THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF COMMUNICATION SCIENCES AND DISORDERS

Predictors of English Verbal Fluency in Mexican Spanish-English Bilinguals and English Monolinguals

RAEGAN HOLZHUTER
SPRING 2024

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Communication Sciences and Disorders
with honors in Communication Sciences and Disorders

Reviewed and approved* by the following:

Dr. Carol Miller
Professor of Communication Sciences and Disorders
Thesis Supervisor

Dr. Carol Miller
Professor of Communication Sciences and Disorders
Honors Adviser

Dr. Krista Wilkinson
Professor of Communication Sciences and Disorders
Faculty Reader

* Electronic approvals are on file.

ABSTRACT

Verbal fluency is a task requiring speakers to produce as many words as possible from a given category within a limited time. This task is widely used – including in research on bilingualism – to understand how lexical retrieval works (Shao et al., 2014). We investigated semantic verbal fluency performance in L2 English speakers as well as monolingual English speakers. Forty-nine participants (23 bilinguals and 26 monolinguals) completed four verbal fluency trials: tools, clothing, animals, and musical instruments. Analyses examined specific category performance, cognate use, cluster use, and maximum cluster size among the clothing, animals, and instruments categories. Correlations were derived to examine predictors of verbal fluency performance. Results indicated that cognate and cluster use do predict higher verbal fluency performance in each group, as well as larger cluster size in the clothing and musical instruments categories. These findings will assist in improving norming data for bilingual testing materials. Future research is needed to expand and generalize these results to a larger population.

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my thesis and honors adviser, Dr. Carol Miller. Your support and guidance throughout my undergraduate career have had an immeasurable impact on my goals and achievements. Thank you for the countless opportunities and valuable mentorship that you have given me -- from the Child Language Lab to PIRE -- and for the endless inspiration that you have instilled for research in my life; I am so grateful to have been a part of such an enriching learning experience. I could not have completed my personal and educational goals without you.

This thesis was also made possible by Dr. Arias-Trejo and the members of the Laboratorio de Psicolingüística at the Universidad Nacional Autónoma de México. Thank you for welcoming me into your space and allowing me the resources and support to complete this study. This thesis would not have been possible without your hospitality and counsel during my time abroad.

I would also like to thank Dr. Krista Wilkinson both for being faculty reader for this thesis as well as being a mentor and motivator during my time at Penn State. Thank you for all that you have done for me and all that you continue to do for this community. Your dedication to this field and to your students is inspirational.

Chapter 1

Introduction

Currently, diagnostic tests of language disorders for bilingual individuals – including Spanish-English bilinguals – are inadequate in distinguishing a language difference from a language disorder. Hemsley et al. (2014) describes this issue, discussing how most speech-language pathologists are monolingual English speakers. This provides a challenge for assessing bilingual children, as an interpreter or bilingual SLP is required to adequately evaluate these clients. Additionally, even when bilingual SLPs are available, a lack of standardized tests that include bilingual children in the norming sample causes language differences from second-language learning to be misdiagnosed as disorders. According to Bedore and Peña (2008, p. 19), these tests developed for languages other than English are "almost always normed on monolingual groups". This causes bilingual children to be wrongfully assessed as not knowing words in their second language although they were not tested in their native language to verify familiarity with these concepts. Not only does this pose an issue for children who are diagnosed for services that they do not need, but also for those who require services but are overlooked during evaluation from these testing materials.

This same problem impacts adults, which provides an avenue to test both populations and determine how language differences manifest across the lifespan so that better norming samples, testing materials, and developmental milestones may be established. Verbal fluency is one task that allows insight into how bilingual individuals process language and – when compared to

monolingual performance – reveals important language differences that can assist in the development of these materials.

During the verbal fluency task, participants are given 60 seconds to name as many items in a category as possible (e.g. kitchen items, colors, etc.) and are scored based on how many of those items are unique and accurate to the category. Shao et al. (2014, p. 1) explains that verbal fluency helps to "measure verbal ability including lexical knowledge and lexical retrieval ability" which has an impact on language production and control. With bilingual participants completing this task in their second language, language experience has a significant impact on word retrieval ability. This is because individuals that have not learned the relevant vocabulary for a given category will score poorly compared to their monolingual peers.

There are multiple components of verbal fluency results that can be analyzed to further reveal details about participants' lexical access. One example of this is the use of and reliance on cognates for bilingual participants during word production. Cognates are "a pair of words that closely resemble each other in phonological form and have the same meaning" (Davis et al. 2018, p. S14). Blumfeld et al. (2016) suggests that the use of cognates is thought to aid participants in word retrieval as these English words have similar pronunciations to their native Spanish counterparts. This allows less cognitive effort while searching for relevant words in each of the verbal fluency categories.

Another component of verbal fluency performance is participant use of clustering and the size of the clusters produced. Brandeker and Thordardottir (2022, p. 1) describe cluster size as "the number of consecutive words that belong to the same category, for example, farm animals or jungle animals". These authors suggest that the use of clusters – specifically larger cluster sizes – may assist participants in word retrieval by breaking each overarching theme into

subcategories that are easier to process during this task. This strategy is thought to improve participant verbal fluency scores.

Understanding this task in adults allows us to address similar issues in a younger population, revealing how bilingual individuals approach this task across the lifespan. In this study we seek to answer the following questions: How do bilingual participants perform on the verbal fluency task compared to monolingual participants? Do Spanish-English cognates yield better verbal fluency performance for bilinguals? Does clustering and cluster size affect the number of responses given by participants? Answering these questions will assist in gaining understanding for how Spanish-English individuals process their second language, helping to improve testing materials and provide more accurate diagnoses of this population in clinical practice.

Chapter 2

Methods

Participants

Recruitment for bilingual participants in this study included advertisements and emails published by the Universidad Nacional Autónoma de México. We screened participants for English fluency, history of traumatic brain injury and speech and language disorders, as well as handedness – all of which are factors that may affect performance. Twenty-three individuals were selected to participate in this study based on these criteria. All tasks were administered in English. According to a self-rated language history survey, the ages of participants ranged from 19 to 37 years. This same survey revealed that their age of acquisition of English ranged from ages 3 to 16 years, and all participants rated themselves as conversationally fluent in English. Out of the 23 participants in this study, 36% of participants were male and 64% were female.

Monolingual participants were recruited from flyers posted at The Pennsylvania State University and screened using the same criteria that was used for bilingual participants. Twenty-seven participants consented to participation in this study, but one participant was excluded during analysis due to missing data. This resulted in a total of 26 participants. The participants selected were between the ages of 18 and 75 years. These participants all reported English as their native language with conversational fluency. Out of the 26 participants, 23% of were male and 77% of were female. One participant reported that they preferred not to answer this demographic question.

Procedure

All data in this study was audio recorded to assure coding accuracy during analysis. During the data collection session, participants completed a battery of language assessment tasks to evaluate skills across linguistic contexts. These included: digit span (forward and backward), sentence repetition tasks, non-word repetition tasks, a lexical decision task, picture naming, word list recall, and verbal fluency. Digit span and verbal fluency were administered by the examiners, while all other tasks were conducted on a laptop during the in-person session.

The verbal fluency task procedure performed in this study included four categories: tools, clothing, animals, and musical instruments (responses for the tools category were very limited and are not reported in this analysis). Each participant was given verbal instructions to name as many unique words pertaining to a given category as possible within a 60 second timeframe. Of these words, none should repeat or contain the same subcategory within the word named (e.g. blue, baby blue, navy blue, etc.). No further prompts were provided.

Data Analysis

Variables

The independent variable in research question one is whether the participant is monolingual or bilingual. The linguistic status of each participant was self-reported in a language history questionnaire. In research question two, the independent variable is bilingual cognate use. Cognates were measured after translating responses into Spanish. For the final research question, the independent variable is cluster size and frequency. Clusters were determined collaboratively.

The dependent variable for this study was verbal fluency performance for all three research questions. This was measured by counting how many unique and relevant words the participant produced in each category. A participant's performance is thought to be impacted by their linguistic experience, whether they use cognates, the presence of semantic clusters, and how long those clusters are.

Transcriptions

Transcriptions were produced for all relevant data from participant audio recordings. This included each verbal fluency category with every response within those categories, as well as the other linguistic tasks not used in the current study.

Data Coding

All participant productions were examined and labeled for the various independent and dependent variable measures (cluster/cluster size, Spanish-English cognates, and verbal fluency performance).

Verbal fluency performance was coded by determining which productions were relevant to the category and excluding repetitions or intrusions. A repetition was considered any word that the participant produced twice in the same category (e.g. "whale, fish, whale" in the animals category) while an intrusion was any irrelevant words produced in a category (e.g. "guitar, piano, wrench" in the musical instruments category). After removing repetitions and intrusions, each participant's data per category was counted to provide a score for verbal fluency performance.

Spanish-English cognates were coded after translation of the English productions into Spanish. Any word that had similar pronunciation and meaning in both languages was considered a cognate (e.g. pants v. pantalones). Cognates were then counted per participant per category.

Clusters were coded by breaking each category into subcategories and counting how many words were in each subcategory as well as how many subcategories there were in total. A group of words with an overarching theme was considered a cluster (e.g. dog, cat, mouse, and rabbit fall under household pets). This data was recorded per participant per category.

Reliability Coding

Members of the Penn State Child Language Lab were trained to complete reliability coding for this study. For total scores, a second rater coded 20% of the participants which were selected at random. Cognates were determined by consensus between two raters who had at least intermediate proficiency in Spanish. Clusters were also coded by consensus between two raters.

Chapter 3

Results

Overall Performance

When analyzing the verbal fluency response rates of monolingual and bilingual participants for the clothing category, monolingual participants had a mean score of 12.4 correct responses while bilingual participants had a mean score of 12.1 correct responses. For the animals category of the verbal fluency task, the monolingual group had a mean score of 15.4 correct responses while the bilingual group had a mean score of 12.5 correct responses. Finally, the monolingual participants had a mean score of 12.5 correct responses in the musical instruments category while the bilingual participants had a mean score of 7.4 correct responses. This data indicates that on average, the monolingual group scored higher in each category than the bilingual group. An independent samples t-test was completed to support this data, providing a Cohen's d score of 0.085 for clothing, 0.718 for animals, and 1.655 for instruments. Additional data from this test can be found in Table 1.

Table 1: Independent Samples T-Test

	t	df	p	Cohen's d	SE Cohen's d
Clothing_Total	0.296	47	0.769	0.085	0.286
Animals_Total	2.509	47	0.016	0.718	0.303
Instruments_Total	5.782	47	< .001	1.655	0.367

Note. **Student's t-test.**

Table 2: Monolingual Descriptive Statistics

Descriptive Statistics

	Clothing Total	Animals Total	Instruments Total
Valid	49	49	49
Missing	0	0	0
Median	12.000	14.000	10.000
Mean	12.245	14.041	10.122
Std. Deviation	3.479	4.320	3.961
Minimum	0.000	3.000	2.000
Maximum	21.000	24.000	17.000

Table 3: Bilingual Descriptive Statistics

Descriptive Statistics

	Clothing Total	Animals Total	Instruments Total
Valid	23	23	23
Missing	0	0	0
Median	12.000	13.000	7.000
Mean	12.087	12.478	7.435
Std. Deviation	3.397	3.788	2.727
Minimum	6.000	6.000	2.000
Maximum	21.000	22.000	13.000

Cognate Use

Spearman's rho correlations were used to compare the total performance score and the use of cognates for bilingual participants in each of the three categories tested. For the clothing category, the Spearman's rho value was 0.48 and the p-value was 0.022. This is a statistically significant value with a positive correlation between the two scores. In the animals category, the Spearman's rho value was 0.74 and the p-value was <.001, which is statistically significant. Similarly, the values for the instruments category were statistically significant with a Spearman's rho value of 0.95 and a p-value of <.001 as well.

Table 4: Bilingual Total v. Cognate

Variable	Spearman's rho	p-value
Clothing Cognate	0.48	0.022
Animals Cognate	0.74	< .001
Instruments Cognate	0.95	< .001

Clustering

Both monolingual and bilingual data were used to compare total performances across categories relating to how many clusters were produced per category. In the clothing category, the monolingual Spearman's rho value was 0.59 and the p-value was 0.002. For animals, the Spearman's rho value was 0.66 and the p-value was < .001. Finally, the instruments Spearman's rho value was 0.68 and the p-value was < .001. All of the values for the monolingual data were statistically significant with positive correlations.

The bilingual data reflected similar outcomes to the monolingual data. For clothing, the Spearman's rho value was 0.71 and the p-value was < .001. In the animals category, the Spearman's rho value was 0.54 and the p-value was 0.008. The instruments Spearman's rho value was 0.51 and the p-value was 0.014. Each of these values reflects statistical significance and positive correlations between the two values.

Table 5: Monolingual Total v. Cluster

Variable	Spearman's rho	p-value
Clothing Cluster	0.59	0.002
Animals Cluster	0.66	< .001
Instruments Cluster	0.68	< .001

Table 6: Bilingual Total v. Cluster

Variable	Spearman's rho	p-value
Clothing Cluster	0.71	< .001
Animals Cluster	0.54	0.008
Instruments Cluster	0.51	0.014

Cluster Size

The maximum cluster size for both the monolingual and bilingual groups were used in this analysis, comparing this data to the total score for each group. The monolingual Spearman's rho value for the clothing category was 0.65 and the p-value was < .001 – a statistically significant value. The animals category values were -0.18 and 0.392 for Spearman's rho and p-value respectively. This data demonstrates a correlation that is not statistically significant. Lastly, the instruments Spearman's rho value was 0.74 and the p-value was < .001 which is statistically significant.

For the bilingual group in the clothing category, the Spearman's rho value was 0.708 and the p-value was < .001. This is a statistically significant value. In the animals category, the Spearman's rho value was 0.13 and the p-value was 0.541. These values indicate that the correlation is not statistically significant but does have a positive correlation. In the final

category for musical instruments the Spearman's rho value was 0.51 and the p-value was 0.014, which is statistically significant.

Table 7: Monolingual Total v. Max Cluster Size

Variable	Spearman's rho	p-value
Clothing Max Cluster	0.65	< .001
Animals Max Cluster	-0.18	0.392
Instruments Max Cluster	0.74	< .001

Table 8: Bilingual Total v. Max Cluster Size

Variable	Spearman's rho	p-value
Clothing Max Cluster	0.71	< .001
Animals Max Cluster	0.13	0.541
Instruments Max Cluster	0.51	0.014

Chapter 4

Discussion

In overall performance, the monolingual participants performed higher than bilingual participants on average in the animals and instruments categories, but not clothing. This was expected before analysis and can be explained by the longer exposure to and mastery of English from monolingual participants. This group completed all tasks in their native language which provided an inherent advantage over the bilingual group, who had varying levels of English mastery which impacted their available vocabulary for verbal fluency.

For bilinguals, cognate use was a statistically significant predictor of higher performance outcomes in each category. This suggests that the more cognates a participant used, the higher their score was for the verbal fluency task. Using cognates is one strategy that eases cognitive effort during task completion for bilingual participants (Blumfeld et al., 2016, p. 192). Less searching for appropriate responses is required because the bilingual participant can draw comparisons between English and Spanish vocabulary to produce examples in each category. This process happens subconsciously and can be referenced when explaining the intrusion of certain words such as “battery” in the instruments category. This common mistake from the present study for bilingual participants comes from a false Spanish-English cognate “batería” which means drum. When searching for words, the participant uses this Spanish foundation and produces a word from English that sounds similar. Although this technique works for words such as “flauta” and “piano”, it does not generalize to all vocabulary used.

Clustering for both monolinguals and bilinguals was a significant predictor in all categories, indicating that an increased use of clusters correlates to increased performance in either group. This strategy is effective in chunking information to make word-finding more

efficient during the verbal fluency task (Brandeker & Thordadottir, 2022, p. 376). The use of a subcategory (cluster) allows the participant to narrow options to sort the vocabulary that they know into digestible segments, making production easier. There are often bridges between these categories where one similar group of animals leads into the next group, maximizing output and raising scores significantly.

In the monolingual and bilingual groups, clothing and instruments categories' maximum cluster size was a significant predictor of verbal fluency performance. However, maximum cluster size was not a significant predictor in the animals category. For clothing and instruments, the correlation of cluster size to higher scores can be explained by the overall clustering strategy as well. Clustering lessens cognitive effort and maximizes efficiency, allowing for long strings of correct responses that follow a similar theme which results in higher overall scoring. The lack of correlation in the animals category may be explained by participants not needing to rely on larger clusters to produce a high number of correct responses because of more exposure to this vocabulary.

This study was completed to investigate predictors of verbal fluency performance for bilingual participants. The data collected helps to further clinical research intended to improve diagnostic testing materials for bilingual language disorders that do not fairly evaluate multilingual populations against testing norms. Biased testing materials result in bilingual children being wrongfully diagnosed with language disorders as well as being undiagnosed for disorders that aren't adequately assessed by these tests. Verbal fluency targets a portion of these language tests that can be modified to tell us more about how bilingual individuals process language.

Limitations and Future Directions

One limitation of this study was that bilingual participants were not tested in both English and Spanish for each task completed. Testing participants in Spanish would allow for comparison between an individual's L1 and L2 performance and whether the same predictors are consistent across languages and categories. This could also further assist in explaining the differences in overall scores between the bilingual and monolingual groups.

Additionally, these results may not be generalizable because of the educational background, geographic limitation, and age range of participants. All participants had completed or were completing higher education and lived in either Pennsylvania or Mexico City. The age range of the monolingual group was also larger than the bilingual group, which may have affected scores. Gathering a larger sample from different areas, education levels, and ages would make this research more generalized.

Future studies should explore testing bilinguals in both their L1 and L2 to investigate differences in performance across both languages. Monolinguals and bilinguals from different geographical locations should also be tested to make conclusions more generalized for a larger population. Finally, this test should be performed with children to gauge performance across the lifespan and to collect data for more accurate norming samples in child language disorder testing materials.

REFERENCES

Bedore, Lisa M., and Elizabeth D. Peña. (2008). Assessment of Bilingual Children For Identification of Language Impairment: Current Findings and Implications for Practice. *International Journal of Bilingual Education and Bilingualism*, 11, 1-29. https://doi.org/10.2167/beb392.0.

Blumenfeld, Henrike K., et al. (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish-English bilinguals' verbal fluency. *International Journal of Speech-Language Pathology,* 18, 190-201. https://doi.org/10.3109/17549507.2015.1081288.

Brandeker, Myrto, and Elin Thordardottir. (2022). Verbal fluency in bilingual school-age children: Looking at switching, clustering, and the effect of language experience. *International Journal of Bilingualism*, 27, 374–393. https://doi.org/10.1177/13670069211055263.

Davis, Andréa K., et al. (2018). Bilingualism, cognates and reading fluency in children." *Journal of Research in Reading*, 41, S12–S29. https://doi.org/10.1111/1467-9817.12263.

Hemsley, Gayle, et al. (2014). Identifying language difference versus disorder in bilingual children. *Speech, Language and Hearing*, 17, 101–11. https://doi.org/10.1179/2050572813y.0000000027.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? predictors of verbal fluency performance in older adults. Frontiers in Psychology, 5. https://doi.org/10.3389/fpsyg.2014.00772.