Goats Don't Wear Coats: An Examination of Semantic Interference in Rhyming Assessments of Reading Readiness for English Language Learners

Sylvia Moreira
Queens College, City University of New York

Maryellen Hamilton
Saint Peter's College

Abstract

Rhyming tests have historically been used in the education system to assess reading readiness. English language learners (ELLs) have consistently scored poorly on these assessment tools. The current article examines a possible reason for this poor performance by ELLs. Specifically, the authors examined the relationship between semantic associations of visual images and performance on rhyming assessments for ELLs. Two groups of students, native English speakers and native Spanish speakers, were tested using a typical rhyming assessment tool. As expected, it was found that the native English speakers outperformed the native Spanish speakers. An analysis of the native Spanish speakers' errors revealed semantic interference. Educational implications of these findings are discussed.

The population of English language learners (ELLs) in our schools continues to increase. It is estimated that by 2040, the number of students who speak a language other than English will more than triple (Edmonston & Passel, as cited in August & Hakuta, 1993). The monumental No Child Left Behind Act (NCLBA) (2002) has insisted that this growing population meet the same state academic standards as native English-speaking children. Furthermore, this federal legislation has mandated the reporting of annual academic progress of each individual student to his or her respective state. The questions that arise are: How will these annual assessments be conducted? Should the same assessment tools be used for both native English speakers and the growing population of ELLs?

Research shows that past practices with standardized assessments have either excluded ELLs completely, or have been conducted without regard to English proficiency (August & Hakuta, 1993; Rivera & Vincent, 1997). Language-minority students consistently show significant deficits in standardized assessments. These assessments infer that ELLs fail to reach the basic levels of performance in the math and reading achievement tests (U.S. Department of Education, 1992). There are many speculations as to why these students have poorer performance, including but not limited to, bias of the test, the students' lack of test-taking skills and the inability to comprehend the information requested on the test because of limited English proficiency (Puckett & Black, 1994). Nevertheless, ELLs and their instructors are being held accountable for ELLs' performance on standardized assessments on a yearly basis (NCLBA, 2002).

One such assessment tool, currently being widely used, is the Early Literacy Profile (ELP). The ELP is a year-to-year ongoing assessment instrument used for identifying and monitoring the progress of literacy readiness of students in the primary grades. Through rhyming, phonemic-awareness and decoding exercises, students are assessed and given a grade-level benchmark (Phelps, 2001). Research has indicated that rhyming may play an important role in the reading process (Walton & Walton, 2002). Specifically, such research has suggested that if preliterate students are aware of rhyming formats, then reading becomes facilitated (Wood, 2000). Therefore, the justification for using the ELP on preliterate students is based on its ability to assess early emergent readers by their performance in rhyming, phonological awareness, alphabet and letter-sound recognition, spelling, and sight-word recognition (Phelps).

The beginning level of the ELP is a rhyming test in which the student is presented with a picture followed by three other pictures in a horizontal row. The teacher instructs the student to listen to the words as they are spoken while pointing to the visual images. The teacher then points to the first picture, says the word, points to the next picture and says that word. This is repeated for each picture (e.g., coat, coat-guitar, coat-goat, coat-shirt). After the teacher is done with naming each item in the row, the student is asked to point to the picture that rhymes with the target item (i.e., point to the picture *goat* since *goat* rhymed with the target item *coat*). Using pictures in conjunction with words is a common practice in assessment measures (Walton & Walton, 2002). However, this becomes a potential problematic issue when dealing with the assessments of ELLs because second-language acquisition is embedded within contextual (semantic) associations to the native language (Olivares, 2002).

Specifically, Olivares' (2002) theoretical framework suggests that spoken language and thought are not always directly correlated, and that abstract concepts, which are not language dependent, can be transferred from the

native language (L1) to the target language (L2) without specific labels. That is, for a second-language learner not every word has a word-to-word translation and this is particularly true at the beginning of new language learning. These learners may associate a word in their new language to a schematic concept, rather than a corresponding word in their L1. Learning L2 is facilitated by the transfer of knowledge from L1 to L2. As the vocabulary and communicative patterns from L2 are developed, academic skills, literacy development, concept formation, subject knowledge and learning strategies are transferred from L1 (Ovando & Collier, 1998). This approach to language learning thus infers that ELLs use their previous knowledge to negotiate information acquired in L2. This previous knowledge is usually semantic information and not based on phonological similarities between the two languages.

When tested with the ELP on the rhyming section, students are shown four pictures representing words. The first picture represents the target word and the students are asked to choose a picture representing the word that rhymes with the target item from the other three pictures. Although rhyming may be a strategy used for literacy in English, this is not a strategy used in all other languages (Vernon & Ferreiro, 1999). Nor has rhyming been commonly taught as a test-taking skill in other languages. Although ELLs might have some experience with rhyming in native languages, it is unclear whether this experience will transfer to a literacy or test-taking strategy in the target language. If these students are not familiar with the concept of *rhyming* as a literacy strategy, or if they lack the vocabulary drilling in L2 (English) to "hear" the same sound, then they might not be rhyming on the ELP. Previous research has also suggested that ELLs who lack English proficiency might not understand what target language assessment tools are asking of them (Puckett & Black, 1994). So what are these students doing on the ELP and how is it negatively impacting their performance on these types of assessment tools?

The purpose of the current research was to empirically examine one possible reason that ELLs are not performing at grade-level scale when taking reading-readiness tests like the ELP. Olivares (2002) has suggested that most ELLs rely on conceptual relations between L1 and L2 when doing a task in their new language. Although Olivares did not explicitly refer to rhyming assessment, we assumed that ELLs would also use this reliance on conceptual relations when performing a rhyming task. Therefore, we hypothesized that ELLs who perform poorly on rhyming assessments in L2 do so by choosing a word semantically associated to the target item. For example, when faced with the three choices given earlier (i.e., coat–guitar, coat–goat, coat–shirt), they would pick the picture *shirt* since not only is *shirt* semantically related to *coat*, but they have the semantic knowledge from L1 to know that goats and guitars don't wear coats.

The current study attempted to empirically test this hypothesis by providing students with an assessment that mimics the ELP used in many elementary schools. The assessment had a picture target item followed by three choices. For each target item there was a picture representing a phonologically associated word (rhyme), a picture representing a semantically associated word that does not rhyme, and a picture representing a non-associated word.[1] It was predicted that ELLs would consistently choose the semantically associated picture to the target picture more than the other choices, when errors are made.

This will be a noteworthy finding since choosing a semantic associate is a much deeper level of cognitive processing than choosing a rhyme (Craik & Lockhart, 1972). Specifically, Craik and Lockhart have proposed that information is processed at different depths. The deeper people process something the more likely they will remember and retain that information. According to this well-accepted cognitive theory, rhyming is one of the lowest levels of processing that a human does, whereas, to identify a semantic associate is among one of the deepest levels of cognitive processing that people can do. Thus, when ELLs are faced with making a decision on an assessment when they do not have the vocabulary in L2 or the training for that particular test-taking skill, and they revert to what is meaningful for them (a semantic association made possible by the use of visual stimuli), they are actually carrying out a higher level of cognitive processing than those students who choose the correct rhyme. Therefore, the current study will examine whether ELLs who are identified as "poor" performers in rhyming when assessed with the ELP in L2, are actually performing, on the rhyming part of the test, at a more advanced cognitive level by using semantic association.

Method

Participants and Setting

Two groups of participants were selected from a first-grade dual language program in a suburban school environment with a high population of minority students. Twenty students spoke English as their native language and 20 students spoke Spanish as their native language. According to the latest statistics on the school district, 54% of the student population are Hispanic, 38% are non-Hispanic Black, 7% are non-Hispanic White, and less than 1% are Asian/Pacific Islander. Thirty percent of the student body is eligible for free lunch programs and 27% of the student body is classified as limited English proficient (New York State Education Department, 2005).

In the dual language program, the native Spanish-speaking students are considered ELLs and both native English speakers and native Spanish speakers are instructed in English and Spanish. The dual language program model used

in this particular setting is as follows: All children receive language arts instruction in their respective native language and receive a period of second-language instruction daily. Both native English speakers and native Spanish speakers remain together for the rest of the day, including lunch and specials (i.e., music, art, health, physical education and library). All content-area instruction is provided in both languages on an alternating basis (one day English and one day Spanish). The native Spanish speakers have had a minimum of one year and a maximum of two years of formal schooling in the United States. According to Cummins' (1979) criteria, the native Spanish-speaking students had acquired basic interpersonal communication skills in English.

Materials and Design

There were two independent variables in this study. One was the type of student tested (either native English speaker or native Spanish speaker). The second independent variable was the relationship of the pictures to the target item (a phonological associate, a semantic associate, or non-associate). The dependent variable was the performance of the students on an assessment instrument created to mimic the rhyming section of the ELP. The assessment instrument created for this study consisted of 80 black-and-white line drawings chosen from Snodgrass and Vanderwart's (1980) set of 260 standardized pictures. These pictures have been normed on four different variables significant to memory and cognitive processing: name agreement, image agreement, familiarity and visual complexity. The 80 pictures were arranged in 20 rows of 4 pictures. The first picture in each row was always the target item. The remaining three pictures consisted of a semantic associate, a phonological associate and a non-associate picture appearing in a random order (see Figure 1 for an example of the materials). These three categories of pictures were purposely chosen to assess the level of processing used by native Spanish speakers in this typical type of early literacy assessment tool. Responses were recorded on an answer sheet duplicating the items listed in Figure 1.

Procedure

Each participant was tested individually. Before beginning the actual test phase, the researcher went over three practice items with the student. The practice items used were the practice items which accompanied the actual ELP test and were very similar in format to the test items (i.e., a target picture accompanied by three alternate picture choices). The practice and test instructions were identical. The student was asked to point to or say the word of the visual image that represents the word that rhymes with the target item. The researcher said the name of the target item and repeated the target item name when saying each name of the picture alternatives. Specifically, as the researchers pointed to each image they said, "Tell me which word rhymes with
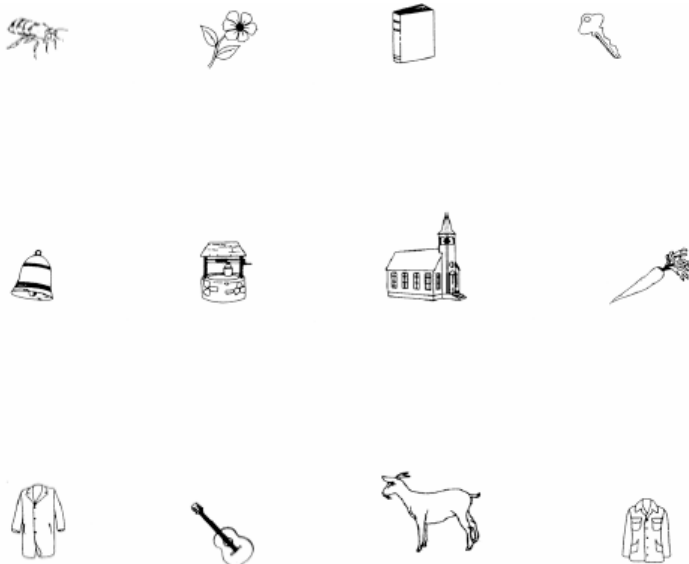
Figure 1. Sample test items.

'coat' (target item)*,* or 'coat/guitar' (non-associate item),'coat/goat' (rhyming item), 'coat/shirt' (semantic associate item)." Responses given by the participant were then recorded on the answer sheet. To ensure that students understood the task requested of them, feedback regarding their choices was provided for the practice items. For example, if the student said *shirt* to the rhyming word *goat*, the researcher said, "No, *shirt* does not rhyme with *goat*, *coat* rhymes with *goat*." No feedback was provided regarding choices during the testing phase.

Results

The mean proportion of correct answers and the mean proportions of semantic errors and non-semantic errors were calculated for the native English speakers ($n = 20$) and native Spanish speakers ($n = 20$). As can be seen in Figure 2, the English speakers ($M = .97, SD = .07$) outperformed the Spanish speakers ($M = .61, SD = .16$). In fact, performance for native English speakers was near perfect. An independent-samples *t* test was conducted to evaluate if there was a significant difference between the mean proportions of correct answers for the two groups. The test was significant, $t(38) = 4.92, p < .05$. Figure 3 illustrates the types of errors made by both native English speakers and native Spanish speakers. As can be seen, native English speakers made very few errors, and when they did there were little differences between the types of errors made, however, native Spanish speakers had more errors for semantic associates than for non-semantic associates.
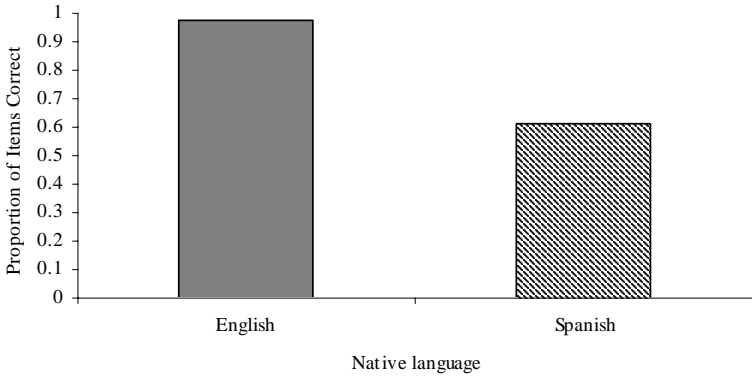
Figure 2. The mean proportion of correct items for native English speakers and native Spanish speakers.
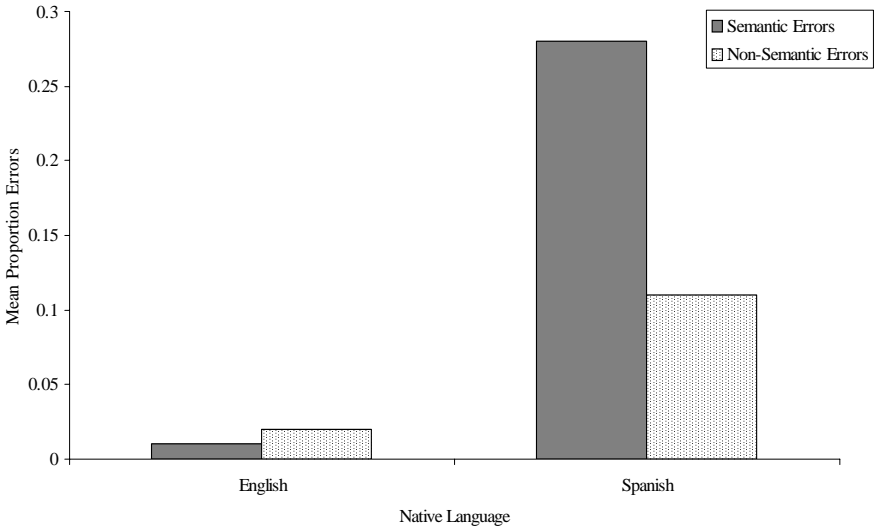

Figure 3. The mean of proportion of errors by type for native English speakers and native Spanish speakers.

These findings were confirmed by a two-way ANOVA which revealed significant main effects for both the language variable $F(1,38) = 23.81, p < .05$, $MSE = .03$ and the types of error $F(1,38) = 6.53, p <. 05, MSE = .02$. In addition there was a significant interaction found between both these variables, $F(1, 38) = 7.32, p < .05, MSE = .02$. However, because the errors were so low for the native English speakers, one needs to be cautious about the interpretation.

Therefore, because of the low error rate for English speakers, and because the main focus of the article was on the types of errors native Spanish speakers would have, we performed a paired-samples t test only on the native Spanish speakers' errors to see if a difference did exist. A significant difference was found between the number of semantic associate errors ($M = .28, SD = .20$) and the number of non-semantic associate errors ($M = .11, SD = .12$) made by the native Spanish speakers, $t(19) = 2.64, p < .05$.

Discussion

As expected, we found a significant difference in performance between the native Spanish speakers and the native English speakers on the rhyming assessment. These results reflect a failing performance by native Spanish speakers when assessed in the strategy of rhyming. On the basis of the knowledge of rhyming words, native English speakers outperformed native Spanish speakers. In a classroom setting, most analysis of assessments would stop here and native Spanish speakers would receive a failing evaluation. This study was implemented to expand on the analysis of the performance of these students and to substantiate some possible explanations for what seems to be, in typical classrooms, the failing results of native Spanish speakers' assessments in rhyming.

The main focus of this paper was to examine the errors made by native Spanish speakers. The native Spanish speakers made more than *twice* as many errors towards the semantic associate than the non-associate. This implies that native Spanish speakers' poor performance may be attributed to the interference of semantic associations with the target rhyming word. The choice of a semantic association over the other two available alternatives supports Olivares' (2002) transfer theory. Specifically, when ELLs were unable to identify the rhyming item, because they lack language or test-taking skills, they relied on their background knowledge to negotiate the new information and chose the items that had meaningful connections to the targets. Thus they chose the semantic associates.

According to Craik and Lockhart's (1972) theory, our findings of ELLs' preference in their errors for semantic associations illustrate a higher-level of processing than rhyming. Although these students performed poorly on the rhyming assessment in this study, when we examine the type of errors they made, we find they are doing a cognitively higher level of processing. The

ELLs are associating two pictures by their semantic meaning— they are not making errors by simply choosing any of the pictures available to them. They are specifically selecting the semantic associate. Therefore, they may be failing the assessment of rhyming, but they are in fact, performing at a higher cognitive level of processing. This pattern was fairly consistent across all the ELLs' errors, regardless of how they performed on the test. That is, both high scorers and low scorers were inclined to make more errors towards the semantic associates. Specifically the 20 native Spanish speakers' performance ranged from 5% to 100% accuracy. For the 17 Spanish speakers who did make errors, the majority of them chose the semantic associate when making errors. This was true whether their total accuracy was at 95% or 5%. As second language learners, they look for clues as to what is being asked of them.

These findings are consistent with a large body of research on miscue analysis (Freeman & Freeman, 1997; Goodman, 1996). Like our findings, the miscue analysis research suggests that not all errors are equal and that some errors are of a higher quality than others (Shaughnessy, 1977). Our findings are also consistent with research that suggests that ELLs use the semantic information of their primary language when dealing with novel situations in their new language (Olivares, 2002). What is unique about our study is that it empirically tests a commonly used assessment tool in our education system and demonstrates that what has been considered poor performance is actually not random errors but systematic higher level cognitive choices.

The implications of these findings are obvious. First, educators need to reconsider what "poor" performance means on these types of rhyming assessments. Evaluators should not just look at the number of items correct but should also look at the types of errors being made. Second, if educators insist on using these types of rhyming assessments with ELLs, these students need to be trained in the strategy of rhyming. López and Greenfield (2004) have suggested that this type of training needs to be done in their L1 in order to be effective. In addition, using visual images to test whether two words "sound" the same may be an adequate testing instrument, but when using them with ELLs, one must take into consideration the findings of this study and use visual images that do not have semantic associations or perhaps not use visual images at all. Lastly, further research should be carried out to investigate how these "poor" performers fare over time in terms of reading acquisition in order to determine whether these rhyming tests are even a valid measure of reading readiness for ELLs.

References

August, D., & Hakuta, K. (1993). *Federal education programs for limited-English proficient students: A blueprint for the second generation.* Stanford, CA: Stanford Working Group.

Craik, F. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimal age question and some other matters. *Working Papers on Bilingualism, 19,* 197–205.

Freeman, Y. S., & Freeman, D. E. (1997). *Teaching reading and writing in Spanish in the bilingual classroom.* Portsmouth, NH: Heinemann and Heinemann.

Goodman, K. (1996). *On reading.* Portsmouth, NH: Heinemann and Heinemann.

López, L. M., & Greenfield, D. B. (2004). The cross-language transfer of phonological skills of Hispanic head start children. *Bilingual Research Journal, 28,* 1–18.

New York State Education Department (2005). *School comprehension information report.* Retrieved March 3, 2003, from http://www.emsc.nysed.gov/repcrd2004/cir/280209030001.pdf

No Child Left Behind Act, Pub. L. No. 107–110 (2002).

Olivares, R. A. (2002). Communication, constructivism and transfer of knowledge in the education of bilingual learners. *International Journal for Bilingual Education and Bilingualism*, *5*(1), 4–19.

Ovando, C. J., & Collier, V. P. (1998). *Bilingual and ESL classrooms: Teaching in multicultural contexts.* Boston: McGraw Hill.

Phelps, S. (2001). *Implementing the early literacy profile through collaboration in-service.* Buffalo, NY: Buffalo State College, Technical Assistance Center.

Puckett, M. B., & Black, J. K. (1994). *Authentic assessment of the young: Celebrating development and learning.* New York: Macmillan.

Rivera, C., & Vincent, C. (1997). *High school graduation testing: Policies and practices in the assessment of English language learners.* Arlington, VA: The George Washington University, Center for Equity and Excellence in Education.

Shaughnessy, M. (1977). *Errors and expectations*. New York: Oxford University Press.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6* (3), 174–215.

U.S. Department of Education. (1992, February 14). *Characteristics and academic achievement: A look at Asian and Hispanic eighth graders in NELS:88.* Retrieved February 2, 2005, from http://nces.ed. gov/pubs92/92479.pdf

Vernon, S. A., & Ferreiro, E. (1999). Writing development: A neglected variable in the consideration of phonological awareness. *Harvard Educational Review*, *69*(4), 395–415.

Walton, P. D., & Walton, L. M. (2002). Beginning reading by teaching in rime analogy: Effects on phonological skills, letter-sound knowledge, working memory, and word-reading strategies. *Scientific Studies of Reading, 6*(1), 79–116.

Wood, C. (2000). Rhyme awareness, orthographic analogy use, phonemic awareness and reading: An examination of relationships. *Educational Psychology*, *20*(1), 5–15.

Endnote

[1] Although the actual Early Literacy Profile did not always contain a semantic associate as a choice, we found that there were semantically associated choices for a majority of test items.

Acknowledgement