. Overall, previous studies have shown important findings about semantic groupings in vocabulary learning, but there have been inconsistencies in design (Papathanasiou 2009) and results (see, for example, Finkbeiner and Nicol 2003; Tinkham 1997).

The Current Study

The present study was designed to investigate the problem of the effect of semantic groupings on word learning while taking previous methodological inconsistencies into account. The overall design was based on that of Finkbeiner and Nicol (2003). A late post-test was added in order to investigate the idea of processing depth. In addition, the study used real Russian words rather than non-words based on English phonotactics (Finkbeiner and Nicol 2003). Third, the procedure simulated classroom L2 vocabulary learning by placing the study in a classroom setting. Finally, in this study, the variable of proficiency level was included. The questions posed in the pilot study were as follows:

1. Does grouping words semantically facilitate or hinder L2 vocabulary learning?
2. Does the effect of semantic groupings diminish as L2 proficiency level increases?

3. Is there evidence of slower learning, and therefore deeper processing, on the late post-tests for semantic groupings?

It was predicted that, in line with previous results, semantic groupings would hinder initial L2 vocabulary learning due to semantic interference. Third-year participants were predicted to be less affected, or unaffected, by semantic interference. Finally, if semantic groupings encourage deeper processing, this would be revealed on the late post-tests, when words that are processed more deeply would be better retained. In that case, there may be no difference between groupings on the immediate post-test.

Methodology

The methodology for this study was based on that of Finkbeiner and Nicol (2003), with a few changes. In their study, they created 32 novel words based on English phonotactics and taught them to L2 speakers of English. They included only an immediate post-test, which included a translation task, and they measured only reaction time. The changes include the target participants (students enrolled in Russian language courses), the use of a real language for the vocabulary items (Russian), the inclusion of proficiency levels as a variable, and the exclusion of the translation task. The procedure and focus of the analysis were also altered. Proficiency level and condition (i.e., whether or not the participant learns the words in semantic groups) were independent variables, while accuracy was a dependent variable. Other independent variables, such as study abroad and previous language study, were taken from the pre-study questionnaire as appropriate for each participant group.

Because the learning phase took place in the participants' Russian language classrooms, the learning phase was structured differently as well. Instruction was based on the format typically used at the students' university and used recommendations on vocabulary instruction given by Barcroft (2012). Throughout the study the words were presented aurally and without orthographic stimuli due to possible confounding, cross-linguistic effects (Hoshino and Kroll 2008). Previous research shows that there may be confounding effects from introducing orthography in initial stages of vocabulary learning (Barcroft 2012); learning the orthographic form of the word may be considered a different stage of word learning.

Participants and materials

The participants were students learning Russian at two levels, the first-year (n=8) and third-year (n=8) of language study[4]. While the students in the respective levels were not tested for proficiency with any tasks during the study, it was assumed that they were of similar levels as their classmates. Therefore, their classification depended on their class placement; the first-year students had been enrolled in the same courses for almost three quarters, and the third-year students had been enrolled in the same courses for at least three quarters. The curriculum at the students' university follows strict descriptors for proficiency levels developed by the American Council on the Teaching of Foreign Languages (ACTFL[5]) and uses them as guidelines for placement in Russian courses. The students were offered extra credit for involvement in the study. At the beginning of the study, before the learning phase, the participants were asked to fill out a questionnaire. Topics included previous language-study, grades in Russian courses, study abroad experience, exposure to Russian outside of class, and total years studying Russian.

The following information was taken from the questionnaires. All of the participants earned either B- or A-averages in their Russian courses. All 16 had studied another language; 7 had studied more than one other language. These languages included: Spanish (12); Latin (3); French (3); German (2); Arabic (2); Czech (1); Italian (1); and Chinese (1). At the third-year level, half (4) of the participants had studied abroad. The average amount of time per week spent on studying Russian was 7.6 hours. All participants were asked to report what area they found most difficult in learning Russian; their responses included: cases (6); grammar (7); vocabulary (4); listening (1); reading (1); syntax (1); and speaking (1). 14 of the 16 participants reported that it was difficult for them to learn the words in the task. They were also asked to report which strategies they used at home to learn vocabulary words. These strategies included: flashcards (7); rewriting (6); using the vocabulary words in sentences (4); repetition aloud (3); and seeing the vocabulary words in use (2).

The stimuli were black and white images[6] depicting the words chosen in three semantic sets: kitchen utensils, sea creatures, and tools (see Appendix for word lists). These items are unlikely to be mentioned in the classes leading up to the third-year; this assertion is based on an analysis

of the books[7] used in first-year, second-year, and third-year Russian courses at the university where the participants study. On the questionnaire, the participants were asked to indicate if they knew any of the words before the task by providing a Russian translation of a given English word. None of the participants included in the final analysis could produce any of the words prior to the experimental part of the study. The words were not included in the participants' other lessons during the time in between the learning phase and the post-tests. In the learning phase the images appeared in a PowerPoint presentation and were accompanied by the aural form of the word. On the post-tests the same images appeared in tests delivered by the computer program, SuperLab.

Procedure

The researcher, who was not an instructor for any of the participants at the time of the study, personally conducted the entire data collection, which required four weeks of contact with the participants. First, the participants completed consent forms and pre-study questionnaires. Second, they performed the learning task over two sessions (two weeks apart) in their Russian classes. This task required the participants to learn words presented to them in class in a manner similar to other classroom vocabulary sessions—from a PowerPoint presentation. They were asked to concentrate on learning all of the items presented to them in the task and not to study the words at home between tests.

An image for each word was shown four times total while the researcher pronounced each word twice every time the image was shown. This format is a close approximation of a classroom lesson at this university, as students in this program are accustomed to learning vocabulary in this way from a variety of teachers. This change from the study of Finkbeiner and Nicol (2003) was incorporated in order to reflect typical classroom learning at the participants' university. There were four trials of each block of words in the learning task. The participants were not asked to repeat the words aloud for the first two trials. This change from the Finkbeiner and Nicol study was made in line with previous studies that show a negative effect of immediate production for novel words (Barcroft 2012). The participants were asked to produce the words during the last two trials (of four) for each block of words.

Each of the participants learned one portion of each set of words (i.e., five of the sea creatures, five of the utensils, five of the tools) in a semantically-related condition during Session 1 (Table 1). These sub-groupings were counterbalanced for length and gender. They learned the other words (4 more of each set) in a mixed condition during Session 2 (Table 2), which also included filler words of similar length to the target words. The filler words were not tested in the post-tests. Words were grouped into blocks of five words each. Each block appeared four times, though never twice in a row. The block design was based on that of Finkbeiner and Nicol (2003).

Table 1. Shows "Condition 1" from Session 1 of the experiment.

Block	Contents
Sea creatures 1	5 sea creatures
Utensils 1	5 utensils
Tools 1	5 tools
Repeat	Repeat above set of three blocks three more times; on the last two times ask the participants to repeat the words.

Table 2. Shows "Condition 2" from Session 2 of the experiment.

Block	Contents
Mixed block 1	These blocks contain 5 words each. 1 from each of the groups (sea creatures, utensils, tools) and 2 filler words. Ex. whale, filler word 1, hammer, spatula, filler word 2.
Mixed block 2	
Mixed block 3	
Mixed block 4	
Repeat	Repeat above set of four blocks three more times, on the last two times ask the participants to repeat the words.

There were fifteen words in the semantic condition, and twenty words in the mixed condition, though the fillers were not tested. This discrepancy is due to the need to add fillers to the second group, along with one word from each of the semantic groups, in order to maintain the same block size in both conditions. While it is possible that having two different numbers of words in the learning phases may skew the results

(especially since learning more words requires more processing effort), this was not the case, as will be discussed in the results section.

Each participant performed the same immediate post-test, which included a production-based task and a recognition-based task, in that order. The test was administered individually to each student in a room where only the researcher was present. The same images used in the training task appeared on the screen in the program SuperLab. None of the participants could recall the words well enough to produce more than one or two in the picture-naming task at either the early or late post-test; because of this, the results of the production task are not reported here. This is not an unexpected result, as passive (recognition) knowledge of vocabulary often precedes active (productive) knowledge (Barcroft 2004). On the recognition task, their responses were recorded for accuracy in SuperLab. All of the items for each group were tested in each task. For the recognition task, the participants heard a word over their headphones and were asked to indicate by pressing one of two keys (i.e., incorrect (L) or correct (S)) whether the word matched the image on the screen. The aural forms of the words for the task were recorded by the same researcher who taught the words in the learning session in order to avoid the possible confounding effect of an unfamiliar voice and accent.

There were three possibilities for word assignment in the recognition task: words assigned correctly to their corresponding images (i.e. "whale" to a whale); words assigned incorrectly to a word within the same semantic grouping (i.e. "walrus" to a whale); or words assigned to an unrelated image from a different grouping (i.e. "spatula" to a whale). The same post-tests were used as late post-tests one week later, with the order of the items used in each test flipped. The order was flipped to avoid possible confounding effects from task familiarity.

The same procedure was used for the second task one week later, when the participants learned the words in mixed groups. After the learning phase, which was the same as the semantic learning phase but with the mixed group of words, the participants performed an identical immediate post-test using the mixed group. All participants then completed the late post-test for the semantically-grouped words. The second session therefore included a learning phase, an immediate post-test for the new set of words, and a late post-test for the first set of words.

All participants were then asked to meet with the researcher one week later for a late post-test on the second, mixed group of words.

Data Analysis

The accuracy of the participants on the recognition task was recorded in all of the post-tests in order to analyze retention. A logistic mixed-effects model (Jaeger 2008) was used to analyze accuracy in the recognition post-tests. It was intended to concentrate on accuracy as a change from previous studies, where the concentration was on latency to the exclusion of accuracy.

For accuracy analysis, a logistic mixed-effects model (Jaeger 2008) was fit to the participants' accuracy data between conditions at both proficiency levels. The model structure is given in Table 3, as well as the output, which is discussed below. Fixed effects include post-test (early or late), manner of word grouping in instruction (semantic or mixed), and year of study (first or third). Other fixed effects included number of other languages studied (one or more), exposure to Russian outside of class (fewer or more than 8 hours, the average amount), and the three possibilities for word assignment in the recognition task (correct (1), semantically related (2), and unrelated (3)). Mixed-effects models are the optimal way to analyze these data because they allow for the analysis of random effects of multiple variables at once. The within-group analysis for the third-year group did not yield significant results and is therefore not reported in the table.

A logistic mixed-effects model was fit to the accuracy data for both groupings and all three conditions in the receptive task at both proficiency levels in R[8]. Within the first-year group of participants, there was a significant negative effect of the second condition (Table 3, A, Condition 2: $B = -0.949$, $z = 2.074$, $p < 0.05$). The second condition was the semantic interference condition, when the participants heard a word that was semantically related to the target image. This suggests a negative effect on accuracy due to semantic interference in the recognition task. This result was also found in the analysis of both proficiency levels across groupings (for both groups of words, mixed and semantic), again for only the first-year participants, and was significant (Table 3, B, Condition 2: $B = 0.936$, $z = 2.655$, $p < 0.05$). There was also a significant positive effect of the interaction of grouping and condition 3 (Table 3, B, $B = 1.926$, $z = 3.025$,

$p < 0.05$). This indicates that for words learned in mixed groupings, participants were more accurate in the third condition (unrelated to target image) on the recognition task. At the third-year level, there were no significant effects in the results for accuracy in either learning condition (semantic or mixed groups) or across conditions.

Table 3: Logistic mixed-effects models fit to accuracy data.

Model	*Fixed Effects*	*Estimate*	*Standard Error*	*z-value*	*p-value*	*Random effects*	*Variance*	*Standard Deviation*
A: *Acc. within 1st-year*	Intercept	2.192	0.402	5.455	**<0.05***	Item	0.184	0.428
	Grouping	-0.632	0.540	1.170	>0.05	Participant	0.035	0.188
	Condition 2	-0.949	0.458	2.074	**<0.05***			
	Condition 3	-0.402	0.539	0.747	>0.05			
	Exposure	-0.000	0.480	0.001	>0.05			
	Language	0.064	0.316	0.206	>0.05			
	Grouping: Condition 2	0.696	0.667	1.044	>0.05			
	Grouping: Condition 3	0.936	0.825	1.133	>0.05			
B: *Acc. across years*	Intercept	2.415	0.383	6.295	**<0.05***	Item	0.129	0.360
	Grouping	-0.693	0.420	1.650	>0.05	Participant	0.134	0.365
	Year	-0.473	0.418	1.130	>0.05			
	Condition 2	-0.936	0.325	2.655	**<0.05***			
	Condition 3	-0.838	0.450	1.823	>0.05			
	Exposure	-0.201	0.427	0.052	>0.05			
	Language	0.019	0.290	0.671	>0.05			
	Grouping: Year	0.026	0.381	0.068	>0.05			
	Grouping: Condition 2	0.534	0.457	1.141	>0.05			
	Grouping: Condition 3	1.926	0.637	3.025	**<0.05***			

Table 3. A gives the output of the logistic mixed-effects models fit to the accuracy data for the first year participant group (between groupings and conditions). B gives the model output for the logistic mixed-effects model fit to the accuracy data between year groups. Significant effects are marked with *. Model fit statistics are as follows: A (AIC = 405.7, BIC = 446.4, log-likelihood = -192.8), B (AIC = 785.2, BIC = 851.8, log-likelihood = -378.6). Intercept values are as follows: A (semantic grouping, correct condition, less exposure), B (semantic grouping, correct condition, less exposure).

To summarize, at the first-year level, participants performed significantly less accurately on the semantically-grouped words in the semantic interference condition on the recognition task. This shows that when the first-year participants learned the words in semantic groups, they performed less accurately in the semantic interference condition for those items. For words learned in the mixed condition, there was a

significant positive effect in the unrelated interference condition. This shows that when the first-year participants learned the words in mixed groups, they performed more accurately on the unrelated condition in the post-test. There were no significant effects of the groupings on the third year participants' performance. There were also no significant effects for late post-tests.

Discussion

The first question was, how does grouping words semantically affect vocabulary learning in the L2? The participants in this study at the first-year level performed significantly less accurately in the semantic interference condition when the words were learned in semantic groups. The participants also performed significantly more accurately in the unrelated interference condition when the words were learned in mixed groups. These results show that semantic groups did not facilitate the word learning; in fact, it negatively affected the learning of the first-year participants and had no effect on the third-year participants' performance. Also, learning words in a mixed group positively affected the first-year students' ability to distinguish between unrelated items on the post-tests.

For the second question on proficiency level, the picture is not as clear. Proficiency level did not significantly predict performance in either condition. Effects were found at the first-year level, but the third-year students did not show any significant effects in the analysis, positive or negative. Therefore, higher-level students may be less susceptible to semantic interference in recognition. This study includes 16 participants; in the future, a replication study with more participants of a higher proficiency level is recommended.

The third question considered processing depth. As stated above, the results show that grouping words semantically negatively affected the accuracy of the first-year responses on those words on post-tests. The first-year participants performed better on the mixed groups of words. These results do not support the idea that semantically-grouped words are processed more deeply because the learner is provided with a large amount of semantic information. This deeper processing should lead to slower learning and longer retention as evidenced by their performance on the late post-tests. The participants performed more accurately on the

items in the mixed group as compared to the items in the semantic group, and there was no significant effect for the late post-tests. If deeper processing facilitated the learning of semantically-grouped items over time, this would not be the case; in fact, these results support the idea that processing may be overloaded by the learner's need to distinguish between semantically related items (Barcroft 2004). The encoding and retention of these items may be negatively impacted due to lack of processing resources.

Finally, further results included those for the performance of the participants on the specific conditions in the recognition tests. There were three conditions: correct word assignment (i.e. "whale" to the image of a whale); semantic interference (i.e. "walrus" to the image of a whale); and unrelated interference (i.e. "spatula" to the image of a whale). Grouping words semantically did not just negatively affect the performance of the first-year students – it negatively affected their performance on the semantic interference condition. Conversely, the first-year participants were more accurate on the unrelated condition for the words they had learned in mixed groups. The implication is that learning words in mixed groups increases the ability of the learner to distinguish between unrelated items, while learning words in semantic groups decreases their ability to distinguish between those semantically related items. This finding supports the idea that learners may use more processing resources to distinguish between semantically-grouped items, lowering their overall level of retention.

These results have important implications for methods of vocabulary instruction in the second-language classroom. This study was conducted in the classroom, with methods that more closely mirror what occurs in the classroom than in typical laboratory studies. First, when new vocabulary items are presented in a semantically related group, it may be more difficult for learners' to encode and distinguish between items in initial stages. In fact, initial receptive knowledge is encouraged when vocabulary items are unrelated. Students may be more able to distinguish between items in initial stages if those items are presented in mixed groupings, or at least in groupings that are not entirely related semantically. Second-language instructors may consider using vocabulary items from more than one semantic category when presenting new words in class—this recommendation is generally applicable across

tasks and topics, as it concerns the method of presentation of vocabulary. At the very least, instructors can be aware of the possible interference and confusion that may occur among semantically-related items if they are first presented at the same time. In the future groupings that fall between related and unrelated, such as the thematic groupings used in Tinkham's study (1997), deserve further research.

Second, a few exposures (e.g., four exposures in the current study) to a target item may not be enough for productive knowledge, but it appears to be enough for initial receptive knowledge of vocabulary. Four exposures were enough for the participants to be able to perform the recognition task in this study, which supports the findings of Barcroft (2012). Multiple and varied presentations of the target word may be necessary for different types of word knowledge. This result can inform instructors' expectations in terms of what learners can be expected to do with new words following initial exposures.

In the future, it is recommended that more participants are included in such research. Students at the third-year level of proficiency were not affected by learning words in semantic groups, showing that this issue needs further investigation at more levels of proficiency. Because the effect of other types of word groups, such as thematic groups, was unclear in previous studies (Tinkham 1997), future research could investigate more types of groups. Also, future research on this topic should be conducted in the classroom rather than the laboratory. By closely mirroring actual classroom methods in empirical research, learning gains and outcomes can be better understood. This study contributes to the existing literature by supporting the findings of previous studies regarding the negative effects of semantic groups on vocabulary learning, and also contributes these data from a new L2: Russian.

Notes

1. In the study Hoshino states "words were selected from within the same theme, according to various picture dictionaries" (Hoshino 2010, 304). They list the following dictionaries: Goodman's (1991) *Let's Learn English Picture Dictionary*; Rosenthal & Freeman's (1987) *Longman Photo Dictionary*; Klevberg's (2005) *The Heinle Picture Dictionary*; Ashworth and Clark's (1997) *The Longman Picture Dictionary American*

English; and Shapiro and Adelson-Goldstein's (1998) *The Oxford Picture Dictionary*.

2. Hyponyms are words whose semantic field is included within another word, their hypernym (Gao and X, 2013).
3. *Golosa* (Robin, Henry, and Robin 1994); *Nachalo* (Lubensky, Ervin, and Jarvis 1996); and *Troika* (Nummikooski 1996), to name a few. It must also be noted that some studies have differentiated between semantic groupings, where all words are the same part of speech (e.g., nouns), and other types of groupings, where this is not the case.
4. The initial number of participants was higher, but over the two week period of the study some participants trickled out for a variety of reasons, including: non-attendance at the second in-class session, exclusion to avoid confounding variables (ex. non-English first language), and inability to attend the third session. There were no heritage speakers of Russian among the participants.
5. For more information on ACTFL guidelines, see: www.actfl.org.
6. The majority pictures were selected from a standardized set of 520 pictures used in the International Picture Naming Project (IPNP) and available for download at http://crl.ucsd.edu/~aszekely/ipnp/1stimuli.html. Those not available from that list were chosen from internet sources and were similar in style and size.
7. In the first and second years, the textbooks used were *Nachalo* Book 1 and *Nachalo* Book 2. In the third year the textbook used is V Puti (see References for more information).
8. For more information see: www.r-project.org. Version used: 3.1.1.

Appendix

Kitchen utensils:	**Tools:**
Spatula – *lopatka*	Drill – *bur*
Whisk – *venchik*	Hammer – *molot*
Frying pan – *skovoroda*	Nail – *gvozd'*
Ladle – *kovsh*	Plow – *plug*
Tongs – *shchiptsy*	Sledgehammer – *trambovka*
Apron – *perednik*	Screwdriver – *otverka*
Potholder – *rukavitsa*	Saw – *pila*
Skewer – *vertel*	Pliers – *shchipchiki*
Rolling pin - *skalka*	Tape measure – *ruletka*
Sea creatures:	**Filler words:**
Whale – *kit*	Iron – *utiug*

Octopus – *sprut*	Ivy – *pliushch*
Sting ray – *skat*	Hairbrush – *rascheska*
Eel – *ugor'*	Ship – *korabl'*
Crab – *rak*	Vest – *zhilet*
Shark – *akula*	Collar – *osheinik*
Jellyfish – *meduza*	Caterpillar – *gusenica*
Walrus – *morzh*	Shopping cart – *telezhka*
Seal – *tiulen'*	

References

Altarriba, J. and K. Mathis. 1997. Conceptual and Lexical Development in Second Language Acquisition. *Journal of Memory and Language, 36:* 550-568.

Abbs, B., P. Gupta, and N. Khetarpal. 2008. Is Overt Repetition Critical to Expressive Word Learning? The Role of Overt Repetition in Word Learning with and without Semantics. *Applied Psycholinguistics, 29*: 627-667.

Barcroft, J. 2002. Semantic and Structural Elaboration in L2 Lexical Acquisition. *Language Learning*, *52.2*: 323-363.

———. (2004). Second Language Vocabulary Acquisition: A Lexical Input Processing Approach. *Foreign Language Annals*, 37(2): 200-08.

———. (2012). *Input-Based Incremental Vocabulary Instruction.* TESOL Publications.

Bousfield, W. A. 1953. The Occurrence of Clustering in the Recall of Randomly Arranged Associates. *Journal of General Psychology*, *49*: 229-240.

Cofer, C. N. 1966. Some Evidence for Coding Processes Derived from Clustering in Free Recall. *Journal of verbal learning and Verbal behavior*, *5*: 188-192.

Cohen, B. H. 1963. Recall of Categorized Word Lists. *Journal of Experimental Psychology*, *66*: 227-234.

Collins, A. and E. Loftus. 1975. A Spreading Activation Theory of Semantic Processing. *Psychology Review, 82*: 407-428.

Cook, V. 2002. Background to the L2 User. In V. Cook (Ed.). *Portraits of the L2 User*. Clevedon, UK: Multilingual Matters LTD.

Craik, F. I. and R. S. Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal learning and Verbal behavior*, *11*: 671-684.

Craik, F. I. and E. Tulving. 1975. Depth of Processing and the Retention of Words in Episodic Memory. *Journal of Experimental Psychology*: *General*, *104*: 268-294.

Erten, I. H. and M. Tekin. 2008. Effects on Vocabulary Acquisition of Presenting New Words in Semantic Sets Versus Semantically Unrelated Sets. *System*, *36*: 407-422.

Finkbeiner, M. and J. Nicol. 2003. Semantic Category Effects in Second Language Word Learning. *Applied Psycholinguistics*, *24*: 369-393.

Gao, C. and B. Xu, B. 2013. The Application of Semantic Field Theory to English Vocabulary Learning. *Theory and Practice in Language Studies, 3(11):* 2030-2035.

de Groot, A. M. B. 2011. *Language and Cognition in Bilinguals and Multilinguals.* New York: Psychology Press.

Grosjean, F. 1998. Studying Bilinguals: Methodological and Conceptual Issues. *Bilingualism: Language and Cognition, 1(2):* 131-49.

Higa, M. 1963. Interference Effects of Intralist Word Relationships in Verbal Learning. *Journal of Verbal Learning and Verbal Behavior*, *2*: 170-175.

Hoshino, Y. 2010. The Categorical Facilitation Effects on L2 Vocabulary Learning in a Classroom Setting. *RELC Journal, 41(3):* 301-12.

Isurin, L. and McDonald, J. 2001. "Retroactive Interference from Translation Equivalents: Implications for First Language Forgetting." *Memory and Cognition 29,* 312-319.

Jaeger, T. F. 2008. Categorical data analysis: Away from ANOVAs (Transformation or Not) and towards Logit Mixed Models. *Journal of Memory and Language*, *59(4)*: 434-446.

Kagan, O., F. Miller, and G. Kudyma. 2005. *V Puti: Russian Grammar in Context, 2nd Edition.* Prentice Hall.

Kintsch, W. and E. Kintsch. 1969. Interlingual Interference and Memory Processes. *Journal of Verbal Learning and Verbal Behavior*, *8*: 16-19.

Kroll, J. F. and E. Stewart. 1994. Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connection between Bilingual Memory Representations. *Journal of Memory and Language*, *33*: 149-174.

Lee, J. F., and B. VanPatten. 2003. *Making Communicative Language Teaching Happen.* New York: McGraw-Hill.

Lewis, M. Q. 1971. Categorized Lists and Cued Recall. *Journal of Experimental Psychology, 87*: 129-131.

Lubensky, S., G. Ervin, L. McLellan, and D. Jarvis. 2002. *Nachalo Book 1.* McGraw-Hill.

Lubensky, S., G. Ervin, and L. McLellan. 2002. *Nachalo Book 2.* McGraw-Hill.

McGeoch, A. and W. T. McDonald. 1931. Meaningful Relation and Retroactive Inhibition. *American Journal of Psychology, 43*: 579-588.

Nation, P. 2000. Learning Vocabulary in Lexical Sets: Dangers and Guidelines. *TESOL Journal, 9*: 6-10.

Papathanasiou, E. 2009. An Investigation of Two Ways of Presenting Vocabulary. *ELT Journal, 63.4*: 313-322.

Schneider, V. I., A. F. Healy, and L. E. Bourne Jr. 2002. What is Learned Under Difficult Conditions is Hard to Forget: Contextual Interference Effects in Foreign Vocabulary Acquisition, Retention and Transfer. *Journal of Memory and Language, 46*: 419-440.

Tinkham, T. 1993. The Effects of Semantic Clustering on the Learning of Second Language Vocabulary. *System, 21*: 371-380.

———. 1997. The Effects of Semantic and Thematic Clustering on the Learning of Second Language Vocabulary. *Second language Research, 13*: 138-163.

Tulving, E. and Z. Pearlstone. 1966. Availability Versus Accessibility of Information in Memory for Words. *Journal of Verbal Learning and Verbal Behavior, 5*: 381-391.

Tulving, E. and J. Psotka. 1971. Retroactive Inhibition in Free Recall: Inaccessibility of Information Available in the Memory Store. *Journal of Experiment Psychology, 87*: 1-8.

Underwood, B. J., B. R. Ekstrand, and G. Keppel. 1965. An Analysis of Intralist Similarity in Verbal Learning with Experiments on Conceptual Similarity. *Journal of Verbal Learning and Verbal Behavior, 4*: 447-462.

Waring, R. 1997. The Negative Effects of Learning Words in Semantic Sets: A Replication. *System, 25*: 261-274.