



Investigating Vocabulary Growth through Total Physical Response Storytelling with Undergraduate Spanish Learners

Eva Jones and Bernard Issa

The University of Tennessee Knoxville

Learning an additional language is a complex process that places myriad demands on learners of any age, but is arguably more difficult for adult learners, resulting in more variable outcomes (see DeKeyser for a detailed discussion). Such variability for adult learners may appear across language skills such as reading, writing, speaking, listening and/or domains like morphosyntax, phonology, the lexicon, etc. Moreover, outcomes may vary by type of learning context; for example, naturalistic or instructed. Research with adult second language learners has shown that instruction has a positive effect on language development (e.g., see Norris and Ortega for a meta-analysis). As a result of these attested benefits, methods for and approaches to teaching additional languages abound, and so, further investigation into the effectiveness of the varied approaches to second language pedagogy is necessary.

Interestingly, there is a divide between approaches for primary and secondary school language learners, on the one hand, and university-level learners on the other. One approach that has taken second language teaching in primary and secondary school contexts by storm is Total Physical Response Storytelling or TPRS, also known as Teaching Proficiency through Reading and

Storytelling (Ray and Asher). This pedagogical approach, first introduced in the 1990s by Blaine Ray, is grounded in the use of the target language, or in other words, the language being learned, via storytelling that requires the constant participation and attention of all learners in the classroom (see the Literature Review for a more detailed description). Although the effects of TPRS have been primarily researched and compared to traditional instructional practices with language learners in the primary and secondary school context, relatively little research has examined the effects of TPRS with language learners in the university context.

A crucial aspect of communication is knowing and using vocabulary to make meaning. In order to promote deep learning of target vocabulary, learners need to acquire multiple dimensions of word knowledge such as the connection between the lexical item and its meaning(s), the spelling of the word, the ability to identify and produce the word in a given context (Brandl 77). Given the multi-faceted nature of developing word knowledge, learners are best positioned to acquire L2 vocabulary when they participate in learning activities that present vocabulary embedded in rich, elaborated, comprehensible input through meaningful and engaging content such as a story (Brandl 77). Indeed,

TPRS is designed to promote this type of learning and has been shown to be effective for younger learners' vocabulary development (e.g., Kara and Eveyik-Aydın; see Lichtman, "Research," for a review; Ray and Asher).

Given that TPRS has been shown to be successful with younger learners and high school-aged learners, additional research examining the effects of this method with learners of a slightly older age group, enrolled in a qualitatively different type of language program such as those in the university context, is needed to more fully understand the impact that this method has on additional language learning. Therefore, this research aims to examine the effectiveness of TPRS on one dimension of linguistic knowledge for adult second language learners in the university classroom: vocabulary development. Examining the effects of TPRS with university students can help deepen our understanding of best practices in teaching additional languages to adult learners and ultimately contribute to second language acquisition and pedagogy researchers' as well as instructors' knowledge about the processes that underlie language acquisition in adults.

1. LITERATURE REVIEW

Total Physical Response Storytelling or TPRS, as the name suggests, is rooted in principles of Total Physical Response or TPR. Developed by James Asher in the mid-60s, TPR was inspired by the process of first language (L1) acquisition. Asher noticed that many utterances by parents were followed by a physical reaction from their infant children. Building off this observation, TPR lessons included a series of verbal cues from the teacher that were then followed by a physical action from the students. As an example, the teacher may say "stand up" in the target language, while also standing up and prompting the students to do so as well. By

having the students recognize and understand the language before having to reproduce it, students can build upon their mental system of the target language by focusing on comprehension first.

While acknowledging the effectiveness of TPR as a teaching method, Blaine Ray observed the difficulty of being able to express more abstract ideas or concepts in the target language through TPR. This difficulty inspired the development of TPRS, and in 1990, Ray and Asher published the first Spanish curriculum with this method titled *Look I Can Talk!* TPRS continues to be a popular and fast-growing method in language teaching today (Lichtman, *Teaching Proficiency* 7). There are several key principles that anchor TPRS as a Communicative Language Teaching (CLT) based approach: (a) the necessity for students to receive a large amount of comprehensible input, (b) the idea that input should be structured so that target forms or vocabulary appear at a high frequency throughout the lesson, (c) the importance of target forms and vocabulary presented in the context in which they are commonly used as opposed to memorizing conjugation charts and (d) the importance of students working mostly on their comprehension skills before focusing on production skills (Lichtman, "Research"; Ray and Asher).

Many of these principles are closely tied to phenomena observed in L1 acquisition. For example, the importance of comprehensible input is parallel to the attested fact that young children gain thousands of uninterrupted hours of language input before beginning to utter words and phrases. As for presenting target forms with a high frequency in TPRS, a similar phenomenon happens in L1 acquisition; for instance, we can imagine a scenario where a guardian might wake their infant up every morning saying, "Good morning! Can you say, 'good morning'?" Additionally, in L1 contexts, much of the input is provided in context with a picture, song, or provided in an environment with

visual stimuli. As an example, a guardian might say “Look at the puppy!” while pointing at a puppy. This type of contextualized input is present in TPRS with the emphasis on presenting target forms and vocabulary in context through storytelling with visual support.

Another influential source for the development of TPRS comes from the fundamental theoretical work for the field of second language acquisition conducted by Stephen Krashen. In his Monitor Model, Krashen notes similarities between the development of our first language to the acquisition of additional languages stating that both “require meaningful interaction in the target language—natural communication—in which speakers are concerned not with the form of their utterances but with the messages they are conveying and understanding” (Krashen 1). This idea that acquisition takes place through exposure to input, along with the idea of *comprehensible input*, which is input that is just a step beyond the learners’ skill level, share core characteristics with the principles of TPRS. Additionally, Krashen’s theory on the *affective filter*, or potential anxiety or negative thoughts associated with language learning, can explain why some learners have exposure to a great amount of input and retain very little. It also can support the use of TPRS since the use of storytelling along with group callbacks to the instructor could potentially lower the affective filter of a nervous language learner (Lightbown and Spada 37). Together, these ideas are present in the principles underlying TPRS and can be used to support its use in second language classrooms.

Turning to the details of how TPRS is implemented, we describe a typical TPRS lesson, which consists of multiple phases, including: (a) an optional pre-story vocabulary/grammar phase, (b) a story phase, and (c) a post-story phase. In the pre-story phase, instructors might include activities or didactic materials designed to establish the meaning of certain target phrases that will be

present in the upcoming story phase (Lichtman, *Teaching Proficiency* 19; Ray and Asher VI).

The main focus of the lesson is the story phase. This phase consists of exposing learners to a story in a verbal and/or auditory modality that is replete with comprehensible input and specific target grammar and vocabulary. These stories can be developed in many ways including: using instructor-written materials, co-constructing a story with the students during the story phase itself, or using a premade story from a TPRS reader. Many of the TPRS storylines follow a similar outline in that there is a problem, a character going to many places to resolve this problem, and an eventual solution. In the story phase, the instructor speaks slowly and with an animated tone to emphasize certain words and to also increase entertainment and engagement of the learners while reading the story. Along with the responsibility of leading the students through the story, the instructor also provides additional input to learners as they read or listen to the story through a technique known as *circling*, which involves repeating comprehension questions, and key vocabulary items as a way to increase frequency of crucial input during this phase. Since learners are not taught grammar structures explicitly before beginning the story phase, if learners inquire about grammar structures they notice in the story, instructors employ a technique called *pop-up grammar*. This technique involves providing explicit instruction regarding grammar concepts as students become aware of them. Throughout the story, students might be asked to read, read along or simply listen to the story. Additionally, students are expected to answer and respond to short-answer comprehension questions and act out any physical gestures to engage further with the story (Lichtman, *Teaching Proficiency* 19; Ray and Asher VI).

In the post-story phase, students engage in debriefing and extension activities to en-

sure understanding of the story. These activities could include additional short-answer questions in a written or oral format, writing down part of the story, retelling the story, or part of the story, to a classmate, or drawing part(s) of the story. Taken together, the phases of a TPRS lesson are designed to target the principles outlined earlier, highlighting the importance of comprehensible input, using target vocabulary and/or grammar with a high frequency, presenting input in context and a focus on receptive language skills (Lichtman, *Teaching Proficiency* 19; Ray and Asher VI).

As a popular teaching method in primary and secondary educational contexts, TPRS has been investigated in second language acquisition research within that population of learners. Generally speaking, the findings indicate that TPRS can (a) lead to linguistic development, (b) improve class enjoyment, (c) help foster positive attitudes towards language learning and (d) promote further language study (see Lichtman, "Research," for a review of empirical research). In the following section, we focus on six empirical investigations that are relevant for the present study as they either targeted vocabulary learning with TPRS or examined the role of storytelling in the language classroom as a whole. Four of these studies included participants with college-aged learners or adults, which is the target age range of the present study.

2. RELEVANT EMPIRICAL STUDIES ON TPRS

In a recent study, Kara and Eveyik-Aydin examined the effects of TPRS on the vocabulary acquisition of nineteen four-year-old English L2 learners in Turkey over the course of seven weeks. This study included 28 unique and unfamiliar vocabulary items identified in children's story books. Each book contained at least two to three items of the target vocabulary. In addition to the story, the researcher used flashcards, props,

and videos depicting the target vocabulary which were used at different stages to facilitate comprehension as well as to keep the young learners' interest and motivation high. The learners' productive and receptive vocabulary knowledge was assessed before the treatment, throughout the treatment at the beginning and end of each week and after the treatment. The productive assessments consisted of the researcher displaying a picture depicting a target word and asking the learner to name it aloud. For the receptive tests, learners were shown three pictures, one target and two distractors. Then, they were asked to point to the picture that showed a certain vocabulary word.

The findings showed a short-term and a long-term effect of TPRS on the learners' L2 vocabulary acquisition. The immediate post-tests following each week's treatment showed significantly higher scores as compared to each week's pre-tests. This suggests a positive short-term effect on L2 receptive and productive vocabulary knowledge. The results also presented a long-term effect of the TPRS treatment on receptive knowledge as the mean scores of delayed post-tests after the seven-week treatment remained the same or increased as compared to the immediate post-tests. However, there was a slight decrease in the mean scores on the delayed post-tests for productive knowledge. Overall, the researcher concluded that TPRS is indeed an effective manner for increasing vocabulary acquisition in young learners, with stronger effects for receptive than productive knowledge.

In another relevant study by Watson, TPRS was compared to another traditional method in a secondary school Spanish classroom with participation from 73 high-school students with English as the L1. The traditional class was focused on conscious learning "with a great deal of the instructional time dedicated to helping students understand grammatical concepts through explanations in English" (Watson 21). Additionally, the traditional class used student-to-

student question and answer exercises, interviews, and games to identify and review vocabulary. Both the TPRS and traditional class had reading assignments outside of class. The traditional class read *Pobre Ana* by Blaine Ray, a TPRS reader, while the TPRS group read both *Pobre Ana* and *Patricia va a California* by Blaine Ray as class readers. Of the 73 participants, twenty-three were enrolled in the traditional class and fifty were enrolled in two sections of a TPRS class. All three classes were taught by the same instructor.

Vocabulary knowledge assessment for this study included a final examination and an oral examination at the end of the semester. The final examination consisted of a listening section that included a true or false answer section based on the audio, a vocabulary and grammar section that included fill in the blank sentences, and a reading section where students were asked to: (a) read sentences and decide if they were probable or improbable and (b) read two passages and answer comprehension questions. For the oral assessment, the author noted, "Students were asked to randomly choose a card with an English word and explain it in two minutes in the target language" (Watson 22). The results of the study showed that the TPRS groups performed nearly identically to each other on both the final examination and the oral exam. The TPRS students outperformed the traditional class, scoring one standard deviation higher than the traditional class in combined scores on the two assessments, suggesting that TPRS had a positive effect on learning.

De Costa's study examined the progress of twenty university students with English as the L1 who were enrolled in a French immersion class. Participants were randomly assigned to either an Experimental group or a Control group. The Experimental group was given TPRS treatment, and the Control group was given a teaching method that "does not incorporate a story context within the lesson" (De Costa 4). All participants received the same pre-test in order to assess

their pre-existing language abilities in writing, listening, and speaking in French. This same pre-test was used as a post-test to determine any changes in their abilities. The study took place over a five-day period. The materials for this study included a vocabulary list of 44 words.

Although the results did not reveal a significant difference between groups, descriptive differences were evident in vocabulary and grammar. The Experimental group made bigger improvements than the Control group in the vocabulary measure. For grammar, the opposite pattern was found in that the Control group made larger descriptive gains than the Experimental group. Despite the lack of significant findings and the contradictory pattern of results, the author notes that there were in fact benefits of TPRS for learners in the study, stating, "It engaged all students in the classroom...and because of daily interaction, the teacher could verify the progress of each student in terms of vocabulary, culture, listening, and comprehension" (De Costa 47). De Costa attributed the improvement in vocabulary knowledge to the repetition inherent in TPRS, estimating that the vocabulary words were repeated upwards of at least 50 times by the researcher and the learners.

In another study with TPRS in the university context, Bustamante examined its effectiveness in a beginner Spanish course at the university level with 19 participants. Fourteen of the participants had English as an L1 and four participants had an L1 of Japanese. The participants were assessed on reading, writing, and were also given the Computer Adaptive Placement Exam (CAPE) exam. Finally, the participants were given a survey that gauged their perception of TPRS and their own learning experience. The study took place over the course of a semester, where participants self-selected to participate by enrolling in the course. The researcher used an eight-chapter textbook specifically designed for a TPRS course, *¡Cuén-*

tame más! (Gaab). The course outline included “four mini-stories, one long story, and one song per chapter, vocabulary and reading comprehension quizzes, a mid-term and a final exam, and two additional activities” (Bustamante 32). Testing measures were given to the participants at several points throughout the semester in order to determine significance. The results of this study were compared to a traditional beginner Spanish course from a previous semester at the same university.

The results of the reading, writing and CAPE language placement tests showed a statistically significant increase in the TPRS course participants’ performance on all measures from the beginning to the end of the semester over the comparison group. Qualitatively, a survey of the participants showed that more than half of the TPRS class felt satisfied with the new learning method. Furthermore, 12 of the 19 participants had taken a Spanish course before and cited that they felt like they were learning Spanish better with the TPRS method.

Pinos-Ortiz and Orbe Guaraca examined the effectiveness of TPRS for vocabulary acquisition in an English university classroom in Ecuador over a five-month period. Additionally, these researchers wanted to investigate the learners’ perception of TPRS’s use in the classroom and its effect on their vocabulary knowledge. The participants were 56 L1 Spanish/L2 English learners who worked with TPRS in their classroom for vocabulary acquisition each day throughout the course of the study, with 30 participants in the experimental group who received the TPRS treatment and 26 in the control group. Pre- and post-testing measures were obtained through the Cambridge vocabulary test, which measures language skills and vocabulary in use. Participants in the experimental group were also given a survey to obtain data on their perspectives of TPRS throughout the process and the perceived overall effect of TPRS on their vocabulary knowledge.

Results of a t-test showed that the experimental TPRS group displayed significantly greater increases across the vocabulary pre- and post-test measures. Additionally, a Likert scale survey of the experimental group participants showed that the experimental group perceived TPRS to be more effective for learning vocabulary than traditional methods, with high levels of overall enjoyment during the process of learning through TPRS.

In order to explore the role of storytelling, a key component of TPRS, in a Chinese second language setting, Nguyen, Stanley and Stanley surveyed 15 instructors and 30 adult learners at the School of Chinese Language in Shaanxi Normal University in Xi’an, China. The survey asked instructors and learners about their “interests, the practice, benefits, and challenges of doing storytelling in the Chinese as a second or foreign language classroom” (Nguyen, Stanley and Stanley 1). The results of the survey showed that participants were interested in storytelling because of its perceived benefits in language learning and comprehension of language. Additionally, participants were interested in its effectiveness for multi-cultural understanding.

In looking at the importance of these six aforementioned studies, they are key in showing the effectiveness of TPRS for vocabulary and grammar development in classrooms of learners of all ages. Kara’s and Eveyik-Ayдын’s study conducted with four-year-old language learners suggests that there are both short-term and long-term effects of TPRS for both receptive and productive vocabulary knowledge. In additional work with a slightly older group of learners, Watson’s study showed that learners can make gains over the course of a semester, highlighting that prolonged exposure to TPRS is also beneficial. While De Costa’s five-day study of TPRS in comparison to a traditional method in a university setting did not prove to have significant overall results in favor of either method, the results did

show that the TPRS group made gains over the Control group in the vocabulary measures on post-testing. Bustamante's semester-long study in the university setting showed relevant significant results as it pertains to the present study and TPRS's usefulness in the university context. Despite its short time frame of five days, Pinos-Ortiz's and Orbe Guaraca's study examining TPRS's effect on vocabulary acquisition in a second language with university-aged learners displayed not only significant gains as compared to the control group, but the results of the qualitative portion of the study also showed that learners were greatly enjoying their experience with TPRS as compared to a traditional language course. Lastly, we examined the qualitative study by Nguyen, Stanley and Stanley investigating the role of storytelling in the second language classroom, which are relevant to the present study insofar as they highlight the enjoyment and potential engagement with TPRS for adult learners.

Motivation for the present exploratory study is rooted in a relative dearth of research on TPRS with undergraduate learners. As noted above we found a total of four studies that have included this population. Results from prior work are promising, and as such we take an additional step in contributing to the ongoing research on TPRS by focusing on undergraduate learners and their development of novel vocabulary items. The current project aims to address the following research questions and hypotheses.

3. RESEARCH QUESTIONS

The research questions that guide this project are:

RQ1: What are the effects of TPRS on the identification of vocabulary words as compared to a more traditional instructional approach for L2 Spanish learners in a university learning context?

RQ2: What are the effects of TPRS on the production of vocabulary words as compared to traditional instruction for L2 Spanish learners in a university learning context?

4. METHODS AND MATERIALS

4.1 Participants

Forty-nine participants were recruited for this study from elementary Spanish classes at a large university in the Southeastern US. Participants were recruited through an email forwarded from their instructors. Only those who agreed to participate and signed the necessary IRB-approved consent forms took part in this research study. Three participants did not complete all sessions of the study and three participants scored above our vocabulary knowledge pre-test cutoff score, which was more than two standard deviations from the mean. Thus, the final sample included 43 participants: 32 female, 10 male and 1 transgender student with ages ranging from 18 to 48. The L1 of all participants was English, and all participants reported on average one year of classroom exposure to Spanish. Participants were randomly divided into two groups: Experimental and Control. Twenty-four participants were assigned to the Experimental group and 19 participants were assigned to the Control group. All required IRB protocols were followed for recruitment of participants, consent, and data storage throughout this study.

4.2 Materials

In order to assess the participants' existing knowledge of targeted vocabulary items and document any changes in this knowledge after the intervention, which is described below, identical pre-test and post-test measures were used for both the Experimental and the Control groups. These

measures consisted of two parts. The first assessed the participants' ability to *identify* the vocabulary. This identification task consisted of a Google form that contained multiple choice items. For each item, a target vocabulary word in Spanish was given and participants were asked to choose the correct English translation equivalent from three options (see Appendices A and C). For the second measure, which was a pencil and paper task, we assessed the participants' ability to *produce* the target words by showing a picture projected as a slideshow. The participants then had to write the word in Spanish that represented the photo (see Appendix B) on a piece of paper. For example, when viewing a picture of a tree, the participant was expected to write "el árbol." Both the production and the identification assessment were graded out of 16, receiving 1 point for each correct response. For the production task, leeway was given to misspellings if the researchers could tell that the student knew the word and either left out an accent mark or were off by a letter or two. Importantly, the same list of sixteen target vocabulary words was used for both the Experimental and Control group (see Appendix C).

4.3 Procedures

This study took place over two sessions. The first session included the reading and signing of the Informed Consent documents, initial pre-test measures, treatment, and immediate post-testing measures and lasted approximately one hour in duration. The second session took place approximately one week later and included the delayed post-testing measures as well as the Language Background questionnaire and took less than 30 minutes. The experimental and control groups were given different materials in the first session for the treatment, all other measures given to the two groups were the identical.

4.3.1 Experimental Group

Participants in the Experimental group were instructed via TPRS. The participants were put into small groups of 3-5 for each session of the experiment. For the first session, once the participants arrived at the classroom where the study took place, they were given a hard copy of the informed consent. Participants were given time to read over the informed consent form and sign it. Then, participants completed both the identification and the production pre-tests.

After pre-testing was complete, the participants listened to the TPRS story that was written and told by the researcher. The title of the short story was "*Harry Styles va a las vacaciones*" (Harry Styles goes on vacation). Before beginning the story, the researcher informed the participants that they were expected to: a) remain engaged, b) repeat words when prompted, and c) answer questions aloud as a group when prompted. A PowerPoint with pictures was used to illustrate the story throughout its telling. For example, at a point in the story when the protagonist was visiting a beach, the PowerPoint contained a picture of the protagonist and a beach to illustrate the scene. Aligning with TPRS models as highlighted above, the researcher told the story very slowly, including repetition from the researcher so that the participants were able to hear correct pronunciation of the target vocabulary. The participants were able to practice their own pronunciation when prompted to repeat target vocabulary or answer questions. An example of this practice in context is shown in the following script:

Researcher: "*Clase, en el cuento, hay un chico. El chico se llama Harry Styles. Clase, ¿Quién es el chico?*" ("Class, in the story there is a boy. The boy's name is Harry Styles. Class, who is the boy?")

Participants: "*¡Harry Styles!*"

Researcher: “Clase, ¡Harry Styles va a la playa! Clase, ¿adónde va Harry Styles?” (“Class, Harry Styles goes to the beach! Class, where does Harry Styles go?”)

Participants: “¡La playa!” (“the beach!”)

The aim of asking the participants questions during the story was to keep them engaged, and it allowed them to be active in the story by repeating target vocabulary up to 8 times throughout the exercise. As another strategy for keeping the students engaged, the researcher used gestures and animated facial expressions to provide physical representations of many parts of the story that might have been difficult for learners to understand. Examples of this included smiling and making a thumbs up when Harry Styles was happy at the end or pretending to trip when Harry Styles tripped on a tree root in the story. The telling of the entire story took about 15 minutes.

After the completion of the story, the researcher handed each student a blank sheet of paper and then called out 5 of the target vocabulary words from the story. The participants were expected to draw the target vocabulary words as they heard them on the blank sheet of paper. For example, the researcher said, “la flor” (“flower”) and the participants would draw a flower. This drawing activity took approximately 5 minutes, for a total of approximately twenty minutes of learning activities for the Experimental group. After the completion of the drawing activity, participants then completed the immediate post-tests which are identical to the identification and production pre-tests given before the treatment.

Approximately one week later, participants returned to take the delayed post-tests which were, again, identical to the pre- and post-tests completed during the first session. Participants also completed a language background questionnaire in order to obtain a better idea of their language learning background and experience with their L1 and L2

and any other languages they might have known.

4.3.2 Control Group

The Control group completed the same pre-treatment steps in session 1 as the Experimental group, which included: reading through and signing the informed consent and completing the identification and production pre-tests. The Control group participants were also put into small groups of 3-5 participants. Participants were then given a list of the target vocabulary in both Spanish and English (see Appendix C). The researcher gave participants a few minutes to look over and study the words, as if they were studying for a vocabulary test.

Next, the researcher had the participants repeat each of the vocabulary words after the researcher read them aloud. The researcher then asked the participants to put the list away while they completed a matching activity, requiring them to match Spanish words and English translation equivalents, that included all of the target vocabulary (see Appendix D). After reviewing the answers to this activity as a group, participants were provided with a fill-in-the-blank activity that allowed them to use the words in context (see Appendix E). Participants were allowed to bring out their bilingual list for this activity, to preserve authenticity of a classroom environment where students often have their books available as a resource. After the researcher went over the answers to the fill-in-the-blank activity with the participants, the participants were then given an additional 2-3 minutes to study the words. The activities chosen for the Control group were meant to replicate those typically found in a language textbook. In total, activities completed for the Control group’s treatment took approximately twenty minutes.

After the completion of the Control group’s treatment, participants completed the immediate identification and production post-tests, which were identical to the pre-

tests. In a parallel fashion to the Experimental group, approximately one week later, participants returned to complete the delayed post-tests and the language background questionnaire.

5. RESULTS

5.1 Identification test

This study aimed to examine the effects of TPRS on the identification and production of vocabulary for university-aged second language students. Descriptive results including means and standard deviations for vocabulary identification are displayed in Table 1 for each group at the Pre-test, Post-test, and Delayed Post-test. As seen in Table 1, both groups make large descriptive gains. In order to examine whether changes over time and between groups were statistically significant, we conducted a repeated measures analysis of variance (ANOVA) with Group (Control or Experimental) as a between-subjects' factor and Time (Pre-,

Post-, or Delayed Post-test) as a within-subjects factor for both the identification and production tasks. Effect sizes were calculated as Cohen's d , which represent the magnitude of the difference between scores on the different tests used here.

For the identification tasks, there was a significant main effect for Time ($p < .0001$), indicating that both groups made significant gains over time. Follow-up parameters estimate tests of significance revealed that both groups made significant gains from pre-test to post-test ($p < .0001$, $d = 1.18$) and pre-test to delayed post-test ($p < .0001$, $d = 0.66$), as indicated in Table 2. The effect sizes we report above indicated a medium to large effect for gains from pre-test to post-test and a small effect for pre-test to delayed post-test, which were interpreted following Plonsky and Oswald's field-specific guide for effect size interpretation. Finally, there was no statistically significant main effect for Group ($p = .07$) and no interaction between Time and Group ($p = .55$), suggesting that both groups performed similarly at all time points.

Table 1

Means, Standard Deviations within groups for Identification task

	Pre-test		Post-test		Delayed post-test	
	M	SD	M	SD	M	SD
Control	.65	.11	.99	.02	.97	.04
Experimental	.64	.12	.96	.03	.92	.07

Table 2

Changes in Means and Standard Deviations within groups for Identification task

	Pre-test to post		Pre-test to delayed		Post-test to delayed	
	M	SD	M	SD	M	SD
Control	.34*	.11	.32*	.10	.02	.05
Experimental	.32*	.21	.28*	.16	-.04	.08

* denotes statistical significance at $p < .0001$

5.2 Production test

Descriptive results from the production test are displayed in Table 3. These data suggest that both groups made large descriptive gains from pre-test to immediate and delayed post-tests.

For the production task, results from a repeated measures ANOVA, with Group as a between-subjects factor and Time as a within-subjects factor, yielded a statistically significant main effect for Time ($p < .0001$), and as revealed by post-hoc parameter estimate tests of significance, both groups made significant gains, with a small effect size, from pre-test to immediate post-test ($p < .0001$, $d = 0.66$) and from pre-test to delayed post-test, with a negligible effect size ($p < .0001$, $d = 0.17$) as shown in Table 4. Additionally, there was a significant main effect for Group ($p = 0.04$), indicating that overall, from the initial pre-test to the delayed post-test, the Control group performed better

than the Experimental group. However, an exploratory post-hoc paired samples t-test revealed that scores significantly increased for the Experimental group from immediate to delayed post-tests with a small effect size ($p = .006$, $d = .27$) whereas the Control group did not improve ($p = .68$, $d = .11$). Lastly, there was no interaction between Time and Group ($p = 0.58$), which indicates that the scores for each group at each individual time were not different from one another. Taken together, the identification and production test results suggest that both groups of learners made gains in *identification* because of their respective treatments, however, the Experimental group's ability to *produce* the vocabulary items increased over time whereas the Control group's productive ability remained the same after one week without instruction.

Table 3

Means, Standard Deviations within groups for Production task

	Pre-test		Post-test		Delayed post-test	
	M	SD	M	SD	M	SD
Control	0.35	0.23	0.85	0.14	0.84	0.17
Experimental	0.29	0.19	0.74	0.19	0.81	0.18

Table 4

Changes in Means and Standard Deviations within groups for Production task

	Pre-test to post test		Pre-test to delayed		Post test to delayed	
	M	SD	M	SD	M	SD
Control	0.5*	0.20	0.49*	0.20	0.01	0.10
Experimental	0.45*	.13	0.52*	0.17	0.07*	0.12

* denotes statistical significance at $p < .0001$

6. DISCUSSION

This study aimed to assess the effectiveness of TPRS as a method for teaching vocabulary to novice learners of Spanish at the undergraduate level. Results generally confirm that TPRS promotes vocabulary learning insofar as learners in the TPRS group made significant gains in their ability to both perceive

and produce target vocabulary items. Interestingly, the Control group also performed well, and at least on the production task, performed significantly better than the Experimental group.

In the receptive dimension of vocabulary learning, scores on our identification test revealed that the Control and Experimental group essentially performed the same on the

pre-test and scored similarly on the immediate post-test and delayed post-tests. This was evidenced by a lack of significant main effect for Group and no interaction between Time and Group. The results pertaining to identification are surprising insofar as the Experimental group did not outperform the Control group on this measure. These findings might be related to the type of assessment used for our post-tests. Learners in the Control group completed learning tasks that in some ways were more similar to the post-tests than the type of tasks that the learners in the Experimental group completed, which might explain their performance. That is, the learners in the experimental group completed a drawing task, and the control group completed a fill-in-the-blank task. This phenomenon is known as Transfer Appropriate Processing or TAP (Craik and Lockhart; Lightbown and Han), which states that learning tasks that are similar in nature to assessment tasks will allow for better performance on said assessments. In our study, the Control group completed written tasks with a focus on translation equivalents and so they were able to transfer that experience to their performance on the post-tests administered, which shared a similar format. If TAP were a plausible explanation for the Control group's performance, it could be considered as an advantage for the Control group and a disadvantage for the Experimental group. However, such a disadvantage for the Experimental group highlights the strengths of TPRS as a method, because even though learners could not transfer their experience directly to the identification task, they still performed well and made significant gains overall.

Turning to results for the Production task, one comparison stands out: the significant increase in productive ability from post-test to delayed post-test for the Experimental group only. We can theorize that the TPRS story and the learning tasks resonated with the students more than the lesson received and tasks completed by the Control group

due to its stark difference from a traditional language class. This sort of resonance may have promoted a deeper level of processing in learners. Depth of processing refers to the amount of cognitive effort, attentional focus and manipulation of learning material deployed by a learner (see Leow and Mercer for an overview), which ultimately could have led to better retention of the vocabulary albeit with a slower emergence as we saw on the delayed post-test. Taking a closer look at the production scores, it is interesting to note that the Control group seemed to plateau after the immediate post-test, as evidenced by the very similar mean scores between the immediate and delayed post-tests (see Table 4). On the other hand, the Experimental group improved significantly from the immediate post-test to the delayed post-test.

In relation to the prior research on TPRS, there are many observations that can be made. Kara and Eveyik-Aydin concluded during their seven-week-long study in which learners were assessed through a pre- and post-test, as well as weekly testing, that TPRS showed positive short-term effects on receptive and productive knowledge and positive long-term effects on receptive knowledge. The present study did not test for long-term effects; however, we did find similar results that suggest that the Experimental group experienced short-term effects from pre-test to post-test and then a week later on the delayed post-test. These differences may be explained by differences in study design and the learner populations. First, the participants in Kara's and Eveyik-Aydin's study were much younger than the participants of the present study, and so, their capacity to benefit from implicit learning was greater than our older learners (e.g., Lichtman, "Research on TPR Storytelling"). Second, this study was conducted over the course of seven weeks, while the present study was conducted in two sessions that took place one week apart.

Watson's study also presented differing results than the present study, showing the

TPRS group outperforming the traditional class by one standard deviation, whereas we found that our two groups generally performed similarly. Again, there were differences in study design, Watson's being a full semester and our study, short-term with only two sessions, as well as age of learners. Watson's study examined secondary school learners while the present study examined university learners. Given the discrepant results, we suggest that further investigation into the effects of TPRS with varying treatment lengths is necessary.

Turning to De Costa's study, which also investigated university learners in a shorter time frame, results are strikingly parallel. For vocabulary, De Costa saw an increase for the TPRS group over the Control group, whereas the Control group outperformed the TPRS group in the grammar category. Similarly, we found that the Control group outperformed our TPRS group in some dimensions of vocabulary knowledge whereas the Experimental group performed better in other areas. De Costa attributes the Control group's success to having received more explicit instruction on the target grammar structures. We have argued above that TAP may account for the comparable findings in our study.

Looking at the studies by Bustamante and by Pinos-Ortiz and Orbe Guaraca, who both examined TPRS in a second language classroom in the university context, we are encouraged by their results in comparison to our own. Both studies looked at an aspect of vocabulary acquisition in the university classroom that we were also interested in examining during our study, however ours on a much shorter time frame than that of these researchers. Both of these studies showed significant gains in favor of TPRS aiding in vocabulary acquisition. While our study did show significant results for the Experimental group, results were not as definitive as the results that are shown in the studies by Bustamante or by Pinos-Ortiz and Orbe Guaraca, which were conducted over several

months. As we will state more thoroughly in the Limitations and Recommendations for Future Research section, we believe that more time would have greatly benefited our participants and thus could have made our results more definitive in supporting TPRS as a valuable tool in the university classroom.

Overall, we believe that the findings from this study highlight the effectiveness of TPRS as a technique within Communicative Language Teaching for promoting vocabulary development. Our learners showed significant improvement in their receptive and productive abilities with the key vocabulary used in our TPRS story. Although the Experimental group did not outperform the Control group, they made progress over the course of the study that should not be overlooked.

7. LIMITATIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Regarding the limitations of this study and our recommendations for future research, we believe it would have been beneficial to make the learning session longer. One way to implement this would be to collect data with learners over the course of several weeks in a typical classroom setting (see Watson, Bustamante, or Pinos-Ortiz and Orbe Guaraca). This type of design potentially could have revealed more robust effects for the Experimental group, as TPRS is a more implicit approach to teaching and as such, may require more time for effects to emerge. Additionally, a longer-term design would allow us to probe the durability of learning gains for both groups over time.

Moreover, we think it is important to note that our study was implemented with only one instructor. Although we chose to have one instructor administer the learning sessions for consistency and control across administrations, we acknowledge that having multiple instructors implement a teaching method has more ecological validity,

representing real world classroom environments more faithfully and as such may have yielded different results. Every instructor teaches in a different manner, even while using TPRS, as there are numerous ways to implement it, and therefore, its effectiveness may vary depending on the instructor and audience.

An additional limitation of this study, as mentioned in the Discussion, is the nature of the assessments and their similarities with the Control group's learning intervention, which may have led to the Control group's outsized performance. We suggest that future research employ multiple types of tasks for each dimension of language knowledge tested that align more closely with the types of learning activities completed by both groups. As an example, in the present study, we only used one measure for each identification and production assessment. For example, one additional way to assess the Experimental group's identification abilities, besides selecting translation equivalents, would be to include a drawing-based assessment or an assessment in which participants choose an image that corresponds to a target word, either of which would be similar to what they completed in the learning intervention. These modifications would adhere to the principle of TAP and therefore put both groups on a more even playing field.

Another recommendation that we offer for future research into TPRS in the university language classroom is based on our observation of student enjoyment in the TPRS group during the learning session of the experiment. During the course of the study, we noticed that the learners in the Experimental group were more engaged and were more actively participatory in the lesson as compared to the Control group, which we believe could be an additional benefit of using TPRS in the university classroom. Anecdotally, at the delayed post-test session, participants in the Experimental group inquired about the possibility of hearing more of the TPRS story, indicating that they enjoyed the

prior learning session centered around the story. In order to assess potential enjoyment and perspectives about participation during the experiment, in a future study, we recommend including a debriefing questionnaire or interview, similar to measures implemented by Nguyen, Stanley and Stanley and by Pinos-Ortiz and Orbe Guaraca in their aforementioned studies. This process could also provide insight into factors such as motivation, anxiety, or desired continuation of learning a second language. Overall, this recommendation could further elucidate the effectiveness of TPRS in promoting learning an additional language among learners in this age group. Furthermore, given our results and these observations, we endorse using TPRS in the university level language classroom as a way to promote comprehensible input and provide instructors with a unique approach to vocabulary presentation, which we believe ultimately provides learners with an opportunity for deeper learning.

8. CONCLUSION

Our aim for this study was to assess the effectiveness of TPRS as a tool to enhance identification and production of vocabulary for university-level learners of Spanish. Results showed that both groups made substantial and significant gains from the pre-test to the post-test, and that those gains were maintained at the delayed post-test one week later. Additionally, the experimental group showed significantly higher scores at the delayed post-test as compared to the immediate, whereas the control group did not. These results have implications for teaching, as they highlight the utility of TPRS for learners and support our claim that TPRS can promote vocabulary growth in the university second language classroom as a pedagogical tool under the CLT approach. Additionally, our project leaves abundant and necessary room for further research to be performed in terms of grammar, culture, and

the implementation of a longer-term longitudinal design. This could provide educators with the opportunity to do their own research and experimentation with TPRS as a teaching method with their own students.

Lastly, we hope that the present study provides inspiration to other university-level instructors to explore methods that go beyond conventional textbook material.

Appendix A

Identification Test

La naturaleza *

natural

nature

grass

el árbol *

tree

air

cloud

Appendix B

Pictures Used for Production Test



Appendix C

Bilingual List for Control Group and List of Target Vocabulary Words

La Naturaleza	Nature
El árbol	Tree
La flor	Flower
Las raíces	Roots
El campo	Countryside
La Playa	Beach
La cueva	Cave
El acantilado	Cliff
El desierto	Desert
El lago	Lake
El bosque	Forest
Las hojas	Leaves
El mar	Ocean
La cascada	Waterfall
El musgo	Moss
Las montañas	Mountains

Appendix D

Matching Activity for Control Group

1. El musgo	Mountains
2. La cueva	Nature
3. El árbol	Moss
4. Las hojas	Tree
5. Las montañas	Cave
6. El acantilado	Waterfall
7. La flor	Beach
8. El desierto	Lake
9. El campo	Ocean
10. El bosque	Cliff
11. La naturaleza	Leaves

Appendix E

Fill in the Blank Activity for Control Group

1. Los árboles tienen _____ y _____.
2. _____ de Smokies están muy cerca de Knoxville.
3. A los estudiantes les gusta ir a _____ para el descanso de primavera.
4. Hace mucho calor en _____.
5. _____ favorita de mi mamá es una rosa.
6. Los murciélagos (bats) viven en _____.
7. _____ de Niagara es la más alta en el mundo.
8. A mis amigos les gusta ir en un barco (boat) en _____ en Tennessee durante el verano.
9. Muchos conejos (rabbits) y pájaros (birds) viven en _____.
10. Mi gato escala (climbs) _____.

Works Cited

- Brandl, Klaus. *Communicative Language Teaching in Action: Putting Principles to Work*. Pearson Prentice Hall, 2021.
- Bustamante, Maria Carolina. "Measuring the Effectiveness of a TPRS Pilot Course in Spanish at the 100 College Level." 2009. U Nebraska at Kearney, Master's Thesis.
- Craik, Fergus I.M., and Robert S. Lockhart. "Levels of Processing: A Framework for Memory Research." *Journal of Verbal Learning and Verbal Behavior*, vol. 11, no. 6, 1972, pp. 671-84.
- De Costa, Rishani Samari Merinnage. "Traditional Methods Versus TPRS: Effects on Introductory French Students at a Medium-sized Public University in the Midwestern United States." 2015. U Minnesota Mankato, Master's Thesis.
- Dekeyser, Robert. "Age Effects in Second Language Learning." *The Routledge Handbook of Second Language Acquisition*, edited by Susan Gass and Alison Mackey. Routledge, 2012.
- Gaab, C. *Cuéntame más*. TPRS Publishing, Inc., 2005.
- Kara, Ketevan, and Evrim Eveyik-Ayдын. "Effects of TPRS on Very Young Learners' Vocabulary Acquisition." *Advances in Language and Literary Studies*, vol. 10, no. 1, 2019, pp. 135-46.
- Krashen, Stephen D. "Principles and Practice." *Learning*, vol. 46, no. 2, 1982, pp. 327-69.
- Leow, Ronald P., and Johnathan D. Mercer. "Depth of Processing in L2 Learning: Theory, Research, and Pedagogy." *Journal of Spanish Language Teaching*, vol. 2, no. 1, 2015, pp. 69-82.
- Lichtman, Karen. "Research on TPR storytelling." *Fluency Through TPR Storytelling*, 7th edition. Edited by Blaine Ray and Contee Seely. Command Performance, 2015, pp. 364-79.
- . *Teaching Proficiency Through Reading and Storytelling (TPRS): An Input-based Approach to Second Language Instruction*. Routledge, 2018.
- Lightbown, Patsy M., and Z. Han. "Transfer Appropriate Processing as a Model for Classroom Second Language Acquisition." *Understanding Second Language Process*, vol. 27, 2008, pp. 27-44.
- Lightbown, Patsy M., and Nina Spada. *How Languages are Learned*, 5th ed. Oxford UP, 2021.
- Nguyen, Kate, Nile Stanley, and Laurel Stanley. "Storytelling in Teaching Chinese as a Second/Foreign Language." *Linguistics and Literature Studies*, vol. 2, no. 1, 2014, pp. 29-38.
- Norris, John M., and Lourdes Ortega. "Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-Analysis." *Language Learning*, vol. 50, no. 3, 2000, pp. 417-528.
- Ortiz, Mónica A. Pinos, and Mariory P. Orbe Guaraca. "The Effects of the TPRS Method on the Students' English Vocabulary Acquisition." *Dominio de las Ciencias*, vol. 4, no. 3, 2018, pp. 264-77.
- Plonsky, Luke, and Frederick L. Oswald. "How Big is 'Big'? Interpreting Effect Sizes in L2 Research." *Language Learning*, vol. 64, no. 4, 2014, pp. 878-912.
- Ray, Blaine, and Asher, J.J. *Look, I Can Talk!: Teacher's Guidebook*. Sky Oaks, 1990.
- Ray, Blaine, and Contee Seely. *Pobre Ana*. Continental Book Company, 1999.
- Ray, Blaine. *Patricia va a California*. Command Performance, 2001.
- Watson, B. "A Comparison of TPRS and Traditional Foreign Language Instruction at the High School Level." *International Journal of Foreign Language Teaching*, vol. 5, no. 1, 2009, pp. 21-24.