UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Paths to characterizing typical and atypical bilingual language development

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Language and Communicative Disorders

by

Irina Potapova

Committee in charge:

San Diego State University

Sonja Pruitt-Lord, Chair
Henrike Blumenfeld
Sarah Mattson

University of California San Diego

Leanne Chukoskie
Seana Coulson
Jeanne Townsend

2019

The Dissertation of Irina Potapova is approved, and is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

San Diego State University

2019

EPIGRAPH

There is a crack in everything—that's how the light gets in

– Leonard Cohen

Глаза боятся, а руки делают

– Russian proverb

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor, Dr. Sonja Pruitt-Lord. You accepted a linguistics undergrad as a summer research assistant and became my introduction to the field of speech-language pathology; you were the first person to suggest that I apply for doctoral programs; and you ultimately assumed the responsibility of being my mentor. As a direct result of meeting you, my trajectory changed. Thank you for your enduring support and encouragement and for always acting in my best interest. I could not have had a better mentor, and I am beyond grateful to learn from your example.

I also wish to thank my committee members, Dr. Henrike Blumenfeld, Dr. Leanne Chukoskie, Dr. Seana Coulson, Dr. Sarah Mattson and Dr. Jeanne Townsend, for their support of an ambitious project. I am fortunate that this talented group of individuals invested their time, thoughts and energy in my research. Thank you for demonstrating the positive impact that faculty may have on a student's dissertation experience.

I would also like to recognize the entire faculty and staff of the Joint Doctoral Program in Language and Communicative Disorders (JDP-LCD) and the School of Speech, Language, and Hearing Sciences at San Diego State University, past and present. I would like to thank Dr. Rachel Mayberry and the late Dr. Beverly Wulfeck, who co-directed the JDP-LCD when I enrolled. You made me feel like I belonged in an environment that would otherwise be intimidating. To Dr. Tracy Love, who currently co-directs the JDP-LCD alongside Dr. Coulson, thank you for going to bat for the students of this department. Thanks also to the many faculty who have mentored me in academic settings; I would not have reached this point without the insight I gained from your courses in speech-language pathology, cognitive science, linguistics, psychology and neuroscience. I would also like to thank Dr. Giang Pham and Dr. Jessica Barlow who, like Dr. Blumenfeld, did me the kindness of inviting me to collaborate in their research. Dr. Pham, I think I have made you laugh with ideas good and bad—I hope it was more often the former than the latter. Dr. Barlow, you became my academic auntie, alongside Dr. Leah Fabiano-Smith—thank you both for everything that entails, from providing guidance in my research to making international conferences as fun as possible. I am so grateful to my clinical supervisors at the San Diego State University Speech-Language Clinic and my off-campus placements, including Deanna Hughes, Margaret Jones, Paul McCrae, Nicole Miller, Karyn Searcy, and especially Charlotte Lopes. I would not understand communication disorders the way I do without the opportunities that you gave me. To Janet Park and Marla Fulton, thank you for each question you answered and problem you solved; to Janet Shin, thank you for taking care of students across two campuses.

To the many members of the Child Language Development, Disorders and Disparities Lab and our collaborators, I am grateful to have worked alongside you. To our research assistants, thank you for transcribing language samples under deadline, wrangling preschoolers, and pushing yourselves to learn new things. To our participants, their families, and their teachers, thank you for your time and for your flexibility. To the graduate students and research assistants who brought my F31 research to life, including Alicia Escobedo, Cristal Toscano, Fabiola Morales, Genesis Maldonado, Izamar Bergado, Jonathan Robinson Anthony, Laura Silva, Leilani Melendrez and Liliana Michel, thank you for each morning you spent with a preschooler and an eye tracker. A special thanks to Daena Taylor and SDSU Speech-Language Clinic Director Carrie Goodwiler for lending their voices to this project. Alicia, I've been the only child for a while (unless we count Adelyn!)—it was a pleasure to share the title of resident doc student with you. To student collaborators Quynh Dam and Jonathan Robinson Anthony, thank you for the opportunity to join forces and learn. Lastly, a shout out to the lab's summer crew—Sydney Barnes, Julia Moluf, Jaclyn Morris and Chelsie Sanders—you are the greatest of all time.

To the students of the JDP-LCD, I am honored to be part of this club with you. Thank you to the students who gifted me with advice and resources, including Stephanie De Anda, Kristi Hendrikson and Kristen Secora; you shape academia for the better. Thank you to Erin Smolak, who provided extra support and motivation as we approached our defenses. An extra special thanks to Natalie Sullivan and Philip Combiths, I couldn't imagine this without you. And to Rachel Bristol and Carson Rigoli of the Cognitive Science Department at the University of California San Diego, thank you for adopting me.

I am full of gratitude for the people who helped me reach this point. Mama and papa, none of my accomplishments would be possible without you. Sanya, you are my greatest role model. TJ, I am deeply grateful for your partnership, thank you for taking care of me. Nims, thank you for keeping my time in graduate school adventurous. And finally, thank you to the many, many educators who cleared my path to graduate school.

Chapter 2, in full, is a reprint of material as it appears in Potapova, I., Kelly, S., Combiths, P. N., & Pruitt-Lord, S. L. (2018). Evaluating English morpheme accuracy, diversity, and productivity measures in language samples of developing bilinguals. *Language, Speech, and Hearing Services in Schools*, *49*(2), 260-276. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of material as it appears in Potapova, I & Pruitt-Lord, S. L. (2019). Spanish-English bilingual children's relative use of English tense and agreement morphemes. *Journal of Monolingual and Bilingual Speech, 1*(1), 118-142. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of material as it appears in Potapova, I., Blumenfeld, H. K., & Pruitt-Lord, S. L. (2016). Cognate identification methods: Impacts on the cognate advantage in adult and child Spanish-English bilinguals. *International Journal of Bilingualism*, *20*(6), 714-731. The dissertation author was the primary investigator and author of this paper.

Chapter 5 is being prepared for submission for publication of the material. Potapova, I. & Pruitt-Lord, S.L. (*in preparation*). Cross-language interactions during word learning in bilingual children with typical and atypical language development. The dissertation author was the primary investigator and the primary author of this paper.

I gratefully acknowledge the following sources of funding throughout graduate school: the NIH Training Grant entitled “Neurocognitive Approaches to Communication Disorders” (Grant Number T3DC007361); the San Diego State University Lipinsky Fellowship; the University of California San Diego Frontiers of Innovation Scholars Program; the San Diego State University Inger P. Davis Doctoral Fellowship; and the NIH Predoctoral Training Grant entitled “Cross-language interactions during word learning in bilingual children with typical and atypical language development” (Grant Number F31DC016194). I would also like to recognize funding that supported research efforts included in this dissertation: an NIH RO3 award (Grant Number DC012141; PI: Pruitt-Lord), a research grant from a local community foundation awarded to Dr. Pruitt-Lord; the University Grants Program at SDSU and the Summer Undergraduate Research Program at SDSU.

VITA

EDUCATION

Ph.D. San Diego State University/University of California San Diego — 2019
Language and Communicative Disorders

B.A. Cornell University — 2012
Linguistics, *summa cum laude*
Phi Beta Kappa

RESEARCH EXPERIENCE

Doctoral student — 2013 – 2019
San Diego State University
Child Language Development, Disorders and Disparities Lab

Lab manager — 2012 – 2013
San Diego State University
Child Language Development, Disorders and Disparities Lab

Undergraduate honors thesis — 2011 – 2012
Cornell University
Linguistics department

Research assistant — 2010 – 2012
Cornell University
Language Acquisition Lab

FUNDING

NIH NIDCD Predoctoral Trainee – F31DC016194 — 2017 – 2019
"Cross-language interactions during word learning in bilingual children with typical and atypical language development"

Fellow, San Diego State University Inger P. Davis Doctoral Fellowship — 2018 – 2019

Recipient, UC San Diego Frontiers of Innovation Scholars Program — 2016 – 2017
"Eye movements during word learning in children with and without Autism Spectrum Disorder"

Fellow, San Diego State University Lipinsky Fellowship — 2016 – 2017

NIH NIDCD Trainee – T32DC007361 — 2013 – 2016
"Neurocognitive Approaches to Communication Disorders"

Travel awards

Symposium for Research in Child Language Disorders Student Travel Award	2018
San Diego State University Center for Cognitive and Clinical Neuroscience Student Travel Award	2017
International Association for the Study of Child Language Conference Travel Award	2017
UC San Diego Graduate Student Association Travel Grant	2016
San Diego State University Graduate Student Travel Fund	2016
Symposium for Research in Child Language Disorders Student Travel Award	2015
International Association for the Study of Child Language Conference Travel Award	2014
College of Health and Human Services Travel Award	2014
San Diego State University Graduate Student Travel Fund	2013

PUBLICATIONS (PEER REVIEWED):

Potapova, I. & Pruitt-Lord, S. (2019). Spanish-English Bilingual Children's Relative Use of English Tense and Agreement Morphemes. *Journal of Monolingual and Bilingual Speech.*

Potapova, I. & Pruitt-Lord, S. (2019). Towards understanding the bilingual profile in typical and atypical language development: A tutorial. *International Journal of Speech-Language Pathology.*

Potapova, I., Kelly, S., Combiths, P., & Pruitt-Lord, S. (2018). Evaluating English morpheme accuracy, diversity, and productivity measures in language samples of developing bilinguals. *Language Speech and Hearing Services in Schools.*

Combiths, P., Barlow, J. A., **Potapova**, I., & Pruitt-Lord, S. (2017). Influences of Phonological Context on Tense Marking in Spanish-English Dual Language Learners. *Journal of Speech Language and Hearing Sciences*.

Potapova, I., Blumenfeld, H., & Pruitt-Lord, S. (2016). Cognate Identification Methods: Impacts on the cognate advantage in adult and child Spanish-English bilinguals. *International Journal of Bilingualism.*

SUBMITTED:

Dam, Q., Pham, G., **Potapova**, I., & Pruitt-Lord, S. (Under Review). First and Second Language Acquisition
in Vietnamese-English Bilingual Children.

Robinson Anthony, J., Blumenfeld, H., **Potapova**, I., & Pruitt-Lord, S. (Under review). Language dominance predicts cognate effects and metalinguistic awareness in preschool bilinguals.

CONFERENCE PRESENTATIONS

* mentee—undergraduate student research assistant
† mentee—master's student research assistant

Robinson Anthony, J.J.D., **Potapova**, I., Blumenfeld, H.K., & Pruitt-Lord, S. Native-language dominance predicts linguistic and cognitive transfer to the non-dominant language in preschool-age bilinguals. Paper presented at the International Symposium for Bilingualism, Alberta, Canada.

Escobedo, A., Gallagher, J., **Potapova,** I. & Pruitt-Lord, S. Grammaticality in Spanish-English preschool-aged bilinguals. Paper presented at the Symposium for Research in Child Language Disorders, Madison, Wisconsin

Escobedo, A., Gallagher, J., **Potapova,** I. & Pruitt-Lord, S. Grammaticality in English-Spanish preschool-aged children. Paper presented at San Diego State University 12th Annual Student Research Symposium, San Diego, CA.

Moluf, J., Escobedo, A., **Potapova,** I. & Pruitt-Lord, S. Getting ready for school: Language performance in young bilinguals. Paper presented at San Diego State University 12th Annual Student Research Symposium, San Diego, CA.

Dam, Q., **Potapova** I., Pruitt-Lord, S. & Pham, G.T. (2018) First and second language acquisition in Vietnamese-English bilingual children. Talk presented at the American Speech-Language and Hearing Association Convention, Boston, MA.

Potapova, I., Dam, Q., Limon-Hernadez, J., Pruitt-Lord, S. & Pham, G.T. (2018) The role of cognates in language learning: Evidence from young bilingual children. Paper presented at the American Speech-Language and Hearing Association Convention, Boston, MA.

Potapova, I. & Pruitt-Lord, S. Tense and agreement morpheme emergence in developing bilinguals with varying language skills. Paper presented at the International Clinical Phonetics and Linguistics Conference, Corinthia San Gorg, Malta.

Potapova, I. & Pruitt-Lord, S. (2018). Relative productivity of English Tense and Agreement morphemes in preschool-aged bilinguals. Paper presented at the Symposium for Research in Child Language Disorders, Madison, Wisconsin.

Potapova, I. (2018). Measures of English morphology for bilingual children. Talk presented at the San Diego State University Speech-Language-Hearing Awareness and Information Day.

Yang, R.*, **Potapova**, I., Pruitt-Lord, S. & Chukoskie, L. Categorizing semantic associations in individuals with Autism Spectrum Disorder. Paper presented at the Symposium for Research in Child Language Disorders, Madison, Wisconsin.

Hurley, M.*, **Potapova,** I., Robinson-Anthony, J.D.D., & Pruitt-Lord, S. (2018). Executive Function Skills with Relation to Language Skills in Bilingual Children. Paper presented at San Diego State University 11th Annual Student Research Symposium, San Diego, CA.

Powell, K.*, Purcell, R.*, Wilder, C.*, Wade, A.[†], **Potapova**, I., & Pruitt-Lord, S. (2018). Does Order Matter? Grammatical Development in Bilingual Preschoolers. Paper presented at San Diego State University 11th Annual Student Research Symposium, San Diego, CA.

Robinson Anthony, J.D.D., **Potapova**, I., Hurley, M. *, & Pruitt-Lord, S. (2018). Language growth predicts executive function in low SES bilingual children. Paper presented at San Diego State University 11th Annual Student Research Symposium, San Diego, CA.

Robinson Anthony, J.J.D., **Potapova**, I., Hurley, M. *, & Pruitt-Lord, S. (2018). Language dominance, language growth, and executive function in preschool-aged bilinguals. Paper presented at the 16th Biennial Conference of the International Association for Cognitive Education and Psychology. Guadalajara, Mexico.

Potapova, I., Kelly, S.[†] & Pruitt-Lord, S. (2017). Measuring Tense and Agreement Morpheme Emergence for English Language Learners: Current Findings & Clinical Training. Talk presented at the American Speech-Language and Hearing Association Convention, Los Angeles, California.

Potapova, I., Cheng, F.[†], Twentyman, O.[†], Yeung, D.[†] & Pruitt-Lord, S. (2017). Morpheme Commission and Omission Error Analysis of Spanish-English Bilingual Preschoolers. Paper presented at the American Speech-Language and Hearing Association Convention, Los Angeles, California.

Potapova, I., Combiths, P., Barlow, J.A. & Pruitt-Lord, S. (2017). Profiling language development in dual language learners. Paper presented at the 14th Annual International Congress for the Study of Child Language. Lyon, France.

Combiths, P., **Potapova**, I., Pruitt-Lord, S. & Barlow, J.A. (2017). Dual Language Influence on the Effective Mass of Phonological Context: Evidence from Third-Person Singular Marking Patterns. Paper presented at the International Symposium on Bilingualism. Limerick, Ireland.

Pruitt-Lord, S., **Potapova**, I., Drexler, E., & Hye, T. (2017). Lessons Learned from a Community-Based Preschool Language Intervention. Paper presented at the Society for Research in Child Development. Austin, Texas.

Potapova, I., Combiths, P., & Pruitt-Lord. S. (2016). Measuring morphological development in preschool-aged Spanish-speaking English Language Learners. Paper presented at the Symposium for Research in Child Language. Madison, Wisconsin.

Combiths, P., **Potapova** I., Pruitt-Lord, S., & Barlow, J. (2016). Phonological effects on grammatical morpheme production at and across the word boundary. Paper presented at the Symposium for Research in Child Language. Madison, Wisconsin.

Augustino, A.[†], Magill, L.[†], **Potapova** I., and Pruitt-Lord, S. (2016). Patterns of morpheme use by preschool-aged Spanish-speaking English Language learners. Paper presented at the San Diego State University Student Research Symposium.

Potapova, I. & Boroditsky, L., (2016). Seeing events in a second language: a cognitive “accent” in event perception. Paper presented at the Events in Language and Cognition Workshop, CUNY Conference on Human Sentence Processing.

Potapova, I., Kelly, S. & Pruitt-Lord, S. (2016). English Tense & Agreement Morpheme Use in Dual Language Learners: Individual & Group Patterns. Talk presented at the American Speech-Language and Hearing Association Convention, Philadelphia, PA.

Potapova, I., Finestack, L.H., & Pruitt-Lord, S. (2016). Exploring Cross-Language Interactions & Inhibitory Control in Preschool-Aged Spanish-Speaking English Language Learners. Paper presented at the American Speech-Language and Hearing Sciences Convention, Philadelphia, PA.

Potapova, I., Pruitt-Lord, S., & Blumenfeld, H. (2015). Exploring profiles of preschool-aged Spanish-English bilinguals that do and do not demonstrate a cognate advantage. Paper presented at the Symposium for Research in Child Language Disorders, Madison, Wisconsin.

Potapova, I., Pruitt-Lord, S., & Blumenfeld, H. (2015). Cognate effects in preschool-aged Spanish-English bilinguals. Talk presented at the American Speech-Language and Hearing Association Convention.

Potapova, I. (2015). Grammatical Aspect and Event Perception: Are you gonna finish that? Talk presented at the Joint Doctoral Program in Language and Communicative Disorders Annual Doctoral Student Colloquium.

Potapova, I. (2015) Are you gonna finish that? Grammatical aspect and event descriptions in Russian-English bilinguals. Talk presented at the 38th Annual SDSU Linguistics Colloquium, San Diego, CA.

Potapova, I. (2015). The impact of Russian-English bilinguals' language experience on event description. Talk presented at the San Diego State University Student Research Symposium, San Diego, CA.

Byers-Straus H.[†], Carbajal, A.*, **Potapova**, I. & Pruitt-Lord, S. (2015). Investigating relationships between the cognate advantage and measures of development. Paper presented at the San Diego State University Student Research Symposium, San Diego, CA.

Potapova, I. (2014). Comparing Cognate Effects in Child and Adult Spanish-English Bilinguals. Talk presented at the Joint Doctoral Program in Language and Communicative Disorders Annual Doctoral Student Colloquium.

Potapova, I. (2014). Cognate Effects in Child and Adult Spanish-English Bilinguals During Receptive Vocabulary Assessment. Talk presented at the California State University Student Research Competition, CSU East Bay.

Potapova, I., Blumenfeld, H., & Pruitt-Lord, S. (2014). Considering Cognate Effects in Child Spanish-English Bilinguals during Receptive Vocabulary Assessment with PPVT-III, Forms A and B. Paper presented at the Convention of the International Association for the Study of Child Language, Amsterdam, The Netherlands.

Potapova, I., Blumenfeld, H. & Pruitt-Lord, S. (2013). Cognate effects in Child and Adult Bilinguals during Receptive Vocabulary Assessment with the PPVT-III. Paper presented at the American Speech-Language and Hearing Sciences Convention, Chicago, Illinois.

Pruitt-Lord, S. & **Potapova**, I. (2013). Tense productivity of English Language Learners. Paper presented at the Symposium on Research in Child Language Disorders, Madison, Wisconsin.

Sherman, J. C., Flyn, S., Whitlock, J., Henderson, C., Costigan, A., Cordella, C., **Potapova** I., Gair, J. & Lust, B. (2012). Dissociating Syntax and Semantics in Language Change in Prodromal AD: Production of Complex Sentences in Amnestic MCI. Paper presented at the Association for Psychological Science Convention, Chicago, Illinois.

Cordella, C., Min, J., **Potapova**, I., O'Neil, A., Bastine, E. & Whitmore, S. (2012). Assessing the nature of linguistic deficits in prodromal Alzheimer's Disease Coordinate Sentences: Pronominal Deletion. Paper presented at the Cornell Undergraduate Research Board Spring Forum.

ACADEMIC HONORS AND AWARDS

San Diego State University Student Success Fee Awardee	2018
ASHA Convention Outstanding Student Oral Presentation Submission	2017
International Association for the Study of Child Language Best Student Poster Finalist	2017
Paul Pfaff Memorial Scholarship recipient	2016
California Speech Language and Hearing Association Graduate Student Scholarship Finalist	2015
Paul Pfaff Memorial Scholarship recipient	2014
California State University Research Competition, Second Place in Behavioral Sciences	2014
San Diego State University Student Research Symposium, President's Award	2014

CLINICAL EXPERIENCE

Speech Language Pathologist - Clinical Fellow 2018 – 2019
San Diego State University Speech and Language Clinic

Student Clinician
Fay Elementary School 2016
Acquired Brain Injury Program, Mesa College 2016
San Diego Regional Center 2016
San Diego State University Speech and Language Clinic 2015 – 2016

TEACHING EXPERIENCE

Instructor of Record 2018, 2019
Foundations Of Speech-Language: Differences And Disorders In Children (SLHS 514)
Online graduate prerequisite course
San Diego State University, College of Extended Studies

Instructor of Record 2017, 2018
Foundations Of Speech-Language: Differences And Disorders In Children (SLHS 514)
In-person undergraduate course
San Diego State University
School of Speech, Language and Hearing Sciences

Teaching assistant, (Course Instructor: Sonja Pruitt-Lord) 2014

Introduction to Language Science (SLHS 300).
In-person undergraduate course
San Diego State University
School of Speech Language and Hearing Sciences

OTHER ACTIVITIES

Graduate Student Representative 2014 – 2019
Joint Doctoral Program in Language and Communicative Disorders, San Diego State University & University of California San Diego

Ad Hoc Peer Review
American Journal of Speech Language Pathology 2019
ASHA convention: Language and Learning in school-aged individuals 2019

Supervised Peer Review 2015 – 2018

Language, Speech and Hearing Services In Schools

International Journal of Bilingual Education and Bilingualism

San Diego Stroke Awareness and Information Day

Organizational Committee 2015 – 2018

Co-chair 2016 – 2018

ABSTRACT OF THE DISSERTATION

Paths to characterizing typical and atypical bilingual language development

by

Irina Potapova

Doctor of Philosophy in Language and Communicative Disorders

San Diego State University, 2019
University of California San Diego, 2019

Sonja Pruitt-Lord, Chair

The realization of appropriate clinical tools for bilingual children rests on an increasingly accurate understanding of bilingual language development. To enhance our understanding of bilingualism, this dissertation investigates typical and atypical bilingual language development under two complementary perspectives. The first approach, a single-language focus, emphasizes careful measurement of performance in a single language to identify clinical tools and establish benchmarks that are appropriate for bilingual children. The second approach, the uniquely bilingual lens, leverages patterns of language use that are specific to the experience of dual language exposure, including cross-language interactions, to support the creation of

clinical tools that are tailored to bilingual speakers. Taken together, these approaches allow for a more comprehensive picture of language development in young bilinguals.

Consistent with a single-language focus, Chapters 2 and 3 utilize measures derived from language samples to characterize young bilinguals' use of English tense and agreement marking, a grammatical feature that is challenging for typically developing children and particularly challenging for children with atypical language development. Participants included Spanish-English bilingual children recruited at a time when language assessment is common, their preschool year. Across these studies, results indicated that clinically relevant information for bilingual children may be obtained from single-language measures provided that the measure is well-matched to the child's level of language development and that comparisons are made with care (i.e., bilingual children are compared to bilingual children).

Chapters 4 and 5 investigate cross-language interactions in Spanish-English bilingual preschoolers, consistent with the uniquely bilingual lens. Results in Chapter 4 demonstrate that young bilinguals may be sensitive to cognates (e.g., *elephant*/*elefante* in English/Spanish), thus demonstrating patterns of cross-language interactions that are widely documented in adult bilinguals. Chapter 5 extends this work to investigate whether Spanish-English bilingual preschoolers with varying levels of language ability demonstrate cognate sensitivity as they encounter novel words in each language. Findings from off-line accuracy measures and fixation patterns indicate that cross-language interactions in development are linked to multiple factors, including language dominance, task requirements and, potentially, language ability.

CHAPTER 1:

Introduction

Introduction

Children readily achieve bilingualism given adequate experience with multiple languages. Maintenance of the native language alongside acquisition of the community language has positive implications for psycho-social development, including enhanced quality of familial relations and sense of identity (Kohnert, 2012; Oh & Fuligni, 2010; Phinney, Romero, Nava & Huang, 2001). Globally, multilingualism is understood to be the norm (e.g., de Zarobe & de Zarobe, 2015), and in the United States, the number of bilingual speakers has steadily risen over time (U.S. Census Bureau, 2016). Within this common experience, the profiles of language performance in bilingual speakers are highly heterogeneous. Bilingual language development is associated with the same factors that predict language performance in monolinguals, such as input quantity and quality, but also factors that are specific to the experience of dual language exposure (e.g., Hoff et al., 2012). Based on combinations of factors like relative exposure, proficiency and dominance in each language; similarities and differences in linguistic structures across the two languages; and the relative status of each language in the speaker's community, children who grow up speaking more than one language may demonstrate widely varying levels of performance in each (Paradis, Genesee & Crago, 2011). That is, two bilingual children may show entirely different levels of performance on the same task when tested in the same language and both still represent typical bilingual language development, with strong underlying language abilities. Speech-language pathologists are thus asked to recognize these language differences as they work to accurately identify multilingual children who also present with patterns associated with language disorders (Oetting, 2018). Understandably, clinicians report difficulty working with children from culturally and linguistically diverse backgrounds (Guiberson & Atkins, 2012; Caesar & Kohler, 2007; Roseberry-McKibbin, Brice & O'Hanlon, 2005).

An increasingly accurate understanding of bilingual language development is needed to support continued improvements in clinical practice for young bilinguals. The research in this dissertation is thus motivated by the overarching question: *how does typical and atypical language development manifest in the context of bilingualism?* This work helps address this question from two complementary perspectives: a *single language focus* and the *uniquely bilingual lens*.

The first approach prioritizes characterizing performance in each language individually. The broad question guiding this *single language focus* is: *what are the meaningful patterns of single language development for children from culturally and linguistically diverse backgrounds, and how do we measure them?* When considering performance in a single language, bilingual children may resemble monolingual counterparts in the developmental trajectory of some language milestones (e.g., non-tense morphemes are acquired more readily than tense morphemes in both monolingual and bilingual English speakers; Paradis, 2005)—but research has made clear that the use of monolingual benchmarks may misrepresent bilingual children's language abilities (e.g., Paradis, 2005; Bialystok, Luk, Peets & Yang, 2010). The full extent of similarities and differences across monolingual and bilingual children remains to be determined and new standards that reflect language development in young bilinguals need to be established. Research with a single language focus delivers this vital information to speech-language pathologists, enabling them to more accurately assess each language in line with best practice. And, as results included in this dissertation will show, detailed analysis of a single language opens doors to consider patterns of language use that result from experience with multiple languages.

This consideration of how knowledge of one language impacts performance in the other is representative of research under a *uniquely bilingual lens*, the second approach to characterizing typical and atypical bilingual language development. Our understanding of bilingualism in adulthood is that the two languages are not separate. Instead, they are jointly activated, allowing for cross-language interactions, as reflected in prominent theories of adult bilingual language processing (e.g., BIA+, Dijkstra & Van Heuven, 2002; Revised Hierarchical Model, Kroll, van Hell, Tokowicz & Green, 2010; Distributed Features Model, van Hell & de Groot, 1998). Moreover, mounting evidence suggests that similar processes may be at play for young bilinguals, including research presented in this dissertation (Sheng, Lam, Cruz & Fulton, 2016; Kelley & Kohnert, 2012; Pérez, et al., 2010; Schelletter, 2002). The question guiding research under this approach is: *what are the patterns of cross-language interactions in young bilinguals, and do they differ as a function of typical vs. atypical language development?* Ultimately, the pairing of a single language focus with the uniquely bilingual lens supports an increasingly complete picture of typical and atypical dual language development.

Overview of the dissertation

Studies included in this dissertation are informed by theoretical frameworks that highlight cross-language interactions in bilingual speakers (e.g., Kan & Kohnert, 2008) and by known areas of weakness for children with atypical language development (e.g., Leonard, 2014; Gutiérrez-Clellen, Simon-Cereijido & Wagner, 2008; Kan & Windsor, 2010; Kapantzoglou, Restrepo & Thompson, 2012). Synthesizing our understanding of bilingualism and of language disorder, the research included in this dissertation examines typical and atypical language development in bilingual children.

Chapters 2 and 3 represent work that is consistent with a single language focus. Each study investigates the use of English tense and agreement (T/A) morphemes, a grammatical skill that is frequently considered in language assessment and is understood to be indicative of language ability in English-speaking monolingual and bilingual children (e.g., Leonard, 2014; Gutiérrez-Clellen et al. 2008). Chapter 2 presents a published study that investigated three approaches to measuring this important skill for preschool-aged Spanish-English bilinguals with typical language development and bilingual peers with low language skills. Measurement approaches included one traditional method associated with *mastery* of T/A marking (Bedore & Leonard, 1998; Gladfelter & Leonard, 2013; Rice & Wexler, 1996; Rice, Wexler, & Cleave, 1995) and two relatively novel measures that capture *emergence* of the T/A marking system (e.g., Hadley & Short, 2005). Results of this study help clinicians identify and employ appropriate assessment measures for children from culturally and linguistically diverse backgrounds. To better bridge this research with clinical practice, Chapter 2 includes case studies that illustrate the use and meaningfulness of each measure.

In an extension of this work, Chapter 3 presents a published study that investigated the developmental trajectory of English T/A morpheme categories in preschool-aged Spanish-English bilingual children with varying levels of language ability. Results of this study help elevate our understanding of morphosyntactic development in bilingual children to better match work available for monolingual peers. In addition to clarifying patterns of English language development that differ between children with dual language experience and children with experience in a single language, this work invites opportunities to consider the influence of one language on development in the other.

Consistent with the uniquely bilingual lens, Chapters 4 and 5 present research that investigates bilingual language development with a focus on cross-language interactions. Chapter 4 presents a published work investigating cognate sensitivity (as measured by performance on a standardized receptive vocabulary task) in both adult and child bilinguals, thereby allowing for a consideration of cross-language interactions from developing bilingualism into mature bilingualism. Finally, Chapter 5 extends this work and investigates cross-language interactions in young bilinguals with increased experimental control and multiple measures of cognate sensitivity, including eye tracking. Importantly, this study includes children with specific language impairment and thus sheds light on how cross-language interactions may or may not manifest given atypical language development. Together, these studies test whether features of the bilingual profile that are well documented in adult bilinguals are similarly expressed earlier in childhood and explore whether this information may be leveraged to provide clinically relevant information.

The dissertation concludes with a general discussion that integrates findings from all research studies to better understand patterns of typical and atypical language development in young bilinguals.

References

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research*, *41*(5), 185–92.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525-531.

Caesar, L. G., & Kohler, P. D. (2007). The state of school-based bilingual assessment: Actual practice versus recommended guidelines. *Language, Speech, and Hearing Services in Schools*, *38*(3), 190-200.

de Zarobe, L. R., & de Zarobe, Y. R. (2015). New perspectives on multilingualism and L2 acquisition: An introduction. *International Journal of Multilingualism*, *12*(4), 393-403.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(03), 175-197.

Hadley, P. A., & Short, H. (2005). The onset of tense marking in children at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research*, *48*(6), 344–62.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1-27.

Gladfelter, A., & Leonard, L. B. (2013). Alternative tense and agreement morpheme measures for assessing grammatical deficits during the preschool period. *Journal of Speech, Language, and Hearing Research*, *56*(2), 542–52.

Guiberson, M., & Atkins, J. (2012). Speech-language pathologists' preparation, practices, and perspectives on serving culturally and linguistically diverse children. *Communication Disorders Quarterly*, *33*(3), 169-180.

Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics*, *29*(1), 3-19.

Kan, P. F., & Kohnert, K. (2008). Fast mapping by bilingual preschool children. *Journal of child language*, *35*(03), 495-514.

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *53*(3), 739-756.

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. Language, Speech, and Hearing Services in Schools, 43(1), 81-96.

Kelley, A., & Kohnert, K. (2012). Is there a cognate advantage for typically developing Spanish-speaking English-language learners?. *Language, Speech, and Hearing Services in Schools*, *43*(2), 191-204.

Kohnert, K., (2012). Language Disorders in Bilingual Children and Adults, 2nd ed. San Diego, CA: Plural Publishing.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. Bilingualism (Cambridge, England), 13(3), 373.

Leonard, L. B. (2014). Specific language impairment across languages. Child development perspectives, 8(1), 1-5.

Oetting, J. B. (2018). Prologue: Toward Accurate Identification of Developmental Language Disorder Within Linguistically Diverse Schools.

Oh, J. S., & Fuligni, A. J. (2010). The role of heritage language development in the ethnic identity and family relationships of adolescents from immigrant backgrounds. *Social Development*, *19*(1), 202-220.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *36*(3), 172-187.

Paradis, J., Genesee, F., & Crago, M. B. (2011). Dual Language Development & Disorder. Baltimore: Paul H. Brookes Publishing Co.

Pérez, A. M., Peña, E. D., & Bedore, L. M. (2010). Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services (San Diego, Calif.)*, *4*(1), 55.

Phinney, J. S., Romero, I., Nava, M., & Huang, D. (2001). The role of language, parents, and peers in ethnic identity among adolescents in immigrant families. *Journal of youth and Adolescence*, *30*(2), 135-153.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1,239–57.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research*, *38*(4), 850–63.

Roseberry-McKibbin, C., Brice, A., & O'Hanlon, L. (2005). Serving English Language Learners in Public School Settings: A National Survey. *Language, Speech, and Hearing Services in Schools*, *36*(1), 48-61.

Schelletter, C. (2002). The effect of form similarity on bilingual children's lexical development. *Bilingualism: Language and Cognition*, *5*(02), 93-107.

Sheng, L., Lam, B. P. W., Cruz, D., & Fulton, A. (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of Experimental Child Psychology*, *141*, 229-238.

U.S. Census Bureau; “Language Spoken at Home”; generated by Irina Potapova using American FactFinder; factfinder.census.gov; (15 June 2016).

van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1, 193–211.

CHAPTER 2:

Evaluating English Morpheme Accuracy, Diversity and Productivity Measures in Language Samples of Developing Bilinguals

LSHSS

Research Article

Evaluating English Morpheme Accuracy, Diversity, and Productivity Measures in Language Samples of Developing Bilinguals

Irina Potapova,[a,b] Sophia Kelly,[a] Philip N. Combiths,[a,b] and Sonja L. Pruitt-Lord[a]

Purpose: This work explores the clinical relevance of three measures of morpheme use for preschool-age Spanish–English bilingual children with varying language skills. The 3 measures reflect accuracy, diversity (the tense marker total), and productivity (the tense and agreement productivity score [TAP score]) of the English tense and agreement system.
Method: Measures were generated from language samples collected at the beginning and end of the participants' preschool year. Participants included 74 typically developing Spanish–English bilinguals and 19 peers with low language skills. The morpheme measures were evaluated with regard to their relationships with other language sample measures, their ability to reflect group differences, and their potential for capturing morphological development at group and individual levels.
Results: Across both groups, the tense marker total and TAP scores were associated with other language measures and demonstrated both group differences and growth over time. The accuracy measure met few of these benchmarks.
Conclusion: The tense marker total and TAP score, which were designed to capture emerging morphological abilities, contribute valuable information to a comprehensive language assessment of young bilinguals developing English. Case examples are provided to illustrate the clinical significance of including these measures in assessment.

Substantial individual variation is characteristic of bilingual language development, including the acquisition and mastery of morphosyntactic skills (Paradis, 2005; Paradis & Crago, 2000; Paradis, Rice, Crago, & Marquis, 2008). Capturing and assessing these emerging skills is made more difficult by the dearth of clinical tools developed for culturally and linguistically diverse populations (e.g., Bedore & Peña, 2008; Caesar & Kohler, 2007; Gillam, Peña, Bedore, Bohman, & Mendez-Perez, 2013). Language sample analysis is a highly recommended assessment approach that is resistant to cultural and linguistic biases that are likely implicit in standardized assessments (Gutiérrez-Clellen, Restrepo, Bedore, Peña, & Anderson, 2000; Heilmann, 2010; Heilmann, Miller, & Nockerts, 2010; Hewitt, Hammer, Yont, & Tomblin, 2005; Rojas & Iglesias, 2009). However, to maximize the benefits of this culturally sensitive approach, appropriate language sample measures must be identified (Oetting et al., 2010; Stockman, 1996).

Presently, we explore the clinical utility of three measures of English morpheme use generated from the spontaneous language samples of preschool-age Spanish–English bilingual children with typical and low language skills. To do this, we examine whether these measures successfully track progress and/or capture differences across children with varying language abilities. Such investigations of English language measures are imperative and meet a practical need, as English is frequently used in the assessment of bilinguals in the United States (Caesar & Kohler, 2007; Gillam et al., 2013). In addition, two case examples are provided to demonstrate these measures in practice.

Broad Language Sample Measures

Language sample analysis is important for evaluating and monitoring the language development of children from nonmainstream backgrounds, as formal assessments are widely regarded as inadequate for these populations (Bedore

[a]San Diego State University, CA
[b]University of California, San Diego
Correspondence to Irina Potapova: ipotapova@sdsu.edu
Editor-in-Chief: Shelley Gray
Editor: Janna Oetting
Received March 3, 2017
Revision received August 7, 2017
Accepted September 28, 2017
https://doi.org/10.1044/2017_LSHSS-17-0026
Publisher Note: This article is part of the Clinical Forum: Toward Accurate Identification of Developmental Language Disorder Within Linguistically Diverse Schools.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

& Pena, 2008; Caesar & Kohler, 2007; Paradis, Nicoladis, Crago, & Genesee, 2011; Stockman, 1996). Transcribed spontaneous language samples can be analyzed for many broad measures of language production (e.g., type–token ratio, percent intelligibility, grammaticality). Mean length of utterance in words (MLUw) and number of different words (NDW) are two traditional language sample measures that are clinically relevant and appropriate for use with culturally and linguistically diverse clients (Rojas & Iglesias, 2009). Mean length of utterance is associated with morphosyntactic development, and MLUw is considered a preferable measure for bilingual children because it is resistant to cross-linguistic differences in morphological richness (Gutiérrez-Clellen et al., 2000). NDW reflects the number of unique uninflected root words in the sample and is a measure of lexical diversity (Golberg, Paradis, & Crago, 2008).

Both MLUw and NDW are utilized in research and clinical settings for characterizing young bilinguals' productive language. These measures have been found to help identify language impairment in both bilingual and monolingual children (e.g., Bedore, Peña, Gillam, & Ho, 2010; Hewitt et al., 2005; Simon-Cereijido & Gutiérrez-Clellen, 2009). Conveniently, both measures can be automatically generated by transcription software after language samples have been transcribed and coded for bound and unbound morphemes. However, neither measure directly captures a child's development of tense and agreement (T/A) morpheme marking, which has been established as a salient indicator of language impairment (Bedore & Leonard, 1998; Gutiérrez-Clellen, Simon-Cereijido, & Wagner, 2008; Rice & Wexler, 1996).

T/A Morpheme Measures

Difficulty with T/A morphology is a hallmark of language impairment in English-speaking children (e.g., Leonard, 2014). However, errors in morpheme marking are also to be expected in typically developing children acquiring English, whether they are young English monolinguals or children acquiring English in addition to another language (e.g., Rice, 2010). An understanding of morpheme marking in bilingual children—as well as adequate tools to measure morpheme use in this population—are needed to avoid mistakenly identifying typically developing bilingual children as having language impairment (Paradis, 2005; Paradis & Crago, 2000). Carefully characterizing morpheme use in bilingual children is also important because it is possible that influence from the native language will cause bilinguals' developmental trajectories to differ from monolingual peers with regard to sequence of acquisition or error types (Armon-Lotem, 2014; Nicoladis, Song, & Marentette, 2012; Paradis & Blom, 2016). To identify clinically relevant measures of morpheme use for preschool-age bilingual children, this study tested whether approaches based on measuring T/A morpheme accuracy, diversity, and productivity aligned with broad language measures, successfully tracked progress at group and individual levels, and reflected differences between children with varying language abilities.

Accuracy of T/A Morpheme Marking

One traditional approach to measuring morpheme mastery is to calculate the accuracy of morpheme marking. In clinical and research settings, accuracy may be determined based on performance during a spontaneous language sample or on a probe designed to elicit targeted morphemes. In language sample analysis, accuracy measures require the transcription of a sample, followed by coding of all obligatory contexts for each morpheme of interest. Obligatory contexts are then manually reviewed to identify successful morpheme productions (e.g., the child says, "he walked" in a past-tense context), as well as morpheme omissions and other errors (e.g., the child says, "he walk" or "he walks," respectively, in a past-tense context). The number of successful morpheme productions is divided by the total number of obligatory contexts to produce a composite measure of morpheme accuracy. Composite accuracy measures thus collapse performance across multiple morphemes and reflect both correct and errored productions.

As a direct measure of morpheme use, accuracy rates have an important role in clinical decision making. Measures of T/A accuracy have been used to differentiate monolingual children with language impairment from typically developing peers (Bedore & Leonard, 1998; Gladfelter & Leonard, 2013; Rice & Wexler, 1996; Rice, Wexler, & Cleave, 1995), though these studies and others have differed in which morphemes are included in the composite and in other methodological considerations (Balason & Dollaghan, 2002). Limited available research also suggests that English T/A accuracy may differentiate bilingual children (4;5–6;5 [years;months]) with typical and atypical language development (Gutiérrez-Clellen et al., 2008).

However, accuracy may not be the most appropriate measure for all stages of language development. To illustrate, Fitzgerald, Rispoli, Hadley, and McKenna (2012) found that 41% of typically developing English monolingual children in a longitudinal sample demonstrated lower accuracy scores at 27 months of age than at 24. This phenomenon, "backtracking," can be explained by inflated accuracy rates at earlier time points due to the production of high-frequency combinations that do not require morphosyntactic processing or knowledge (e.g., *that's, it's, what's*; Guo, Spencer, & Tomblin, 2013; Rispoli, Hadley, & Holt, 2009). As bilingual children acquiring English may demonstrate similar acquisition patterns to younger monolinguals (e.g., Nicoladis et al., 2012; cf. Paradis & Blom, 2016; Rice, 2010), it is important to consider measures of morpheme accuracy in this population.

Furthermore, bilinguals' morpheme accuracy is also characterized by greater individual variability relative to age-matched monolingual peers (Gutiérrez-Clellen et al., 2008; Paradis, 2005; Paradis & Crago, 2000; Paradis et al., 2008). In addition, parallels in morpheme accuracy have been found between typically developing bilingual children and monolingual peers with language impairment, a group whose mastery of T/A marking is also delayed relative to typically developing monolinguals (Paradis, 2005). Altogether, there is motivation to investigate measures of morpheme use for bilingual children.

Diversity and Productivity of T/A Morpheme Use

The tense marker total and T/A productivity score (TAP score) were designed to better capture morpheme use for children in early stages of English language development. Introduced by Hadley and Short (2005), the two measures reflect contrastive uses of five morpheme categories that have been extensively studied in the literature on language impairment: (a) third-person singular (-3s: *drive/3s*), (b) past tense (-ed: *walk/ed*), (c) forms of copula *BE* (cop *BE: She is fast*), (d) forms of auxiliary *BE* (aux *BE*: *I am going*), and (e) forms of auxiliary *DO* (aux *DO*: *Do you like it?*).[1]

The tense marker total awards points for different surface forms for the five morphemes of interest: -3s, -ed, aux *DO* (*do, does, did*), cop *BE* (*is, am, are, was, were*), and aux *BE* (*is, am, are, was, were*). Higher scores thus indicate an ability to use an increasing number of unique surface forms. The TAP score awards points for each T/A morpheme provided that the child demonstrated sufficiently different productions of each one. Higher TAP scores indicate an ability to use T/A morphemes in increasingly unique syntactic contexts. Both measures were designed to capture onset of the T/A system (Hadley & Short, 2005); in addition, the tense marker total can be thought of as a measure of diversity or breadth of the T/A system, whereas the TAP score can be thought as a measure of productivity or depth. In contrast with measures of T/A accuracy, the scoring protocols for both the tense marker total and TAP score were designed to award points for morpheme uses that meet specific productivity criteria to safeguard against artificially inflated scores (Hadley & Short, 2005). Further scoring details are provided in the Method section.

The tense marker total and TAP score are valuable for measuring early English T/A development. For English monolinguals, these focused measures are correlated with broad language measures that include a wider variety of grammatical forms. Hadley and Short (2005) found that the tense marker total and TAP score were correlated with traditional language sample measures (e.g., mean length of utterance) for monolingual children ages 2;0–3;0 with low language and those at risk for language impairment. Furthermore, higher TAP scores predicted progress toward T/A mastery as measured by spontaneous language samples and standardized probes (Hadley & Short, 2005; Hadley, Rispoli, Holt, Fitzgerald, & Bahnsen, 2014; Rispoli et al., 2009). In addition, these measures are clinically relevant for monolinguals. Young children (2;0–3;0) at risk for specific language impairment had lower TAP scores than their peers, and their scores increased at a slower rate over time (Hadley & Holt, 2006). The tense marker total and TAP score have also been used to differentiate between typical and atypical language development in older English monolingual children (3;0–5;6; Gladfelter & Leonard, 2013; Guo & Eisenberg, 2014). Gladfelter and Leonard (2013) found that the tense marker total correctly identified 85.19% of the children aged 4;0–4;6 (23/27 participants) and that the TAP score correctly identified 82.14% of the children aged 5;0–5;6 (23/28 participants). In summary, the tense marker total and TAP score appear to be meaningful measures for children who are in the stage of development between first use of T/A morphemes and mastery of the T/A system.

However, these promising measures had not yet been considered in the context of dual language exposure. This study thus investigates whether these measures of diversity (i.e., the tense marker total) and productivity (i.e., the TAP score) can serve similar purposes for developing bilingual children who, like English monolinguals, undergo the process of acquiring English T/A morphemes.

[1]Each surface form of the auxiliary and copula verb paradigms (e.g., *am, is, are, was, were,* and *be*) is its own morpheme, and as such, it is appropriate to refer to these paradigms as morpheme classes or categories (Hadley et al., 2014). However, for brevity, we use the term "morpheme" to refer to -3s, -ed, cop *BE,* aux *BE,* and aux *DO.*

Present Study

This research was motivated by the need for appropriate measures of English language development for young bilinguals in the United States (Bedore & Peña, 2008; Bedore et al., 2018; Gillam et al., 2013). The present goal was to consider measures of English morpheme use in preschool-age Spanish–English developing bilinguals with typical language and with low language skills. Three morpheme measures from spontaneous language samples were considered: a composite T/A accuracy measure, the tense marker total, and the TAP score. First, we evaluate these measures on the basis of convergence with established measures; next, we explore their clinical potential to capture group differences and growth over time.

We ask, for preschool-age developing bilinguals with varying language skills, the following:

1. Do morpheme measures reflecting accuracy, diversity, and productivity relate to broad language sample measures?
2. Do morpheme measures reflecting accuracy, diversity, and productivity capture differences across groups and over time?
3. Do morpheme measures reflecting accuracy, diversity, and productivity successfully capture growth at the individual level (i.e., minimize backtracking)?

Given that the tense marker total and TAP score were designed to capture early stages of English T/A development and were proven relevant for assessing language in preschool-age monolinguals, we expected these measures to also be appropriate for preschool-age developing bilinguals who are learning English. Specifically, we expected the tense marker totals and TAP score to be higher for the typically developing bilingual group than for the low language group. We also expected scores to be higher at the end of the school year than at the beginning. Finally, we expected these measures to result in minimal backtracking when examined at the individual level (Rispoli et al., 2009).

Method

Participants

Preschool-age Spanish–English developing bilinguals were identified from an ongoing community-based research project. For inclusion in this study, each participant was required to (a) be exposed to Spanish at home at least 30% of the time (Pearson, Fernandez, Lewedeg, & Oller, 1997), (b) score within normal limits on a nonverbal cognition measure, (c) complete language samples at the beginning and end of his or her preschool year, and (d) produce at least 10 complete and intelligible utterances in each language sample. In total, 93 children (mean age = 4;2, *SD* = 5.05 months) met these criteria and were included in the study.

Per parent report, participants were exposed to Spanish 72.35% of the time (*SD* = 20.11, range 40–100) at home, on average. Scores from the Figure Ground and Form Completion subtests of the Leiter International Performance Scale–Revised (Roid & Miller, 1997), a non-verbal cognition measure, were in the normal range, with an average score of 11.60 (*SD* = 1.93, range 7–16). In addition, maternal education, reported by the parents of 72 participants, was 10.01 years (*SD* = 2.90, range 3–16) on average. This maternal education range could be considered indicative of the entire sample as all of the children were enrolled at the same preschool site and the school setting required below-poverty standards to participate.

Eligible participants were then assigned groups based on language ability: developing bilingual children with typical language development (BiTD) and developing bilingual children with low language skills (BiLL). Group membership was determined by parent report (Gutiérrez-Clellen & Kreiter, 2003; Restrepo, 1998). Parents completed language questionnaires in their preferred language to provide information regarding their child's language experience and development, answering questions such as, "Do you or did you ever have any concerns about your child's speech and/or language?" Children were considered for the BiTD group if the parent reported no concerns. Conversely, participants with reported concerns were considered for the BiLL group. The BiTD group included 74 children (39 boys, 35 girls; mean age = 4;2, *SD* = 5.2 months); the BiLL group included 19 children (12 boys, seven girls; mean age = 4;1, *SD* = 4.5 months).

Subsequent comparison of language sample performance revealed expected group differences across a variety of English language sample measures at the beginning of the year. The BiTD group outperformed the BiLL group on MLUw and NDW, as well as total number of utterances, number of complete and intelligible utterances, type–token ratio, and percent intelligibility (all $ps < .038$, as evidenced by one-tailed t tests; see Table 1). By the end of the year, the BiTD group continued to demonstrate significantly higher MLUw. Importantly, the two groups were comparable on a number of factors that may be relevant to performance, including age, Spanish exposure at home, and maternal education (all $ps > .464$; see Table 1).

Procedure

Data were collected in coordination with a community-based research study under the direction of the final author. Information about the study, consent forms, and language questionnaires were sent home with each child in English and Spanish through collaborative efforts with teachers and classroom personnel at a local preschool. Children whose parents returned signed consent forms were eligible for the larger study, which included participation in onsite data collection at the beginning and end of the academic year. To administer an assessment battery for the larger project, including collecting the language samples used for this research, multiple sessions were planned for each participant. Session length was determined by child engagement, with an upper limit of 40 min. During sessions dedicated to language sample collection, no other assessments or measures were completed. Data were collected by graduate students in speech-language pathology who were trained to administer the standardized assessments accurately, to collect spontaneous language samples, and to monitor child engagement. All children were tested individually, and child assent was obtained before each session. Each wave of data was collected in the span of 2–4 weeks.

Measures

All measures of interest were generated from language samples collected at the beginning and end of an academic year (Time 1 and Time 2, respectively). Each language sample was elicited following a set play protocol, using a toy car, garage, and picnic sets, as well as a standard set of pictures for story retells. The digitally recorded language samples were orthographically transcribed and coded by trained research assistants following Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2012) conventions. In addition, the use of Spanish, as well as its potential impact on the measures of interest, was considered. Using a generous criterion (e.g., including utterances with only a single Spanish element; e.g., "put all your cintos"), only 6.05% of all complete and intelligible utterances included Spanish. Critically, the presence of Spanish utterances did not impact the calculation of the three English morpheme measures. On average, Time 1 samples included 99.32 complete and intelligible child utterances (*SD* = 59.43), and Time 2 samples included 145.35 (*SD* = 70.34). All measures were computed for all language samples at both testing points.

Broad Language Sample Measures

MLUw and NDW were automatically generated using SALT. The use of MLUw, as opposed to mean length of utterance in morphemes, is consistent with related research in bilingual children (e.g., Blom, Paradis, & Duncan, 2012; Paradis & Kirova, 2014). In addition, both MLUw and NDW are considered culturally sensitive and have been recommended for use with Spanish–English bilingual children (Rojas & Iglesias, 2009).

Language, Speech, and Hearing Services in Schools • Vol. 49 • 260–276 • April 2018

Table 1. Participant characteristics for BiTD and BiLL participant groups.

		Background			Broad language sample measures									
					MLUw		NDW		Total utterances		Complete and intelligible utterances		% Intelligibility	
Group		Age	% Spanish heard	Maternal education	Time 1*	Time 2*	Time 1*	Time 2	Time 1*	Time 2	Time 1*	Time 2	Time 1*	Time 2
BiTD	*M*	50.36	73.12	9.95	2.51	3.38	93.81	133.07	145.26	176.61	105.66	144.66	86.20	91.29
	SD	5.18	20.22	2.98	0.76	0.84	46.77	48.21	77.24	85.30	60.80	73.61	13.37	13.20
BiLL	*M*	49.47	69.35	10.27	2.16	2.95	68.89	130.16	109.00	183.42	74.63	148.05	80.16	92.21
	SD	4.54	19.97	2.66	0.73	0.74	41.75	43.39	63.32	73.55	47.42	57.48	11.75	6.45

Note. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills; MLUw = mean length of utterance in words; NDW = number of different words.
$^{*}p < .05$.

Accuracy of T/A Marking

A composite measure of T/A accuracy (Pruitt & Oetting, 2009; Rice & Wexler, 1996, 2001; Rice, Wexler & Hershberger, 1998) was calculated to reflect productive marking in obligatory contexts. To most appropriately match the tense marker total and TAP score, the composite accuracy measure was calculated based on use of -3s, -ed, cop *BE,* aux *BE,* and aux *DO* in obligatory contexts, with overregularizations considered successful uses (Gladfelter & Leonard, 2013). Thus, the composite T/A accuracy measure was calculated by dividing the total number of correct uses and overregularizations by the total number of correct uses, overregularizations, omissions, and other errors. This proportion was multiplied by 100 to yield a percentage.

To facilitate scoring of morpheme accuracy, SALT morpheme codes were used to extract utterances containing obligatory contexts for the five morphemes of interest. Trained graduate research assistants categorized each obligatory context as a correct use (e.g., *play/ed* in a past-tense context), an overregularization (e.g., *break/ed* in a past-tense context), an omission (*e.g., play* in a past-tense context), or as an "other error" (including agreement, such as *they play/3s,* and tense errors, such as *he play/3s* in a past-tense context).

Diversity and Productivity of T/A Morpheme Use

Tense marker totals and TAP scores were generated following protocol outlined in Hadley and Short (2005), awarding points for contrastive uses of -3s, -ed, cop *BE,* aux *BE,* and aux *DO.* The tense marker total awards 1 point for each possible surface form of the five morphemes of interest, for a maximum score of 15. The TAP score awards up to 5 points for sufficiently different uses of each T/A morpheme, for a maximum score of 25. For -3s and -ed, sufficiently different uses are determined by the production of different lexical verbs (e.g., *want/3s* and *need/3s*). For the copula and auxiliary verbs, sufficiently different uses are characterized by the presence of different subjects (e.g., *the baby is* and *the mommy is*) or different surface forms (e.g., *the baby is* and *the baby was*). In other words, children are not able to earn points for repeated productions (e.g., *the baby is* produced multiple times).

For both the tense marker total and TAP score, points are awarded for correct uses (e.g., *play/ed* in a past-tense context) and for overregularizations (e.g., *break/ed* in a past-tense context), as both are indicative of productive use. Conversely, no points are awarded for other errors, including morpheme omissions, T/A errors, and productions with null subjects. Furthermore, scoring restrictions for copula and auxiliary verbs were established to ensure that the scored productions reflect grammatical encoding, as opposed to direct activation of common forms (Hadley & Short, 2005; Rispoli & Hadley, 2011). Contracted copula and auxiliary verbs are scored when used with nouns (e.g., *baby/'s hungry*), but not with pronouns (e.g., *she/'s hungry*; Hadley & Short, 2005; Rispoli et al., 2009). Uncontracted forms (e.g., *baby is hungry; she is hungry*) are always eligible for scoring.

Scoring procedures for the tense marker total and TAP score are demonstrated—and contrasted with morpheme accuracy—using the abbreviated language sample in Appendix A. The tense marker total for this abbreviated sample is 6: 1 point each for Utterances 1 (cop *BE, is*), 3 (aux *BE, is*), 7 (-3s, *looks*), 11 (cop *BE, am*), 12 (aux *DO, does*), and 13 (-ed, *play/ed*). The TAP score, which awards additional points for sufficiently different uses of the same surface form, is 8: 1 point for each of the utterances above, as well as additional points for Utterances 5 (auxiliary *BE, is*) and 15 (-ed, *break/ed*). Note that Utterances 6 and 14 did not contribute to the TAP score, as neither meets the criterion for sufficiently different use: Utterance 6, which includes an aux *BE* form contracted to a noun, repeats a subject/surface form *(Daddy/'s* was awarded a point in Utterance 5), and Utterance 14 repeats a lexical verb (*play/ed* was awarded a point in Utterance 13). Following scoring procedures for both measures, no points were awarded for errored productions (Utterances 4 and 9) or for forms contracted to pronouns (Utterance 2).

In contrast to the tense marker total and TAP score, a measure of morpheme accuracy would take into consideration all utterances in the abbreviated sample. The number of correct productions and overregularizations, 11, would be divided by the total number of obligatory contexts, 15, and multiplied by 100 to yield an accuracy rate of 73.33%. Unlike the diversity and productivity measures, this approach both rewards productions in repeated contexts (e.g., Utterances 5 and 6; Utterances 13 and 14) and reflects both successful and errored productions.

To facilitate scoring, SALT codes were used to extract utterances with relevant T/A morphemes. The samples were hand scored by trained research assistants and the first and second authors to identify contrastive uses of the five target morphemes.

Reliability

Steps to ensure data reliability were taken at each level of transcription, coding, and scoring. All research assistants received training relevant to their assignment (transcription, coding, and/or scoring) and completed sample tasks to a satisfactory criterion prior to contributing to data processing. An adapted consensus procedure was utilized for transcription and coding (e.g., Eisenberg, Guo, & Germezia, 2012). After a trained research assistant transcribed a language sample, a second research assistant independently reviewed the transcript while listening to the corresponding audio file. All transcribers were instructed to mark an utterance as unintelligible if they were not able to transcribe the utterance after listening to the audio three times. Research assistants trained in coding protocol then coded the agreed-upon transcriptions for bound and unbound morphemes following established lab procedures. As with transcription, each sample was then independently reviewed for coding conventions by a second trained research assistant. Disagreements were resolved by referencing training materials or by appealing to a third transcriber/coder (the first or final authors) if needed. Coding was further reexamined during subsequent scoring procedures, as calculating the composite accuracy measure, tense marker total, and TAP score required the manual review of utterances containing T/A morphemes (Hadley & Short, 2005). Over 50 samples across the two testing points were independently scored for the composite T/A accuracy measure; average reliability was 94.75%. Over 25 samples across the two testing points were independently scored for tense marker totals and TAP scores; average reliability was 95.73%. All research assistants were blind to group status and children's performance on other measures.

Results

Do morpheme measures reflecting accuracy, diversity and productivity relate to culturally sensitive broad language sample measures?

Correlational analyses were conducted to test for convergence between the three morpheme measures of interest (composite accuracy, tense marker total, and TAP score; see Table 2) and the two culturally sensitive broad language sample measures (MLUw and NDW) at each time point for each group. Correlation coefficients of .2, .4, .6, and .8 were considered benchmarks for weak, moderate, strong, and very strong relationships, respectively (Evans, 1996).

There was little evidence of convergence for the composite accuracy measure and broad language skills across the two groups (see Table 3). For all analyses including the composite accuracy measure, children with fewer than three obligatory contexts for the five morphemes of interest were excluded, as accuracy could not be reliably calculated (Balason & Dollaghan, 2002). As a result, 53 BiTD and 10 BiLL participants were included in analyses involving the composite accuracy measure at Time 1; 72 BiTD and all 19 BiLL participants were included at Time 2. The only relationship to reach significance at Time 1 was between the composite accuracy score and MLUw for BiTD participants ($r = .281$, $p < .05$). At Time 2, accuracy was only significantly related to NDW for BiTD participants ($r = .419$, $p < .01$). For BiLL participants, accuracy was not significantly correlated with MLUw or NDW at either time point.

Analyses for the tense marker total and TAP score included all participants in each group. Both measures demonstrated consistent and significant convergence with broad language measures (see Table 3). At Time 1 for BiTD participants, the diversity and productivity measures demonstrated strong positive correlations with MLUw and NDW ($rs = .658–.757$, $ps < .01$). At Time 2 for this group, moderate to very strong positive relationships were demonstrated ($rs = .500–.826$, $ps < .01$). For BiLL participants, the diversity and productivity measures were moderately to strongly correlated with MLUw and NDW at Time 1 ($rs = .531–.682$, $ps < .05$) and Time 2 ($rs = .537–.758$, $ps < .05$).

Furthermore, the three measures of morpheme use were also related at both time points. Accuracy was related to the tense marker total and TAP score for BiTD participants at Time 1 and Time 2 ($rs = .328–.520$, $ps < .001$) and for BiLL participants at Time 2 ($rs = .555–.637$, $ps < .05$). The strength of these relationships was greater at the second time point.

Do morpheme measures reflecting accuracy, diversity, and productivity capture differences across groups and over time?

To address our second question, 2 × 2 analyses of variance that included participant group (BiTD vs. BiLL) as a between-subjects factor and time (Time 1 vs. Time 2) as a within-subject factor were conducted for each morpheme measure. Participants with fewer than three obligatory contexts for the T/A morphemes were again excluded from analyses for the composite accuracy measure (Balason & Dollaghan, 2002), but not for the tense marker total or TAP score.

For the accuracy-based measure of morpheme use, no significant main effect of group or time emerged, nor was the interaction significant ($ps > .156$). That is, accuracy rates were comparable across BiTD and BiLL participants, and scores were not indicative of growth over the course of the academic year (see Figure 1).

For the diversity and productivity measures, both main effects were significant (see Figures 2 and 3): Tense marker totals were higher for BiTD participants than for BiLL peers, $F(1, 91) = 4.621$, $p = .034$, $\eta_p^2 = .04$, and scores increased from Time 1 to Time 2, $F(1, 91) = 92.603$, $p < .001$, $\eta_p^2 = .408$ (see Figure 1). Similarly, TAP scores were higher for BiTD participants than for BiLL

Table 2. Performance on the three morpheme measures at Time 1 and Time 2 for BiTD and BiLL participants.

		Composite accuracy measure†		Tense marker total		TAP score	
Group		Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
BiTD	*M*	63.52	66.62	2.43	4.95	4.41	9.31
	SD	20.30	18.65	2.20	2.53	4.63	5.77
BiLL	*M*	56.25	58.65	1.37	3.79	2.26	6.89
	SD	29.64	19.34	2.11	2.35	3.96	5.03

Note. BiTD $n = 74$; except †, $n = 53$. BiLL $n = 19$; except †, $n = 10$. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills; TAP score = tense and agreement productivity score.

Table 3. Pearson correlations between MLUw, NDW, tense marker total, TAP score, and composite accuracy measure at Time 1 (above the diagonal; in roman) and Time 2 (below each diagonal; in italics) for BiTD and BiLL participants.

Group	Measure	MLUw	NDW	Tense marker total	TAP score	Composite accuracy measure†
BiTD	MLUw	—	.758**	.708**	.658**	.281*
	NDW	*.498***	—	.754**	.757**	.055
	Tense marker total	*.556***	*.825***	—	.923**	.328*
	TAP score	*.500***	*.793***	*.912***	—	.390**
	Composite accuracy measure†	*.092*	*.419***	*.520***	*.517***	—
BiLL	MLUw	—	.833**	.531*	.579**	.544
	NDW	*.598***	—	.643**	.682**	.449
	Tense marker total	*.542**	*.794***	—	.951**	.569
	TAP Score	*.537**	*.758***	*.929***	—	.533
	Composite accuracy measure†	*.292*	*.246*	*.555**	*.637***	—

Note. BiTD n = 74; except †, n = 53 for Time 1 and 72 for Time 2. BiLL participants = 19; except †, n = 10 for Time 1. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills; MLUw = mean length of utterance in words; NDW = number of different words; TAP score = tense and agreement productivity score.

*$p < .05$. **$p < .01$.

peers, $F(1, 91) = 4.283$, $p = .041$, $\eta_p^2 = .045$, and scores increased from Time 1 to Time 2, $F(1, 91) = 44.870$, $p < .001$, $\eta_p^2 = .330$. The interactions were not significant for either measure.

As an additional test for group-level patterns, performance for each target morpheme was considered. As the TAP score is composed of scores ranging from 0 to 5 for -3s, -ed, cop *BE,* aux *BE,* and aux *DO,* it is possible to compare productive uses of each morpheme at the beginning and end of the year. Five 2 × 2 analyses of variance that included participant group (BiTD vs. BiLL) as a between-subjects factor and time (Time 1 vs. Time 2) as a within-subject factor were conducted for each morpheme. Productive use of each morpheme increased from the beginning to the end of the school year (see Figure 4), as evidenced by a main effect of time: -3s, $F(1, 91) = 7.561$, $p = .007$, $\eta_p^2 = .077$; -ed, $F(1, 91) = 22.225$, $p < .001$, $\eta_p^2 = .196$; aux *DO,* $F(1, 91) = 6.902$, $p = .010$, $\eta_p^2 = .070$; cop *BE,* $F(1, 91) = 23.511$, $p < .001$, $\eta_p^2 = .205$; and aux *BE,* $F(1, 91) = 38.034$, $p < .001$, $\eta_p^2 = .295$. In addition, BiTD children outperformed BiLL peers on productions of -3s, $F(1, 91) = 4.244$, $p = .042$, $\eta_p^2 = .045$, and aux *BE,* $F(1, 91) = 5.202$, $p = .025$, $\eta_p^2 = .054$. No significant interactions emerged for any morpheme.

The composite accuracy measure may also be separated into accuracy measures for each target morpheme. However, such analyses were not feasible for the present data set due to the limited obligatory contexts (e.g., the average number of obligatory contexts for -3s, -ed, and aux

Figure 1. Composite accuracy measure rates at Time 1 and Time 2 for both participant groups. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills.

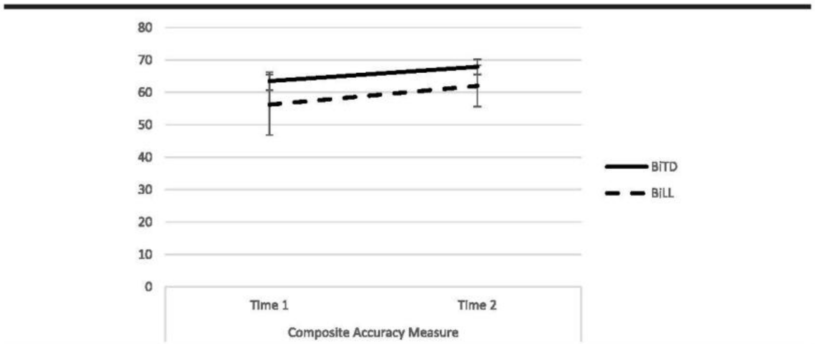
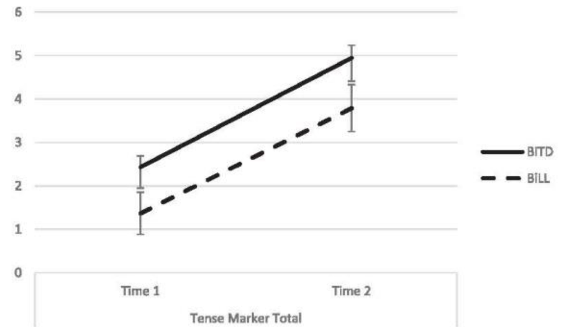


Figure 2. Tense marker totals at Time 1 and Time 2 for both participant groups. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills.



DO at Time 1 for BiTD participants was 2.38, 1.58, and 1.73, respectively; see Table 4).

Do morpheme measures reflecting accuracy, diversity, and productivity successfully capture growth at the individual level (i.e., minimize backtracking)?

Recall that backtracking is a phenomenon in which participants demonstrate lower scores at later time points. This pattern has been identified in the accuracy rates of young monolinguals acquiring the English T/A system (Fitzgerald et al., 2012; Rispoli et al., 2009). To evaluate backtracking in the present sample, each participant's performance on the three morpheme measures was compared at Time 1 and Time 2.

Results indicated that backtracking was common with the accuracy measure. Of the 53 BiTD participants whose samples included at least three obligatory contexts at Time 1, 42% had lower composite accuracy scores at Time 2. For the 10 BiLL children who met the criterion for calculating accuracy at Time 1, 50% demonstrated decreased accuracy rates at Time 2. Neither of these proportions significantly differed from chance ($ps > .27$), suggesting that, at the individual level, the accuracy measure did not reliably capture growth.

For the diversity and productivity measures, backtracking occurred less frequently. Of the 74 BiTD participants, 22% had lower scores on the tense marker total at Time 2, and the same percentage demonstrated backtracking on the TAP score. Of the 19 BiLL participants, only one child (5%) earned a lower tense marker total, and only two children (11%) earned lower TAP scores at Time 2. Each

Figure 3. TAP scores at Time 1 and Time 2 for both participant groups. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills; TAP score = tense and agreement productivity score.

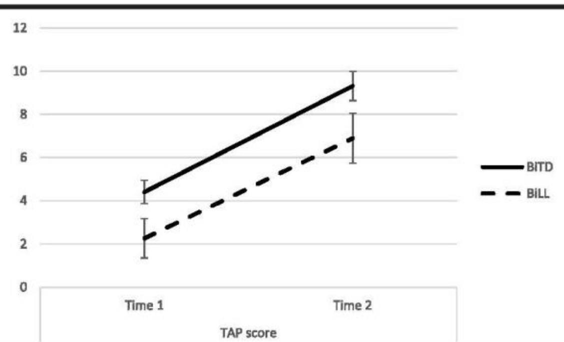
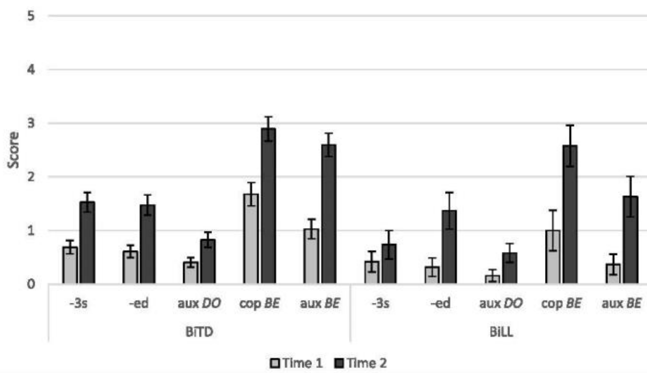


Figure 4. TAP score subscores for individual morphemes at Time 1 and Time 2 for each participant group. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills; TAP score = tense and agreement productivity score; -3s = third-person singular; -ed = past tense; cop *BE* = copula *BE*; aux *BE* = auxiliary *BE*; aux *DO* = auxiliary *DO*.



of these proportions was below chance (*ps* < .001), indicating limited backtracking. Indeed, fewer children demonstrated backtracking for these two measures than for accuracy (*ps* < .05, *V*s = .0083–.0592).[2]

Discussion

The present work sought to identify appropriate measures of English T/A morpheme use in preschool-age Spanish–English developing bilinguals with varying language skills. We considered three morpheme measures that can be derived from spontaneous language samples. One measure, a composite score capturing T/A accuracy, is frequently used for assessing language in clinical and research settings for monolingual and bilingual children (Balason & Dollaghan, 2002; Bedore & Leonard, 1998; Gladfelter & Leonard, 2013; Gutiérrez-Clellen et al., 2008; Rice et al., 1998). However, evidence suggests that there are drawbacks to accuracy measures when working with children whose English T/A systems are emerging (Fitzgerald et al., 2012; Rispoli et al., 2009). Composite accuracy measures do not safeguard against repetitions (e.g., multiple instances of -3s and -ed with the same lexical verb) or potentially formulaic constructions (e.g., *he's, it's, there's*). As a result, this measure may overestimate abilities at early stages of T/A development. Furthermore, limited obligatory contexts may make for an unreliable measure at certain stages of development (Balason & Dollaghan, 2002). The second two measures, the tense marker total and TAP score, were designed to compensate for those weaknesses and have demonstrated clinical utility for monolingual children with emerging T/A systems (Gladfelter & Leonard, 2013; Guo & Eisenberg, 2014; Hadley & Holt, 2006; Hadley & Short, 2005; Hadley et al., 2014; Rispoli et al., 2009; Rispoli, Hadley, & Holt, 2012). This study tested whether these measures are appropriate for young developing bilinguals, as they, too, are likely to be in emerging stages of English T/A development.

Table 4. Number of obligatory contexts per morpheme for BiTD and BiLL participants at Time 1.

Group		-3s	-ed	cop *BE*	aux *BE*	aux *DO*
BiTD	*M*	2.38	1.58	10.11	6.47	1.73
	SD	3.28	2.26	11.23	8.03	2.37
	Range	0–14	0–9	0–44	0–36	0–9
	Mdn	1	0	6	3	1
BiLL	*M*	2.05	1.42	6.63	2.53	0.47
	SD	3.78	3.10	8.88	5.09	0.91
	Range	0–11	0–11	0–29	0–20	0–3
	Mdn	0	0	3	0	0

Note. BiTD = bilingual with typically developing language; BiLL = bilingual with low language skills; -3s = third-person singular; -ed = past tense; cop *BE* = copula *BE*; aux *BE* = auxiliary *BE*; aux *DO* = auxiliary *DO*.

[2]An additional set of analyses considered individual stability for the three measures. Stability is characterized by consistent relative rankings over time, such that children with strong performance relative to peers at one point demonstrate a similarly high ranking at a second test point. This pattern is identified by significant correlations between performance on the same measure at multiple time points (Bornstein, Brown, & Slater, 1996). For BiTD participants, the diversity and productivity measures at Time 1 and Time 2 were found to be moderately positively correlated: tense marker total, $r = .454$, $p < .001$; TAP score: $r = .403$, $p < .001$. Results were comparable for BiLL participants: tense marker total, $r = .543$, $p = .016$; TAP score: $r = .518$, $p = .023$. Conversely, accuracy scores at pre- and posttesting were significantly associated only for BiTD participants: $r = .361$, $p = .008$.

Identifying meaningful measures of morpheme use for bilingual children has important clinical implications, as T/A morpheme marking is an area of particular weakness for English-speaking monolingual and bilingual children with language impairment (Bedore & Leonard, 1998; Gutiérrez-Clellen et al., 2008; Rice & Wexler, 1996).

T/A Morpheme Measures and Broad Language Sample Measures

In order to determine whether the three measures of morpheme use were appropriate for preschool-age developing bilinguals, we first established whether they were associated with culturally and developmentally sensitive measures of language development. Both the tense marker total and TAP score were positively associated with MLUw and NDW for both groups at both testing points. These relationships were generally strong, indicating that the tense marker total and TAP score convey information that is relevant to language development (Ebert & Pham, 2017).

The accuracy composite scores were not found to be consistently correlated with MLUw or NDW. Each may be considered a broad measure, as accuracy takes into account all utterances with obligatory contexts for the target morphemes and MLUw and NDW are calculated with reference to all complete and intelligible utterances. And yet, it was the streamlined diversity and productivity measures that correlated with MLUw and NDW. These results support the use of the tense marker total and TAP score for children with emerging morphological skills. Conversely, the composite accuracy measure is associated with morpheme mastery and may thus be better suited for capturing later stages of morphosyntactic development (Fitzgerald et al., 2012; Rispoli et al., 2009).

The three morpheme measures were also significantly related to one another. This is to be expected, as the three measures all seek to capture the same expressive language skill. Notably, however, these relationships were stronger at the end of the school year for both groups. The tense marker total and TAP score appear to be appropriate at both testing points, as evidenced by their consistent correlations with MLUw and NDW. The increased correspondence between the tense marker total and TAP score with accuracy at Time 2 may indicate that accuracy has become an increasingly reliable measure as the young bilinguals develop their English T/A system.

Overall, results point to the relevance of the tense marker total and the TAP score for measuring morpheme use in developing bilinguals acquiring English. Support for a composite measure of T/A accuracy was less consistent—particularly at earlier stages of T/A morpheme acquisition (in the case of this study, at Time 1). These findings parallel those found for younger monolingual children, who, like the participants in this study, are acquiring the English T/A system (Fitzgerald et al., 2012; Hadley & Short, 2005; Rispoli et al., 2009).

T/A Morpheme Use Across Groups

This study included two groups of preschool-age Spanish–English developing bilinguals: those with typical development and those with low language skills. The groups were comparable in language exposure, age, and maternal education. As expected, the typically developing group outperformed their peers with reported low language skills on numerous language sample measures, including the developmentally sensitive MLUw and NDW. The diversity and productivity measures—but not the composite accuracy measure—reflected these group differences. BiLL participants used fewer surface forms, as evidenced by lower tense marker totals, and they used the target T/A morphemes less contrastively, as evidenced by lower TAP scores. These observed differences indicate that the tense marker total and TAP score may produce information that is relevant to language assessment in preschool-age children acquiring English.

Conversely, accuracy rates were comparable across children in the typically developing and low language groups. This finding diverged from prior research that measured English T/A accuracy in bilingual children. Gutiérrez-Clellen et al. (2008) found that accuracy did differentiate between typical and atypical language development in young Spanish–English bilinguals. Important differences in participant characteristics may explain this discrepancy. Participants in the Gutiérrez-Clellen et al. study included preschoolers, kindergarteners, and first graders with relatively strong English skills (i.e., received minimum parent ratings of 3 for English use on a scale of 0–4, with "substantial difficulty" speaking Spanish, Gutiérrez-Clellen et al., 2008, p. 8). Participants in this study were generally younger and had greater exposure to Spanish. Overall, the participants in Gutiérrez-Clellen et al. likely had more advanced English language skills, making accuracy a more reliable measure. In fact, their typically developing bilingual participants were 84% accurate on English T/A morphemes, performing well above our groups. The present pattern of results for accuracy may thus be related to our participants' relatively early stage of English morphological development. Variable morpheme marking in bilinguals (e.g., Paradis et al., 2008) and unreliable measures of accuracy due to limited obligatory contexts (e.g., Balason & Dollaghan, 2002) are also relevant. Language assessment measures must be appropriate for a child's background and stage of development. The present results highlight the appropriateness of the tense marker total and TAP score measures for preschool-age bilinguals who are learning the English T/A system.

T/A Morpheme Use Over Time

Both tense marker totals and TAP scores increased from the beginning to the end of the school year for both participant groups. Furthermore, significant increases in the productivity of -3s, -ed, cop *BE,* aux *BE,* and aux *DO* were captured for both BiTD and BiLL participants. This is consistent with significant improvements in both MLUw and NDW from Time 1 to Time 2. The ability to monitor T/A acquisition in terms of both overarching measures (i.e., tense marker totals and TAP scores) and specific morphemes provides versatility that is valuable in clinical contexts.

Conversely, group composite accuracy rates did not improve across the two testing points for BiTD or BiLL participants. We argue that these results are likely not indicative of the children's true abilities and that the children's current morphological development must be considered. The limited number of obligatory contexts in these samples is consistent with our understanding that these Spanish–English bilinguals have emerging morphological skills in English. Likewise, they have yet to meet standard criteria for morpheme mastery (i.e., 80%–95% accuracy; e.g., Brown, 1973). Potentially, at this stage, measuring morpheme use with accuracy—as opposed to the more appropriate diversity and productivity measures—underestimates these children's gains. Under unfortunate circumstances, this could contribute to known problems with overidentifying language disorders in bilingual children (Artiles, Rueda, Salazar, & Higareda, 2005; Bedore & Peña, 2008).

Individual Patterns of Growth in T/A Morpheme Use

For both the tense marker total and the TAP score, individual-level findings mirrored group patterns: Most BiTD and BiLL children earned higher tense marker totals and TAP scores by the end of the year than at Time 1, with relatively little evidence for backtracking. For the composite accuracy measure, however, backtracking was common for both groups: Time 2 accuracy rates were lower for 50% of BiLL children and 42% of BiTD children who met the criterion for three obligatory T/A contexts at Time 1. Given the profile of growth demonstrated by the tense marker total, TAP score, MLUw, and NDW—and considering comparable backtracking in accuracy demonstrated by monolinguals acquiring the English T/A system (Rispoli et al., 2009)—the backtracking in accuracy rates observed here is likely due to the relatively poor fit of this measure for this population at this time.

Case Examples and Clinical Implications

Adding to the findings discussed above, the value of the tense marker total and TAP score is illustrated with two case examples. Isabella and Ruby were two participants matched on age, Spanish exposure, and MLUw in English at Time 1. Critically, Ruby's parent report indicated concern with her language development, but Isabella's did not (see Appendix B).

Parent concern, or lack thereof, was reflected in the children's tense marker totals and TAP scores. At Time 1, Isabella used the T/A morphemes contrastively nine times, whereas Ruby did not produce any. By Time 2, Isabella demonstrated considerable productivity, earning a TAP score of 16. Ruby also made improvements by the second testing point—but still used fewer surface forms in fewer contexts than her typically developing peer did at Time 1.

The diversity and productivity measures captured growth for both children while complementing other language sample measures. Based on MLUw, it might appear that the children had comparable language skills at Time 1. Alternately, referring to composite accuracy scores might lead to concern regarding Isabella's language development. Isabella's Time 2 composite accuracy rate (57.83%) demonstrated backtracking from her Time 1 accuracy rate (75.61%) and was lower than her peer's Time 2 accuracy rate (72.41%). However, these observations are inconsistent with Isabella's relatively high TAP score at Time 2, her notable improvement in productivity from Time 1, and her parents' lack of concern regarding language development. A closer look at the two language samples continued to reveal differences between Isabella and Ruby. Isabella's Time 2 transcript included 83 obligatory contexts, whereas Ruby's relatively high Time 2 accuracy rate was based on only 29 obligatory contexts and was paired with a low tense marker total and TAP score. In this case, calculating a percentage- or proportion-based measure like composite accuracy masked absolute counts and could be misleading if taken on its own, particularly when obligatory contexts are limited. Designed to capture emerging morphological skills, the tense marker total and TAP score help characterize Isabella's and Ruby's productive language.

Valuable information is clearly presented in the scoring tables for the diversity and productivity measures (see Appendix B). In reviewing the tense marker total table, one quickly sees which surface forms are missing from Isabella's language sample. Similarly, a review of her TAP score table allows us to ascertain the depth of her knowledge of each morpheme. This criterion-based approach may be useful for identifying areas of strength and weakness (Stockman, 1996) and for tracking progress in specific areas. Comparing Time 1 and Time 2 scoring tables makes clear which new forms have been demonstrated and whether gains in productivity had been made. In contrast, traditional measures do not provide this type of detailed information. For example, Isabella's composite accuracy measures do not indicate which morphemes or surface forms were used or with what degree of success. The tense marker total and TAP score allow clinicians to quickly access meaningful and specific information about a child's morphological development that may be relevant to assessment, treatment, and progress monitoring.

The diversity and productivity measures complement one another. For example, by noting that Isabella's relatively high Time 1 TAP score (9) is paired with a lower tense marker total (3), a scorer is able to deduce that she used few surface forms, but she used them contrastively and in a variety of contexts. Although Isabella's tense marker total reveals that the only surface form of cop *BE* she produced was *is*, her TAP score indicates that she used that form highly productively. Indeed, her language sample included five different subjects: "*There is a frog.*" "*Where's up?*" "*What story is this?*" "*Is it off?*" and "*That is cars.*" By Time 2, we see relatively high scores for both novel measures, suggesting that we should see contrastive uses across multiple morphemes. Reviewing a child's tense marker total and TAP score provides a clinician with concrete information regarding the child's morpheme use over and above what may be captured with other measures.

This valuable information is acquired through streamlined scoring procedures, making the tense marker total and TAP score a practical means of assessing language skills. These measures focus only on productive morpheme uses (i.e., correct productions and overregularizations) of five morphemes and impose a clear ceiling rule. Furthermore, a number of forms are exempt from scoring, including verbs produced with no subject, repeated subject/surface form combinations, and, for copula and auxiliary verbs, contractions onto pronouns (see Hadley & Short, 2005). Therefore, the number of utterances that a clinician must review is substantially reduced relative to other measures. And yet, this focused approach does not appear to detract from the measures' meaningfulness: The tense marker total and TAP score provided substantial and relevant information about T/A morpheme use in bilingual children.

Future Research and Limitations

We encourage other researchers and clinicians to investigate the tense marker total and TAP score with other young bilinguals. Of note, the participants in this study were preschool-age Spanish–English developing bilinguals from low socioeconomic backgrounds in Southern California. Other bilingual groups, including children being raised in additive bilingual communities or those from higher socioeconomic backgrounds, may demonstrate differing tense marker totals and TAP scores. Relatedly, it would be important to consider the diversity and productivity of T/A forms in developing bilinguals as a function of factors relevant to bilingualism, including relative exposure to the two languages. The present sample was characterized by greater exposure to Spanish (e.g., in each group, the modal reported Spanish exposure at home was 100), limiting our ability to investigate the impact of this important variable. The specifics of our sampled population of English learners notwithstanding, the tense marker total and TAP score were designed to capture initial stages of morpheme emergence and productivity; as such, we anticipate that these measures would similarly track morpheme development in children acquiring English under different conditions, though absolute scores may differ. In addition, the lack of ceiling effects suggests that these measures may be explored in older bilingual children. Much could yet be gained from work, extending use of these measures to bilingual children with different profiles.

In the present study, we identified group differences across typically developing and low language groups, which is indicative of diagnostic potential and is consistent with related research in monolinguals with typical and atypical language development (e.g., Gladfelter & Leonard, 2013; Guo & Eisenberg, 2014). We did not compare children with and without confirmed language impairment. However, parent concern—the criterion used to determine group status in this study—is a valuable indicator of language status (Gutiérrez-Clellen & Kreiter, 2003). Nevertheless, future research should explore the ability of these measures to differentiate bilingual groups with and without confirmed language impairment. Similarly, comparisons with other culturally sensitive measures, such as grammaticality (Bedore et al., 2010; Ebert & Pham, 2017), may serve to bolster the relevance of these measures.

This study would also be strengthened with a larger sample of children with low language skills, particularly considering that the composite accuracy measure could only be calculated for subsets of each group. However, a number of findings indicate that the differential findings for the accuracy measure and the diversity and productivity measures are not a result of reductions in sample size for accuracy analyses. As one example, at Time 2, when all BiLL participants were eligible for accuracy analyses, significant correlations emerged between broad language sample measures and the diversity and productivity measures, but not the composite accuracy measure. In addition, group-level effects (e.g., growth over time) persisted when analyses for the tense marker total and the TAP score were repeated using the reduced participant groups imposed by the accuracy measure. That accuracy could not be reliably calculated for a number of our participants may be interpreted as an indicator that this measure is less appropriate for bilingual children at this stage of English T/A development.

Yet another exciting direction for future research is to consider more closely the use of each individual T/A morpheme. In work completed by Hadley, Rispoli, and colleagues, a robust onset pattern has emerged for young monolinguals: cop *BE* increases in productivity most rapidly, followed by -3s, -ed, and aux *DO*, with aux *BE* demonstrating growth in productivity most slowly (e.g., Rispoli et al., 2012). In the present data, we see relatively high productivity of aux *BE*, a trajectory that diverges from the monolingual data but is consistent with previous work that demonstrated "precocious" use of this morpheme in bilingual children in terms of accuracy (Paradis & Blom, 2016). These findings point to potential qualitative differences in the development of the English T/A system in bilingual children relative to monolinguals, though there may be broad similarities in how the two groups develop morphosyntactic skills gradually (e.g., Rice, 2010). Identifying areas of similarity and contrast across bilingual and monolingual trajectories in acquiring the English T/A system is important for establishing appropriate reference points for clinical settings. Present results also revealed that, in addition to lower tense marker totals and TAP scores, the BiLL group showed lower productivity for -3s and aux *BE* relative to their typically developing peers. Future research can investigate whether these morphemes are particularly sensitive to varying language skills in young bilinguals.

Summary

Results indicate that English morphological development in Spanish–English developing bilinguals can be meaningfully measured with reference to morpheme diversity and productivity. Specifically, the tense marker total and TAP score (Hadley & Short, 2005) converged with traditional language measures, were sensitive to varying language skills, and demonstrated growth over time. Several weaknesses of accuracy measures were identified, suggesting that the diversity and productivity measures may be an important complement to language assessment in bilingual children, particularly for children whose English language skills are emerging. When used appropriately, English language measures may have a valuable place in bilingual language assessment (Bedore et al., 2018; Gillam et al., 2013; Gutiérrez-Clellen et al., 2008), particularly when combined with a parent report and assessment of the native language (e.g., Bedore & Peña, 2008; Gillam, Peña, & Miller, 1999).

Acknowledgments

The first and third authors were supported by the Lipinsky Fellowship at San Diego State University. In addition, funding for the project was provided by a National Institute on Deafness and Other Communication Disorders RO3 grant (Grant DC012141) and a Price Family Philanthropies research grant awarded to the final author. We would like to thank the participants, their families, and the teachers at Rosa Parks Elementary School. We would also like to thank the members of the Child Language Development, Disorders and Disparities Lab for research assistance and helpful discussions of this work. We extend sincere appreciation to Pamela Hadley for graciously sharing training materials for the tense marker total and the TAP score and for her thoughtful feedback on the current research.

References

Armon-Lotem, S. (2014). Between L2 and SLI: Inflections and prepositions in the Hebrew of bilingual children with TLD and monolingual children with SLI. *Journal of Child Language, 41*(1), 3–33.

Artiles, A. J., Rueda, R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English language learners in urban school districts. *Exceptional Children, 71*(3), 283–300.

Balason, D. V., & Dollaghan, C. A. (2002). Grammatical morpheme production in 4-year-old children. *Journal of Speech, Language, and Hearing Research, 45*(5), 961–969.

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*(5), 1185–1192.

Bedore, L. M., & Peña, E. D. (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), 1–29.

Bedore, L. M., Peña, E. D., Anaya, J. B., Nieto, R., Lugo-Neris, M. J., & Baron, A. (2018). Understanding disorder within variation: Production of English grammatical forms by English language learners. *Language, Speech, and Hearing Services in Schools, 49*, 277–291.

Bedore, L. M., Peña, E. D., Gillam, R. B., & Ho, T. H. (2010). Language sample measures and language ability in Spanish–English bilingual kindergarteners. *Journal of Communication Disorders, 43*(6), 498–510.

Blom, E., Paradis, J., & Duncan, T. S. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular *-s* in child L2 English. *Language Learning, 62*(3), 965–994.

Bornstein, M. H., Brown, E., & Slater, A. (1996). Patterns of stability and continuity in attention across early infancy. *Journal of Reproductive and Infant Psychology, 14*(3), 195–206.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Caesar, L. G., & Kohler, P. D. (2007). The state of school-based bilingual assessment: Actual practice versus recommended guidelines. *Language, Speech, and Hearing Services in Schools, 38*(3), 190–200.

Ebert, K. D., & Pham, G. (2017). Synthesizing information from language samples and standardized tests in school-age bilingual assessment. *Language, Speech, and Hearing Services in Schools, 48*(1), 42–55.

Eisenberg, S. L., Guo, L. Y., & Germezia, M. (2012). How grammatical are 3-year-olds? *Language, Speech, and Hearing Services in Schools, 43*(1), 36–52.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing.

Fitzgerald, C., Rispoli, M., Hadley, P., & McKenna, M. (2012). Productivity scoring as a metric of early finiteness marking. *Poster presentation at the Symposium on Research in Child Language Disorders*.

Gillam, R. B., Peña, E. D., Bedore, L. M., Bohman, T. M., & Mendez-Perez, A. (2013). Identification of specific language impairment in bilingual children: I. Assessment in English. *Journal of Speech, Language, and Hearing Research, 56*(6), 1813–1823.

Gillam, R. B., Peña, E. D., & Miller, L. (1999). Dynamic assessment of narrative and expository discourse. *Topics in Language Disorders, 20*(1), 33–47.

Gladfelter, A., & Leonard, L. B. (2013). Alternative tense and agreement morpheme measures for assessing grammatical deficits during the preschool period. *Journal of Speech, Language, and Hearing Research, 56*(2), 542–552.

Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics, 29*(1), 41–65.

Guo, L. Y., & Eisenberg, S. (2014). The diagnostic accuracy of two tense measures for identifying 3-year-olds with language impairment. *American Journal of Speech-Language Pathology, 23*(2), 203–212.

Guo, L. Y., Spencer, L. J., & Tomblin, J. B. (2013). Acquisition of tense marking in English-speaking children with cochlear implants: A longitudinal study. *Journal of Deaf Studies and Deaf Education, 18*(2), 187–205.

Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics, 24*(2), 267–288.

Gutiérrez-Clellen, V. F., Restrepo, M. A., Bedore, L., Peña, E., & Anderson, R. (2000). Language sample analysis in Spanish-speaking children: Methodological considerations. *Language, Speech, and Hearing Services in Schools, 31*(1), 88–98.

Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics, 29*(1), 3–19.

Hadley, P. A., & Holt, J. K. (2006). Individual differences in the onset of tense marking: A growth-curve analysis. *Journal of Speech, Language, and Hearing Research, 49*(5), 984–1000.

Hadley, P. A., Rispoli, M., Holt, J. K., Fitzgerald, C., & Bahnsen, A. (2014). Growth of finiteness in the third year of life: Replication and predictive validity. *Journal of Speech, Language, and Hearing Research, 57*(3), 887–900.

Hadley, P. A., & Short, H. (2005). The onset of tense marking in children at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*(6), 1344–1362.

Heilmann, J. J. (2010). Myths and realities of language sample analysis. *SIG 1 Perspectives on Language Learning and Education, 17*(1), 4–8.

Heilmann, J. J., Miller, J. F., & Nockerts, A. (2010). Using language sample databases. *Language, Speech, and Hearing Services in Schools, 41*(1), 84–95.

Hewitt, L. E., Hammer, C. S., Yont, K. M., & Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders, 38*(3), 197–213.

Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Miller, J., & Iglesias, A. (2012). *SALT: Systematic Analysis of Language Transcripts. Software for the analysis of oral language*. Middleton, WI: SALT Software.

Nicoladis, E., Song, J., & Marentette, P. (2012). Do young bilinguals acquire past tense morphology like monolinguals, only later? Evidence from French–English and Chinese–English bilinguals. *Applied Psycholinguistics, 33*(3), 457–479.

Oetting, J. B., Newkirk, B. L., Hartfield, L. R., Wynn, C. G., Pruitt, S. L., & Garrity, A. W. (2010). Index of Productive Syntax for children who speak African American English. *Language, Speech, and Hearing Services in Schools, 41*(3), 328–339.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools, 36*(3), 172–187.

Paradis, J., & Blom, E. (2016). Do early successive bilinguals show the English L2 pattern of precocious BE acquisition? *Bilingualism: Language and Cognition, 19*(3), 630–635.

Paradis, J., & Crago, M. (2000). Tense and temporality: A comparison between children learning a second language and children with SLI. *Journal of Speech, Language, and Hearing Research, 43*(4), 834–847.

Paradis, J., & Kirova, A. (2014). English second-language learners in preschool: Profile effects in their English abilities and the role of home language environment. *International Journal of Behavioral Development, 38*(4), 342–349.

Paradis, J., Nicoladis, E., Crago, M., & Genesee, F. (2011). Bilingual children's acquisition of the past tense: A usage-based approach. *Journal of Child Language, 38*(3), 554–578.

Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics, 29*(4), 689–722.

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18*(1), 41–58.

Pruitt, S., & Oetting, J. (2009). Past tense marking by African American English–speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research, 52*(1), 2–15.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1398–1411.

Rice, M. L. (2010). Evaluating maturational parallels in second language children and children with specific language impairment. *Applied Psycholinguistics, 31*(2), 320–327.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research, 39*(6), 1239–1257.

Rice, M. L., & Wexler, K. (2001). *Rice/Wexler Test of Early Grammatical Impairment*. San Antonio, TX: Psychological Corporation.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*(4), 850–863.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1412–1431.

Rispoli, M., & Hadley, P. (2011). Toward a theory of gradual morphosyntactic learning. In I. Aron & E. V. Clark (Eds.), *Experience, variation, and generalization: Learning a first language* (pp. 15–34). Amsterdam, the Netherlands: John Benjamins.

Rispoli, M., Hadley, P. A., & Holt, J. K. (2009). The growth of tense productivity. *Journal of Speech, Language, and Hearing Research, 52*(4), 930–944.

Rispoli, M., Hadley, P. A., & Holt, J. K. (2012). Sequence and system in the acquisition of tense and agreement. *Journal of Speech, Language, and Hearing Research, 55*(4), 1007–1021.

Roid, G. H., & Miller, L. (1997). *Leiter International Test of Intelligence–Revised.* Chicago, IL: Stoelting.

Rojas, R., & Iglesias, A. (2009). Making a case for language sampling: Assessment and intervention with (Spanish-English) second language learners. *The ASHA Leader, 14*(3), 10–13.

Simon-Cereijido, G., & Gutiérrez-Clellen, V. F. (2009). A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. *Applied Psycholinguistics, 30*(2), 315–337.

Stockman, I. J. (1996). The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech, and Hearing Services in Schools, 27*(4), 355–366.

Appendix A

Abbreviated Language Sample

Utterance no.	Utterance	Morpheme	Correct/error
1	This is a girl.	cop *BE*	Correct
2	And now we're gonna sit down	aux *BE*	Correct
3	The mama's going up at the car.	aux *BE*	Correct
4	The baby going under.	aux *BE*	Omission
5	Daddy's driving down.	aux *BE*	Correct
6	Daddy's driving!	aux *BE*	Correct
7	That looks like french fries.	-3s	Correct
8	The boy eat french fries.	-3s	Omission
9	They is hungry.	cop *BE*	Agreement error
10	Yesterday he cry.	-ed	Omission
11	But I am tired now.	cop *BE*	Correct
12	He doesn't know.	aux *DO*	Correct
13	I played basketball.	-ed	Correct
14	He played basketball	-ed	Correct
15	It breaked!	-ed	Overregularization

Tense Marker Total: Diversity of the Tense and Agreement System

		aux *DO*			cop *BE*					aux *BE*					
-3s	-ed	do	does	did	is	am	are	was	were	is	am	are	was	were	Total
Utterance 7	Utterance 13		Utterance 12		Utterance 1	Utterance 11				Utterance 3					6

Tense and Agreement Productivity Score: Productivity of the Tense and Agreement System

	-3s	-ed	aux *DO*	cop *BE*	aux *BE*	Total
Instance 1	Utterance 7	Utterance 13	Utterance 12	Utterance 1	Utterance 3	8
Instance 2		Utterance 15		Utterance 11	Utterance 5	
Instance 3						
Instance 4						
Instance 5						

Note. -3s = third-person singular; -ed = past tense; cop *BE* = copula *BE*; aux *BE* = auxiliary *BE*; aux *DO* = auxiliary *DO*.

Appendix B

Case Examples: Isabella and Ruby

Participant	Age (years; months)	Language exposure at home	MLUw		Tense marker total		TAP score		Composite accuracy measure	
			Time 1	Time 2	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Isabella (BiTD)	3;10	100% Spanish	2.36	3.68	3	7	9	16	75.61%	57.83%
Ruby (BiLL)	3;10	100% Spanish	2.19	3.14	0	2	0	4	42.86%	72.41%

Note. MLUw = mean length of utterance in words; TAP score = tense and agreement productivity score; BiTD = developing bilingual children with typical language development; BiLL = developing bilingual children with low language skills.

Isabella's Time 1 Tense Marker Total

		aux *DO*			cop *BE*					aux *BE*					
-3s	-ed	do	does	did	is	am	are	was	were	is	am	are	was	were	Total
✓					✓					✓					3

Isabella's Time 1 Tense and Agreement Productivity Score

	-3s	-ed	aux *DO*	cop *BE*	aux *BE*	Total
Instance 1	✓			✓	✓	9
Instance 2	✓			✓		
Instance 3	✓			✓		
Instance 4				✓		
Instance 5				✓		

Isabella's Time 2 Tense Marker Total

		aux *DO*			cop *BE*					aux *BE*					
-3s	-ed	do	does	did	is	am	are	was	were	is	am	are	was	were	Total
✓	✓		✓	✓	✓			✓		✓					7

Isabella's Time 2 Tense and Agreement Productivity Score

	-3s	-ed	aux *DO*	cop *BE*	aux *BE*	Total
Instance 1	✓	✓	✓	✓	✓	16
Instance 2	✓		✓	✓	✓	
Instance 3			✓	✓	✓	
Instance 4			✓	✓	✓	
Instance 5				✓		

Note. -3s = third-person singular; -ed = past tense; cop *BE* = copula *BE*; aux *BE* = auxiliary *BE*; aux *DO* = auxiliary *DO*.

Chapter 2, in full, is a reprint of material as it appears in Potapova, I., Kelly, S., Combiths, P. N., & Pruitt-Lord, S. L. (2018). Evaluating English morpheme accuracy, diversity, and productivity measures in language samples of developing bilinguals. *Language, Speech, and Hearing Services in Schools*, *49*(2), 260-276. The dissertation author was the primary investigator and author of this paper.

CHAPTER 3:

Spanish-English Bilingual Children's Relative Use of English Tense and Agreement Morphemes

Journal of Monolingual and Bilingual Speech

JMBS (PRINT) ISSN 2514–8370
JMBS (ONLINE) ISSN 2514–8389

Article

Spanish-English Bilingual Children's Relative Use of English Tense and Agreement Morphemes

Irina Potapova
San Diego State University & University of California, San Diego

Sonja L. Pruitt-Lord
San Diego State University

Abstract

Best practice for bilingual speakers involves considering performance in each language the client uses. To support this practice for young clients, a comprehensive understanding of how bilingual children develop skills in each language is needed. To that end, the present work investigates relative use of English tense and agreement (T/A) morphemes – a skill frequently considered as part of a complete language assessment – in Spanish-English developing bilingual preschoolers with varying levels of language ability. Results indicate that developing bilingual children with both typical and weak language skills demonstrate greater use of copula and auxiliary *BE* relative to third person singular, past tense and auxiliary *DO*. Findings thus reveal a relative ranking of T/A morphemes in developing bilingual children that differs from that of English monolingual children, who demonstrate relatively later emergence and productivity of auxiliary *BE*. In turn, findings demonstrate the importance of utilizing appropriate comparisons in clinical practice.

Affiliation

San Diego State University, School of Speech, Language, and Hearing Sciences, 5500 Campanile Drive, San Diego, CA 92182-1518, USA
Email: ipotapova@sdsu.edu

JMBS VOL 1.1 2019 118–142
doi: https://doi.org/10.1558/jmbs.11125

Irina Potapova and Sonja L. Pruitt-Lord

Keywords: bilingualism; tense and agreement morphology; morphological development; precocious BE; typical bilingual language development; low language skills

Introduction

Approaches to bilingual language assessment and treatment are improving in the field of speech-language pathology (e.g. Rojas & Iglesias, 2009; Kohnert, 2010; Gillam, Peña, Bedore, Bohman & Mendez-Perez, 2013; Fabiano-Smith & Hoffman, 2018). One important advance has been the recognition that language assessment must account for performance in each language the client uses. To support this practice for young clients, the field requires more detailed information about how bilingual children develop skills in each language, as we cannot assume that monolingual norms are appropriate for bilingual children (Paradis, 2005). With the use of appropriate bilingual reference groups and careful measurement of abilities in a single language, we can arrive at clinically relevant information (Gillam et al., 2013; Potapova, Kelly, Combiths & Pruitt-Lord, 2018) which, in combination with other culturally and linguistically sensitive approaches (e.g. parent interview; Paradis, Emmerzael & Duncan, 2010), contributes to a holistic bilingual language evaluation.

In characterizing how bilingual children develop skills in each language, it is important and practical to focus on areas that are prevalent in language assessment. For monolingual and bilingual preschool-age and early school-age children who speak English, this includes use of English tense and agreement (T/A) morphemes (monolingual: Gladfelter & Leonard, 2013; Bedore & Leonard, 1998; Rice & Wexler, 1996; bilingual: Gutiérrez-Clellen, Simon-Cereijido & Wagner, 2008; Potapova et al., 2018), including third person singular (-3s; she *walks*), past tense (-ed; *he jumped*), copula *BE* (cop *BE*; *I am fast*), auxiliary *BE* (aux BE; *they are going*), and auxiliary DO (aux *DO*; *Do they run?*). Across languages, weaknesses in morphosyntax are characteristic of developmental language disorder (DLD; also known as specific language impairment, primary language impairment or language impairment; e.g. Leonard, 2014); in English, T/A morphemes have been identified as presenting a particular challenge for monolingual and bilingual children with DLD compared to other grammatical features (e.g. Rice, Wexler & Cleave, 1995; Gutiérrez-Clellen et al., 2008). While the development of English T/A marking has been relatively widely studied in monolingual children, developmental patterns are less clearly understood in young developing bilinguals.

As such, the present study aims to clarify the relative use of these key morphemes in preschool-aged Spanish-English developing bilinguals with varying levels of language ability. Results will support clinical decision-making by providing information about English-language performance in developing bilingual children and, in turn, allowing us to determine whether bilingual trajectories in T/A use differ from those of monolingual peers. Presently, we use a productivity-based approach to measuring morpheme use that is appropriate for monolingual and bilingual children who are developing skills in English T/A use (Hadley & Short, 2005; Potapova et al., 2018).

English tense and agreement morphemes in monolingual and bilingual children

Typically developing monolingual English children begin using T/A morphemes – -3s, -ed, cop *BE*, aux *BE* and aux *DO* – around two years of age (Hadley & Short, 2005; Rice, 2010). English T/A marking in obligatory contexts increases over the next two years of life, as children demonstrate increasingly adult-like productions (Goffman & Leonard, 2000; Rice, Wexler & Hershberger, 1998). It has long been recognized that use of each morpheme category is not mastered simultaneously (e.g. Brown, 1973). Instead, converging evidence indicates that T/A development in English monolinguals is characterized by relatively early emergence and productive use of cop *BE*, and relatively later emergence and productive use of aux *BE*.

Developmental trends in morpheme use have been clarified with a productivity-based approach offered by Hadley, Rispoli and colleagues (e.g. Hadley & Short, 2005; Hadley & Holt, 2006; Rispoli, Hadley & Holt, 2009; Hadley, Rispoli, Holt, Fitzgerald & Bahnsen, 2014). The T/A Productivity Score (TAP score), first described in Hadley and Short (2005), is a type-based measure that awards points to children for contrastive uses of each of the five T/A morpheme categories in spontaneous language sample (see the present Methods section and Hadley & Short, 2005, for further details on this measure). Critically, this approach has been shown to mitigate limitations of accuracy-based measures for children in the early stages of English acquisition and provide diagnostically relevant information about a child's morphological development (Rispoli et al., 2009).

Using sub-scores for each morpheme category taken from the TAP score, Rispoli, Hadley & Holt (2012) captured initial instances of T/A marking, or T/A emergence, in monolingual English-speaking children followed longitudinally between 21 and 33 months. Across monolingual children with typical language development, use of cop *BE* consistently emerged first, with aux *BE* trailing behind the remaining morphemes. This relative ranking was maintained across multiple testing points, and their pattern of emergence may be summarized as: cop *BE* > -3s, -ed, aux *DO* > aux *BE*. T/A use may also be measured with regard to proportion of correct use, or accuracy rates. For example, Paradis, Rice, Crago and Marquis (2008) found that monolingual

English-speaking children in a similar age range (2;6–3;8) produced cop and aux *BE* more successfully than aux *DO* in the context of a structured probe, with no difference in use from -3s or -ed. Though this set of findings is consistent with the understanding that different morphemes are used with differing degrees of success during language development, this analysis did not allow for a direct comparison between each morpheme category: cop and aux *BE* were collapsed into a single category, as were -3s and -ed.

In older monolingual children, a productivity-based approach continues to provide clinically relevant information about language development and helps capture use of specific morpheme categories (e.g. Guo & Eisenberg, 2014). As children surpass the initial stages of T/A use, the TAP score helps capture productivity, or T/A use in increasingly varied morphosyntactic contexts. Like Rispoli et al. (2012), Gladfelter and Leonard (2013) used sub-scores from the TAP score to compare relative use of the five T/A morpheme categories in children aged 4;0–4;6 and 5;0–5;6 with typical language development and peers with DLD. Results revealed the same relative rankings for cop and aux *BE*: for both groups, cop *BE* was the most frequently used morpheme, and aux *BE* was among the least productively used. Children in the typically developing group further demonstrated greater productivity of -3s and aux *DO* relative to aux *BE* and -ed, such that their pattern of use could be summarized as: cop *BE* > -3s, aux *DO* > aux *BE*, -ed. Children with DLD demonstrated greater productivity of -3s than aux BE, such that their pattern might be described as: cop *BE* > -3s > aux *DO*, aux *BE*, -ed. Thus, the two groups evinced patterns that were remarkably similar, with the children with DLD demonstrating an expected delay relative to their typically developing counterparts (e.g. Rice, 2010). These findings confirmed the relative ranking of use between the morpheme categories demonstrated by younger monolingual children. Taken together with earlier work, these results demonstrated that additional significant distinctions between morpheme categories may emerge as children develop.

Recent work indicates that the TAP score, a cumulative score summing up performance across five English T/A categories, is also an appropriate measure for preschool-aged English-speaking developing bilinguals (Potapova et al., 2018). However, the developmental trajectory of individual T/A morpheme use has yet to be investigated in developing bilingual children utilizing this approach. Instead, our understanding of T/A morpheme use in young bilinguals acquiring English more frequently relies on accuracy rates and grammaticality judgements in structured probes. In this work, developmental patterns for typically developing bilingual children learning a second language

(L2) have been shown to diverge from the monolingual trajectory.* Specifically, a pattern of 'precocious BE' has been identified in typically developing children learning English as a second language from a large array of language backgrounds (Paradis, 2010). For example, where Paradis et al. (2008) found that monolingual children performed comparably on a composite of cop *BE* and aux *BE* and a composite of -3s and -ed, the same study revealed that typically developing bilingual children with English as an L2 (4;2–7;10) and with eight different native languages produced cop and aux *BE* more successfully than -3s and -ed. Emerging evidence further indicates that precocious *BE* is mirrored in bilingual children with English as an L2 and with weak language skills – Paradis (2008) found that two bilingual children with DLD consistently demonstrated higher performance on a composite of cop *BE* and aux *BE* relative to a composite of -3s and -ed over time (see also Paradis 2010, 2016).

Other work has utilized error analyses in language samples to identify patterns of T/A use in bilingual children. Using this approach, Ionin and Wexler (2002) provide further evidence for preferential use of cop and aux *BE*: Russian-speaking children (3;9–13;10) who were learning English omitted cop and aux *BE* less frequently than -3s and -ed in spontaneous language samples, though comparable analyses for aux *DO* were unavailable. Gutiérrez-Clellen et al. (2008) importantly demonstrated that accuracy of T/A marking differed across Spanish-English bilingual children (4;5–6;5) with typical development and those with DLD, and provided accuracy rates for each English T/A morpheme category. Consistent with related work, cop *BE* was observed to be among the most successfully used morphemes and -3s was among the least in a sample that included both English-dominant bilingual children and Spanish-speaking English language learners. However, statistical analyses were not employed to test differences across categories.

Altogether, available research suggests a tentative ranking of cop *BE*, aux *BE* > -3s, -ed, aux *DO* for developing bilinguals. However, given methods employed in prior work, it is not yet clear if further distinctions between morpheme categories may exist. Moreover, explorations of relative English T/A use in bilingual children have yet to employ productivity-based measures, which have been shown to be valuable for children who are acquiring English. Presently, we build upon efforts in Potapova et al. (2018) and use the TAP score to measure productivity of individual T/A morpheme categories and test for specific distinctions between them in Spanish-English bilingual children with varying language skills. This approach both complements existing work in T/A use in bilingual children acquiring English and facilitates comparisons

* Of note, simultaneous bilingual children have been shown to demonstrate morphosyntactic development that closely resembles that of monolingual peers (see De Houwer, 2009).

with work featuring monolingual children. Recognizing the relevance of morphosyntactic performance – including T/A use in English speakers – to developmental language disorder, it is important to understand young bilinguals' patterns of performance to inform clinical practice.

Present study

The present study aims to enhance our understanding of English T/A morpheme use in developing bilingual children. Participants included Spanish-dominant preschoolers with typical language abilities, as well as bilingual peers with relatively weak language skills. Our first aim was to identify a ranking of relative use of the five English T/A morphemes in young bilinguals using a productivity-based approach. Thus, our first research question was (1) *What is the relative productivity of English T/A morphemes in Spanish-English developing bilingual children with varying language skills?* To answer this question, we compared use of each target morpheme category in spontaneous language samples using the TAP score (Gladfelter & Leonard, 2013; Potapova et al., 2018) at the beginning and end of the academic year. Next, we considered a more conservative test of relative rankings between the target morphemes and asked (2): *What is the pattern of emergence of T/A morphemes in Spanish-English developing bilingual children with varying language skills?* To answer this, we identified children who began the year with no observed uses of any target English T/A morphemes and evaluated their T/A performance at the end of the school year. By focusing on children who began the school year with relatively limited productive English skills, results may better reflect an earlier stage of development of English morphological skills, or T/A emergence.

It was predicted that differences in productivity and emergence would be identified across morpheme categories. Specifically, cop *BE* and aux *BE* were anticipated to be the most productive, consistent with prior work in bilingual children learning English (Paradis & Blom, 2016; Ionin & Wexler, 2002) and unlike the established monolingual trajectory, in which aux *BE* emerges last (Rispoli et al., 2012; Gladfelter & Leonard, 2013).

Method

Participants

Participants were preschool-aged Spanish-English developing bilinguals identified from an ongoing community-based research project under direction of the second author. Data were analysed for children who met the following criteria: Spanish exposure of at least 30% of the time in the home (Pearson, Fernandez, Lewedeg & Oller, 1997); non-verbal cognition scores within or above normal limits; language samples completed at the beginning and end of the school year; and caregiver questionnaires completed during data collection.

In total, 132 Spanish-English developing bilingual preschoolers (mean age = 4;2, *SD* = 5.33 months; 59 females) were included in the study. Per parent questionnaires, participants were exposed to Spanish 72.04% of the time at home, on average (*SD* = 20.32, range = 33.33–100). All participants were recruited from the same preschool site in an English-speaking school setting that required below-poverty standards to participate, with data collection for the first testing point completed within two months of the start of the academic year. Average performance on the Figure Ground and Form Completion subtests of the *Leiter International Performance Scale – Revised* (Roid & Miller, 1997), a non-verbal cognition measure, was within normal limits (M = 11.67, *SD* = 1.97, range = 7–16.5). Maternal education was reported by 78 participants and was 9.97 years (*SD* = 2.86, range = 3–16), on average.

Participants were subsequently considered for inclusion in one of two groups, developing bilingual children with typical language development (BiTD) or developing bilingual children with low language skills (BiLL), on the basis of caregiver report (Gutiérrez-Clellen & Kreiter, 2003; Restrepo, 1998). Caregivers provided information about their child's language background and history via questionnaire in their preferred language. Children were assigned to the BiTD group if the caregiver report indicated no concerns with language development. This group included 100 children (mean age = 4;1, *SD* = 5.6 months, range = 3;0-5;6). Conversely, children whose caregivers

Table 1. Participant characteristics for bilingual children with typical development (BiTD) and bilingual children with low language (BiLL) at the beginning (Time 1) and end (Time 2) of the academic year

			Background			*Broad language sample measures*					
						MLUw		**NDW**		**Complete and intelligible utterances**	
		Age	**% Spanish heard**	**% English heard**	**Maternal education** (in years)	Time 1*	Time 2*	Time 1*	Time 2	Time 1*	Time 2
BiTD	*Mean*	4;1	73.30	26.70	9.95	2.64	3.30	95.07	132.92	104.66	144.58
	SD	5.6 months	20.08	20.08	2.92	.94	.87	53.71	48.79	64.11	67.73
BiLL	*Mean*	4;2	68.08	31.92	10.06	2.28	2.96	76.91	129.7	77.25	138.43
	SD	4.5 months	20.89	20.89	2.70	.79	.75	42.84	41.43	43.74	53.72

*$p < .05$

did report concerns with language performance were considered for the BiLL group. This group included 32 children (mean age = 4;2, *SD* = 4.5 months, range = 3;5–5;5). Group differences in language ability were confirmed with consideration of language sample performance. Children in the BiTD group demonstrated significantly higher performance than counterparts in the BiLL group in mean length of utterance in words at the beginning and end of the year ($ps < .03$, as evidenced by one-tailed t-tests; Rojas & Iglesias, 2009). Predicted group differences were also found at the beginning of the academic year for number of different words, number of complete and intelligible utterances and type–token ratio (all $ps < .042$, as evidenced by one-tailed t-tests; see Table 1). Notably, BiTD and BiLL participants were comparable in characteristics that may be relevant to language performance, including age, Spanish and English exposure at home, and maternal education (all $ps > .226$; see Table 1).

Procedure

With support from teachers and classroom personnel, caregivers received information about the larger study, consent forms, and caregiver questionnaires. Children whose caregivers provided signed consent forms were then eligible to complete an assessment battery associated with the larger research project. Data collection took place at the preschool at both the beginning (Time 1) and end (Time 2) of the academic year, with each wave of data collection completed in two to four weeks and multiple sessions planned for each participant. Children were tested individually, and session length was determined by child engagement. Examiners included supervised graduate and undergraduate students in speech-language pathology trained to administer the assessment battery, including collecting spontaneous language samples.

Measures

Morpheme use was measured with the TAP score, a productivity-based measure that provides sub-scores for each morpheme category and which has been utilized to answer similar research questions for monolingual children (Hadley & Short, 2005; Rispoli et al., 2012). TAP scores and, accordingly, morpheme category sub-scores, were derived from language samples collected at Time 1 and at Time 2. Language samples were elicited following a set play protocol, using toy car, garage, and picnic sets, and a standard set of pictures for story retells. Digitally recorded language samples were orthographically transcribed and coded by trained research assistants following conventions for the Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2012) software. Time 1 samples included 97.92 complete and intelligible child utterances (*SD* = 60.75), on average, and Time 2 samples included 145.16 (*SD* = 64.63), on average.

Morpheme category sub-scores

TAP scores were determined following Hadley and Short (2005), with points awarded for contrastive, or *sufficiently different*, uses of -3s, -ed, cop *BE*, aux *BE* and aux *DO*. To meet this criterion, children are required to use different lexical verbs for the bound morphemes (-3s, -ed) to earn points. For the copula and auxiliary verbs, children are required to produce different combinations of subjects (e.g. *the sister is* and *the brother is*) or surface forms of the morpheme (e.g. *the sister is* and *the sister was*). Correct uses and over-regularizations are eligible for scoring, but, critically, repeated uses of lexical verbs (e.g. *he eat/3s*, *she eat/3s*) or subject and surface form combinations (*the baby is tired*; *the baby is hungry*) do not contribute to the child's score. Similarly, no points are awarded for errored productions (e.g. morpheme omissions, tense and agreement errors), nor for forms that are less likely to indicate grammatical processing, including contracted copula and auxiliary verbs used with pronouns (e.g. *he's going;* see Hadley & Short, 2005 for detailed scoring criteria; see also Rispoli & Hadley, 2011).

Children may earn up to five points per morpheme category, yielding a TAP score ranging between zero and 25 points and morpheme category sub-scores ranging from zero to five points. Given the present goal of identifying patterns of use across morpheme categories, the measures of interest were the morpheme sub-scores. Sub-scores and total TAP scores were derived by trained graduate research assistants for each participant at each time point. SALT codes were used to extract relevant utterances and facilitate scoring. Reliability, completed for over 10% of the samples, was 93.85% for TAP scores.

Results

What is the relative productivity of T/A morphemes in Spanish-English developing bilingual children?

The relationship between productivity (i.e. morpheme sub-score) and morpheme category (-3s, -ed, cop *BE*, aux *BE*, aux *DO*) was investigated with linear mixed effects models run separately for performance at Time 1 and at Time 2 using R (R Core Team, 2013) and the package 'lme4' (Bates, Maechler, Bolker & Walker, 2015). Fixed effects included group status (BiTD, BiLL) and morpheme category, allowing for their interaction. When morpheme category was found to be a significant predictor of morpheme sub-score, specific relationships between the five target morphemes were determined using the package 'multcomp' (Hothorn, Bretz & Westfall, 2008). In addition, models included fixed effects for participant characteristics expected to be associated with English-language performance: percentage exposure to English at home and chronological age. Subject intercepts were entered as random effects, and p-values for tested effects were obtained by likelihood ratio tests of the model with the targeted effect included against a reduced model without that effect.

In service of identifying group-level patterns, outlier analyses were conducted. Outlying sub-scores for each morpheme category were identified as observations that fell 1.5 times the interquartile range above the third quartile, or 1.5 times below the first quartile within each group. For each time point, there were a total of 660 observations (five observations for each of 132 participants). At Time 1, a total of 5.6% of all observations were excluded across the two groups; at Time 2, 3.3% of observations were excluded.* If a participant had an outlying data point in a morpheme category, linear mixed effects modelling allowed for their remaining observations to contribute to analyses; that is, the exclusion of individual observations did not result in the exclusion of that participant. The resulting morpheme sub-scores for each participant group are provided in Table 2.

Table 2. Morpheme sub-scores for developing bilingual children with typical development (BiTD) and developing bilingual children with low language (BiLL) at Times 1 and 2

		Time 1					*Time 2*				
		cop *BE*	**aux *BE***	**-3s**	**-ed**	**aux *DO***	**cop *BE***	**aux *BE***	**-3s**	**-ed**	**aux *DO***
BiTD	*Mean*	1.93	1.18	.47	.39	.33	2.94	2.62	1.43	1.46	.47
	SD	1.99	1.70	.73	.63	.55	1.98	1.87	1.51	1.60	.68
BiLL	*Mean*	1.47	.29	.29	.47	0	2.50	1.72	.57	.58	.49
	SD	1.81	.60	.53	.80	0	1.72	1.63	.97	.58	.69

Time 1: Morpheme productivity

All fixed effects were found to be significantly predictive of morpheme sub-scores. BiTD children outperformed BiLL participants, $F(1, 123.91) = 6.275$, $p = .014$. As expected, morpheme productivity increased with greater English exposure, $F(1, 123.56) = 6.957$, $p = .009$, and age, $F(1, 121.69) = 9.431$, $p = .003$. Critically, the effect of morpheme category was significant, $F(4, 488.41) = 31.237$, $p < .001$, indicating that there were differences in sub-scores across

* *BiLL group, Time 1*: -3s, .6% of observations; -ed, 0%; cop *BE*, 0%; aux *BE*, 2.5%; aux *DO*, 3.1%; *BiTD group, Time 1*: -3s, 2%; -ed, 0%; cop *BE*, 2.4%; aux *BE*, 0%; aux *DO*, 1%; *BiLL group, Time 2*: -3s, 1.3%; -ed, 3.8%; cop *BE*, 0%; aux *BE*, 0%; aux *DO*, 1.9%; *BiTD group, Time 2*: -3s, 0%; -ed, 0%; cop *BE*, 0%; aux *BE*, 0%; aux *DO*, 2.2%.

the five target morphemes and that it would be possible to identify a relative ranking between them.

The interaction between morpheme category and group was not significant (p = .086). Relationships between the multiple levels of morpheme category were thus compared using a more parsimonious model which included the significant fixed effects and no interaction. Results indicated that cop *BE* was significantly more productive than all other morpheme categories (ps < .001 for all comparisons) and that aux *BE* was significantly more productive than -3s, p = .002; -ed, p < .001; and aux *DO*, p < .001. The remaining morphemes (-3s, -ed and aux *DO*) did not significantly differ in scores (ps > .474). The pattern of relative productivity at Time 1 may be summarized as: cop *BE* aux *BE* > -3s, -ed, aux *DO* (see Figure 1).

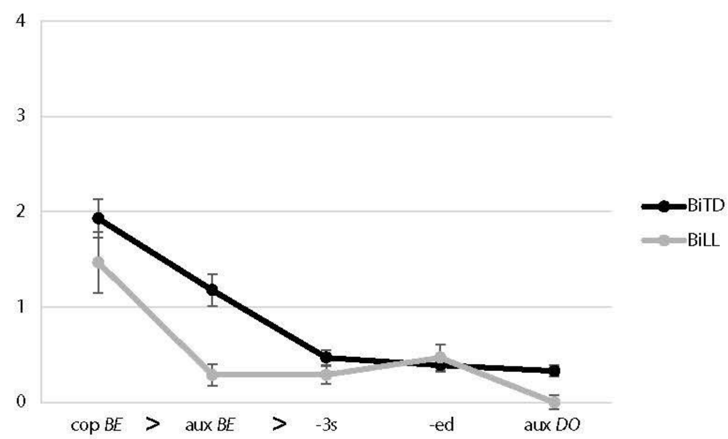


Figure 1. Relative productivity of English T/A morphemes for BiTD and BiLL groups at the beginning of the school year (Time 1). Error bars represent one standard error of the mean. Significant differences in use across morpheme categories are represented with '>'

Time 2: T/A morpheme productivity

At the end of the academic year, BiTD children continued to outperform BiLL participants, $F(1, 131.50) = 10.349$, $p = .002$, and productivity increased with age, $F(1,127.20) = 24.798$, $p < .001$. Again, morpheme category significantly predicted morpheme scores, $F(4, 508.95) = 48.600$, $p < .001$, indicating that children earned higher scores for some morphemes than for others. Neither English exposure, as indexed by caregiver report at the beginning of the year, nor the interaction between morpheme category and group were significant (ps > .187).

As such, analyses to determine relative productivity of the target T/A morphemes were conducted on a reduced model which excluded the morpheme category by group interaction and English exposure. Again, cop *BE* was significantly more productive than all other morpheme categories (*p*s < .001 for all comparisons, except *p* = .041 for aux *BE*) and aux *BE* was significantly more productive than the remaining morphemes (*p*s < .001). Productivity for -3s and -ed did not significantly differ, and each morpheme was significantly more productive than aux *DO* (*p*s < .001). The resulting relative ranking of productivity was cop *BE* > aux *BE* > -3s, -ed > aux *DO*. See Figure 2.

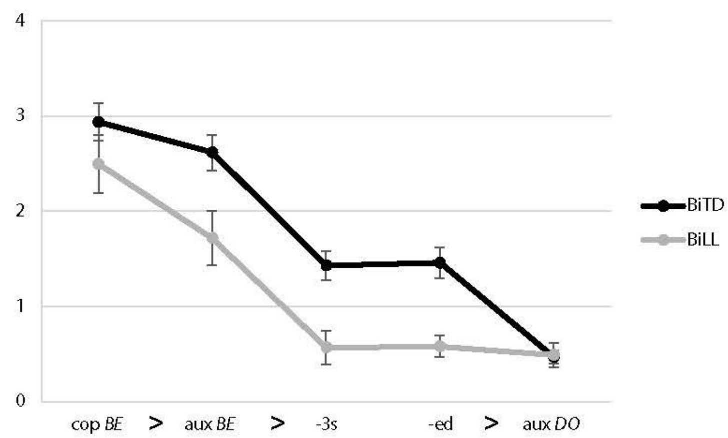


Figure 2. Relative productivity of English T/A morphemes for BiTD and BiLL groups at the end of the school year (Time 2). Error bars represent one standard error of the mean. Significant differences in use across morpheme categories are represented with '>'

What is the pattern of emergence of T/A morphemes in Spanish-English developing bilingual children?

To better isolate patterns of T/A emergence, the larger data set with all observations included (i.e. prior to outlier removal) was refined to select a subset of children who demonstrated no productive uses of the target morphemes at the beginning of the year. It was reasoned that, by highlighting these participants and considering their morpheme use at Time 2, we would more stringently test for relative rankings between morpheme categories. Participants who met the criteria for this subset included 14 BiLL children and 30 BiTD children. All analyses paralleled those for the first research question.

As with the larger sample, we began with outlier analyses to clarify group-level patterns within the subset. Of the 220 observations available for 44 participants, 7.7% of observations were excluded at Time 2;* as before, exclusion of an observation from a participant did not require the removal of that participant from analyses. Analyses were conducted solely for Time 2 data, as all participants, by design, demonstrated equivalent performance at the beginning of the year (i.e. sub-scores of zero for each morpheme category). The resulting morpheme sub-scores are provided in Table 3.

Results indicated that, at the end of the academic year, morpheme category significantly predicted morpheme sub-scores, $F(4, 161.897) = 17.098$, $p < .001$, and this effect did not interact with language group ($p = .562$). The effect of language exposure and group were not significant ($p = .542$, $p = .099$, respectively), but chronological age approached significance ($p = .073$). Thus, an updated model reflected the significant predictors from the Time 2 model for the full data set: fixed effects of language group, morpheme category, and chronological age. Morpheme category continued to be a significant predictor, $F(1, 160.833) = 22.147$, $p < .001$, and both language group and chronological age were marginally significant ($p = .074$ and $p = .072$, respectively).

Table 3. Morpheme sub-scores at Time 2 for the subset of children demonstrating no productive uses of target T/A morphemes at the beginning of the year, including developing bilingual children with typical development (BiTD) and developing bilingual children with low language (BiLL)

				Time 2		
		cop *BE*	**aux *BE***	**-3s**	**-ed**	**aux *DO***
BiTD	*Mean*	1.8	1.8	0.62	0.46	0
	SD	1.97	1.67	0.77	0.74	0
BiLL	*Mean*	1.33	1.14	0	0.36	0
	SD	0.98	1.29	0	0.5	0

Subsequent analyses were run on this reduced model to characterize relative emergence of the target T/A morphemes. Results indicated that cop *BE* and aux *BE* were more productive than the remaining morphemes (*p*s < .001), and the remaining morphemes did not significantly differ in productivity (*p*s > .307). Patterns of relative emergence may be summarized as: cop *BE*, aux *BE* > -3s, -ed, aux *DO* (see Figure 3). This pattern was maintained when a fully reduced model, including only morpheme category as a fixed effect, was used: cop and aux *BE* were more productive than the remaining morphemes

* *BiLL group form subset, Time 2*: -3s, 4.3% of observations; -ed, 4.3%; cop *BE*, 2.9%; aux *BE*, 0%; *DO*, 2.9%; *BiTD group from subset, Time 2*: -3s, .7%; -ed, 1.3%; cop *BE*, 0%; aux *BE*, 0%; *DO*, 2.7%.

($ps < .001$), and the remaining morpheme sub-scores did not differ from one another ($ps > .296$).

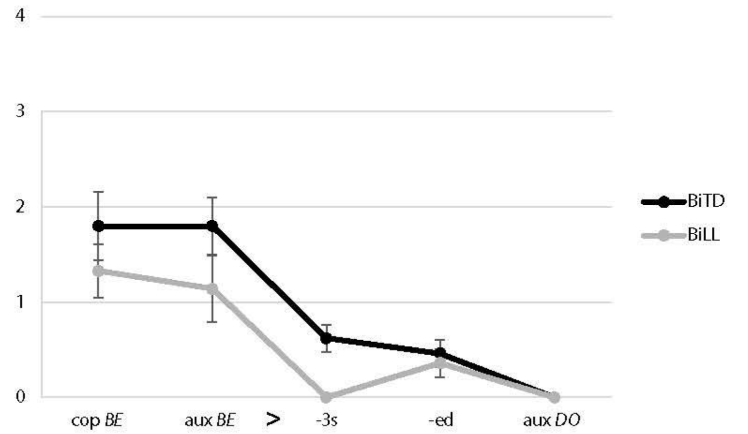


Figure 3. Relative emergence of English T/A morphemes for BiTD and BiLL groups at the end of the school year (Time 2). Error bars represent one standard error of the mean. Significant differences in use across morpheme categories are represented with '>'

Discussion

Bilingual language development is characterized by notable variability (Paradis, 2005) – and clinicians are tasked with evaluating these variable profiles to accurately identify language disorders within patterns of differences in language use associated with bilingualism (Oetting, 2018). Though there are instances of common ground in language development between monolingual and bilingual children (e.g. both groups tend to acquire non-tense morphemes such as plural -s and progressive -ing more readily than T/A morphemes; Paradis, 2005), parity in developmental patterns between bilinguals and monolinguals cannot be assumed. Accordingly, we must critically investigate developmental patterns within bilingual children to help provide an evidence base for clinical practice. Further, a careful consideration of how bilinguals perform in one language offers opportunities to consider issues of theoretical relevance, including cross-linguistic influence, a phenomenon observed across multiple linguistic domains for child and adult bilingual speakers (e.g. Potapova & Pruitt-Lord, 2019; De Houwer, 2018; Kehoe, 2018; Liceras, Fuertes & de la Fuente, 2012; Paradis, 2001; Nicoladis, 2006; MacWhinney, 2005).

In the present work, we aimed to provide detailed information on how a key marker of language impairment in English, T/A morpheme use, is demonstrated in developing bilingual children at an age when language assessment is common (National Institute on Deafness and Other Communication Disorders, 2016). Our analyses allowed us to determine a relative ranking of use between the morpheme categories and, in turn, compare this bilingual pattern with established monolingual patterns. Recall that substantial evidence indicates that, for English-speaking monolingual children, the developmental trajectory of T/A morphemes is cop *BE* > -3s, -ed, aux *DO* > aux *BE*. This pattern has been identified in the earliest stages of T/A acquisition, or emergence, in monolingual English-speaking children (Rispoli et al., 2012), with a similar pattern persisting for monolinguals aged 4;0–5;6 (Gladfelter & Leonard, 2013). Meanwhile, available work indicates that bilingual children learning English demonstrate a different pattern of typical English T/A development, with enhanced use of both *BE* forms (Paradis & Blom, 2016; Ionin & Wexler, 2002). We confirm this finding here in Spanish-English developing bilingual preschoolers, with consistently high use of both forms of *BE* relative to the remaining categories: cop *BE*, aux *BE* > -3s, -ed, aux *DO*. Further, our work indicates that young developing bilinguals' patterns of English T/A use may grow to include additional distinctions between morpheme categories: cop *BE* > aux *BE* > -3s, -ed > aux *DO*.

The current findings contribute to our understanding of English T/A development in bilingual children in several ways. First, our inclusion of a large sample of children with weak language skills facilitates the extension of such investigations into clinical practice. In addition, this work both includes all five target morphemes and maintains distinctions between each morpheme category throughout analyses. Our use of spontaneous language samples is not only consistent with best practice for culturally and linguistically diverse children (Stockman, 1996), it also allows for measurement of forms that may not be included in structured probes. Further, while use of productivity-based measures have been shown to be appropriate for bilingual children in the early stages of learning English (Potapova et al., 2018), the present paper utilizes this approach to consider relative productivity of individual morpheme categories, expanding the methods used to investigate T/A development in bilingual children. Finally, we offer analyses of performance at two time points, with an effort to capture both relative productivity of T/A marking as well as T/A emergence. Our work thus supplements methods employed in studies of developing bilingualism and supports comparisons with recent work in monolingual children.

Relative use of English T/A morphemes in Spanish-English bilingual children

Across analyses, a similar pattern emerged for the five target morphemes, with additional levels of granularity offered by analyses at multiple time points and with varying analysis groups for our two research questions. One distinction between morpheme categories was evinced across all three sets of analyses (the complete sample at Time 1, the complete sample at Time 2, and the subset at Time 2): cop and aux *BE* were used more productively than the remaining morphemes, -3s, -ed and aux *DO* (cop *BE*, aux *BE* > -3s, -ed, aux *DO*).

Observed here in the participants' spontaneous language samples, this distinction between both forms of *BE* and the remaining morphemes is consistent with prior work in developing bilingual children that utilized structured probes (e.g. Paradis 2008; Paradis et al., 2008; Paradis, 2010; Paradis & Blom, 2016). One possibility is that this distinction is the most robust, as no other distinctions were evinced in analyses for our second research question. However, this analysis utilized a subset of children who, by design, were in the particularly early stages of acquiring English T/A morphemes. Not only were these participants selected on the basis of no productive uses at the beginning of the school year, but our sample as a whole was characterized by greater exposure to Spanish at home. Further, it is broadly representative of developing bilinguals in the United States that the first substantial exposure to English occurs upon entering the school environment (Bedore & Peña, 2008). This distinction in morpheme categories may thus be interpreted as the earliest to emerge. Additionally, the subset represented one-third of our larger sample, potentially resulting in limited power to capture further distinctions across morpheme categories. Indeed, factors that consistently predicted morpheme use in analyses for the complete sample (i.e. participant group and chronological age) were only marginally significant for the subset. Moreover, because of the subset selection criteria, these participants may have demonstrated limited performance at Time 2 which restricted the predictive potential of those factors. Despite using a smaller sample characterized by relatively low productivity in English, the results of our second research question indicated that there was a meaningful difference in Spanish-English bilingual children's productive use of cop and aux *BE* relative to the remaining morphemes as T/A skills are emerging.

Within our larger sample, further distinctions between morpheme categories were identified, particularly when followed over time. By maintaining categorical differences between cop *BE* and aux *BE* in our analyses, we were able to observe that cop *BE* was significantly more productive than aux *BE* at both the beginning and the end of the academic year (cop *BE* > aux *BE* > -3s, -ed, aux *DO*) for the complete sample. At Time 2, an additional distinction was realized: aux *DO* significantly trailed behind all other morpheme categories (cop *BE* > aux *BE* > -3s, -ed > aux *DO*). Of note, the participants were clearly below ceiling on morpheme sub-scores (and overall TAP scores) at both testing points, indicating that, as a whole, participants were in the process of developing morphological skills (see Gladfelter & Leonard, 2013, for similar findings in preschool-aged monolinguals). The variability in morpheme productivity within each group (see Rispoli et al., 2012) further indicates that the participants were in a stage of developing English language skills, with some children demonstrating higher productivity than others. Indeed, children from Spanish-dominant homes entering school with English as the language of instruction can be expected to be in the process of acquiring English language skills more broadly and to move towards mastery of those skills as they advance in school (Rojas & Iglesias, 2013; Pham & Kohnert, 2014). As such, analyses including the entire sample may be also be considered reflective of developmental patterns. Present findings thus indicate that morphological development in young bilinguals is both simultaneous and sequential, with a relative ranking between morphological categories within a period of general morphosyntactic and linguistic development.

Cross-linguistic factors

Though the focus of the present paper was to characterize English language use in bilingual children, the findings are also consistent with a perspective of bilingual language development that is highly interactive and dynamic (Kan & Kohnert, 2008; Kohnert 2010). Under this perspective, use of a target language is understood to be impacted by knowledge of the non-target language. This potential for cross-linguistic influence helps explain why bilingual children may differ from monolingual peers in T/A use. For example, the relatively high use of both cop and aux *BE* in English may be explained by positive transfer associated with a relatively similar structure in the participants' native language, Spanish. In English, cop and aux *BE* are unbound morphemes that frequently precede nouns, adjectives or verbs; the translation equivalents in Spanish, *ser* and *estar*, are similar in morphological structure and syntactic function. As precocious use of both cop and aux *BE* has been recognized in English-speaking bilingual children from a variety of native language backgrounds (see Paradis, 2010), transfer effects likely do not fully explain patterns of performance. Potentially, these unbound morphemes are less marked than their inflected counterparts, with children who are acquiring grammatical skills in multiple languages being particularly sensitive to this difference in markedness (Paradis, 2010). Further, we also found differences in use between the *BE* categories. As in monolingual English speakers, cop *BE* was the most productively used morpheme category, suggesting that still other factors may be at play, including a morpheme category's input frequency (Rispoli et al., 2012; Hadley, Rispoli, Fitzgerald & Bahnsen, 2011). And yet transfer may also be related to the relatively low use of -3s, -ed and aux *DO* in the current study. Both -3s and -ed may require children to produce complex consonant clusters in word-final position (e.g., *jumps*). Such a structure is incompatible with the phonotactics of Spanish, which allow only a limited selection of word-final consonants (see Combiths, Barlow, Potapova & Pruitt-Lord, 2017). Similarly, there is no clear counterpart of aux *DO* in Spanish. Thus, it is possible that there is interference as children attempt to produce each of these forms. Future investigations may serve to identify the contributions of these various factors.

Age, exposure and language ability effects

In addition to addressing our research questions, our findings also offer opportunities to consider factors relevant to language development more broadly, including chronological age, relative language exposure and language ability. At Time 1 and Time 2, morpheme use in our larger data set increased with age in months, with a similar trend for our subset in our second research question. This is to be expected, as monolingual and bilingual children alike increase in language skills over time. Similarly, increased exposure to English at home significantly predicted English morpheme use, reflecting an association between access to input and language performance (e.g. Hoff, 2003; Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991). However, this relationship between input and morpheme sub-scores was only significant at Time 1 for the entire sample. At Time 2, for both the complete sample and for the subset, the factor was not significant. Though this might be interpreted as exposure playing less of a role over the course of the year, we understand this pattern of findings to reflect measurement methods. Specifically, Spanish and English exposure were reported at the beginning of the year. As a result, this descriptor did not reflect the children's exposure to both languages by the second testing point, when exposure and dominance patterns have likely shifted. Potentially, an updated measure of language exposure collected at Time 2 would better predict morpheme use.

Morpheme use was also associated with language group: BiTD participants outperformed BiLL participants, with significant group differences for the complete data set at Time 1 and Time 2, and a trend in the expected direction for the subset at Time 2. Though bilingual children may differ from monolingual children in myriad ways, research indicates that bilingual children with DLD struggle in similar areas of language as monolingual children with DLD, including T/A marking in English (Gutiérrez-Clellen et al., 2008). In the current work, this is reflected in terms of T/A productivity, with BiLL children using individual morpheme categories less contrastively than BiTD peers (see also Potapova et al., 2018). As language ability was not found to interact with morpheme category in any analysis, results suggested that bilingual children with relatively low language skills demonstrated a comparable ranking of morpheme categories as their typically developing peers, despite having lower relative scores. This not only mirrors findings in monolingual children, it reinforces our understanding that weak language skills are not characterized by deviant patterns of language use, but by delayed patterns of acquisition and mastery (Rice, 2010). However, given restrictions with interpreting null results, it is important to continue to consider this relationship in future research.

Clinical implications

These findings serve as a reminder of the importance of using appropriate reference groups when considering language performance – neither bilingual group in the present study demonstrated the pattern of morpheme use observed in English-speaking monolingual children. Consequently, expectations rooted in an understanding of monolingual performance would be inappropriate for bilingual children, regardless of language ability. In addition, present findings may offer guidelines for what clinicians may expect in English performance when assessing Spanish-English bilingual children. To illustrate, our findings suggest that delayed onset of aux *DO* is more expected in a Spanish-English bilingual child than difficulty with producing either cop or aux *BE;* importantly, established monolingual norms would fail to predict bilingual children's relative success with aux *BE.*

Limitations and future directions

The current work aimed to understand the relative use of English T/A morphemes in Spanish-English bilingual children to support clinical decision-making for bilingual children. Additional steps may be taken to strengthen and expand this work.

Bilingual experiences and contexts differ largely across the world, with potentially far-reaching implications for language development. Consequently, it is important that the specific context of this study be noted. This study included children from largely Spanish-dominant homes, a minority language in Southern California, enrolled in a preschool programme where English, the majority language, was the language of instruction. As is the case for many children from culturally and linguistically diverse backgrounds in the United States, this transition likely resulted in increased overall exposure to English (Bedore & Peña, 2008); moving forward, it is likely that these participants will experience a change in dominance from their heritage language to the majority language (Pham & Kohnert, 2014; Rojas & Iglesias, 2013). The current measures thus likely reflect the profiles of children who are in the relatively early stages of acquiring English skills. Accordingly, the patterns of morphosyntactic development observed here are consistent with those observed for children who may be described as dual language learners, English language learners, or sequential bilinguals (e.g. Paradis et al., 2008), and not simultaneous bilinguals (e.g. De Houwer, 2009). Future work that systematically investigates the impact of language status, language dominance, and age of acquisition would enhance our understanding of English T/A development in bilinguals. Similarly, it will be important to include bilingual children from other linguistic and cultural backgrounds. By extending the productivity-based measures of English T/A use to children whose native language is not English or Spanish, the roles of the native language (i.e. cross-language influence) and the second language (i.e. input and language structure) may be clarified. In particular, it would be telling to consider productivity of cop and aux *BE* in bilingual children whose native language does not share a structurally similar morpheme or to measure use of -3s and -ed in children whose native language allows complex word-final consonant clusters. Finally, extensions of productivity-based measures to languages beyond English would allow for a more complete picture of morphosyntactic development in bilingual children, including the role of cross-linguistic influence.

The current study would also be improved with more frequent measurement points, specifically in efforts to capture the emergence of T/A morphemes. Presently, we strived to reflect emergence by identifying children that began the academic year with no productive uses of the target morphemes and evaluating their use of the target morphemes at the end of the year. However, by the end of the year, there was evidence of use of multiple morphemes; as such, we are unable to determine which morpheme category emerged first with certainty. With language samples collected more regularly, this question would be better addressed.

There are also considerations with regard to how morpheme use is measured. Original scoring protocol prevents children from earning points for copula and auxiliary verbs contracted to pronouns (e.g. *he's*) to avoid inflated scores for potentially rote forms (Rispoli & Hadley, 2011). It is possible that such contracted forms are indicative of grammatical processing in the current sample; if this is the case, that would reveal only greater differences between the forms of *BE* and the bound morphemes, for which scores would remain unaffected. Similarly, future extensions may also consider raising the scoring maximum of five points per morpheme (see Rispoli et al., 2012). However, as current participants demonstrated T/A use that was clearly below ceiling, such a scoring adjustment was not expected to alter present findings. Indeed, the fact that our participants did not reach ceiling on this measure suggests that similar analyses with young school-age bilinguals may be appropriate to further inform our understanding of how T/A skills develop in the context of dual language learning.

Summary

This research was motivated by the need to better understand developing bilingual children's performance in each language to support best practice for children from culturally and linguistically diverse backgrounds. Using language sampling and a productivity-based measure of morpheme use, we confirmed and clarified patterns of English use in developing bilingual children reported in previous works. Specifically, we found a clear distinction in productive uses between copula and auxiliary *BE* relative to third person singular, past tense and auxiliary *DO*; in addition, findings suggested further distinctions, with more productive uses of copula *BE* than auxiliary *BE*, and with particularly low rates for auxiliary *DO*. This relative ranking between the morpheme categories differs from the trajectory identified for monolingual English speakers, for whom auxiliary *BE* emerges last and is used least productively. Though multiple influential factors likely impact bilingual children's use of grammatical features in each language, the current findings are consistent with patterns of cross-language influence, which have been identified in all aspects of language use for bilingual speakers. In addition, these results illustrate the importance of utilizing appropriate reference groups during language assessment. Ultimately, detailed information about bilingual children's development in a single language bolsters a clinician's toolkit as they utilize a combination of culturally and linguistically appropriate measures to arrive at a holistic language assessment.

Acknowledgements

We would like to thank the participants, their families and the teachers who took part in our community-based collaboration. We would also like to thank the members of the Child Language Development, Disorders and Disparities Lab for research assistance and helpful discussions about this work. In particular, we would like to recognise the efforts of Sophia Kelly and Allegra Schiff. We continue to be grateful to Dr Pamela Hadley, who graciously shared training materials for the TAP score.

Funding

The first author was supported by an NIH NIDCD F31 (grant number DC016194). In addition, funding for the project was provided by an NIH

NIDCD RO3 (grant number DC012141) and a research grant from a local community foundation.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research*, *41*(5), 1185–92.

Bedore, L. M., & Peña, E. D. (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), 1–29.

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Combiths, P. N., Barlow, J. A., Potapova, I., & Pruitt-Lord, S. (2017). Influences of phonological context on tense marking in Spanish–English dual language learners. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2199–216.

De Houwer, A., (2009). *Bilingual first language acquisition*. Bristol: Multilingual Matters.

De Houwer, A. (2018). Input, context and early child bilingualism: Implications for clinical practice. In A. Bar-on & D. Ravid (Eds.), *Handbook of communication disorders: Theoretical, empirical and applied linguistic perspectives* (pp. 599–616). Berlin: Walter de Gruyter.

Fabiano-Smith, L., & Hoffman, K. (2018). Diagnostic accuracy of traditional measures of phonological ability for bilingual preschoolers and kindergarteners. *Language, Speech, and Hearing Services in Schools*, *49*(1), 121–34.

Gillam, R. B., Peña, E. D., Bedore, L. M., Bohman, T. M., & Mendez-Perez, A. (2013). Identification of specific language impairment in bilingual children. I. Assessment in English. *Journal of Speech, Language, and Hearing Research*, *56*(6), 1813–23.

Gladfelter, A., & Leonard, L. B. (2013). Alternative tense and agreement morpheme measures for assessing grammatical deficits during the preschool period. *Journal of Speech, Language, and Hearing Research*, *56*(2), 542–52.

Goffman, L., & Leonard, J. (2000). Growth of language skills in preschool children with specific language impairment: Implications for assessment and intervention. *American Journal of Speech-Language Pathology*, *9*(2), 151–61.

Guo, L. Y., & Eisenberg, S. (2014). The diagnostic accuracy of two tense measures for identifying 3-year-olds with language impairment. *American Journal of Speech-Language Pathology*, *23*(2), 203–12.

Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, *24*(2), 267–88.

Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics*, *29*(1), 3–19.

Hadley, P. A., & Holt, J. K. (2006). Individual differences in the onset of tense marking: A growth-curve analysis. *Journal of Speech, Language, and Hearing Research*, *49*(5), 984–1000.

Hadley, P., Rispoli, M., Fitzgerald, C., & Bahnsen, A. (2011). Predictors of morphosyntactic growth in typically developing toddlers: Contributions of parent input and child sex. *Journal of Speech, Language, and Hearing Research*, *54*(2), 549–66.

Hadley, P. A., Rispoli, M., Holt, J. K., Fitzgerald, C., & Bahnsen, A. (2014). Growth of finiteness in the third year of life: Replication and predictive validity. *Journal of Speech, Language, and Hearing Research*, *57*(3), 887–900.

Hadley, P. A., & Short, H. (2005). The onset of tense marking in children at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research*, *48*(6), 1344–62.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–78.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–63.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236.

Ionin, T., & Wexler, K. (2002). Why is 'is' easier than '-s'? Acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research*, *18*(2), 95–136.

Kan, P. F., & Kohnert, K. (2008). Fast mapping by bilingual preschool children. *Journal of Child Language*, *35*(3), 495–514.

Kehoe, M. (2018). Cross-linguistic interaction in early bilingual phonology: A critical review. In E. Babatsouli & D. Ingram (Eds.), *Phonology in protolanguage and interlanguage* (pp. 49–75). Sheffield, UK: Equinox Publishing.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*, *43*(6), 456–73.

Leonard, L. B. (2014). Specific language impairment across languages. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and Language Impairments in Children* (pp. 129–44). London/New York: Psychology Press.

Liceras, J. M., Fuertes, R. F., & de la Fuente, A. A. (2012). Overt subjects and copula omission in the Spanish and the English grammar of English–Spanish bilinguals: On the locus and directionality of interlinguistic influence. *First Language*, *32*(1–2), 88–115.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

Miller, J., & Iglesias, A. (2012). SALT: Systematic Analysis of Language Transcripts. Software for the analysis of oral language.

National Institute on Deafness and Other Communication Disorders. (2016). Specific Language Impairment. Retrieved 3 August 2016 from https://www.nidcd.nih.gov

Nicoladis, E. (2006). Cross-linguistic transfer in adjective–noun strings by preschool bilingual children. *Bilingualism: Language and Cognition*, *9*(1), 15–32.

Oetting, J. B. (2018). Prologue: Toward accurate identification of developmental language disorder within linguistically diverse schools. *Language, Speech, and Hearing Services in Schools*, *49*(2), 213–7.

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*, *5*(1), 19–38.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *36*(3), 172–87.

Paradis, J. (2008). Tense as a clinical marker in English L2 acquisition with language delay/impairment. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective*. Amsterdam: John Benjamins.

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics*, *31*(2), 227–52.

Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research*, *59*(1), 171–82.

Paradis, J., & Blom, E. (2016). Do early successive bilinguals show the English L2 pattern of precocious BE acquisition? *Bilingualism: Language and Cognition*, *19*(03), 630–5.

Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, *43*(6), 474–97.

Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics*, *29*(04), 689–722.

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*(1), 41–58.

Pham, G., & Kohnert, K. (2014). A longitudinal study of lexical development in children learning Vietnamese and English. *Child Development*, *85*(2), 767–82.

Potapova, I., Kelly, S., Combiths, P. N., & Pruitt-Lord, S. L. (2018). Evaluating English morpheme accuracy, diversity, and productivity measures in language samples of developing bilinguals. *Language, Speech, and Hearing Services in Schools*, *49*(2), 260–76.

Potapova, I., & Pruitt-Lord, S. L. (2019). Towards understanding the bilingual profile in typical and atypical language development: A tutorial. *International Journal of Speech-Language Pathology*, 1–11.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, *41*(6), 1398–411.

Rice, M. L. (2010). Evaluating maturational parallels in second language children and children with specific language impairment. *Applied Psycholinguistics*, *31*(02), 320–7.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1239–57.

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech, Language, and Hearing Research, 38*(4), 850–63.

Rice, M. L., Wexler, K., & Hershberger, S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1412–31.

Rispoli, M., & Hadley, P. (2011). Toward a theory of gradual morphosyntactic learning. In I. Aron & E. V. Clark (Eds.), *Experience, variation, and generalization: Learning a first language* (pp. 15–34). Amsterdam, the Netherlands: John Benjamins.

Rispoli, M., Hadley, P. A., & Holt, J. K. (2009). The growth of tense productivity. *Journal of Speech, Language, and Hearing Research, 52*(4), 930–44.

Rispoli, M., Hadley, P. A., & Holt, J. K. (2012). Sequence and system in the acquisition of tense and agreement. *Journal of Speech, Language, and Hearing Research, 55*(4), 1007–21.

Roid, G. H., & Miller, L. (1997). *Leiter International Performance Scale-Revised (Leiter-R) manual.* Wood Dale, IL: Stoelting.

Rojas, R., & Iglesias, A. (2009). Making a case for language sampling: Assessment and intervention with (Spanish-English) second language learners. *The ASHA Leader, 14*(3), 10–13.

Rojas, R., & Iglesias, A. (2013). The language growth of Spanish-speaking English language learners. *Child Development, 84*(2), 630–46.

Stockman, I. J. (1996). The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech, and Hearing Services in Schools, 27*(4), 355–66.

Chapter 3, in full, is a reprint of material as it appears in Potapova, I & Pruitt-Lord, S. L. (2019). Spanish-English bilingual children's relative use of English tense and agreement morphemes. *Journal of Monolingual and Bilingual Speech, 1*(1), 118-142. The dissertation author was the primary investigator and author of this paper.

CHAPTER 4:

Cognate Identification Methods: Impacts on the Cognate Advantage in Adult and Child Spanish-English Bilinguals

Article

International Journal of Bilingualism

International Journal of Bilingualism
2016, Vol. 20(6) 714–731

DOI: 10.1177/1367006915586586
ijb.sagepub.com
SAGE

Cognate identification methods: Impacts on the cognate advantage in adult and child Spanish-English bilinguals

Irina Potapova
Joint Doctoral Program in Language and Communicative Disorders, San Diego State University & University of California, San Diego, USA

Henrike K Blumenfeld and Sonja Pruitt-Lord
San Diego State University, USA

Abstract

Objectives: The purpose of this study was to determine whether four different cognate identification methods resulted in notably different classifications of cognate status for *Peabody Picture Vocabulary Test-Third Edition (PPVT-III)* test items and to investigate whether differences across criteria would impact findings of cognate effects in adult and preschool-aged Spanish-English bilingual speakers.

Methodology: We compared four cognate identification methods: an objective criterion based on phonological overlap; two subjective criteria based on a translation elicitation task; and a hybrid criterion integrating objective and subjective standards. We then used each criterion to investigate cognate effects on the *PPVT-III* in 26 adult and 73 child Spanish-English bilinguals.

Data and analysis: The test items identified as cognates by each criterion were compared (Experiment 1). Then, cognate advantage magnitudes, cognate accuracy rates, non-cognate accuracy rates, and number of individuals demonstrating the cognate advantage were investigated in both adult (Experiment 2) and child bilinguals (Experiment 3).

Conclusions: Objective and subjective cognate identification methods were found to select notably different subsets of test items as cognates. Further, the methods led to differences in cognate effects, as well as in cognate and non-cognate accuracy rates, for both child and adult bilinguals.

Originality: Although the cognate advantage has been widely studied in adult bilinguals, research on the cognate advantage in child bilinguals is limited and methods of identifying cognates are inconsistent across studies. The present study provides information about cognate effects in a young population and is the first comparison of objective and subjective approaches to cognate identification.

Implications: This study extends previous work on cognate word processing in both child and adult bilinguals. Further, results offer an evaluation of methodologies that are critical for

Corresponding author:
Irina Potapova, School of Speech, Language, and Hearing Sciences, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-1518, USA.
Email: ipotapov@ucsd.edu

investigating the cognate advantage. This both facilitates interpretation of previous findings and can be used to guide methodological decisions in future research.

Keywords
Spanish-English bilinguals, adult bilinguals, child bilinguals, cognate advantage, cognate identification

Introduction

Bilingual speakers provide unique opportunities to examine the mental lexicon, in particular regarding the influence of experience in one language on processing in another language. One topic that has been particularly well-researched is adult bilinguals' processing of cognate words—cross-linguistic translation equivalents that are similar in phonology, orthography, or both (e.g. *triangle/triángulo* in English and Spanish). The term "cognate advantage" refers to bilinguals' relative ease in processing these words as compared to non-cognates, translation equivalents that lack cross-linguistic form overlap (e.g. *apple/manzana* in English and Spanish; see Sánchez-Casas & García-Albea, 2005, for a review). Although the distinction between cognates and non-cognates may appear self-evident, operational criteria for assigning cognate status differ considerably across studies. Thus, it is important to evaluate whether these methodological differences impact findings of cognate effects in bilinguals. This may be especially important for young bilinguals, who show cognate effects less consistently than adult bilinguals.

The cognate advantage is well-attested in adult bilinguals. Bilingual adults respond to cognates with greater accuracy and speed on various linguistic tasks, including categorization (Dufour & Kroll, 1995), translation (De Groot & Poot, 1997), word association (Van Hell & De Groot, 1998), and word learning (De Groot & Keijzer, 2000; Van Assche, Duyck, & Brysbaert, 2013). Performance on the *Boston Naming Test* also indicates enhanced naming ability for cognate relative to non-cognate items (Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007; Rosselli, Ardila, Jurado, & Salvatierra, 2012). Although adult bilinguals consistently demonstrate a cognate advantage, there are fluctuations in the effect. For example, cognate effects appear to be greater for non-balanced than balanced bilinguals and when speakers are tested in their weaker language (Caramazza & Brones, 1979; Gollan et al., 2007; Rosselli et al., 2012, but see Davis et al., 2010).

Much less is known about cognate sensitivity in child bilinguals—there are relatively few studies, and not all age groups are yet represented. Nevertheless, some research suggests that children show similar patterns to adults. For example, Dutch-speaking fifth- to ninth-grade English Language Learners (ELLs) showed a cognate advantage in both reaction times and accuracy on an English lexical decision task (Brenders, Van Hell, & Dijkstra, 2011).

However, many studies in children have found limited cognate effects. These studies have frequently been in the context of standardized language assessments—perhaps because such measures are routinely administered to large groups of children. For example, 8–13-year-old Spanish-speaking ELLs were found to have higher accuracy rates for cognate than non-cognate test items on the *Peabody Picture Vocabulary Test-Third Edition* (*PPVT-III*, Kelley & Kohnert, 2012). Despite these group-level effects, only 60% (18 of 30 participants) demonstrated the advantage and older children were more likely to show cognate effects. These findings suggest that the cognate advantage may be weaker in child than adult speakers, and that the advantage may develop over time. Nevertheless, initial findings do suggest that cognate effects emerge even in young children. Pérez, Peña, and Bedore (2010) found higher accuracy rates for cognate items on the Picture Vocabulary Subtest of the *Test of Language Development-Primary, Third Edition* in Spanish-English bilingual kindergarteners and first graders, but only for those with high Spanish exposure (60–80%). These results are consistent with adult research indicating greater cognate effects for unbalanced bilinguals tested in their weaker language (e.g. Rosselli et al., 2012). These findings also support the idea that the cognate advantage is more tenuous in children than adults.

In fact, some argue that child bilinguals do not demonstrate cognate effects, as revealed in two studies of first-grade (Umbel, Pearson, Fernández, & Oller, 1992) and first-, third-, and sixth-grade Spanish-English bilinguals (Umbel & Oller, 1994) on the *PPVT-Revised* and the *Test de Vocabulario en Imágenes Peabody—Adaptación Hispanoamericana* (*TVIP*; Dunn, Padilla, Lugo, & Dunn, 1986). Due to the limited number of studies addressing cognate effects in children, it is difficult to determine the reason behind these discrepant findings. It is likely that a number of factors mediate the presence of cognate effects. Pérez et al.'s (2010) findings suggest the presence of cognate effects in children may be a matter of relative exposure; yet, Umbel et al.'s (1992) sample included children in the same age range with arguably similar language profiles, with exposure to Spanish at home and English in school. Kelley and Kohnert's (2012) study suggests an important factor may be age; yet, cognate effects have also been reported in younger children (Pérez et al., 2010) and were not reported for sixth-grade bilinguals by Umbel and Oller (1994).

In summary, much is left to be understood about the factors that determine cognate effects in child bilinguals. There are also practical applications to understanding cognate processing in children. For example, successful cognate identification (subsequent to cognate awareness training) has been associated with greater English reading comprehension in fifth- and sixth-grade Spanish-English bilinguals (Nagy, García, Durgunoglu, & Hancin-Bhatt, 1993) and in deducing English word meanings (Dressler, Carlo, Snow, August, & White, 2011). It would be beneficial to learn to what extent such strategies may be available to bilingual children prior to explicit training. Finally, for clinical purposes, it is important to learn if bilingual speakers are impacted by the cognate status of items on standardized measures.

To methodically pursue this line of research, it is essential that we reflect on how cognate status is assigned. Although it is widely agreed that cognates are translation equivalents with similar phonology and orthography, the exact criterion for sufficient similarity varies considerably. Objective methods of assigning cognate status have included using a cognate dictionary and a minimum criterion of three shared phonemes (Pérez et al., 2010), as well as more fine-grained phonological comparisons, such as the *Crosslinguistic Overlap Scale for Phonology* (COSP; Kohnert, Windsor, & Miller, 2004; also see DeGroot & Keijzer, 2000).

Although objective approaches provide an efficient and consistent method of determining cognate status, it is possible that factors beyond phonology ought to be considered, especially in the context of investigating the cognate advantage. Namely, it may be valuable to include a measure that reflects cross-linguistic similarities that are salient to speakers. This is possible with subjective methods of cognate identification, which have also been widely used (e.g. Brenders et al., 2011; Friel & Kennison, 2001; Gollan et al., 2007; Nagy et al., 1993; Rosselli et al., 2012). For example, monolingual English speakers could be asked to provide similarity ratings for translation equivalents in English and another language or to guess the English meaning of a foreign word. High ratings or successes in translations would suggest that the cross-linguistic overlap is salient (Friel & Kennison, 2001). Such prominent similarity may ultimately be what is reflected in behavioral findings of cognate advantages. In fact, theoretical accounts of cognate processing posit stronger associative links between form-similar translation equivalents, with bilinguals' translation of cognates faster and more accurate than for non-cognates (Boada, Sanchez-Casas, Gavilan, Garcia-Albea, & Tokowicz, 2013; Friel & Kennison, 2001; Kroll & Stewart, 1994). Despite these potential advantages of subjective approaches, speaker judgment tasks are more time intensive relative to objective approaches. They may also be less stable, as similarity ratings or translation elicitations may change as different individuals are surveyed and when local variations in vocabulary are encountered.

Thus, while vast methodological variability is present in the current literature on the cognate advantage, little is known about the differences between these approaches (cf. Friel & Kennison, 2001). Crucially, it is not yet clear whether objective and subjective approaches assign cognate status to the same sets of words or if different methodologies may impact findings of cognate effects. Comparing approaches is particularly important for child bilinguals, who show relatively attenuated and inconsistent cognate effects, as differences in methodology may impact whether or not a cognate advantage is found.

Present study

In Experiment 1, we compare four cognate identification methods that represent both objective and subjective approaches. In Experiments 2 and 3, we compare the implementation of these criteria in evaluations of cognate effects in adult and child bilinguals, respectively. Both groups completed the *PPVT-III* (Dunn & Dunn, 1997), a task that is appropriate for children and adults and has been previously used in examining Spanish-English cognate effects (Kelley & Kohnert, 2012; Umbel et al., 1992). To our knowledge, this is the first study of cognate effects to include both adult and child speakers, as well as the first cognate processing study in children that are preschool-aged. In summary, we ask the following:

1) Are there differences in the quantity and quality of *PPVT-III* test items selected by objective and subjective cognate identification criteria? (Experiment 1)
2) Do cognate identification criteria influence the magnitude and consistency of cognate effects for adult bilinguals? (Experiment 2)
3) Do cognate identification criteria influence the magnitude and consistency of cognate effects for child bilinguals? (Experiment 3)

Experiment 1

In our first experiment, we compared four approaches of determining cognate status, including one objective criterion, two subjective criteria, and one hybrid criterion, in order to ascertain if these methods select notably different subsets of *PPVT-III* test items as cognates. The objective selection criterion consisted of the previously used COSP scale (Kelley & Kohnert, 2012). To implement the subjective cognate identification criteria, English monolingual participants were asked to back-translate Spanish translation equivalents of the English *PPVT-III* test items (described in detail below), with accurate back-translations only possible in the presence of salient form overlap between English and Spanish translations (e.g. Friel & Kennison, 2001). It was predicted that the four cognate selection criteria would identify quantitatively and qualitatively different sets of cognate words from among the 204 *PPVT-III* test items.

Method

Participants. Students in a large undergraduate class were offered extra credit to complete a translation elicitation task together with a brief language background questionnaire. Out of 118 participating students, the 12 monolingual native English speakers with no reported experience in any other spoken language were selected for the subjective cognate identification task. These selected students (mean age = 21.25 years, age reported by 67%) rated their proficiency in Spanish on a scale from 0 (*none*) to 10 (*perfect*), with a mean proficiency rating of0.19 (*SD* = 0.39), suggesting minimal familiarity with Spanish.

Materials. To examine the influence of cognate identification criteria on the selection of cognate and non-cognate word sets, items on the *PPVT-III* were examined. The *PPVT-III* is a measure of receptive vocabulary that asks individuals to match pictures to target items. The 204 targets are divided into 17 sets, with each subsequent set designed to increase in difficulty. Following testing protocol, participants' age determines their starting set, and individuals continue until eight errors are made in a single set. Thus, even adult participants are typically not familiar with all words presented within the last 2–3 sets of the test.

To identify cognate items on the *PPVT-III*, four Spanish-English bilingual research assistants independently translated each item from English into Spanish. Translation equivalents were selected from the pooled responses, with preference first given to Spanish words that also appeared on the *TVIP.* For the remaining items, preference was given to translations that faithfully represented the corresponding test image (e.g. the noun *mosca* instead of the verb *volar* for the English noun *fly*) and to translations that were provided by multiple translators. Translations with orthographic overlap with the target word were also given preference (e.g. selecting *vehículo* instead of *coche* for *vehicle*).

Procedure and analyses. The established English-Spanish translation equivalents of items on the *PPVT-III* were then analyzed for cognate status in four different ways:

Firstly, we adopted the COSP (Kohnert et al., 2004) as an *objective criterion* of cognate status. Degree of overlap was measured in four domains: word-initial sound, number of syllables, percentage of overlapping consonants, and percentage of overlapping vowels. Total COSP scores ranged from 0 (e.g. *knight/caballero*) to 10 (e.g. *cupola/cúpula)*. Words that received scores of six or higher (e.g. *selecting/seleccionar)* were considered cognates (Kelley & Kohnert, 2012).

We also used two *subjective criteria*, a *50% Translation criterion* and a *75% Translation criterion.* Using a translation task similar to that described by Friel and Kennison (2001) (see also De Groot & Nas, 1991; Kroll & Stewart, 1994), the English monolingual participants in the current experiment received a typed list of words containing Spanish translation equivalents of the *PPVT-III* items and were asked to guess each word's meaning. Participants were instructed to write an English word alongside each Spanish word. Successful back-translations by the English monolingual participants indicated salient similarities between the Spanish and English translation equivalents (i.e. cognate status). Test items were printed in a randomized order to avoid a gradual increase in difficulty across the task, and each participant back-translated only half of the test items.

Successful back-translations included exact matches, root matches (e.g. *decorate* for target *decorated*, from Spanish translation equivalent *decorado*), and synonyms (e.g. *video camera* for target *camcorder*, from Spanish translation equivalent *videocámara*). For each test item, the percentage of successful translations was calculated. Words were classified as cognates by the 50% Translation criterion if half or more of the monolinguals produced a successful response (Friel & Kennison, 2001). A higher subjective standard—the 75% Translation criterion—was also established to select particularly transparent cognates.

Finally, we combined objective and subjective components to create a *Hybrid criterion*. With this method, a word qualified as a cognate if (1) at least 50% of the monolingual speakers back-translated the word correctly AND the word received a COSP score of 6 or higher, *or* (2) at least 75% of the speakers back-translated the word correctly AND the word received a COSP score of 4 or higher. The first condition selected words that are highly similar phonologically with a relatively low speaker recognition requirement (*painting/pintar*, COSP = 9, correctly back-translated by

50%); the second criterion lowers the objective requirement to accommodate words that are relatively more salient for speakers (*gigantic/gigante*, COSP = 4, correctly back-translated by 83%). The goal was to select words that are both objectively similar and subjectively recognizable.

Results and discussion

Of the 204 test items, 30 words were identified as cognates by all four definitions (e.g. *closet/clóset, accident/accidente*) and 90 were selected by none (e.g. *drinking/tomar, empty/vacío*, see Table 1). Potentially, these 120 items are particularly clear exemplars of cognates and non-cognates. The remaining 84 words were selected either by one, two, or three of the available criteria (see Table 2). These items may be more ambiguous in their cognate status, as they fail to meet at least one objective or subjective criterion. The disagreements across criteria suggest that the four methods do select different subsets of test items as cognates.

Indeed, the approaches differed in the quantity of test items selected as cognates (see Table 3). The COSP criterion selected a significantly higher proportion than the next most generous approach, the 50% Translation criterion, $\chi^2(1) = 17.174$, $p < .001$. The 50% Translation and the Hybrid criteria did not differ in the proportion of cognates selected, $p > .5$. In turn, the Hybrid criterion identified significantly more cognates than the 75% Translation criterion, $\chi^2(1) = 4.09$, $p = .04$. Thus, the COSP criterion selected more, and the 75% Translation criterion selected fewer, cognates than any other approach, and quantitative analyses revealed that the four cognate selection criteria differ in their assignment of cognate status.

Further, qualitative observations also revealed important differences across cognate identification criteria. The COSP method is advantageous in that any word pair can be analyzed by any trained individual with high reliability across scorers. However, lacking human intuition, the COSP criterion both overlooked important similarities (e.g. *camcorder/videocámara*, COSP = 3, correctly back-translated by 83% of English monolinguals; *helicopter/helicóptero*, COSP = 4, correctly back-translated by 100%) and imposed correspondences where speakers did not perceive them (e.g. *measuring/medir*, COSP = 8, no correct back-translations). The objective COSP criterion identified significantly more cognates than any other approach. It also selected more test items from latter test sets than the subjective criteria did, which, given the test's design, reflects that the objective criterion selected more high-difficulty words. In the final three sets of the *PPVT-III*, 20 of 36 words were identified as cognates by the COSP criterion (e.g. *dromedary/dromedario*, COSP = 8, no correct back-translations). In contrast, only two words in those sets met even the 50% Translation criterion.

Relative to objective measures, subjective approaches require more time and resources, and they may produce less consistent results. In this study, speaker judgments occasionally resulted in surprising evaluations of cognate status. Monolingual speakers sometimes failed to recognize

Table 1. Number of *Peabody Picture Vocabulary Test-Third Edition* items selected by cognate identification criteria.

Number of criteria	Number of cognates selected	Examples
1	58	*island/isla*
2	2	*bus/autobus*
3	24	*helicopter/helicóptero*
4	30	*fragile/frágil*

Note: Of the 204 test items, 90 were not selected as cognates by any criteria.

Table 2. Number of criteria that identified each item as a cognate.

PPVT-III Form A item	Number of criteria that selected the item as a cognate	*PPVT-III* Form A item	Number of criteria that selected the item as a cognate	*PPVT-III* Form A item	Number of criteria that selected the item as a cognate
1	2 (50%, 75%)	86	3 (C, 50%, H)	148	1 (C)
6	1 (50%)	87	3 (C, 50%, H)	149	4 (C, 50%, 75%, H)
7	4 (C, 50%, 75%, H)	88	1 (C)	150	3 (C, 50%, H)
9	3 (C, 50%, H)	89	4 (C, 50%, 75%, H)	151	1 (C)
10	3 (50%, 75%, H)	90	1 (C)	153	1 (C)
17	3 (C, 50%, H)	91	3 (C, 50%, H)	154	3 (C, 50%, H)
24	1 (C)	92	3 (50%, 75%, H)	155	1 (C)
28	1 (C)	94	4 (C, 50%, 75%, H)	156	1 (C)
29	1 (C)	96	1 (C)	158	3 (C, 50%, H)
31	1 (C)	97	4 (C, 50%, 75%, H)	159	1 (C)
32	1 (C)	98	1 (C)	160	4 (C, 50%, 75%, H)
33	4 (C, 50%, 75%, H)	100	1 (C)	162	1 (C)
34	4 (C, 50%, 75%, H)	101	1 (50%)	163	1 (C)
35	4 (C, 50%, 75%, H)	102	4 (C, 50%, 75%, H)	164	4 (C, 50%, 75%, H)
37	1 (C)	103	3 (C, 50%, H)	165	3 (C, 50%, H)
38	1 (C)	104	1 (C)	168	3 (50%, 75%, H)
39	3 (C, 50%, H)	106	4 (C, 50%, 75%, H)	169	1 (C)
42	1 (C)	109	1 (C)	171	1 (C)
43	4 (C, 50%, 75%, H)	110	3 (C, 50%, H)	172	4 (C, 50%, 75%, H)
51	4 (C, 50%, 75%, H)	111	3 (C, 50%, H)	174	1 (C)
53	4 (C, 50%, 75%, H)	114	1 (C)	175	1 (C)
55	1 (50%)	116	4 (C, 50%, 75%, H)	176	1 (C)
59	4 (C, 50%, 75%, H)	117	4 (C, 50%, 75%, H)	177	1 (C)
61	4 (C, 50%, 75%, H)	120	4 (C, 50%, 75%, H)	179	1 (C)
62	3 (C, 50%, H)	121	3 (C, 50%, H)	180	1 (C)
65	3 (50%, 75%, H)	124	1 (C)	181	1 (C)
67	4 (C, 50%, 75%, H)	125	4 (C, 50%, 75%, H)	184	1 (C)
68	1 (C)	126	1 (C)	185	1 (C)
70	4 (C, 50%, 75%, H)	127	1 (C)	187	1 (C)
71	4 (C, 50%, 75%, H)	129	3 (C, 50%, H)	188	3 (C, 50%, H)
73	3 (50%, 75%, H)	131	1 (C)	192	1 (C)
75	4 (C, 50%, 75%, H)	135	1 (C)	195	1 (C)
79	4 (C, 50%, 75%, H)	138	3 (C, 50%, H)	196	1 (C)
80	3 (C, 50%, H)	139	4 (C, 50%, 75%, H)	197	1 (C)
81	1 (50%)	140	1 (C)	199	1 (C)
82	2 (50%, 75%)	141	1 (C)	200	1 (C)
83	1 (50%)	143	1 (C)	203	1 (C)
85	4 (C, 50%, 75%, H)	146	4 (C, 50%, 75%, H)	204	1 (C)

In parentheses: (C) = *Crosslinguistic Overlap Scale for Phonology*; (50%) = 50% Translation criterion; (75%) = 75% Translation criterion; (H) = Hybrid.
PPVT-III: Peabody Picture Vocabulary Test-Third Edition.

words with high objective similarities (e.g. *cascade*/*cascada*, COSP = 9, no correct back-translations). They also successfully back-translated three Spanish words that lack phonological overlap with their English counterparts but are apparently familiar even to monolinguals (e.g. *heart*/*corazón*,

Table 3. Number of cognates and non-cognates identified by each of four cognate identification criteria.

Criterion	Number of cognates	Number of non-cognates
COSP	102	102
50% Translation	61	143
75% Translation	37	167
Hybrid	54	150

COSP: *Crosslinguistic Overlap Scale for Phonology.*

COSP = 2, correctly back-translated by 67%). These differences suggest that unlike the objective approach, speaker judgments are sensitive to factors like word frequency. For example, while the COSP criterion treated *closet/clóset* and *dromedary/dromedario* as equivalent (COSP = 8), more English speakers successfully back-translated the former than the latter. Potentially, this ability to reflect subjectively salient—as opposed to objective—overlap is an advantage of speaker-based approaches.

The Hybrid criterion successfully combined both objective and subjective factors in selecting cognates. For example, both *measuring/medir*, which met the COSP criterion, and *heart/corazón*, which met both subjective criteria, were classified as non-cognates by the Hybrid criterion. Thus, words that were similar in terms of only phonology or only speaker judgment were deemed non-cognates. Words that fell slightly short of the COSP cognate identification threshold but were highly recognizable (e.g. *helicopter/helicóptero*) were identified as cognates.

Because differences across criteria were established, their potential impact on behavioral cognate advantage measurements needed to be evaluated. We addressed our second and third research questions by implementing the four methods to investigate cognate effects in adult (Experiment 2) and child (Experiment 3) Spanish-English bilinguals.

Experiment 2

Because the cognate advantage has been so widely demonstrated in adult bilinguals, this population provides a good initial opportunity to compare different cognate identification methods. Adult Spanish-English bilinguals were administered the *PPVT-III*, and responses across test items were grouped and analyzed according to each of the four cognate selection criteria. It was expected that, given the robust nature of the cognate advantage in adults (e.g. Sánchez-Casas & García-Albea, 2005), all four cognate identification criteria would yield cognate effects in this population. In addition, we predicted that cognate identification criteria would modulate the magnitude of cognate effects.

Method

Participants. Twenty-six Spanish-English adult bilinguals (two males; mean age = 21.77 years, *SD* = 3.17) were selected from a sample of 75 bilinguals. Selection criteria were that participants had early language histories that were highly similar to the child bilinguals in Experiment 3, with Spanish reported as their native language, English acquired after Spanish but before age 6, and with no knowledge of other spoken languages. Language histories, including when participants were first exposed to each language and their current exposure and proficiency in each language, were collected based on participants' self-reports using the *Language Experience and Proficiency Questionnaire* (Marian, Blumenfeld, & Kaushanskaya, 2007). At the time of testing, participants reported being exposed to Spanish 37.08% of the time (*SD* = 15.34, Range = 5–65) and to English 62.65% of the time (*SD* = 15.08, Range = 35–95). Participants reported an average spoken Spanish proficiency of 8.56 (*SD* = 1.26, Range = 5–10) and an average spoken English proficiency of 8.6 (*SD* = .84, Range = 7–10) on a proficiency scale of 0 (*none*)–10 (*perfect*). Mean matrix reasoning *t*-scores from the *Wechsler Abbreviated Scale of Intelligence* were 51.25 (*SD* = 9.69, Range = 23–66) and mean raw scores from Subtest 7 ("Numbers Reversed") of the *Woodcock-Johnson III Tests of Cognitive Abilities* were 16.92 (*SD* = 3.47; Range = 13–25). Scores were thus within the normal range, and no participants reported a history of language, learning, or hearing disabilities. At the time of testing, participants had completed 15.29 years of education (*SD* = 3.41, Range = 6–20). Average standard *PPVT-III* scores were 101.65 (*SD* = 10.39, Range = 66–126).

Procedures and analyses. Participants completed Form A of the *PPVT-III* (see *Materials* in Experiment 1). The adult bilinguals were administered the *PPVT-III* in a laboratory setting as part of an English-only session of a larger study. Although the test was administered according to published guidelines for most participants, eight adult participants completed the entire *PPVT-III* for research purposes and all responses contributed to analyses.

The 204 *PPVT-III* test items were each classified as either a cognate or non-cognate using each of the four cognate identification criteria developed in Experiment 1. Then, participants' responses for all completed test items were coded as correct or incorrect. This provided mean accuracy rates for cognate and non-cognate test items for each participant under each criterion. Because the approaches differed in their cognate assignment, each participant had different accuracy rates and cognate advantage magnitudes under each criterion. The cognate advantage for each criterion was calculated by subtracting the non-cognate accuracy rate from the corresponding cognate accuracy rate (e.g. Kelley & Kohnert, 2012). A positive number indicated that the speaker demonstrated a cognate advantage. We investigated whether each method led to findings of a cognate advantage, and then examined underlying differences in cognate accuracy rates, non-cognate accuracy rates, magnitude of the cognate advantage, and number of individuals demonstrating the effect.

Results

Cognate versus non-cognate performance. As expected, adults performed significantly more accurately on cognate than non-cognate *PPVT-III* items under all four criteria: COSP: $t(25) = 4.78, p < .001, d = .94$; 50% Translation: $t(25) = 9.07, p < .001, d = 1.78$; 75% Translation: $t(25) = 6.20, p < .001, d = 1.22$; Hybrid: $t(25) = 8.68, p < .001, d = 1.70$. To investigate underlying differences across criteria, we examined both the magnitude of the detected cognate advantages and the number of individuals demonstrating the effect.

Magnitude of the cognate advantage. Larger differences between cognate and non-cognate accuracy rates indicate greater cognate advantages. Cognate advantage magnitudes captured by the four methods (see Figure 1) differed significantly according to a Friedman test: $\chi^2(3) = 35.90, p < .001$. Planned post hoc analyses were conducted using Wilcoxon signed-rank tests with a Bonferroni correction, resulting in a significance level threshold of $p < 0.008$. The cognate advantage detected with the COSP criterion was significantly smaller than that detected by any other approach: 50% Translation ($z = -3.59, p < .001$); 75% Translation ($z = -2.91, p = .004$); Hybrid criteria ($z = -3.59, p < .001$). Differences between the magnitudes detected by the 50% and 75% Translation criteria approached significance ($z = -2.63, p = .009$). The remaining comparisons did not differ significantly ($ps > .012$). Overall, the COSP criterion detected the smallest cognate advantage for the adult bilinguals.

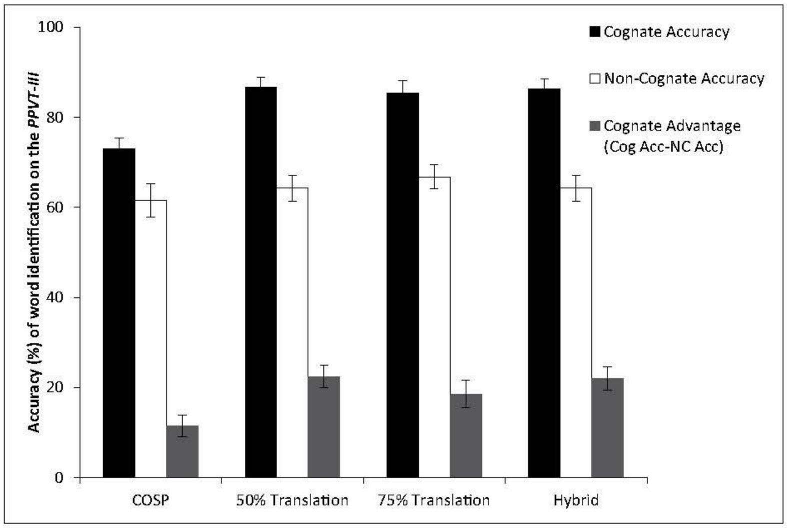


Figure 1. Adult Spanish-English bilinguals' cognate accuracy rates, non-cognate accuracy rates, and cognate advantages (cognate minus non-cognate accuracy) for the four cognate identification criteria.

Recall that the cognate advantage is calculated by subtracting the non-cognate accuracy rate from the cognate accuracy. Thus, one potential explanation for the observed differences in cognate advantage magnitudes is underlying differences across criteria in cognate and/or non-cognate accuracy rates (see Figure 1). Indeed, significant differences across criteria were found for cognate accuracy rates: $\chi^2(3) = 44.30, p < .001$. Post hoc analyses revealed that the cognate accuracy was significantly lower with the COSP than with the 50% Translation ($z = -3.98, p < .001$), 75% Translation ($z = -3.78, p < .001$), or the Hybrid criteria ($z = -3.98, p < .001$). Significant differences were also found for non-cognate accuracy rates: $\chi^2(3) = 30.35, p < .001$. The 75% Translation criterion resulted in significantly higher non-cognate accuracy than the COSP ($z = -3.24, p = .001$), 50% Translation ($z = -4.46, p < .001$), or Hybrid criteria ($z = -4.18, p < .001$). Thus, the least selective criterion (COSP) resulted in relatively low cognate accuracy rates while the most selective criterion (75% Translation) resulted in relatively high non-cognate accuracy rates.

Number of individuals demonstrating a cognate advantage. To examine the consistency of cognate effects, we considered the number of individuals that presented a cognate advantage. Under the COSP criterion, 20 of 26 participants showed cognate effects. Under the remaining three criteria, 25 of 26 adult bilinguals demonstrated a cognate advantage. Both proportions were found to be greater than chance levels with a combined sign test: $p = .005$ and $p < .001$, respectively. However, the proportion of individuals found to show a cognate advantage under the COSP criterion was significantly lower than any of the remaining criteria's proportions: $\chi^2(1) = 4.13, p = .04$.

Discussion

Adult Spanish-English bilinguals demonstrated higher accuracy rates for cognate than non-cognate *PPVT-III* test items. This held true with all four approaches of assigning cognate status. The 50%

Translation and the Hybrid criteria yielded the most similar results, with no differences in cognate advantage magnitude, cognate accuracy rates, non-cognate accuracy rates, or number of individuals demonstrating the cognate advantage. However, the COSP and 75% Translation criteria generated distinctive results.

Using the COSP criterion resulted in lower cognate accuracy rates relative to the remaining criteria, which may explain the relatively small cognate advantage captured by this approach. Potentially, both patterns were a result of the criterion identifying half of the test items as cognates, including many of the more challenging and lower-frequency items in the latter test sets. Participants likely struggled more with these higher-difficulty items, resulting in lower cognate accuracy rates, and thereby decreasing the difference between cognate and non-cognate performance. The relatively low cognate accuracy rate and cognate advantage magnitude under this criterion suggest that *PPVT-III* test items were over-identified as cognates. This is further supported by the significantly lower proportion of individuals found to demonstrate the cognate advantage under this approach.

In contrast, the 75% Translation criterion, which selected the fewest cognates, resulted in relatively high non-cognate accuracy rates. This finding suggests that some test items labeled as non-cognates were actually saliently similar across languages to at least some adult speakers, as this would explain the enhanced non-cognate accuracy rate. Like over-identification of cognates, under-identification could decrease cognate advantage magnitudes. In fact, magnitude differences between the 50% and the 75% Translation criteria approached significance. It is possible that the 75% Translation criterion was too restrictive in its assignment of cognate status.

Despite differences in cognate advantage magnitudes and underlying accuracy rates, all four methods detected cognate advantages for adult bilinguals. These findings attest to the robustness of cognate effects in adult bilinguals. Extending this investigation to child bilinguals provides an opportunity to examine the influence of cognate identification criteria on cognate advantages in a population where these effects appear less robust.

Experiment 3

Research on the cognate advantage in children suggests that the effects may be less robust than in adult speakers, with not all children demonstrating the effect (e.g. Kelley & Kohnert, 2012; Pérez et al., 2010). Moreover, there is some disagreement in the literature over whether children show a sensitivity to cognates. In Experiment 3, we address whether findings of cognate effects in child bilinguals can vary as a result of methodological decisions. Age-appropriate sets of the *PPVT-III* were administered to children, and it was expected that, as in Experiment 2, cognate identification criteria would modulate the magnitude of the cognate effect.

Method

Participants. Child participants included 73 Spanish-English bilinguals (40 males; mean age = 54.12 months, *SD* = 7.28). Questionnaires were sent home and completed by the children's parents. Parents were asked to detail the child's percentage of exposure to Spanish and English, and reported that the children were exposed to Spanish 70.08% of the time (*SD* = 20.32, Range = 33–100), and to English 29.90% of the time (*SD* = 20.31, Range = 0–67).[1] Selection criteria required that the child's home language was Spanish and that the child was exposed to Spanish at least 30% of the time (see Pearson, Fernandez, Lewedeg, & Oller, 1997). Thus, the adult bilinguals in Experiment 2 and the child bilinguals in the current experiment had highly similar early-life language exposure profiles. In addition, via the same questionnaire, parents reported that average maternal education, available for 56 participants, was 10.55 years (*SD* = 3.07, Range = 3–16). Scores from the Figure

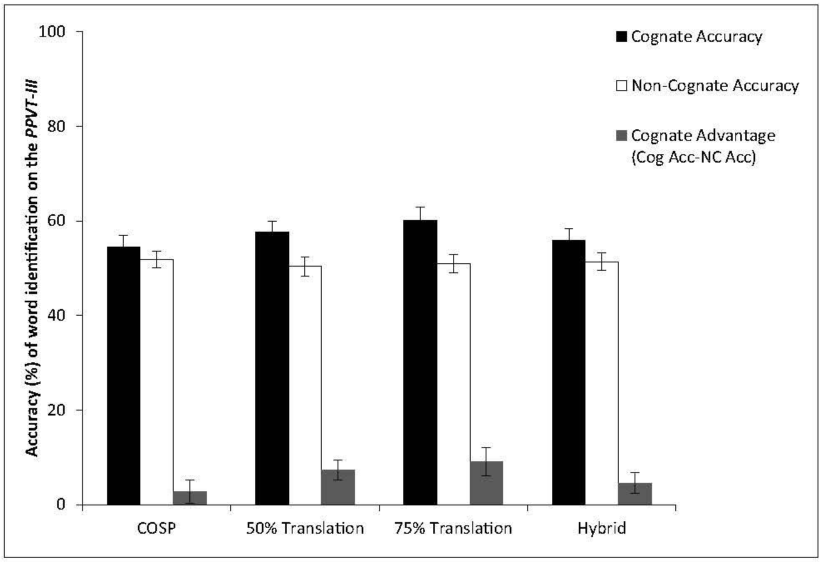


Figure 2. Child Spanish-English bilinguals' cognate accuracy rates, non-cognate accuracy rates, and cognate advantages (cognate minus non-cognate accuracy) for the four cognate identification criteria.

Ground and Form Completion subtests of the *Leiter International Performance Scale–Revised*, available for 72 participants, were in the normal range, with an average score of 11.73 (SD = 2.53, Range = 4–21). Parents and teachers reported that all children were developing typically. Consistent with previous findings that English-learning children perform below monolingual norms on the *PPVT-III* (e.g. Bialystok, Luk, Peets, & Yang, 2007; Gutiérrez-Clellen, 1999; Millett, Atwill, Blanchard, & Gorin, 2008, Umbel et al., 1992), the average *PPVT-III* score for this sample was 73.56 (SD = 18.03, Range = 40–108).

Procedure and analyses. Child participants were administered Form A of the *PPVT-III* (see *Materials* in Experiment 1), according to testing guidelines, at a local elementary school during an English-only session as part of a larger research project. As in Experiment 2, each participant's responses were coded as correct or incorrect. Cognate and non-cognate accuracy rates were calculated for each participant under all four cognate identification criteria. As in Experiment 2, we investigated whether each method detected a cognate advantage for the sample and then examined differences in cognate accuracy rates, non-cognate accuracy rates, magnitude of the cognate advantage, and number of individuals demonstrating the effect.

Results

The child bilinguals showed higher accuracy rates for cognates than non-cognates with three of the cognate selection criteria (see Figure 2): 50% Translation criterion: $t(72) = 3.46, p = .001, d = .41$; 75% Translation criterion: $t(72) = 3.01, p = .003, d = .35$; Hybrid criterion: $t(72) = 2.056, p = .043, d = .24$. Under the COSP criterion, no significant difference was found between cognate and non-cognate accuracy rates ($p > .25, d = .13$).

Further, cognate advantage magnitudes differed significantly across criteria, according to a Friedman test: $\chi^2(3) = 9.85, p = .02$. Post hoc analyses with Wilcoxon signed-rank tests were conducted with a Bonferroni correction, resulting in a significance level threshold of $p < 0.008$. The 50% Translation criterion resulted in a significantly higher cognate advantage magnitude than the Hybrid ($z = -2.94, p = .003$), and the remaining pairs were not significantly different from each other (ps > .03).

Because the cognate advantage is the difference between cognate and non-cognate accuracy rates, those accuracies were also analyzed. Cognate accuracy rates did not differ across criteria. However, non-cognate accuracy rates did: $\chi^2(3) = 14.22, p = .003$. Post hoc analyses revealed that accuracies for non-cognates were significantly higher for the Hybrid than the 50% Translation criterion ($z = -3.42, p = .001$). Potentially, this pattern contributes to the larger cognate advantage found with the 50% Translation criterion relative to the Hybrid.

Number of child bilinguals showing a cognate advantage. The four cognate identification methods found differing numbers of children as showing a cognate advantage. Under the COSP criterion, 35 of 73 children had a higher cognate than non-cognate accuracy rate; under the 50% Translation, 48; under the 75% Translation, 44; and the Hybrid, 41. However, these proportions only differed significantly between the 50% Translation and COSP criteria ($\chi^2(1) = 4.72, p = .03$). Only the 50% and 75% Translation criteria found a proportion of children showing the cognate effect that differed from chance, according to a combined sign test: 50% Translation, $p = .005$; 75% Translation, $p = .04$.

Discussion

Although the COSP criterion failed to detect a cognate advantage, the remaining three criteria did yield cognate effects for preschool-aged Spanish-English bilinguals. Thus, it appears that methodological decisions regarding cognate identification methods may impact findings of cognate effects in young bilinguals. Specifically, subjective criteria—which selected fewer test items as cognates—detected larger and more consistent cognate effects in young bilinguals. Of the four approaches, only the 50% and 75% Translation criteria identified a proportion of child bilinguals with cognate effects that differed from chance. Potentially, introducing objective phonological criteria may mask cognate effects in child bilinguals. Not only did the COSP criterion lead to no significant differences between cognate and non-cognate accuracy rates (i.e. no cognate advantage), but the introduction of an objective standard alongside a subjective one with the Hybrid criterion also appeared to attenuate cognate effects. For example, the cognate advantage detected by the Hybrid criterion was significantly smaller than that detected by the subjective 50% Translation criterion.

Recall that the objective criterion selected significantly more test items than any other approach, and effectively cast a wider net for what qualifies as a cognate. The difference in findings of cognate effects for the COSP and the subjective approaches in child bilinguals suggests that young bilinguals may need more than pure phonological overlap to make use of cross-linguistic similarities. In other words, it appears that young bilinguals have not yet learned to make use of all available cues of cross-linguistic similarity and are especially reliant on salient perceptual similarities.

Overall, these results largely align with previous work on cognate effects in child bilinguals. Relative to reported findings of cognate effects in adults, cognate effects in these young bilinguals appeared to be present but were inconsistent—only subsets of the sample demonstrated the advantage, and effects were captured with varying degrees of success by different cognate identification criteria. Interestingly, effects were found despite this sample's young age.

General discussion

In Experiment 1, we compared four methods of determining cognate status: an objective phonological criterion (COSP), subjective 50% and 75% Translation criteria, and a Hybrid criterion that combined objective and subjective elements. We found that these methodological differences gave rise to differing assignments of cognate status for *PPVT-III* test items. In our second and third experiments we found that, in turn, differences emerged in cognate accuracy rates, non-cognate accuracy rates, and in the magnitudes of cognate advantages across the four cognate selection criteria. Unlike prior research on the cognate advantage, we investigated effects in both child and adult participants. Inclusion of data from adults and children allowed us to examine cognate identification criteria and cognate effects in two populations that appear to differ in the robustness of the cognate effect.

Findings from our first experiment suggested that the COSP criterion identified more cognates than any other approach (102 of 204) and the 75% Translation criterion selected the fewest (37). There was also substantial disagreement between methods: Only 30 of the 204 test items were unanimously selected as cognates, but 84 words were selected by one, two, or three of the available criteria (see Table 2). Objective and subjective criteria differed markedly, as evidenced by the fact that 53 words selected as cognates by the COSP were selected as non-cognates by the remaining methods. Differences were particularly pronounced in the final *PPVT-III* test sets, which include the most challenging items.

The goal of our second experiment was to test whether these detected differences between cognate selection criteria would impact findings of cognate effects in adult bilinguals. Indeed, the COSP criterion resulted in a lower cognate accuracy rate, a smaller cognate advantage, and a smaller proportion of individuals with a cognate advantage than the remaining methodologies. These patterns can be explained by the large number of cognates identified by this approach, including high-difficulty items that may have lacked easily recognizable cross-linguistic overlap or that the speakers may not have recognized in one or either language (e.g. *incarcerating*/*encarcelar*, COSP = 6, no correct back-translations). In contrast, the highly selective 75% Translation criterion resulted in particularly high *non*-cognate accuracy rates. Test items with partial cross-linguistic form overlap that were classified as cognates by the remaining criteria appeared in this criterion's non-cognate subset, potentially inflating adult speakers' non-cognate accuracy rate. Although the COSP and 75% Translation criterion may be less optimal for measuring the cognate advantage in adults, cognate effects were robust in this group and were found under each criterion.

Finally, in our third experiment, we extended the investigation to young bilinguals. While the cognate advantage was consistently present in adults, cognate effects in children were found with three criteria but not with the COSP criterion. Further, in children, an important new distinction emerged: only the subjective criteria identified a substantial proportion of individuals with a cognate advantage that differed from chance. The lack of cognate effects with the COSP criterion and the overall lower proportions of individuals who showed the advantage suggest that the children in this sample were limited in their sensitivity to cross-linguistic overlap relative to the adult bilinguals. In interpreting these findings, it is important to recall the specific profiles of our bilingual groups (see *Method* sections for Experiments 2 and 3). As illustrated in the current sample, young bilingual children, who are at various stages of learning English, may by definition show a wide range of English proficiencies. These individual differences pose challenges for language assessment. Further examination of cognate effects, and their relation to language development, may open another avenue for determining typical language development in these young bilinguals. Specifically, future work can consider why some, but not all, child bilinguals demonstrated a cognate advantage, and can search for relationships between the cognate advantage and bilingual profile characteristics (e.g. relative exposure) or measures of language development (e.g. standardized language performance, language sample measures). Additional work with other young bilingual populations (e.g. a sample with relatively high English exposure) would be needed to help determine whether the relatively attenuated cognate sensitivity found in this sample is a dependable pattern.

Overall, these findings support and add to previous research on the cognate advantage. Consistent with previous work, adults consistently demonstrated higher cognate than non-cognate accuracy rates (e.g. Sánchez-Casas & García-Albea, 2005). Because a cognate advantage was present under all four criteria, it appears that adult bilinguals showed facilitated processing both for near-identical cognates (those selected by all four criteria; for example, *closet/clóset*) and for cognates with less cross-linguistic overlap (those selected by only one criterion; for example, *confiding/confiar*). In contrast, cognate effects in child bilinguals were less resilient to methodological differences, as indicated by the absence of a cognate advantage under the COSP criterion. Consistently, prior work on the cognate advantage in children has used objective cognate selection criteria and found limited effects (e.g. Kelley & Kohnert, 2012; Pérez et al., 2010). Our findings suggest that perhaps the advantage would be more consistent across child participants if a subjective criterion for cognate status were used.

In addition to informing methodological considerations, examination of child and adult performance also provided for exploratory considerations of cognate effects across the lifespan. We acknowledge that the current data provide only a preliminary developmental comparison since our child and adult participants differ in current exposure to Spanish and English (with child bilinguals being Spanish-dominant and adult bilinguals being English-dominant), but this difference actually reflects common developmental trajectories of heritage Spanish speakers in the United States. After spending years in an English school system, the child bilinguals in this study will likely become more balanced in their language exposure and will thus resemble the current adult sample. Importantly, only Spanish-English bilingual adults whose early language profiles resembled those of the current child participants were included in this study. Thus, considering the two groups' patterns of cognate effects suggests that cognate sensitivity may emerge and become more robust over time. This finding is further supported by comparing across previous studies of cognate effects in adult and child bilinguals.

The identified differences in cognate effects in children and adults also indicate that not all assumptions about adult bilingual lexical processing may extend to child speakers. Current models of adult bilingual lexical processing (e.g. *Bilingual Language Interaction Network for Comprehension of Speech*, Shook & Marian, 2013; the *Bilingual Interactive Activation Model*, Dijsktra & Van Heuven, 2002; Dijkstra, Van Heuven, & Grainer, 1998) all posit a partially integrated lexicon in which high cross-linguistic overlap—as seen in cognates—explains advantages in performance. However, if child bilinguals show different patterns of performance, models may need to be adapted to account for child behavior and the developmental trajectory of cognate effects. For example, distinctions may be drawn between translation pairs with high and low phonological overlap and between high- and low-frequency words (that may or may not be known in both languages). Consistent with the assumption of an integrated lexicon, we predict that children initially show processing advantages for highly form-similar cognates that are frequent in both languages, thus providing robust cross-linguistic scaffolding. As children's overall vocabulary and metalinguistic awareness expand, they may learn to make use of phonological overlap with less frequent cognate pairs, and they may become more sensitive to cognate pairs with incomplete cross-linguistic form overlap. Additional research is needed to address these predictions, to control for patterns of language exposure across age groups where possible, and to plot the developmental trajectory of the cognate advantage.

Limitations of this study include working with a predetermined word list, precluding us from controlling for word frequency or difficulty. Potentially, the COSP criterion would not have diverged from the remaining criteria as strongly if there were fewer high-difficulty items. Further, due to test administration protocol, not all participants completed the same sets or number of items. Both of these factors may have impacted child bilinguals in particular, whose vocabularies are not yet fully developed and who therefore completed fewer items than the adults. Nevertheless, this method offered the opportunity to conduct initial investigations of methodological impacts on cognate effects in child and adults bilinguals using a format appropriate to both age groups. Further, the current findings can inform interpretation of performance on the *PPVT-III* in Spanish-English bilinguals. To better investigate cognate effects, factors like word frequency and phonological overlap would need to be experimentally manipulated. Carefully designed cognate probes would be an exciting way to explore why some, but not all, children demonstrate a cognate advantage.

Conclusions

The present work may be used to guide methodological decisions in future research on the cognate advantage. For example, when working with high-difficulty, low-frequency words (as in the latter *PPVT-III* test sets), objective phonological criteria may over-identify cognates. In this study, the COSP criterion detected the weakest effects in adult speakers and failed to identify cognate effects in child speakers. Thus, we concluded that this objective approach is not the most advantageous for assessing cognate effects, especially for child bilinguals or when difficult items are included as stimuli. This is useful information for those investigating the cognate advantage in languages such as English and Spanish, for which many cognates are lower-frequency words (Schepens, Dijkstra, Grootjen, & Van Heuven, 2013). Participants' age should also play a role in methodological decisions. For child speakers in particular, subjective criteria may be valuable, as we found that an objective criterion overlooked cognate effects. For adult speakers, however, the 75% Translation criterion seemed to under-identify cognates. In future work, we recommend either setting a lower threshold (i.e. 50%), or using the higher threshold but refining the non-cognate set to filter out near-cognates (e.g. *inhale*/*inhalar*, correctly back-translated by 67%).

Ultimately, a balance between objective and subjective cognate identification approaches is desired. The weak cognate effects with the COSP criterion suggest that the cognate identification process is well-served by a subjective component. However, objective analysis is necessary to ensure that speaker judgments reflect true form similarities and not other factors, like familiarity (e.g. *heart*/*corázon*). Thus, we expected the Hybrid criterion to emerge as the superior cognate identification method. In fact, this was a suitable approach for adult bilinguals, as evidenced by the lack of extreme accuracy rates like those found with the COSP and the 75% Translation criteria. However, for child bilinguals, even this limited inclusion of phonological criteria appeared to dilute cognate effects.

To conclude, although the general characterization of a cognate word is widely known, this study provides compelling evidence that the operational criteria for cognate status warrant attention. The use of different methods can yield meaningful differences in stimulus selection. Further, the criteria can impact bilinguals' performance on cognates, especially in the case of child speakers. Our findings also attest to the robustness of cognate effects in adult speakers and suggest that some preschool-aged bilinguals are also sensitive to cross-linguistic overlap, particularly for highly transparent translation equivalents.

Acknowledgements

The authors would like to thank Alaina Kelley and Kathryn Kohnert for sharing their COSP rating criteria, as well as members of the Language Development, Disorders and Diversity Lab and the Bilingualism and Cognition Lab at San Diego State University for research assistance and helpful discussions of this work.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The first author was supported by an NIH training grant [grant number T32 DC007361]. In addition, funding for the project was provided by San Diego State University Grants Program Grant [grant number 242338] and an ASH Foundation New Investigator Grant awarded to the second author and an NIH NIDCD RO3 [grant number DC012141] and a Price Charities research grant awarded to the third author.

Note

1. One family did not report exposure percentages but listed Spanish as the language the child heard and used in the home. The child's teacher confirmed English use in the classroom.

References

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2009). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*, 525–531.

Boada, R., Sanchez-Casas, R., Gavilan, J. M., Garcia-Albea, J. E., & Tokowicz, N. (2013). Effect of multiple translations and cognate status on translation recognition performance of balanced bilinguals. *Bilingualism: Language and Cognition*, *16*, 183–197.

Brenders, P., Van Hell, J., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology*, *109*, 383–396.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, *13*, 212–214.

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, *95*, 94–103.

Davis, C., Sánchez-Casas, R., García-Albea, J. E., Guasch, M., Molero, M., & Ferré, P. (2010). Masked translation priming: Varying language experience and word type with Spanish-English bilinguals. *Bilingualism: Language and Cognition*, *13*, 137–155.

De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*, 1–56.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, *30*, 90–123.

De Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning*, *47*, 215–226.

Dijkstra, A., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175–197.

Dijkstra, T., Van Heuven, W., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica*, *38*, 177–196.

Dressler, C., Carlo, M., Snow, C., August, D., & White, C. (2011). Spanish-speaking students' use of cognate knowledge to infer the meaning of English words. *Bilingualism: Language and Cognition*, *14*, 243–255.

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, *23*, 166–180.

Dunn, L., & Dunn, L. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Dunn, L., Padilla, E., Lugo, D., & Dunn, L. (1986). *Test de vocabulario en imágenes peabody—adaptación hispanoamericana*. Circle Pines, MN: American Guidance Service.

Friel, B. M., & Kennison, S. M. (2001). Identifying German–English cognates, false cognates, and noncognates: Methodological issues and descriptive norms. *Bilingualism: Language and Cognition*, *4(3)*, 249–274.

Gollan, T. H., Fennema-Notestine, C., Motoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society*, *123*, 197–208.

Gutierrez-Clellen, V. F. (1999). Language choice in intervention with bilingual children. *American Journal of Speech-Language Pathology*, *8*, 291–302.
Kelley, A., & Kohnert, K. (2012). Is there a cognate advantage for typically developing Spanish-speaking English language learners? *Language, Speech, and Hearing Services in Schools*, *43*, 191–204.
Kohnert, K., Windsor, J., & Miller, R. (2004). Crossing borders: Recognition of Spanish words by English-speaking children with and without language impairment. *Applied Psycholinguistics*, *25*, 543–564.
Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.
Marian, V., Blumenfeld, K. H., & Kaushanskaya, M. (2007). The Language Proficiency and Experience Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*, 940–967.
Millett, J., Atwill, K., Blanchard, J., & Gorin, J. (2008). The validity of receptive and expressive vocabulary measures with Spanish-speaking kindergarteners learning English. *Reading Psychology*, *29(6)*, 534–551.
Nagy, W., García, G., Durgunoglu, A., & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Literacy Research*, *25*, 241–259.
Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, *18*, 41–58.
Pérez, A., Peña, E., & Bedore, L. (2010). Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Child Services*, *4*, 55–67.
Rosselli, M., Ardila, A., Jurado, M., & Salvatierra, J. (2012). Cognate facilitation effect in balanced and non-balanced Spanish–English bilinguals using the Boston Naming Test. *International Journal of Bilingualism*, *0*, 1–14.
Sánchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory: Can cognate status be characterized as a special kind of morphological relation? In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism; psycholinguistic approaches* (pp. 226–250). New York, NY: Oxford University.
Schepens, J., Dijkstra, T., Grootjen, F., & Van Heuven, W. J. B. (2013). Cross-language distributions of high frequency and phonetically similar cognates. *PLoS ONE*, *8(5)*, 1–15.
Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for comprehension of speech. *Bilingualism: Language and Cognition*, *16*, 304–324.
Umbel, V., & Oller, D. K. (1994). Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning*, *44*, 221–242.
Umbel, V., Pearson, B., Fernández, M., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development*, *63*, 1012–1020.
Van Assche, E., Duyck, W., & Brysbaert, M. (2013). Verb processing by bilinguals in sentence contexts: The effects of cognate status and verb tense. *Studies in Second-Language Acquisition*, *35*, 237–259.
Van Hell, J., & De Groot, A. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*, 193–211.

Author biographies

Irina Potapova is in the Joint Doctoral Program in Language and Communicative Disorders at San Diego State University and the University of California, San Diego. Her research focuses on the scaffolding and other interactions across bilingual speakers' two languages.

Henrike K Blumenfeld is an assistant professor in the School of Speech, Language, and Hearing Sciences at San Diego State University. Her research focuses on lexical processing and the relationship between linguistic and cognitive processes in young adult and aging bilinguals.

Sonja Pruitt-Lord is an associate professor in the School of Speech, Language, and Hearing Sciences at San Diego State University. Her research focuses on child language development within the context of language variation and poverty and efficacy of prevention models for "at-risk" populations.

Chapter 4, in full, is a reprint of material as it appears in Potapova, I., Blumenfeld, H. K., & Pruitt-Lord, S. (2016). Cognate identification methods: Impacts on the cognate advantage in adult and child Spanish-English bilinguals. *International Journal of Bilingualism*, *20*(6), 714-731. The dissertation author was the primary investigator and author of this paper.

CHAPTER 5:

Cross-Language Interactions During Word Learning in Bilingual Children with Typical and Atypical Language Development

Abstract

Mature bilingualism is characterized by cross-language interactions, or the capacity for knowledge of one language to impact performance in the other. Emerging evidence suggests that similar cross-language interactions are demonstrated by typically developing child bilinguals, as indicated by their enhanced performance for cognate (i.e., *elephant/elefante* in English/Spanish) relative to non-cognate targets (e.g., *bird/pájaro*). The current study advances research in this area by testing for cognate sensitivity in the context of a novel word learning task and by utilizing measures of eye gaze alongside off-line accuracy rates. Participants included three groups of Spanish-English bilingual preschoolers: children with typical language who were enrolled in a Spanish-language classroom; children with typical language who were enrolled in an English-language classroom, and children with atypical language development who were enrolled in an English-language classroom.

Given the challenging task of learning novel words following relatively few exposures, resulted indicated that only typically developing bilingual children with exposure to Spanish at home and in school demonstrated cognate sensitivity, as evidenced by enhanced accuracy for and fixations to cognate targets. Conversely, bilingual children with atypical language skills who were enrolled in English language classrooms did not demonstrate cognate sensitivity and trended towards enhanced performance for learning non-cognates relative to cognates. Results thus suggest that cognate sensitivity may be associated with language experience and, potentially, language ability. Continued efforts to understand how bilingual children with various profiles scaffold information across their two languages will inform clinical approaches for children with dual language exposure, including the use of cognates in bilingual language assessment and treatment.

Introduction

The number of bilinguals in the United States has steadily increased (U.S. Census Bureau, 2016), and it is projected that 40 percent of children in the United States will learn English as their second language by 2030 (U.S. Department of Education & National Institutes of Child Health and Human Development, 2003). Accordingly, speech-language pathologists can expect to see children from culturally and linguistically diverse backgrounds on their caseloads (e.g., Caesar & Kohler, 2007). Accurate assessment of language abilities in bilingual children requires practitioners to recognize language differences, or non-clinical patterns of language use that reflect a child’s language experience, and then determine if patterns of disorder are present within those language differences (Oetting, 2018). Best clinical practice for bilingual clients thus includes assessing performance in each language separately and comparing a child’s performance in each language to appropriate bilingual comparison groups or norms (e.g., Potapova, Kelly, Combiths & Pruitt-Lord, 2018; Gillam et al., 2014; Gutierrez-Clellen, Simon-Cereijido & Wagner, 2008).

To complement these methods and support clinical approaches tailored to bilingual speakers, it is important to recognize that the two languages are not fully independent, and, instead, there is potential for knowledge of one language to impact performance in the other (e.g., Potapova & Pruitt-Lord, 2019). While there is substantial research in such cross-language interactions in adult bilinguals, and a growing body of work in typically developing child bilinguals, this potential is not yet well understood in the context of atypical language development. To address this gap, the present study aims to better understand cross-language interactions in preschool-aged bilinguals with typical development and bilingual peers with specific language impairment (SLI). Here, we test whether children are sensitive to cross-language similarity as they encounter novel words in each language. Results will inform our understanding of typical and atypical language development in the context of dual language exposure and have implications for clinical approaches for bilingual children.

Cross-language interactions in bilingual speakers

Typically developing child bilinguals and healthy adult bilinguals alike have been shown to scaffold information from one language into the other, even when there is no indication that the non-target language is relevant to the task at hand (see Kroll, Dussias, Bogulski & Valdes-Kross, 2010a; Potapova & Pruitt-Lord, 2019). Frequently, cross-language interactions are assessed through sensitivity to cognates, or translation equivalents that share sound and meaning across languages (e.g., *elephant* and *elefante* in English and Spanish) relevant to non-cognates (e.g., *bird/pájaro*). For example, adult bilinguals have been found to learn novel cognates more successfully than non-cognates (Lotto & De Groot, 1998; De Groot & Keizjer, 2000). Indeed, adult bilinguals' sensitivity to cross-language similarity has demonstrated in a wide variety of language tasks, including categorization (Dufour & Kroll, 1995), naming (Gollan, Sandoval & Salmon, 2007; Rosselli, Ardila, Jurado & Salvatierra, 2012), translation (De Groot & Poot, 1997) and word association (Van Hell & De Groot, 1998) and with a wide variety of measures, including accuracy (e.g., Rosselli, et al., 2012), reaction time (e.g., Lotto & De Groot, 1998), eye gaze (e.g., Blumenfeld & Marian, 2005) and event related potentials (e.g., Strijkers, Costa & Thierry, 2010). These robust findings have informed models of adult language processing, which reflect the possibility of cross-language interactions (e.g., BIA+, Dijkstra & Van Heuven, 2002; Revised Hierarchical Model, Kroll, van Hell, Tokowicz & Green, 2010; Distributed Features Model, van Hell & de Groot, 1998).

Comparable research in young bilinguals is emerging (see Potapova & Pruitt-Lord, 2019), with mounting evidence that typically developing child bilinguals also demonstrate cognate effects, or sensitivity to cross-language similarity. A substantial number of studies in this area have repurposed standardized language assessments by identifying test items as either cognates or non-cognates and found that preschool-age and school-age children demonstrate higher accuracy on cognate items relative to non-cognate items (Potapova, Blumenfeld & Pruitt-Lord, 2016; Kelley & Kohnert, 2012; Pérez, Peña & Bedore, 2010). Gradually, cognate sensitivity in young bilinguals is being considered with more rigorous methods and measures that more closely resemble work available for adult bilinguals. For example, Sheng, Lam, Cruz and Fulton (2016) controlled for the phonological structure, frequency, length and age of acquisition of English and Spanish cognate and non-cognate targets in a task that required 4- to 7-year old children to name objects in English. While Spanish-English bilingual participants had greater success with cognate items relative to non-cognate items, neither English monolinguals nor Mandarin-English bilinguals showed the same effect, indicating Spanish-English bilinguals were sensitive to cognate status, or information that was not accessible to either of the other groups.

Cognate sensitivity and cross-language interactions have been demonstrated across development and into adulthood, but the effect is not uniform. The degree of observed cognate sensitivity may be impacted by characteristics of the target items, such as their degree of cross-language similarity. While the terms “cognate” and “non-cognate” suggest a categorical split, evidence indicates that cognate status is a continuum, with graded levels of salient overlap. In both child and adult bilinguals, increasing that degree of overlap in cognate targets has been shown to enhance performance. Dutch-Frisian bilingual children followed longitudinally between ages 5 and 8 with relatively low Frisian exposure were found to have higher accuracy for identical cognates, such as *poes/poes* in Frisian/Dutch (‘cat’) than for non-identical cognates, such as *easten/oosten* (‘east’; Bosma, Blom, Hoekstra & Versloot, 2016). The same pattern has also been demonstrated by adult bilinguals, with a greater degree of observed cognate effects when cognate pairs have a higher degree of overlap, as evidenced by faster response times (Duyck et al. 2007; Dijkstra et al. 2010; Van Assche et al., 2011).

The magnitude of cognate effects has also been linked to individual differences, such as age. Older school-age bilinguals have been found be more likely to demonstrate cognate sensitivity than younger school-age bilinguals in a receptive vocabulary task (Kelley & Kohnert, 2012). Similarly, Potapova et al., (2016) found that where nearly all adult Spanish-English bilingual participants demonstrated cognate sensitivity on a standardized English receptive vocabulary test, the effect was less robust in preschool-aged Spanish-English bilinguals that completed the age-appropriate sections of the same standardized assessment. Available evidence thus suggests that, while child bilinguals have a potential for cognate sensitivity, this effect also grows with development. Language dominance has also been recognized as an important factor in observed cognate sensitivity, particularly when considered alongside the language of task administration. In both child and adult bilinguals, greater cognate effects are found when a participant is asked to perform tasks in their non-dominant language (Pérez et al., 2010; Rosselli et al., 2012). Thus, cognate effects may be better captured when speakers recruit information from their dominant language into their relatively weaker language.

Our understanding of cross-language interactions in young bilinguals is growing to match our understanding of this phenomenon in adult bilinguals—but it has yet to be extended to children with specific language impairment (SLI), a communication impairment characterized by weaknesses in both expressive and receptive language. Empirical efforts to understand cognate sensitivity for children with weak language skills are needed, considering that cognates have been recommended as treatment targets for bilingual children with an expectation that they can support cross-language transfer and generalization (e.g., Kohnert, 2010; Kambanaros, Michaelides & Grohmann, 2016). More broadly, continued efforts to expand methods and measures used to investigate cognate sensitivity in child bilinguals will bolster our understanding of cross-language interactions in development.

Cross-language interactions in children with SLI

While cross-language interactions may be considered characteristic of bilingualism, this understanding is, at present, largely informed by research in typically developing child bilinguals and healthy adults, with some evidence from adults with impaired language skills (Potapova & Pruitt-Lord, 2019; Kroll et al., 2012; Kohnert, 2004). To our knowledge, two studies have considered cognate sensitivity in children with SLI, including one study featuring English monolinguals. Kohnert, Windsor and Miller (2004) tested whether form overlap between Spanish and English facilitated recognition of Spanish words in 8- to 13-year-old English monolinguals with typical development and monolingual peers with SLI. In this study, participants were asked to match an unfamiliar Spanish label with one of two target pictures. Typically developing monolinguals demonstrated higher accuracy rates and faster response times for Spanish targets that resembled their English translation equivalents than did peers with SLI, suggesting that they were benefiting from cross-language similarity to a greater extent. Accordingly, the authors highlighted the potential contribution of cognate sensitivity to language assessment. Results of this study also revealed that within each group, performance for Spanish targets with no discernable overlap with English was lower than for targets with overlap, indicating some degree of cognate sensitivity within each group. A similar pattern of results was found by Grasso, Peña, Bedore, Hixon and Griffin (2017) in 5- to 9-year-old Spanish-English bilingual children with typical development and bilingual peers with SLI. In this study, the measure of interest was accuracy on cognate and non-cognate targets identified from Spanish and English standardized expressive vocabulary measures. While typically developing participants outperformed participants with SLI in overall accuracy, each group named cognate targets more frequently than non-cognate targets across the two languages, and authors highlighted the value of cognates in bilingual language therapy. Together, findings suggest that children with SLI do demonstrate difficulties with cognates relative to typically developing peers, but that sensitivity to cognates may nevertheless be present.

In the present study, we expand upon this research by investigating cross-language interactions in development with an approach that has improved our understanding of deficits associated with SLI: novel word learning. Children with SLI have relative difficulty in various lexical-semantic tasks, including word learning, with evidence particularly readily available for monolinguals (see Kan & Windsor, 2010, for a meta-analysis on word learning in monolingual children with primary language impairment). Difficulties have been identified both during fast mapping, which includes quickly forming form-meaning associations between novel labels and objects following limited exposure to the novel form, and during extended word learning, which includes lexical acquisition given further exposure to the novel items (e.g., Rice et al., 1994; Gray, 2005). To illustrate, Rice and colleagues (1994) used video story presentations to introduce four high difficulty nouns and four high difficulty verbs to 5-year-old English monolingual children with SLI, as well as age-matched and language-matched control groups. Given three exposures to the target words, the typically developing age-matched control group demonstrated significant growth in comprehension of the new words, while children with SLI failed to demonstrate word learning. Similarly, bilingual children with SLI have also been found to have difficulty with this task. For example, 4- and 5-year-old Spanish-speaking children with SLI demonstrated weaker comprehension for three novel Spanish words than did typically developing age-matched bilingual peers following a scripted structured play session that introduced the labels and referents (Kapantzoglou, Restrepo & Thompson, 2012).

Altogether, research in word learning and other lexical-semantic tasks suggests that children with SLI demonstrate relatively weak semantic associations. In these works, novel words have been presented in a single language; as such, only weaknesses in *within*-language semantic associations have been identified. The present study will introduce novel words in each of the bilingual children's languages and manipulate the targets' cognate status, allowing for an investigation of semantic associations *across* languages. An understanding of cross-language associations in typical and atypical bilingual language development will support clinical practices tailored to the bilingual experience. Meaningful differences in cognate sensitivity between typically developing bilingual children and bilingual peers with SLI will point to the usefulness of cognate tasks in language assessment, while relative success with cognates in children with SLI would suggest that translation equivalents with high degrees of cross-linguistic form overlap may provide useful scaffolds in language therapy.

Present study

This study investigated typical and atypical bilingual language development by considering cross-language interactions in preschool-aged Spanish-English developing bilingual children, including a group of bilingual children with SLI. Cognate sensitivity was measured through a novel word learning task that featured cognate and non-cognate targets. Word learning tasks have previously been used to both capture cognate effects in adult bilinguals (e.g., Lotto & De Groot, 1998) and to characterize typical and atypical language skills in children (Kan & Windsor, 2010; Kapantzoglou et al., 2012). The current research capitalizes on this overlap to test for cognate sensitivity in bilingual children with varying levels of language ability. Further, this method allows for cognate effects to be tested independently of a child's previous experience, as exposure to novel labels is controlled within the study; this contrasts with previous work investigating cognate sensitivity in child bilinguals, which typically features labels and objects that participants have pre-existing knowledge of. The central question driving this research was: *Do bilingual children with typical and atypical language development demonstrate cross-language interactions (i.e., cognate effects) as they encounter novel words in each of their languages?*

As in related work, cognate sensitivity was measured by comparing accuracy rates for cognate and non-cognate targets during a receptive comprehension task, with higher accuracy rates for cognates relative to non-cognates reflecting cross-language interactions. In addition, cognate sensitivity was measured via eye tracking during the same receptive comprehension task. Eye tracking offers an opportunity to identify effects not present in overt behavioral responses, such as selecting an image by pointing (Conklin & Pellicer-Sánchez 2016; Lai et al., 2013; Hendrickson, Mitsven, Poulin-Dubois, Zesiger & Friend, 2015). This relatively nuanced measure has been utilized in adult studies of cognate sensitivity (e.g., Blumenfeld & Marian, 2005) and to detect partial word knowledge in young children (Hendrickson et al., 2015).

Given the lexical-semantic weaknesses associated with the profile of SLI, bilingual children with SLI were expected to differ from bilingual children with typical development in terms of cognate sensitivity, as measured by both off-line accuracy and fixation patterns.

Method

Participants

Spanish-English developing bilinguals were recruited from a local preschool site as part of an on-going community-based research project conducted by the second author. The preschool site required below-poverty standards to participate and featured both English-language and Spanish-language classrooms. Inclusionary criteria for the present study were that children were at least four years of age during preliminary data collection (see below); that Spanish exposure in the home was at least 30 percent; that non-verbal cognition scores were within or above normal limits; that spontaneous language samples were completed at the beginning of the academic year in both Spanish and English; and that both caregiver and teacher questionnaires were completed during data collection.

Bilingual participants were identified as having typical language (BiTD) or specific language impairment (BiSLI) using a combination of direct and indirect language measures in each of their languages described in the following section. For inclusion in the BiSLI group, participants were required to meet at least three of the four following criteria relevant to identifying atypical language development in children from culturally and linguistically diverse backgrounds: (a) caregiver concern regarding language development provided via questionnaire; (b) below average performance on spontaneous language measures in both Spanish and English, as measured by mean length of utterance in words; (c) presence of an Individualized Education Plan; and (d) average teacher ratings of less than 4.18 on the *Inventory to Assess Language Knowledge* (ITALK; Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2018), a questionnaire reflecting English and Spanish performance across language domains. Conversely, for children to be considered for either BiTD group, there could be no caregiver concern regarding language development and no more than one of additional three indicators could be present.

In total, four participants met the criteria for the BiSLI group (mean age = 4;3, SD = 3.30 months; 1 female). Each of children that met the criteria for the BiSLI group was enrolled in an English language classroom. An additional 22 participants met the criteria for the BiTD group, including four children enrolled in Spanish-language classrooms (BiTD-Span; mean age = 4;2; SD = .96 months; 4 females) and 18 children enrolled in English-language classrooms. To best allow for group comparisons, four BiTD children enrolled in English language classrooms were matched with children from the BiSLI and BiTD-Span on the basis of age and Spanish exposure at home, resulting in a BiTD-Eng group (mean age = 4;4; SD = 2.2 months; 3 females). Thus, a total of 12 children were included in the present study.

All participants spoke Spanish as a native language, were exposed to Spanish at home and were comparable in age; see Table 1 for group characteristics and Appendix A for individual participant characteristics. In addition, non-verbal cognition was within or above normal limits for all participants, as measured by the *Leiter International Performance Scale–Revised* (Roid & Miller, 1997), a non-verbal cognition measure (mean = 13.17, SD = 2.27, Range = 9.5–17). Maternal education was reported to be 10.83 years (SD = 3.1, Range = 6–16), on average. Despite these similarities in demographic variables relevant to language performance (all *ps* > .600, except for the *Leiter-R*, *p* = .083), children in the BiSLI group significantly differed from both BiTD groups in terms of language performance, as would be expected. Not only did children in the BiSLI group demonstrate significantly lower performance on measures that were used to separate the groups (mean length of utterance in words in a spontaneous English language sample, mean length of utterance in words in a Spanish spontaneous language sample, and teacher ITALK ratings; *p*s from one-tailed t-tests = .037, .003 and .001, respectively), this group also had a significantly lower number of complete and intelligible utterances in their spontaneous language samples in Spanish and English, a smaller number of different words in spontaneous language samples in Spanish and English, and lower Language Index scores on the *Bilingual English-Spanish Assessment* (*BESA*; Peña, et al., 2018; *p* values ranged from .014 to .048). Meanwhile, the BiTD-Span and BiTD-Eng groups were similar in language performance in each of those language measures (*p* values ranged from .235 to .905). Thus, the BiSLI group demonstrated lower levels of language performance than children in the BiTD groups, while children in the two BiTD groups were seemingly comparable in language abilities.

Table 1. Participant characteristics for BiTD-Span, BiTD-Eng and BiSLI groups.

		Age	% Spanish Heard	% English Heard	Maternal Education (in years)	Leiter-R	MLUw – English*	MLUw – Spanish†	ITALK†	BESA Language Index†
BiTD-Span	*Mean*	4;2	76.39	23.61	11.75	15.13	2.5	3.26	4.6	101.75
	SD	.96 months	27.36	27.36	.50	2.37	1.08	1.17	.69	12.77
BiTD-Eng	*Mean*	4;4	82.50	17.50	10.50	12.88	2.89	3.65	4.55	108.75
	SD	2.22 months	23.63	23.63	3.11	1.93	.70	.62	.09	5.17
BiSLI	*Mean*	4;3	81.55	18.45	10.25	11.50	1.66	1.63	2.85	78.25
	SD	3.30 months	13.52	12.52	4.65	2.48	.24	.23	.79	11.71

*$p < .05$ for comparisons between the BiSLI and two BiTD groups
†$p < .01$ for comparisons between the BiSLI and two BiTD groups

Procedure

Through partnerships with teachers and classroom personnel, caregivers received information about a larger study led by the second author, information about the present study, consent forms, and caregiver questionnaires regarding child language history. Children whose parents provided signed consent forms completed language assessment batteries associated with the larger research project at the beginning of the year. As part of this preliminary data collection, participants were tested individually across multiple planned sessions on-site at the preschool. All children completed the standard language assessment battery, including spontaneous play-based English language samples. In addition, children whose parents reported use of Spanish at home completed spontaneous play-based language samples in Spanish, as well as the Morphosyntax and Semantics subtests of the *BESA.* Data collection for all participants was completed within the span of 2–4 weeks. Information collected at this time determined eligibility for the present study and informed group assignment for the present study, as described above. Examiners included supervised graduate and undergraduate students in speech-language pathology trained to administer the assessment batteries; examiners were fluent in the language corresponding to the task.

Data for the present study were collected at the midpoint of the academic year, also on-site at the preschool. Participants were tested individually by trained graduate and undergraduate students in speech-language pathology who were fluent in both Spanish and English. The present study included multiple planned sessions in which participants were introduced to novel words in Spanish and English in the context of child-friendly stories and then tested for their receptive comprehension of novel target labels in both languages (see Table 2 and the following sections). In Session 1, participants viewed a Spanish story and completed comprehension tests in Spanish (Comprehension Test 1) and English (Comprehension Test 2); examiners were trained to use Spanish with the participant until the Spanish comprehension test had been completed. In Session 2, completed on a separate day, participants viewed the English version of the story and then completed comprehension tests in English (Comprehension Test 3) and Spanish (Comprehension Test 4); examiners were trained to use English with the participants until the English comprehension test had been completed.

Table 2. Summary of session format and tasks.

	Session 1 One			**Session Two**		
Task	Story	Comprehension Test 1	Comprehension Test 2	Story	Comprehension Test 3	Comprehension Test 4
Language	Spanish	Spanish	English	English	English	Spanish
Content	Paired four novel visual referents with novel Spanish labels - Two cognate labels - Two non-cognate labels	24 total trials - Two cognate labels - Two non-cognate labels	24 total trials - Two cognate labels - Two non-cognate labels	Paired the same visual referents with novel English labels - Two cognate labels - Two non-cognate labels	24 total trials - Two cognate labels - Two non-cognate labels	24 total trials - Two cognate labels - Two non-cognate labels

Stimuli

Novel words. Each participant was introduced to a total of eight novel labels, or four pairs of Spanish-English translation equivalents matched with four novel visual referents (e.g., Kan & Kohnert, 2008; Rice et al., 1994). Spanish labels were presented in a Spanish story in Session 1; English labels were presented in an English story in Session 2. The key manipulation in the novel labels was the cognate status of each target. Two of the translation-equivalent pairs were cognates across English and Spanish (e.g., *codon/codón*) and two pairs were non-cognates (e.g., *tamiz/fathom*). Both cognate and non-cognate novel labels were in fact high-difficulty two-syllable English and Spanish words. Targets were novel in the sense that they were unfamiliar to preschool-aged children and referred to novel objects. The use of true word forms in Spanish and English ensured that each item reflected the phonotactics of its respective language.

Potential cognate and non-cognate stimuli were identified from scientific materials (e.g., glossaries for specialized professional fields). From this list of potential stimuli, targets were selected that allowed for (1) highly salient overlap between cognate pairs across English and Spanish; (2) minimal overlap between non-cognate pairs across English and Spanish; and (3) balance between cognates and non-cognates targets within each language. All potential words were required to have a structure of CVCVC (e.g., *codon*) or CVCCVC (e.g., *gasket*). To be considered a potential cognate target, translation equivalents in English and Spanish were required to share one hundred percent orthographic overlap, excluding accent marks (e.g., *codon/codón* in English/Spanish), resulting in high degrees of phonological overlap. Conversely, non-cognate translation equivalents across English and Spanish were not allowed to share any consonants in the same word position. Thus, a visual referent matched with the Spanish non-cognate *tamiz* could not also be paired with an English label with word initial /t/, word medial /m/ or word final /s/ or /z/; accordingly, the translation equivalent for *tamiz* in the present study was *fathom*. To further ensure that overlap existed between cognate pairs only, no labels aside from the intended cognate pair were allowed to share consonants in the same word position. Thus, to include the cognate pair *codon/codón*, no other cognate or non-cognate labels in either English or Spanish could have word initial /k/, word medial /d/ or /ð/ or word final /n/.

To create balance between the two cognate and non-cognate targets within each language, word length and phonotactic probability were considered. First, both cognate and non-cognate targets included one CVCVC pair (e.g., cognate: *codon/codón;* non-cognate: *gullet/matiz*) across English and Spanish and one CVCCVC pair (e.g., *torpor/torpor;* non-cognate: *fetlock*/*cincel*). Finally, cognate and non-cognate targets were balanced in terms of total biphone positional frequencies and single phoneme positional frequencies using the English and Spanish CLEARPOND Databases (Marian, Bartolotti, Chabal & Shook, 2012). Across cognate and non-cognate targets, all measures of phonotactic probability were comparable within each language (*p* values ranged from .217 to .896). To further ensure that observed effects were related to cognate status and not to characteristics of specific target labels, two combinations of cognate and non-cognate targets were created that met all criteria described above. For a summary of the cognate and non-cognate translation equivalents used in the present study, see Table 3. Participants were randomly assigned to one combination of target words (Combination 1) or the other (Combination 2) prior to assignment of group status.

Though the manipulation of cognate status was key to the current study, this information was not explicitly provided to participants. Children were instructed that they would hear “nuevas palabras” in Session 1 and “new words” in Session 2, but no mention of cross-language similarity was provided in either session.

Table 3. Novel cognate and non-cognate labels.

		Combination 1				Combination 2			
		Cognates		Non-cognates		Cognates		Non-cognates	
		CVCVC	CVCCVC	CVCVC	CVCCVC	CVCVC	CVCCVC	CVCVC	CVCCVC
Spanish	Label	codon	torpor	matiz	cincel	radon	vector	tamiz	dintel
	Phonological transcription	/koðon/	/toɾpoɾ/	/matis/	/sinsel/	/raðon/	/bektoɾ/	/tamis/	/dintel/
	Individual Position Frequency	1.43	1.46	1.41	1.42	1.39	1.54	1.38	1.46
	Total biphone frequency	1.08	1.04	1.05	1.04	1.03	1.05	1.03	1.07
English	Label	codon	torpor	gullet	fetlock	radon	vector	fathom	gasket
	Phonological transcription	/koʊdɑn/	/tɔɹpɚ/	/gʌlɪt/	/fɛtlɑk/	/ɹeɪdɑn/	/vɛktɚ/	/fæðəm/	/gæskɪt/
	Individual Position Frequency	1.29	1.30	1.30	1.28	1.25	1.37	1.24	1.41
	Total biphone frequency	1.01	1.01	1.02	1.01	1.01	1.02	1.01	1.03

Child-friendly stories. Child friendly stories were designed to introduce participants to novel labels and visual referents. In the story, participants were introduced to two characters on a space adventure whose rocket ship had broken. The characters searching for tools to fix their rocket ship provided the context to introduce four novel objects and their corresponding (cognate and non-cognate) labels.

The English story script was created by the first author and subsequently translated into Spanish by native Spanish-speaking research assistants. English and Spanish stories were matched in content and provided six exposures to each novel label in total. Two versions of the story were created in each language, differing only in the use of target labels from Combination 1 or Combination 2. The story scripts in English and Spanish were then recorded by native speakers of English and Spanish in sound-treated booths.

Story images included a total of 12 panels with vector images edited in Adobe Illustrator. The novel visual referents for the target labels were selected from the Novel Objects and Unusual Names Database (Horst & Hout, 2016). Each novel objects and its corresponding label were introduced in a manner that highlighted their pairing—the object was prominently displayed in the panel by a character with no other distracting objects and its name was repeated in the context of the story (e.g., *Andrea looked around the ship and she found a codon. "Look, I have a codon, maybe a codon can fix our rocket ship." Andrea thought that the codon could help.)*

Participants were seated comfortably in front of 15.6-inch portable computer monitors to view the stories. Each story panel was preceded by neutral panels with no images or audio, allowing the examiner to monitor the participant's engagement prior to initiating the presentation of each story panel.

Comprehension tests. To test for the children's learning of each target label and for cross-language interactions, participants completed comprehension tests in each language following the story. Comprehension tasks were presented on the same portable monitor as the stories and required participants to match a target label presented via pre-recorded audio to one of four visual referents. Audio for the comprehension tasks was recorded in sound-treated booths by the same native English and Spanish speakers that narrated the stories. Eye gaze data was collected at a rate of 120 Hz throughout the task with a remote Tobii x3-120 eye tracker mounted on the portable monitor. Prior to completing the comprehension test, participants completed an engaging calibration task using Tobii Pro Lab software to support accurate eye tracking.

Preceding each experimental trial was a blank screen with a centrally located star that participants were instructed to fixate; once the participant's attention had been confirmed by the examiner, the test trial would begin. The four novel objects would appear in the quadrants of the screen, and participants heard “Show me the _____” in English tests and “Enseñame el_____” in Spanish tests. Target word onset occurred 500ms after the images appeared on screen. While gaze data was being collected, participants also indicated their selected match to the auditory target using a child-friendly pointer. The pointer allowed for children to indicate a choice without physically approaching the screen and impeding eye tracking. To ensure that participants used the pointers successfully, participants completed a training activity with familiar stimuli (e.g., a car, a dog) that matched the structure of the comprehension task prior to beginning the comprehension tests. In addition, examiners reminded participants of how to use the pointer between test trials, as needed.

In total, participants completed 24 trials per comprehension test, including 12 trials with cognate-targets and 12 trials with non-cognate targets (i.e., six trials for each of the four novel labels per language). Two presentation orders were created, Order A and Order B. Participants were randomly assigned to presentation order irrespective of group status. The order that a participant received was counterbalanced across languages and across sessions. To illustrate, a participant that received Order A for Comprehension Test 1 in Spanish then received Order B in Comprehension Test 2 in English during Session 1. Then, in Session 2, that same participant completed Comprehension Test 3 in English in Order A and then Comprehension Test 4 in Spanish test in Order B. Within each order, trials were pseudo-randomized such that the same word was not the target for more than one trial in a row. In addition, images were counterbalanced such that the correct answer appeared in each quadrant an equal number of times and such that each label appeared as the target in each quadrant at least once and no more than twice.

Each of the four comprehension tests was later analyzed for accuracy and fixation patterns to test for cognate sensitivity. Importantly, each test offered different opportunities for participants to demonstrate cognate effects. Comprehension Test 1 was administered in Spanish immediately following the Spanish story that introduced the novel words (see Table 2); as such, cognate status was not yet evident, and no cognate effects were expected. Comprehension Test 2 was administered in English immediately following Comprehension Test 1 and challenged the children to pair the same visual referents with entirely unfamiliar English labels. Comprehension Test 2 thus offered the first opportunity for cognate effects, as children may have recognized and benefitted from the similarity between the Spanish and English labels for cognate targets. The presentation of the English story in Session 2 formally provided the pairing between the visual referents and their English labels. Thus, cognate effects were again expected in Comprehension Test 3, administered in English, and in Comprehension Test 4, administered in Spanish.

Results

Preliminary analyses

The goal of the present work was to understand whether bilingual children differed in their sensitivity to cross-language similarity during word learning as a function of language ability. To investigate this, it is helpful to first summarize children's performance on the comprehension test. Overall, all children in the study demonstrated above-chance recognition of at least one word across comprehension tests (i.e., accuracy higher than .25 given a field of four), with some children reliably recognizing each of the four labels by Comprehension Tests 3 and 4. Given this pattern of performance, the task was understood to be challenging but within an appropriate range for this population. For group-level performance, see Table 4; for performance summaries at the individual level, see Appendix A.

Table 4. Number of targets with above-chance performance by group (out of a maximum of four).

		Comprehension Test			
		1	**2**	**3**	**4**
BiTD-Span	*Mean*	2.50	1	2	1.75
	SD	.87	.71	.71	1.48
BiTD-Eng	*Mean*	2.75	1.25	2.5	2
	SD	1.08	1.09	1.5	1
BiSLI	*Mean*	1.75	2	1.75	2
	SD	.93	.71	1.30	.71

In addition, analyses were completed to ensure that performance did not differ across that two combinations of novel labels (see Table 3). In total, six participants were exposed to novel labels in Combination 1 and six participants were exposed to novel labels from Combination 2. As a result of random assignment to order combinations, four BiSLI participants, one BiTD-Eng participant and one BiTD-Span participant received Combination 2; the remaining participants, including three BiTD-Eng and three BiTD-Span participants, received Combination 1. Total percent accuracy did not differ between children who were exposed to words in Combination 1 as compared to words in Combination 2 for any of the four comprehension tests (*p* values ranged from .143 to .585). As all BiSLI participants received the same combination and language status had the potential to affect performance on the word learning task (e.g., Rice et al., 1994), a second comparison between Combination 1 and Combination 2 included only BiTD-Span and BiTD-Eng participants. Again, overall performance across word combinations did not significantly differ for any of the four comprehension tests (*p* values ranged from .129 to .765). As such, the two combinations were considered comparable in difficulty.

Cognate and non-cognate accuracy

Recall that the comprehension tasks allowed multiple opportunities for the participants to identify each novel label. Children's performance for cognates and non-cognates was thus summarized as a proportion of items correct (e.g., if a child identified the correct visual referent for five trials with cognate target labels out of 12 trials with cognate targets, her accuracy rate was calculated as 5/12 = .42). Because proportions are on a bounded scale (0-1), the proportions of accurate responses for cognates and non-cognates were each adjusted with a logit transformation.

Children's accuracy in recognizing cognate and non-cognate novel words was analyzed with linear mixed effects models run separately for each comprehension test using R (R Core Team, 2013) and the package "lme4" (Bates, Maechler & Bolker, 2015). Fixed effects included group (BiTD-Span, BiTD-Eng and BiSLI) and cognate status (cognate, non-cognate), allowing for the interaction of these variables. Models also included fixed effects for participant characteristics expected to be associated with cognate sensitivity, including percent exposure to Spanish at home and chronological age (e.g., Pérez et al., 2010; Kelley & Kohnert, 2012). Significant interactions were investigated using the package "lsmeans" (Lenth, 2016) to test for differences between levels of one factor at specific levels of the second factor. Subject intercepts were entered as random effects, and *p*-values for tested effects were obtained by likelihood ratio tests of the model with the targeted effect included against a reduced model without that effect. Finally, to support a more reliable comparison of cognate and non-cognate accuracy, children who did not demonstrate above-chance performance (i.e., reliable recognition) for any of the four target labels were excluded from analyses for that corresponding comprehension test. Across four points of analysis for 12 participants and a total of 48 comprehension tests, four children were excluded from four tests: two children (one BiTD-Span and one BiTD-Eng) were excluded from analyses for Comprehension Test 2; one child was excluded from analyses for Comprehension Test 3 (one Bi-SLI) and one child was removed from analyses for Comprehension Test 4 (one BiTD-Sp; see Appendix A).

Comprehension Test 1. For the first comprehension test, administered in Spanish immediately after the Spanish story introduced the four novel objects with Spanish cognate and non-cognate labels, the only significant predictor of performance was chronological age, $F(1, 12) = 11.621$, $p = .005$, with older children performing with higher accuracy than younger children. Accuracy was not significantly predicted by group, cognate status or Spanish exposure, nor was there a significant interaction of group and cognate status (see Figure 1). Results thus indicated that overall accuracy was comparable across cognate status and the three groups. Recall that no cognate effects were anticipated at this testing point, as both the story and the test were administered in Spanish; prior to the introduction of English labels via comprehension test or story, cognate status was not yet evident.

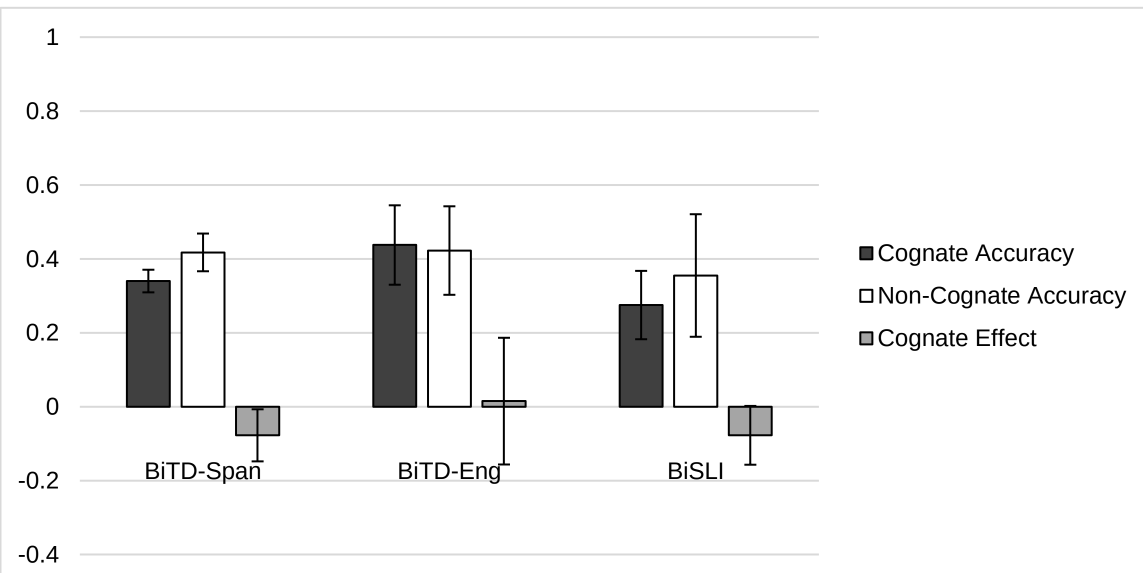


Figure 1. Proportion of accurate responses for cognate and non-cognate targets for BiTD-Span, BiTD-Eng and BiSLI groups in the first comprehension test. Cognate effects reflect the difference between cognate and non-cognate proportions.

Comprehension Test 2. The second comprehension test was administered in English immediately following the Spanish story and Spanish comprehension test. As half of the English targets were cognates with their Spanish translation equivalents, cognate status was potentially accessible to participants. Enhanced accuracy for cognates relative to non-cognates would indicate that children used information from word learning experiences in Spanish to scaffold performance in a challenging receptive English task.

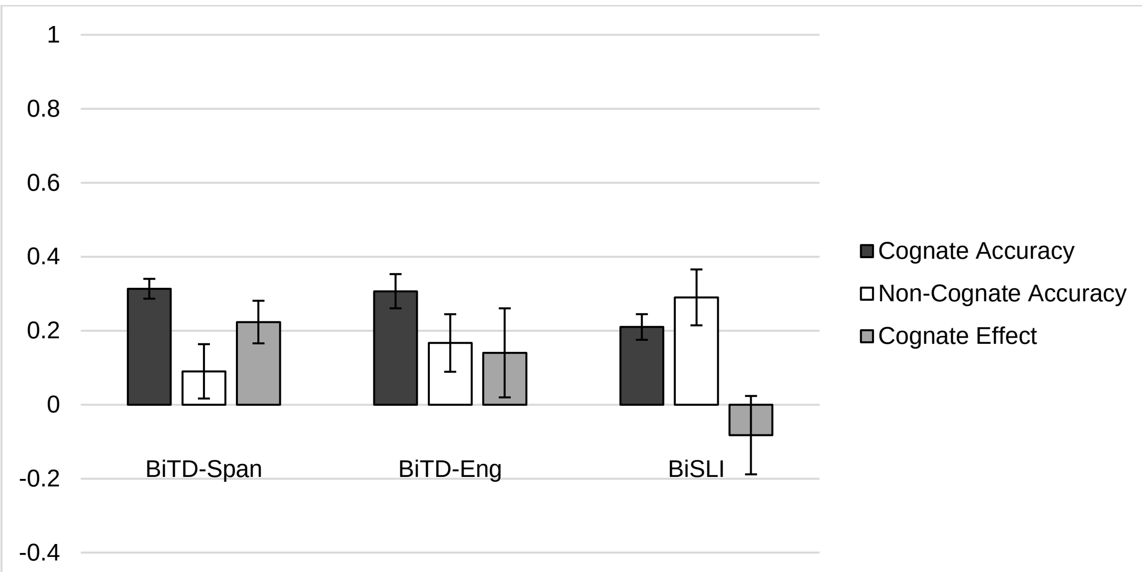


Figure 2. Proportion of accurate responses for cognate and non-cognate targets for BiTD-Span, BiTD-Eng and BiSLI groups in the second comprehension test. Cognate effects reflect the difference between cognate and non-cognate proportions.

Neither age nor Spanish exposure significantly predicted accuracy on the second comprehension test. Conversely, there was a main effect of cognate status, $F(1, 20) = 6.64$, $p = .018$, as well as an interaction between group and cognate status, $F(2, 20) = 4.29$, $p = .028$. Post hoc tests revealed that for the BiTD-Span group, performance on cognate targets (mean proportion correct = .313) was greater than their accuracy for non-cognate targets (mean = .090; $p = .025$; see Figure 2 and Appendix A). No other post hoc comparisons reached significance; both the BiTD-Eng and BiSLI groups demonstrated comparable accuracy rates for cognates and non-cognates when tested in English in CT2.

Comprehension Test 3. The third comprehension test was administered in English immediately following the English story in Session 2. By this testing point, participants had encountered each of the target words in both Spanish and English; as a result, all translation equivalents had been presented and cognate status was accessible.

No significant main effects emerged. However, there was a significant interaction between group and cognate status, $F(2, 11) = 7.054$, $p = .011$. Post hoc tests for this model revealed that the BiTD-Span group demonstrated significantly higher performance on cognate targets than for non-cognate targets (mean proportions correct = .408 and .188, respectively; $p = .013$), consistent with a positive cognate effect (see Figure 3 and Appendix A). Conversely, no other pairwise comparisons reached significance and there were comparable accuracy rates for cognate and non-cognate targets for both the BiTD-Eng and BiSLI groups.

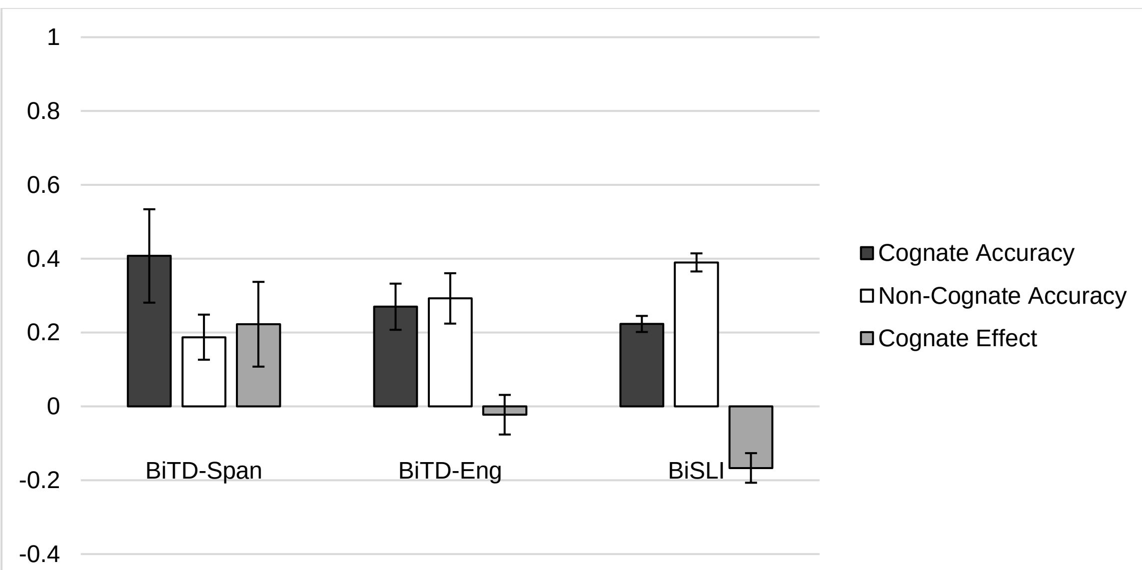


Figure 3. Proportion of accurate responses for cognate and non-cognate targets for BiTD-Span, BiTD-Eng and BiSLI groups in the third comprehension test. Cognate effects reflect the difference between cognate and non-cognate proportions.

Comprehension Test 4. The final comprehension test was administered in Spanish immediately following the English story and English comprehension test. Higher accuracy rates for cognates relative to non-cognate labels would indicate scaffolding form English into Spanish.

Children with greater Spanish exposure demonstrated higher performance on the last comprehension test, $F(1, 22) = 14.829$, $p = .001$. There was also a significant main effect of group, $F(2, 22) = 4.254$, $p = .027$, with a significant interaction between group and cognate status, $F(2, 22) = 3.571$, $p = .045$. Post hoc analyses reveal that the BiTD-Span group's proportion of correct responses for cognates (mean = .507) was significantly higher than both the BiTD-Eng group (mean =.153; $p = .039$) and the BiSLI group (mean = .168; $p = .016$; see Figure 4 and Appendix A).

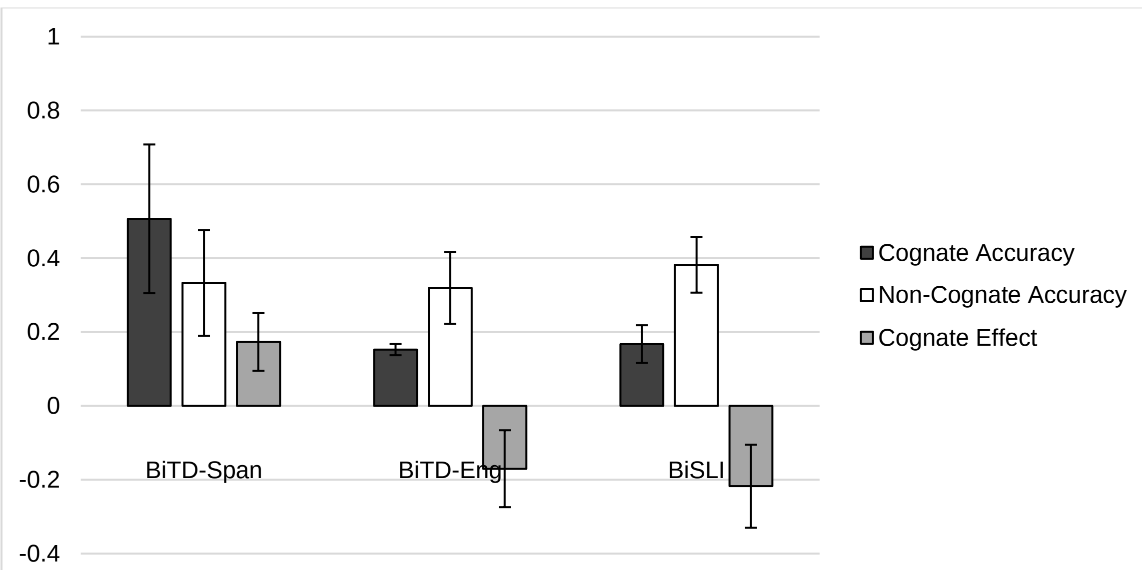


Figure 4. Proportion of accurate responses for cognate and non-cognate targets for BiTD-Span, BiTD-Eng and BiSLI groups in the third comprehension test. Cognate effects reflect the difference between cognate and non-cognate proportions.

Fixations to cognate and non-cognate targets

Recall that throughout the comprehension test, children's eye gaze was continuously tracked with a remote eye tracker mounted on the portable computer monitor. The measure of interest was proportion of fixations to the correct visual referent for trials with cognate targets and those with non-cognate targets starting 200ms post target word onset in the pre-recorded audio (e.g., Blumenfeld & Marian, 2005). Analysis procedures largely mirrored those of our first research question, including fixed effects of group (BiTD-Span, BiTD-Eng and BiSLI) and cognate status (cognate, non-cognate), allowing for their interaction; Spanish exposure and chronological age were also entered as fixed effects, and participant intercepts were entered as random effects. Post hoc tests were again conducted using the "lsmeans" package when testing significant interactions. In the absence of a cognate status by group interaction, main effects of group, a variable with three levels, were tested using the package "multcomp" (Hothorn, Bretz & Westfall, 2008) to determine which groups significantly differed from one another in performance. As for the accuracy analyses, the outcome measure was a proportion and was adjusted with a logit transformation.

Because eye tracking was utilized to capture patterns of performance that may not have been evident in offline behavioral responses, analyses were not restricted to participants who demonstrated above chance performance in accuracy. Instead, data was processed on the basis of relatively reliable eye tracking performance: children with 50 percent track loss or greater within a single comprehension test were removed from analyses for that corresponding comprehension test. Across 12 participants who completed a total of 48 comprehension tests, a total of five participants were excluded from 10 comprehension tests. One BiSLI participant was excluded from analyses for Comprehension Tests 1, 2 and 4; one BiTD-Span participant was excluded from analyses for Comprehension Tests 2, 3 and 4; one BiTD-Eng participant was excluded from analyses for Comprehension Tests 3 and 4; one BiTD-Eng participant was excluded from analyses for Comprehension Test 2 and one BiTD-Eng participant was excluded from analyses for Comprehension Test 3 (see Appendix A).

Comprehension Test 1. Recall that cognate effects were not expected to emerge in the first comprehension test, as the novel words' cognate status was not yet evident. Indeed, there was no main effect of cognate status nor an interaction between cognate status and group. Instead, the main effects of both age and group were significant, F(1,11 = 6.062, p = .032 and F(2,11) = 4.683, p = .034, respectively. Older children fixated the correct visual referent for a larger proportion of the trial than younger children. In addition, BiTD-Span participants demonstrated a significantly higher proportion of looks to the correct target image (mean = .334 across cognate and non-cognate trials) than BiSLI participants (mean = .219 across cognate and non-cognate trials; p = .033). BiTD-Eng participants were not found to differ from their BiTD peers in Spanish classrooms, and they trended towards more successful target fixation than their BiSLI peers, with an average fixation proportion of .342 across cognate and non-cognate trials (p = .080).

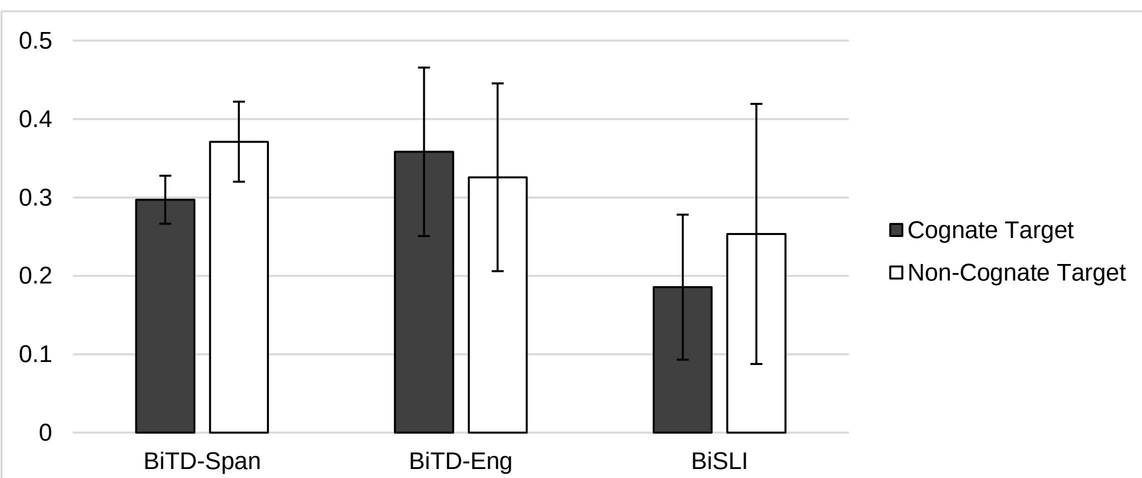


Figure 5. Proportion of fixations to correct visual referent during cognate and non-cognate trials for BiTD-Span, BiTD-Eng and BiSLI groups in the first comprehension test.

Critically, there was neither a main effect of cognate status nor a significant interaction between cognate status and group, suggesting that cognate and non-cognate words were well-matched in difficulty for all groups (see Figure 5).

Comprehension Test 2. In the second comprehension test, participants were challenged to match entirely unfamiliar English words with visual referents encountered in the Spanish story. As two of the four English words were cognate translation equivalents of the presented Spanish labels, it was possible that participants could benefit from that cross-language similarity. Indeed, analyses revealed a main effect of both group and cognate status, with a significant interaction between the two, $F(2,18) = 12.226$, $p < .001$. Post-hoc testing revealed that both BiTD-Span and BiTD-Eng participants looked at the correct target referent for a greater proportion of the trial when given cognate labels relative to non-cognate labels: $p = .003$ and $p = .033$, respectively. BiTD-Span participants' average proportion of fixations was .230 for trials with cognate targets and .112 for trials with non-cognate targets; BiTD-Eng participants' average proportion of fixations was .290 for trials with cognate targets and .191 for trials with non-cognate targets (see Figure 6 and Appendix A).

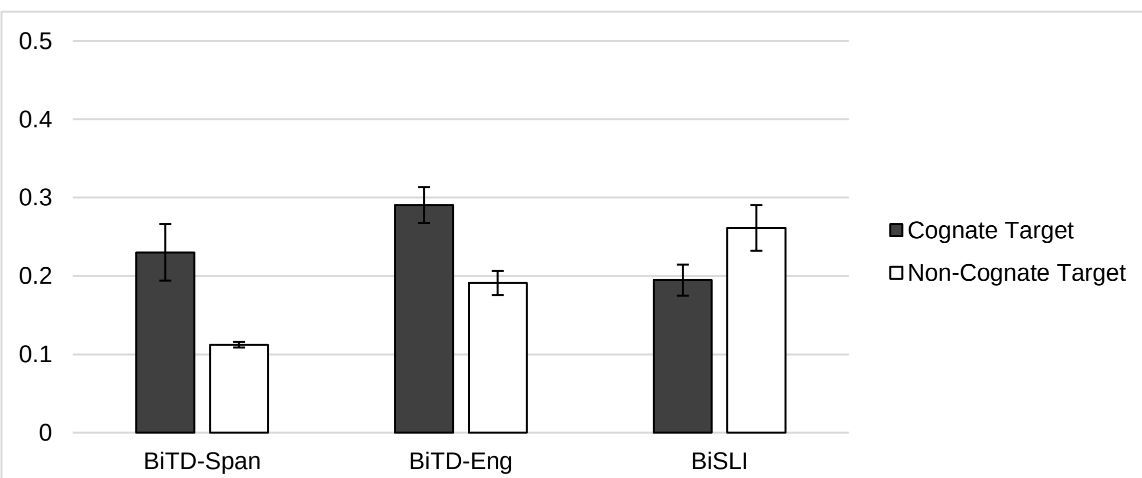


Figure 6. Proportion of fixations to correct visual referent during cognate and non-cognate trials for BiTD-Span, BiTD-Eng and BiSLI groups in the second comprehension test.

Comprehension Test 3. The third comprehension test was administered in Session 2, after participants had been formally exposed to novel English labels in the context of a story. Again, results indicated a significant interaction between cognate status and group, $F(2, 9) = 12.940$, $p = .002$ and no other significant effects. Post hoc analyses revealed that BiTD-Span participants fixated on the correct target for larger proportion of cognate trials as compared to non-cognate trials, with average fixation proportions of .328 and .180, respectively ($p = .009$). In addition, BiSLI participants fixated on the correct visual referent for a significantly larger proportion of the trial when the target was a non-cognate label than a cognate label, with average fixation proportions of .359 and .227, respectively ($p = .014$; see Figure 7 and Appendix A).

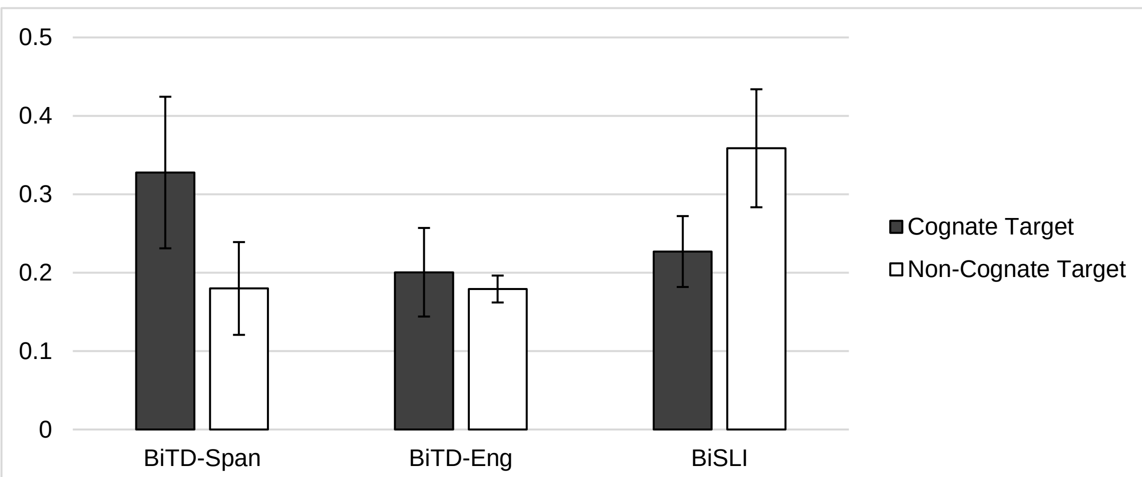


Figure 7. Proportion of fixations to correct visual referent during cognate and non-cognate trials for BiTD-Span, BiTD-Eng and BiSLI groups in the third comprehension test.

Comprehension Test 4. The last comprehension test was administered in Spanish during Session 2. Fixation to target was significantly predicted by the participant’s exposure to Spanish at home, with higher exposure predicting higher fixation proportions. In addition, there was a significant interaction between cognate status and group, $F(2, 18) = 6.005$, $p = .010$ (see Figure 8 and Appendix A). However, post hoc analyses did not reveal significant pairwise differences.

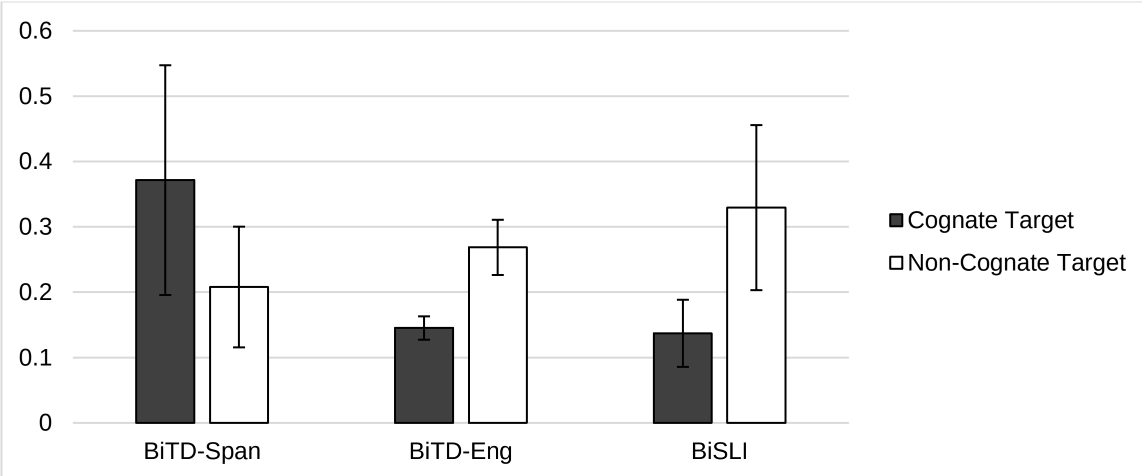


Figure 8. Proportion of fixations to correct visual referent during cognate and non-cognate trials for BiTD-Span, BiTD-Eng and BiSLI groups in the fourth comprehension test.

Discussion

A better understanding of cross-language associations in bilingual language development has the potential to support clinical practice, facilitating the development of assessment and treatment tools tailored to bilingual children. In the present work, cross-language interactions in preschool-aged Spanish-English developing bilinguals were investigated through a novel word learning task. With carefully selected stimuli designed specifically to test cognate effects, a task that minimizes the effect of prior experience, and use of eye tracking, this work represents a new approach to studying cognate sensitivity in young bilinguals with typical and atypical language development. Altogether, this work expands available research in child bilinguals to better mirror available work for adult bilinguals and provides much needed information on how children with relatively weak language systems may respond to cross-language similarity. Results help us understand factors associated with children's sensitivity to cross-language similarity and ability to scaffold information from one language into the other as they learn new words in each.

Participants in this study included Spanish-English bilingual preschoolers with typical language development and bilingual peers with SLI. All participants were recruited from the same preschool, but there was a distinction in their academic experiences: half of the children with typical language were enrolled in classrooms that primarily used Spanish. In prior research, education programs that support the child's native language have been associated with positive outcomes in both the child's native language and in the community or majority language (e.g., Thomas & Collier, 2002; Alanis, 2000; Paradis, Genesee & Crago, 2011). In the present study, these different academic experiences aligned with differences in cross-language interactions: significant group by cognate status interactions were found in each of comprehension tests where cognate status was accessible (Comprehension Tests 2, 3, and 4), with these effects routinely driven by enhanced performance for cognates by children who were both typically developing and enrolled in Spanish-language preschool classrooms.

Cognate sensitivity as measured by accuracy

Given exposure to novel labels in the Spanish story, children in the BiTD-Span group demonstrated an ability to benefit from cross-language similarity when tested in English in the same session. In the second comprehension test, administered as part of Session 1, participants were asked to match entirely unfamiliar English labels to the visual referents presented in the Spanish story. In this challenging task, BiTD-Span participants accurately paired English cognate labels with their visual referents nearly one third of the time. A review of individuals' performance further revealed that for relatively simple cognate targets with CVCVC structure, the correct visual referent was selected as frequently as half of the time, despite no prior exposure demonstrating an association between that target label and visual referent (i.e., the English story had not yet been presented). BiTD-Span participants continued to show enhanced accuracy for cognates relative to non-cognates in Session 2, after being exposed to the English labels in the context of the English story. In the third comprehension test, this group's average proportion of accurate responses for cognate targets was .408. Together, results from Comprehension Tests 2 and 3 suggest that, when tested in English, the typically developing children with regular exposure to Spanish at home and in school benefitted from cross-language similarity. Finally, in the last comprehension test, the significant group by cognate status interaction in accuracy of responses was driven by a difference between groups, with BiTD-Span participants demonstrating higher accuracy rates for cognates than both BiTD-Eng and BiSLI participants. The interplay of language dominance and language of task administration may explain this finding. With exposure to Spanish at home and in school, BiTD-Span participants were likely Spanish dominant and, thus, would not scaffold from English into Spanish or demonstrate a cognate effect when tested in their relatively stronger language (Bosma et al, 2016; Pérez et al., 2010). The fact that Spanish exposure was a significant predictor of accuracy in the fourth comprehension test—but not for either comprehension test conducted in English—further suggests that differences in patterns of performance may be expected between comprehension tests administered in the two different languages.

As reflected by accuracy on the comprehension tests, neither BiTD-Eng nor BiSLI participants demonstrated cognate sensitivity. Though BiTD-Eng and BiSLI children differed in their levels of language ability, they shared an academic environment, as children in each group were enrolled in English-language classrooms. Taken together with the findings from the BiTD-Span group, results are consistent with previous work that points to language dominance as an important predictor of observed cognate effects (Bosma et al., 2016; Pérez et al., 2010). Though differences in accuracy rates for cognates and non-cognates did not reach significance for the BiTD-Eng group, as was originally expected, it is worth noting that, descriptively, the BiTD-Eng group consistently demonstrated a level of cognate sensitivity that fell between the extremes of the BiSLI and BiTD-Span groups. For example, in the third comprehension test, the difference between cognate and non-cognate accuracy rates for BiTD-Span group was .223, with a preference for cognates and the difference between cognate and non-cognate accuracy was .167 for the BiSLI group, favoring the non-cognates (see Figure 3). Meanwhile, the BiTD-Eng group demonstrated nearly equivalent performance on the two types of targets, with an average accuracy of .270 for cognates and .293 for non-cognates. While it is possible that bilingual children enrolled in English classrooms truly do not differ in cognate sensitivity across levels of language ability, it is also possible that the present sample size did not allow us to capture existing differences between these two groups.

Cognate sensitivity as measured by eye tracking

Eye tracking was utilized in the present study as a more sensitive measure of cognate effects during word learning. In large part, these results mirrored those of the offline accuracy measure. As with accuracy, there were no significant differences in proportion of fixations to visual referents with cognate vs. non-cognate target labels in the first comprehension test. There were also significant cognate status by group interactions in the remaining comprehension tests that indicated that the three groups responded to cognate and non-cognate targets in differing manners. Specifically, eye tracking analyses captured the same cognate sensitivity in the BiTD-Span group as found with the accuracy measures, with a higher proportion of looks to the correct visual referent for trials with cognate targets relative to non-cognate targets in Comprehension Tests 2 and 3. With this approach, evidence also emerged that BiTD-Eng participants patterned more closely with BiTD-Span peers than with BiSLI participants.

First, a main effect of group for the first comprehension test revealed that both of the BiTD groups demonstrated relative success with fixating the correct visual referent. While the BiTD-Span and the BiSLI group differed across two important dimensions (classroom language and language status), the BiTD-Eng and BiSLI groups differed only with respect to language status. Thus, even if a Spanish comprehension task was relatively more challenging for the two groups of children enrolled in English classrooms, this resulted in a lower proportion of fixations to the correct visual referent for the BiSLI group only. Potentially, these results indicate that the initial task was most challenging for children with the weakest language skills.

Next, analyses for the second comprehension task revealed that both BiTD groups demonstrated a larger proportion of fixation to the correct visual referent when the target label was a cognate. Thus, while evidence of cognate sensitivity emerged in both accuracy measures and fixation patterns for BiTD-Span participants, sensitivity to cross-language similarity emerged only in fixation proportions for the BiTD-Eng group. Of note, BiTD-Eng participants did trend towards higher accurate rates for cognate targets (mean accuracy = .307) than for non-cognates (mean accuracy = .167) in this comprehension test, though this difference did not reach significance. Potentially, eye tracking provided an opportunity to capture covert cross-language facilitation that was not evident in off-line accuracy (Hendrickson et al., 2015).

Finally, in the third comprehension task, the BiSLI group was found to look to the correct visual referent for a significantly larger proportion of the trial when targets were non-cognates, unlike either of the BiTD groups. This aligns with their accuracy rates, which were higher for non-cognate targets (mean accuracy = .390) than for cognate targets (mean accuracy = .223), though this difference did not reach statistical significance. As before, it possible that eye tracking measures revealed some preference for target labels that were dissimilar across Spanish and English (i.e., targets that were non-cognates). In this case, fixation data reveals a pattern that diverges from positive cognate effects demonstrated by participants in the BiTD-Span group and potentially the BiTD-Eng group in this study.

Across comprehension tests, evidence from fixation proportions supports and potentially extends patterns identified via offline accuracy measures. Continued use of sensitive, varied measures across various language tasks will continue to shed light on cross-language interactions in typical and atypical bilingual language development.

Cognate sensitivity in development

In the present study, the clearest examples of cognate sensitivity were observed when typically developing, Spanish-dominant preschoolers were tested in English. Results thus indicated that cognate sensitivity and cross-language interactions may be present in early in development, particularly when children are typically developing and have the opportunity to transfer information from their stronger language into their weaker language. Previous studies of child bilingualism demonstrated this with tasks that involved naming or recognizing relatively familiar objects (i.e., standardized assessments or naming tasks require participants to have some knowledge of the referents and labels). Accordingly, this work detected cognate sensitivity for targets that bilingual children had knowledge of prior to the task, likely following many exposures over time (cf. Kohnert et al., 2004). The present study extended this demonstration to a new and challenging context: word learning. Results indicated that some bilingual children demonstrated cross-language interactions as they encountered novel labels in a relatively naturalistic context, with no explicit instruction. As such, it is possible that cognate effects may be present from the earliest stages of word learning. Mature bilingualism is similarly characterized by interactions across the two languages (e.g., Kroll et al., 2012). Accordingly, available models of adult bilingual language processing also account for cross-language interactions (e.g., Dijkstra & Van Heuven, 2002) and recognize differences in processing between the native or stronger language and the second language (e.g., Kroll et al., 2010). A growing body of research, including the present study, suggests that similar factors are at play in young bilinguals, supporting the extension of these models into development.

The present results also serve to expand our understanding of cognate effects in the context of weak language skills, an area for which research is particularly limited. While prior work found that bilingual children with SLI continued to show relatively enhanced performance for cognates relative to non-cognates (Grasso et al., 2018), participants in the present study did not show this effect. Descriptively, this group consistently performed less accurately on cognate targets relative to non-cognate targets across comprehension tests (see Figures 2-4); moreover, eye tracking analyses revealed a significantly higher proportion of looks for referents with non-cognate labels in the third comprehension test, a pattern of negative cognate effects that differs from prior research and from the BiTD groups in the present study. Importantly, Grasso et al. (2018) used a standardized expressive vocabulary measure that was designed to include at least some familiar objects in portions of the test administered to the participants. A word learning task such as the one in the current study places different demands on the child. Potentially, under the relatively challenging context of learning new words given limited exposures, children with weaker language skills used strategies other than sensitivity to cross-language similarity to process the novel labels. Given continued exposure to the labels or training in attending to cognate awareness, it is possible that this group would show enhanced performance for cognate relative to non-cognate targets. Together with previous work, the current results suggest that cognate effects not only vary as a function of target characteristics (e.g., Bosma et al, 2016) and individual differences across participants (e.g., Kelley & Kohnert, 2012), but, potentially, also as a result of the task at hand.

Clinical implications

Accurately diagnosing specific language impairment can be a challenging process, and clinical practice benefits from efforts to characterize patterns of performance in children with typically and atypically developing language. Such efforts may be particularly necessary for children from culturally and linguistically diverse backgrounds, for whom tools are less readily available and who may be particularly susceptible to misdiagnosis of SLI (Paradis, 2005; Paradis & Crago, 2000). The present findings align with broader research in bilingualism that language performance is associated with language experience, including language dominance (e.g., Bedore, Peña, Griffin & Hixon, 2016). As such, bilingual children's performance on language tasks ought to be considered in light their language background, as should the type of support and scaffolding they receive.

In addition, results of this study offer insight on the potential role of cognates in clinical and educational approaches for bilingual speakers. One key finding is that while some bilingual children demonstrated cognate sensitivity in terms of accuracy, others did not—including children with typical language development. As such, it cannot be assumed that children with recognize or benefit from cross-language similarity without instruction, though, as discussed above, expectations for cognate effects may shift on the basis of task and individual child factors. In addition, qualitative descriptions of the data in the present study alongside fixation patterns suggest that bilingual children with SLI demonstrated relative difficulty with cognates during novel word learning as compared to their typically developing peers. Should further work confirm this pattern, sensitivity to cognates—particularly in the context of a challenging task—may prove to be a helpful contribution to bilingual language assessment. For example, English-speaking clinicians could introduce bilingual children to novel words in English that are either cognates or non-cognates with the child's native language and test for the child's ability to scaffold from the native language. Once a general protocol for introducing novel cognates is created, it would be possible to identify cognate pairs for additional language pairings and adjust this approach for speakers from a wide array of language backgrounds.

It remains an open question how young bilinguals and bilinguals with SLI respond to cognates and non-cognates once given explicit instruction and practice with recognizing cross-language similarity. Available research indicates that typically developing school-age bilingual children benefit in academic settings from training in cognate awareness (e.g., Dressler, Carlo, Snow, August & White, 2010; Nagy, García, Durgunoğlu, & Hancin-Bhatt, 1993). Future work in this area would shed light on the use of training in cognate awareness for dynamic assessment, as well as for bilingual language therapy.

Limitations and future directions

While this study found clear delineations in performance between children in the BiTD-Span group and children in the other two groups, expected differences between children with typical language skills and peers with SLI emerged in fixations patterns but did not reach significance in off-line accuracy measures. It is difficult to draw definitive conclusions from these results, particularly given the small sample size in the present study. As such, this work would benefit from replications that include larger samples of children. In particular, it would be informative to test how children with SLI enrolled in Spanish classrooms perform on the present tasks. If language experience in academic settings was the primary factor driving cognate effects, this group would be expected to pattern with the current BiTD-Span group; if language ability plays an additional role, then these children would be expected to demonstrate weaker cognate effects relative to the BiTD-Span group. In addition, future work with children from different language backgrounds and with different profiles of language experience is needed to make strong claims about the nature of cross-language interactions in development.

While there are many novel word learning tasks available for monolingual children, such work is not equally available for young bilinguals, particularly when both languages are considered. Indeed, only two studies have taught bilingual children in both languages (Kan & Kohnert, 2012, 2008) and, in contrast to the current work, these studies did not include children with SLI, nor did they manipulate the cross-language overlap of the novel items. To the best of the authors' knowledge, this is the first study to provide exposure to novel labels and test for learning in multiple languages in young bilinguals with varying levels of language ability. Though efforts were made to draw from related studies, the present work is nevertheless exploratory in some respects. As such, it is worth considering whether the designed tasks were appropriate for the participants and for the goals of the study and to consider possibilities for improving experimental design for future work.

Overall, results suggest that the present design was largely successful in meeting its aims. Each child in the study demonstrated some success in recognizing novel labels giving relatively limited exposures, with some children showing notable success, suggesting that the task was designed with an appropriate level of difficulty. In addition, results indicated that the key manipulation in the study was successful. While there was evidence that some children were benefiting from cross-language similarity in the second, third and fourth comprehension tests, it is important to note that there was no difference in cognate and non-cognate performance during the first comprehension test. Critically, this comprehension test was administered in Spanish after Spanish labels had been presented in Spanish story, but before children saw any pairing between the novel visual referents and English labels. Consequently, the Spanish labels are effectively neither cognates nor non-cognates from the perspective of the participants. That accuracy rates and fixations were comparable for labels that were designed to be cognates and non-cognates indicates that experimental stimuli were well balanced and that subsequent differences in performance may be attributed to the key manipulation, and not to differences in other factors (e.g., phonotactic probability).

Nevertheless, future studies that seek to advance research in this area may continue to expand and improve upon the present design. For example, it would be important to test whether the pattern of effects relies the first language of presentation. In the current study, participants' first exposure to novel labels occurred in Spanish. Potentially, reversing the order of presentation would have resulted in other children demonstrating cognate effects; indeed, perhaps starting with English stories would have better differentiated between the BiTD-Eng and BiSLI groups in the present study, as participants in each of these groups were enrolled in English classrooms. This study also presented the Spanish and English stories on separate days with different narrators in each language, mimicking the bilingual experience of encountering multiple labels for the same object in separate contexts. However, it is also the case that bilingual children may be exposed to code-switching or the use of both languages in one interaction, in which case novel labels would be provided in both languages in relatively rapid succession. Future studies may thus consider the impact of providing exposures in both languages within the same session. Better understanding the process of word learning when children are given exposures in one language as compared to both would have far-reaching implications for clinical practice as well as broader educational settings. Another key aspect of the present study was that each story provided six exposures to each item in each story. Number of exposures has implications for the relative ease of difficulty of this task (e.g., Rice et al., 1994). Potentially, a different number of exposures would better differentiate between children with typical development and peers with SLI. Indeed, it is notable that, unlike in other studies, BiTD and BiSLI participants did not differ in overall accuracy rates from one another, and further restricting the number of exposures (e.g., Rice et al. 2004) may better serve this goal. Ultimately, the present study offers one approach to structuring a novel word learning experience for bilingual children in two languages. Numerous aspects of this design may be adjusted to result in valuable experimental manipulations that shed light on cross-language interactions and word learning in bilingual children.

Finally, it is important to note that eye tracking with young participants is a methodological challenge. While steps were taken to ensure that tasks were engaging (e.g., calibration included animated images; story images were colorful), the comprehension task may have been perceived as challenging or not sufficiently engaging by participants. We are encouraged by the substantial overlap in findings between measures of off-line accuracy and eye gaze, but recognize that revisions to this task, including creating more stimulating comprehension tests or providing more training opportunities for participants to practice attending to the screen, may improve future studies utilizing eye tracking to measure cognate sensitivity.

Summary

This study utilized innovative methods to investigate cross-language interactions in young bilinguals with typical and atypical language development. Together with available research in this area, present findings suggest that cognate sensitivity is associated with internal factors, such as language experience and language ability, in combination with external factors, such as level of task difficulty and language of task administration. Cognate sensitivity was most clearly observed when typically developing, preschool-aged bilinguals with Spanish exposure at home and in school completed a challenging receptive task in English. Cognate effects were less clearly observed in Spanish-English bilingual children enrolled in English-language classrooms, regardless of language ability. However, for children in these classrooms, fixation patterns and trends in accuracy suggested some preference for cognate targets in young bilinguals with typical language development, whereas peers with specific language impairment demonstrated some preference for non-cognate targets. Continued efforts to understand the circumstances under which bilingual children with typical and atypical language development scaffold information across their two languages will inform clinical approaches for bilingual children, including the use of cognates in assessment and treatment.

References

Alanís, I. (2000). A Texas two-way bilingual program: Its effects on linguistic and academic achievement. *Bilingual Research Journal*, *24*(3), 225-248.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bedore, L. M., Peña, E. D., Griffin, Z. M., & Hixon, J. G. (2016). Effects of age of English exposure, current input/output, and grade on bilingual language performance. *Journal of Child Language*, *43*(3), 687-706.

Blumenfeld, H. K., & Marian, V. (2005). Covert bilingual language activation through cognate word processing: An eye-tracking study. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 27, No. 27).

Bosma, E., Blom, E., Hoekstra, E., & Versloot, A. (2016). A longitudinal study on the gradual cognate facilitation effect in bilingual children's Frisian receptive vocabulary. *International Journal of Bilingual Education and Bilingualism*, *22*(4), 371-385.

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research*, *32*(3), 453-467.

De Groot, A., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1-56.

De Groot, A., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language learning*, *47*(2), 215-264.

Dijkstra, T., K. Miwa, B. Brummelhuis, M. Sappelli, and H. Baayen. 2010. "How Cross-language Similarity and Task Demands Affect Cognate Recognition." Journal of Memory and Language 62: 284–301.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(03), 175-197.

Dressler, C., Carlo, M. S., Snow, C. E., August, D., & White, C. E. (2011). Spanish-speaking students' use of cognate knowledge to infer the meaning of English words. *Bilingualism: Language and Cognition*, *14*(02), 243-255

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, *23*(2), 166-180.

Duyck, W., E. Van Assche, D. Drieghe, and R. J. Hartsuiker. 2007. "Visual Word Recognition by Bilinguals in a Sentence Context: Evidence for Non-selective Lexical ACCESS." Journal of Experimental Psychology: Learning, Memory, and Cognition 33: 663–679.

Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science*, *22*(9), 1155-1164.

Grasso, S. M., Peña, E. D., Bedore, L. M., Hixon, J. G., & Griffin, Z. M. (2018). Cross-Linguistic Cognate Production in Spanish–English Bilingual Children With and Without Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *61*(3), 619-633.

Gray, S. (2005). Word Learning by Preschoolers With Specific Language Impairment: Effect of Phonological or Semantic Cues. *Journal of Speech, Language, and Hearing Research*, *48*(6), 1452-1467.

Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied psycholinguistics*, *29*(1), 3-19.

Hendrickson, K., Mitsven, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2015). Looking and touching: What extant approaches reveal about the structure of early word knowledge. *Developmental science*, *18*(5), 723-735.

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior research methods*, *48*(4), 1393-1409.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–63.

Kambanaros, M., Michaelides, M., & Grohmann, K. K. (2017). Cross-linguistic transfer effects after phonologically based cognate therapy in a case of multilingual specific language impairment (SLI). *International journal of language & communication disorders*, *52*(3), 270-284.

Kan, P. F., & Kohnert, K. (2008). Fast mapping by bilingual preschool children. *Journal of Child Language*, *35*(3), 495-514.

Kan, P. F., & Kohnert, K. (2012). A growth curve analysis of novel word learning by sequential bilingual preschool children. *Bilingualism: Language and Cognition*, *15*(3), 452-469.

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*.

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech, and Hearing Services in Schools*.

Kelley, A., & Kohnert, K. (2012). Is there a cognate advantage for typically developing Spanish-speaking English-language learners?. *Language, Speech, and Hearing Services in Schools*.

Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and language*, *91*(3), 294-302.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of communication disorders*, *43*(6), 456-473.

Kohnert, K., Windsor, J., & Miller, R. (2004). Crossing borders: Recognition of Spanish words by English-speaking children with and without language impairment. *Applied Psycholinguistics*, *25*(4), 543-564.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Valdes-Kroff, J. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. Ross (Ed.), The Psychology of Learning and Motivation, Volume 56 (pp. 229-262). San Diego: Academic Press.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*(3), 373-381.

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., ... & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational research review*, *10*, 90-115.

Length, R.V. (2016). Least-Squares Means: The R Package lsmeans. Journal of Statistical Software, 69(1), 1-33.

Lotto, L., & De Groot, A. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language learning*, *48*(1), 31-69.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PloS one*, *7*(8), e43230.

Nagy, W. E., García, G. E., Durgunoğlu, A. Y., & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Reading Behavior*, *25*(3), 241-259.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language. *Language, Speech, and Hearing Services in Schools*.

Paradis, J., & Crago, M. (2000). Tense and temporality: A comparison between children learning a second language and children with SLI. *Journal of Speech, Language, and Hearing Research*, *43*(4), 834-847.

Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). Bilingual children with specific language impairment: How do they compare with their monolingual peers. *Journal of Speech, Language and Hearing Research*, *46*, 1-15.

Paradis, J., Genesee, F., & Crago, M. B. (2011). Dual Language Development & Disorder. Baltimore: Paul H. Brookes Publishing Co.

Peña, E. D., Gutierrez-Clellen, V., Iglesias, A., Goldstein, B., & Bedore, L. M. (2018). *BESA: Bilingual English-Spanish Assessment*. Paul H. Brookes Publishing Company.

Pérez, A. M., Peña, E. D., & Bedore, L. M. (2010). Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services (San Diego, Calif.)*, *4*(1), 55.

Potapova, I., Blumenfeld, H. K., & Pruitt-Lord, S. (2016). Cognate identification methods: Impacts on the cognate advantage in adult and child Spanish-English bilinguals. *International Journal of Bilingualism*, *20*(6), 714-731.

Potapova, I., & Pruitt-Lord, S. L. (2019). Towards understanding the bilingual profile in typical and atypical language development: A tutorial. *International Journal of Speech-language Pathology*, 1-11.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *37*(1), 106-122.

Roid, G. H., & Miller, L. (1997). *Leiter International Performance Scale-Revised (Leiter-R) manual. Wood Dale, IL: Stoelting.*

Rosselli, M., Ardila, A., Jurado, M. B., & Salvatierra, J. L. (2014). Cognate facilitation effect in balanced and non-balanced Spanish–English bilinguals using the Boston Naming Test. *International Journal of Bilingualism*, *18*(6), 649-662.

Sheng, L., Lam, B. P. W., Cruz, D., & Fulton, A. (2016). A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of experimental child psychology*, *141*, 229-238.

Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of memory and language*, *65*(4), 345-362.

Thomas, W. P., & Collier, V. P. (2002). A national study of school effectiveness for language minority students' long-term academic achievement.

U.S. Census Bureau; "Language Spoken at Home"; generated by Irina Potapova using American FactFinder; factfinder.census.gov; (15 June 2016).

U.S Department of Education & National Institute of Child Health and Human Development. National symposium on learning disabilities in English language learners. Symposium summary. Washington, DC: Authors; 2003.

Van Assche, E., W. Duyck, R. J. Hartsuiker, and K. Diependaele. 2009. "Does Bilingualism Change Native-Language Reading? Cognate Effects in a Sentence Context" Psychological Science 20 (8): 923–927.

Van Hell, J. G., & De Groot, A. M. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and cognition*, *1*(03), 193-211.

Appendix A: Individual participant information.

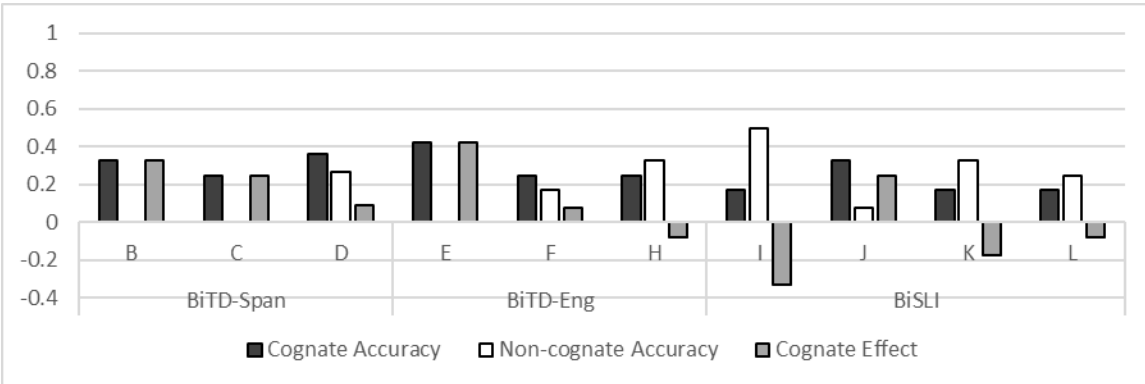
Individual participant characteristics.

Group	ID	Age (in mos.)	Sex	% Spanish Heard	% English Heard	Maternal Education (in years)	Leiter-R	MLUw – English	MLUw – Spanish	ITALK	BESA Language Index
BiTD-Span	A	51	F	55.56	44.44	11.00	15.00	3.62	3.38	5.00	94.00
	B	49	F	100.00	0.00	12.00	14.50	1.61	4.40	5.00	123.00
	C	50	F	50.00	50.00	12.00	14.00	3.53	1.33	5.00	90.00
	D	49	F	100.00	0.00	12.00	17.00	1.25	3.93	3.40	100.00
BiTD-Eng	E	53	M	80.00	20.00	6.00	13.00	3.89	3.28	4.60	113.00
	F	51	F	100.00	0.00	11.00	15.50	2.17	4.61	4.60	112.00
	G	55	F	100.00	0.00	12.00	12.00	2.30	3.73	4.40	110.00
	H	50	F	50.00	50.00	13.00	11.00	3.19	2.98	4.60	100.00
BiSLI	I	49	M	71.43	28.57	7.00	10.00	1.53	1.54	1.60	64.00
	J	51	F	71.43	28.57	12.00	11.50	1.36	1.58	3.00	93.00
	K	56	M	83.33	16.67	16.00	9.50	1.76	1.39	3.80	86.00
	L	49	M	100.00	0.00	6.00	15.00	1.99	2.00	3.00	70.00

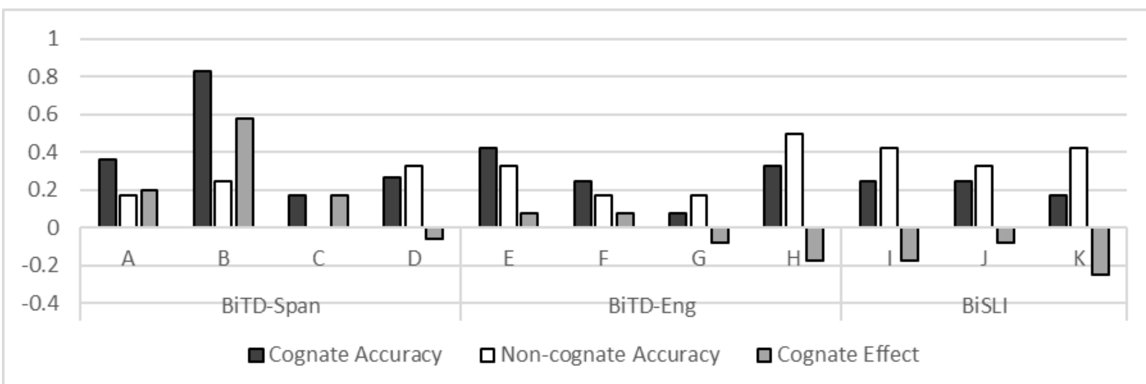
Individual participants' number of words with above-chance accuracy rates on receptive comprehension test (maximum = 4).

		Comprehension Test			
Group	**Participant**	**1**	**2**	**3**	**4**
BiTD-Span	A	2	0	2	2
	B	2	1	3	4
	C	2	1	1	1
	D	4	2	2	0
BiTD-Eng	E	4	1	4	1
	F	3	1	1	3
	G	1	0	1	3
	H	3	3	4	1
BiSLI	I	2	3	3	3
	J	1	1	3	2
	K	3	2	1	2
	L	1	2	0	1

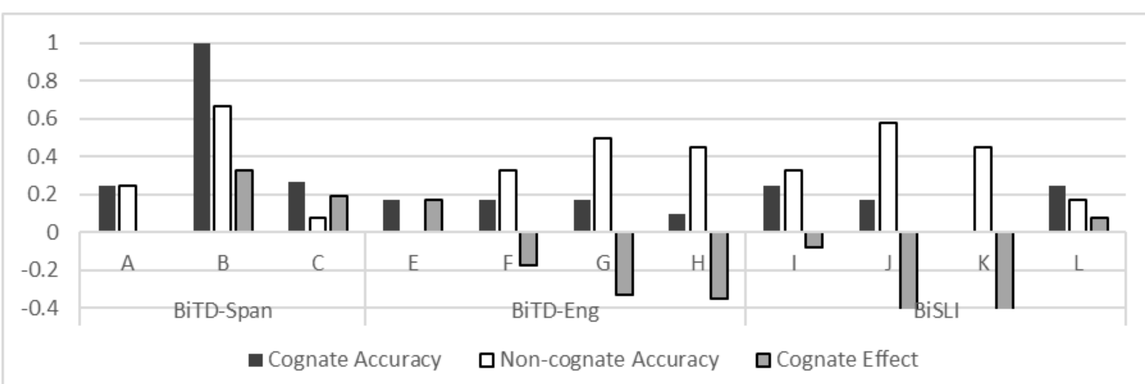
Comprehension Test 2 – individual participants' average cognate accuracy, non-cognate accuracy and cognate-effect



Comprehension Test 3 – individual participants' average cognate accuracy, non-cognate accuracy and cognate-effect



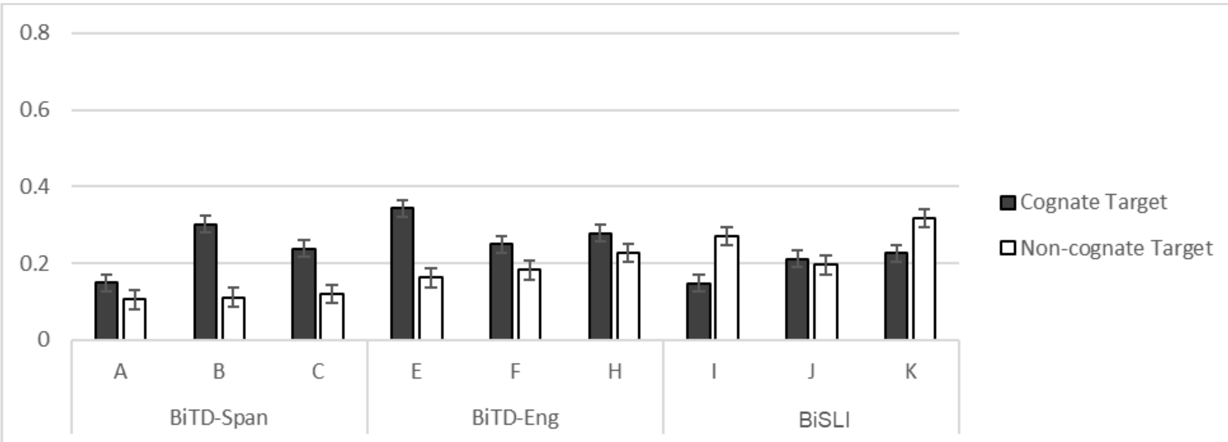
Comprehension Test 4 – individual participants' average cognate accuracy, non-cognate accuracy and cognate-effect



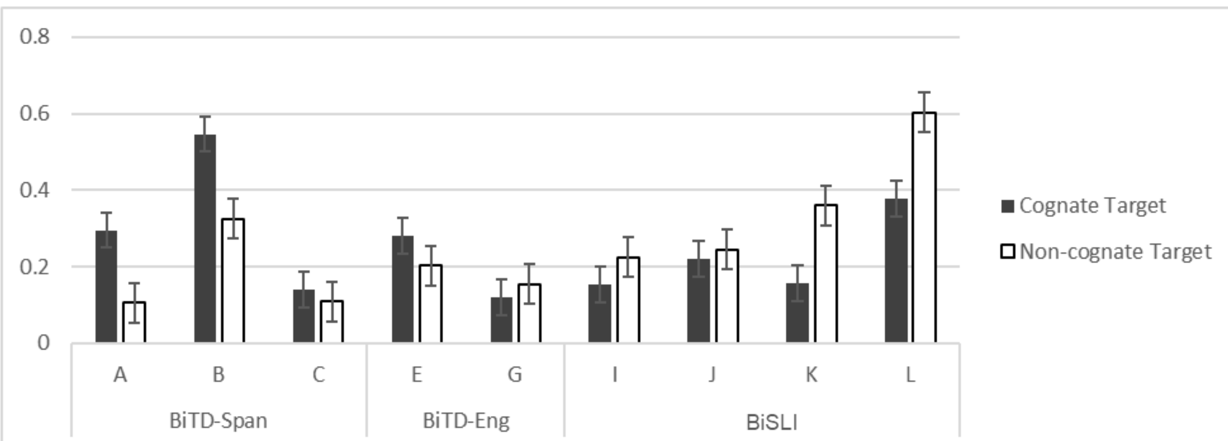
Individual participants' percentage trackloss

Group	Participant	Comprehension Test 1	2	3	4
BiTD-Span	A	14.00%	26.65%	16.70%	24.33%
	B	16.42%	14.11%	15.90%	32.27%
	C	15.42%	14.61%	23.57%	36.12%
	D	49.03%	76.63%	62.27%	53.83%
BiTD-Eng	E	29.36%	38.67%	22.77%	26.33%
	F	12.00%	15.14%	51.32%	23.41%
	G	16.92%	52.39%	21.92%	22.46%
	H	24.07%	19.67%	83.53%	91.32%
BiSLI	I	22.16%	27.13%	29.88%	20.87%
	J	24.20%	16.13%	26.32%	30.56%
	K	13.75%	10.10%	36.48%	40.61%
	L	54.18%	54.36%	46.94%	82.45%

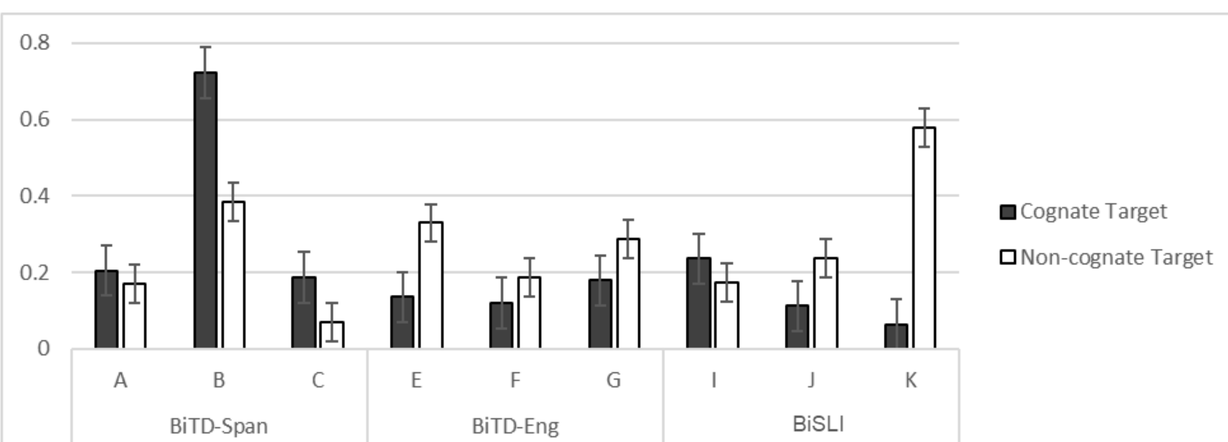
Comprehension Test 2 – Individual participants' proportion of fixations to target.



Comprehension Test 3 – Individual participants' proportion of fixations to target.



Comprehension Test 4 – Individual participants' proportion of fixations to target.



Chapter 5 is being prepared for submission for publication of the material. Potapova, I. & Pruitt-Lord (*in preparation*). Cross-language interactions during word learning in bilingual children with typical and atypical language development. The dissertation author was the primary investigator and the primary author of this paper. Preparation of this manuscript was supported, in part, by National Institute on Deafness and Other Communication Disorders Grants (NIDCD) F31 DC016194. We would like to thank the participants, their families and the teachers who participate in our community-based collaboration. We would also like to thank the members and collaborators of the Child Language Development, Disorders and Disparities Lab for their contributions to this project. In particular, we would like to recognize the efforts of: Alicia Escobedo, Cristal Toscano, Daena Taylor, Carrie Goodwiler, Fabiola Morales, Genesis Maldonado, Izamar Bergado, Jonathan Robinson Anthony, Laura Silva, Leilani Melendrez and Liliana Michel.

CHAPTER 6:

General Discussion

General Discussion

The goal of this dissertation was to contribute to our understanding of typical and atypical bilingual language development to support clinical practice for children from culturally and linguistically diverse backgrounds. In service of this discussion, it is helpful to first briefly review typical and atypical language development in monolingual children.

In the process of mastering their native language, children who are typically developing and monolingual omit obligatory tense and agreement markers (e.g., *He want to go),* make word choice errors (e.g., overgeneralize *tiger* to other big cats) and otherwise produce utterances unlike those of mature speakers. When children show more persistent difficulties in these areas, they may be diagnosed with specific language impairment, developmental language disorder, language impairment, primary language impairment, or receptive/expressive language disorder. These differing labels may reflect differences in theoretical perspectives, differences in inclusionary or exclusionary criteria, and changing preferences in terminology, but generally converge on a group of children with weaknesses in both language production and comprehension.

Among the most widely studied characteristics of specific language impairment is weaknesses in morphosyntax. For English-speaking monolingual and bilingual children, this includes difficulty in marking tense and agreement (T/A) morphemes, including third person singular (*she snacks*); past tense (*she snacked*), copula *BE* (*she is hungry),* auxiliary *BE* (*she is snacking)* and auxiliary *DO* (*do you have the popcorn?*). Difficulties in morphosyntax are observed cross-linguistically (Leonard, 2014), though the difficulty is not uniformly expressed in T/A marking. For example, children with atypical language development acquiring Romance languages demonstrate persistent weaknesses in producing direct object pronouns that precede the verb (e.g., *Ella lo come,* 'she eats it'). Thus, the specific presentation of errors differs in relation to the structure of the ambient language (Leonard, 2014). Cross-linguistically, errors made by children with specific language impairment are characterized by omissions of obligatory morphemes. Though typically developing children also demonstrate these errors for a period of development, children with specific language impairment demonstrate these error patterns for a protracted period (e.g., Rice, 2010). Thus, difficulties in morphosyntax for children with atypical language are not characterized by patterns of aberrant use, but by a delay within a delay (Rice, 2003; see also Rice, 2004).

Children with atypical language development have also been found to have weaknesses across various lexical-semantic tasks. As a group, these children have delayed onset of first words, perform below typically developing peers on standardized and informal measures of vocabulary, and demonstrate difficulty in word learning (e.g., Gray, Plante, Vance & Henrichsen, 1999; Watkins, Kelly, Harbers, & Hollis, 1995; Kan & Windsor, 2010; Kapantzoglou, Restrepo & Thompson, 2012). Weaknesses in semantic representations have thus been proposed as another characteristic of language disorder (Mainela-Arnold, Evans & Coady, 2010; McGregor, Newman, Reilly & Capone, 2002).

Though there is agreement regarding expected weaknesses and linguistic profiles of monolingual children with language disorder, diagnosis remains imprecise. In addition to a lack of consensus in the field regarding diagnostic nomenclature, criteria for language disorder are inconsistent in both research and clinical practice (Spaulding, Plante, & Farinella, 2006). Further, approaches that were designed to compensate for the limitations of static, standardized assessments, including dynamic assessment approaches and subjective measures of learning and performance, have their own challenges in implementation (Tomblin, Records, & Zhang, 1996).

Difficulties in identifying clinically low language performance are compounded for bilingual children, whose performance in each language may vary as a result of various non-clinical factors, including relative exposure, proficiency and dominance in each language; similarities and differences in linguistic structures across the two languages; and the relative status of each language in the speaker's community (e.g., Paradis, Genesee & Crago, 2011). Differences in language use resulting from these factors may overlap with clinical markers of language disorder for monolingual children. For example, developing bilinguals who are acquiring English also make errors in English T/A marking (e.g., Gutiérrez-Clellen, Simon-Cereijido & Wagner, 2008) and demonstrate performance on morphosyntactic tasks that would, upon cursory assessment, seem consistent with an impaired language system (e.g., Paradis, 2005). Researchers and clinicians alike must thus recognize these patterns of language differences in order to successfully identify patterns that are indicative of language disorder in children from culturally and linguistically diverse backgrounds (Oetting, 2018).

The studies included in this dissertation aimed to inform our understanding of typical and atypical bilingual language development using two complementary perspectives. Chapters 2 and 3 represent work under a *single language focus*. To better understand English T/A marking—an aspect of language development relevant to language assessment and treatment—and how it may be measured in bilingual children with typical language development and peers with low language skills, the research presented in Chapter 2 compared the clinical utility of three measures of English T/A morpheme use. Results indicated that there were disadvantages to measuring Spanish-English bilingual preschoolers' English T/A marking with a traditional accuracy-based measure that is commonly utilized for monolingual children of a similar age. Conversely, two morpheme measures that were designed to capture the onset of T/A marking successfully captured growth in the use of T/A morphemes over the course of the academic year and reflected expected differences in levels of language performance between the two bilingual groups. Results thus revealed that tools that are known to be meaningful and valid for monolingual children may not reflect the same quality of information for bilingual peers. Conversely, when language measures are selected to match the child's stage of language development, single-language measures may provide clinically relevant information about a bilingual child's language abilities (see also Gillam, Peña, Bedore, Bohman & Mendez-Perez, 2013).

Chapter 3 further explored English morphosyntactic development in bilingual children by testing the relative productivity of the five T/A morphemes. Results indicated that Spanish-English bilingual preschoolers with typical language development and peers with low language skills used copula *BE* and auxiliary *BE* more successfully than third person singular, past tense, or auxiliary *DO*. This work helped clarify patterns of language differences demonstrated by bilingual children, as well as patterns associated with low language skills in the context of dual language exposure. With respect to language differences, the relative ranking observed in this study differs from patterns identified in English monolinguals, who demonstrate a relatively later onset and productivity of auxiliary *BE* (e.g., Gladfelter & Leonard, 2013). This work thus provides a specific demonstration of how monolingual benchmarks may not accurately predict bilinguals' performance, even for bilingual children who are typically developing. With respect to patterns that are indicative of atypical language development, typically developing bilingual children in this study outperformed their peers with low language skills, as is to be expected—but each group demonstrated the pattern of precocious *BE*. Results are thus consistent with the understanding that atypical language development is characterized by delay (Rice, 2010). Finally, results of this work support a dynamic understanding of bilingual language development, where there is an interplay between the two languages (e.g., Kan & Kohnert, 2008): bilingual children demonstrated relative success for English morphemes that structurally aligned with their Spanish counterparts and demonstrated relative difficulty with auxiliary *DO*, a morpheme that has no clear Spanish equivalent, and two bound morphemes (third person singular, past tense) that result in consonant clusters incompatible with Spanish phonotactics (see also Combiths, Barlow, Potapova & Pruitt-Lord, 2017).

Chapters 4 and 5 focused on this potential for cross-language interactions to investigate typical and atypical bilingual language development through a uniquely bilingual lens. A substantial body of research has demonstrated that adult bilinguals consistently activate each of their languages, even when context specifically calls for them to use one language over the other (e.g., Bialystok, 2010). Work presented in Chapter 4 revealed that cognate sensitivity may be present in preschool-aged bilinguals, a younger group of children than had been represented in the literature at time of publication (cf. Kelley & Kohnert, 2012; Pérez, Peña & Bedore, 2010). In this study, Spanish-English bilingual children were found to identify cognate targets on a standardized English receptive vocabulary measure more successfully than non-cognate targets, suggesting that structural and semantic cross-linguistic overlap facilitated performance in the target language.

Chapter 5 extended upon this work to test whether cross-language interactions—a uniquely bilingual language phenomenon—could be leveraged to identify strong and weak underlying language abilities. This exploratory study utilized innovative methods to investigate cross-language interactions in development. First, the use of a novel word learning task across both of the children's languages allowed for the measurement of cross-language sensitivity in a relatively naturalistic context. Bilingual children do face the task of acquiring labels for the same objects in each of their languages, and there are clinical and education implications to understanding how knowledge of each of their languages may support this process. In addition, use of eye tracking allowed for confirmation and tentative extensions of results based on off-line accuracy. In this work, cross-language facilitation was most clearly captured when typically developing bilingual children who were exposed to Spanish at home and in school were tested in English; these children demonstrated enhanced performance in learning cognates, as measured by accuracy and eye gaze patterns during a receptive task. The finding that language dominance plays a role in how cross-language interactions surface is consistent with previous work in older bilingual children and in bilingual adults (Pérez et al., 2010; Rosselli, Ardila, Jurado & Salvatierra, 2012). In contrast to patterns demonstrated by typically developing bilingual children, results did not indicate cognate sensitivity in bilingual preschoolers with specific language impairment. Continued research is needed to clarify the role of task difficulty, relative language dominance, and language ability in the observation of cross-language interactions in bilingual language development.

Together, the four studies in this dissertation contribute to our understanding of typical and atypical language development in young bilinguals. Findings indicated that bilingual children with low language skills demonstrated weaknesses in aspects of language use that mirror those of monolingual children with atypical language development. At the same time, their language use also reflected language differences associated with dual language exposure. The studies included in Chapters 2 and 3 demonstrated that bilingual children with low language skills performed below their typically developing bilingual peers in morphosyntax, just as would be expected in monolingual children with atypical and typical language development—but across language ability groups, these bilingual children also demonstrated a pattern of English tense and agreement marking that differed from established monolingual trajectories. In addition, tentative evidence in Chapter 5 suggests that bilingual children with specific language impairment were less sensitive to lexical-semantic information across their two languages than typically developing peers. Potentially, just as monolingual children with language disorder are understood to have relatively sparse within-language semantic representations, bilingual children with atypical language development may demonstrate these weaknesses both within and across languages. Altogether, evidence presented in this dissertation indicates that for appropriate assessment, bilingual children's language performance must be considered in light of their linguistic experience.

Given the heterogeneity of bilingual speakers, there are limitations to how present results may be generalized. All research included in this dissertation features preschool-aged children who speak Spanish, a minority language in the United States. Parent report indicated that participants were generally exposed to more Spanish than English at home, an experience that is shared with many bilingual children in the United States, for whom enrolling in preschool marks a period of increased use of the community language (Bedore & Peña, 2008). Participants included in these studies were recruited at a local preschool through an on-going community-based research project. As indicated by their enrollment in a school that required below-poverty status for participation, as well as average maternal education rates, these children also come from low socioeconomic backgrounds. Broadly, these children may be representative of many bilingual children in the United States—but this is not the sole bilingual experience.

To thoroughly understand typical and atypical bilingual language development, continued work is needed in children with from other cultural and linguistic backgrounds; children from other socioeconomic backgrounds; and children with more varied profiles of dominance and proficiency. Indeed, Chapter 5 demonstrated how two groups of typically developing bilingual children—recruited from the same neighborhood, comparable in numerous demographic variables, but enrolled in either English-language or Spanish-language classrooms—demonstrated different degrees of cognate sensitivity, revealing how a seemingly small difference in language experience may impact language performance. This finding is consistent with growing efforts to consider individual-level factors to better understand bilingual language development. Shifting away from simple comparisons between bilingual children and monolingual children that collapse wide-ranging experiences into a single group, current research efforts treat bilingualism as a continuous variable (Luk & Bialystok, 2013) and focus on differences within bilingual groups to identify links between bilingual experiences and language performance. In addition to linguistic factors like exposure and proficiency, emerging findings suggest that future work may also consider the role of individual differences in broader cognitive functions. Namely, skills in executive function—a set of cognitive processes including inhibition, selective attention, monitoring, task-switching and working memory—have been linked to cognate sensitivity in adults (Linck, Hoshino & Kroll, 2008) and to the number of translation equivalents across languages in bilingual toddlers (Crivello et al, 2016). Research in typical and atypical bilingual language development will benefit from continued efforts to account for wide-ranging individual-level factors among children from increasingly diverse cultural and linguistic backgrounds.

As research in this area continues, the combined insights gained from work under a single language focus and work under the uniquely bilingual lens will provide a more complete picture of language development in children from culturally and linguistically diverse backgrounds, and, in turn, support clinical practice for these populations.

References

Bedore, L. M., & Peña, E. D. (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), 1-29.

Bialystok, E. (2010). Bilingualism. *Wiley interdisciplinary reviews: Cognitive science*, *1*(4), 559-572.

Combiths, P. N., Barlow, J. A., Potapova, I., & Pruitt-Lord, S. (2017). Influences of Phonological Context on Tense Marking in Spanish–English Dual Language Learners. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2199-2216.

Crivello, C., Kuzyk, O., Rodrigues, M., Friend, M., Zesiger, P., & Poulin-Dubois, D. (2016). The effects of bilingual growth on toddlers' executive function. *Journal of Experimental Child Psychology*, *141*, 121-132.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(03), 175-197.

Gillam, R. B., Peña, E. D., Bedore, L. M., Bohman, T. M., & Mendez-Perez, A. (2013). Identification of specific language impairment in bilingual children: I. Assessment in English. *Journal of Speech, Language, and Hearing Research*.

Gladfelter, A., & Leonard, L. B. (2013). Alternative tense and agreement morpheme measures for assessing grammatical deficits during the preschool period. *Journal of Speech, Language, and Hearing Research*, *56*(2), 542–52.

Gray, S., Plante, E., Vance, R., & Henrichsen, M. (1999). The diagnostic accuracy of four vocabulary tests administered to preschool-age children. *Language, Speech, and Hearing Services in Schools*, *30*(2), 196-206.

Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics*, *29*(1), 3-19.

Kan, P. F., & Kohnert, K. (2008). Fast mapping by bilingual preschool children. *Journal of Child Language*, *35*(3), 495-514.

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *53*(3), 739-756.

Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. Language, Speech, and Hearing Services in Schools, 43(1), 81-96.

Kelley, A., & Kohnert, K. (2012). Is there a cognate advantage for typically developing Spanish-speaking English-language learners?. *Language, Speech, and Hearing Services in Schools*.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. Bilingualism (Cambridge, England), 13(3), 373.

Leonard, L. B. (2014). Specific language impairment across languages. Child development perspectives, 8(1), 1-5.

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *The mental lexicon*, *3*(3), 349-374.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605-621.

Mainela-Arnold, E., Evans, J. L., & Coady, J. A. (2010). Explaining lexical–semantic deficits in specific language impairment: The role of phonological similarity, phonological working memory, and lexical competition. *Journal of Speech, Language, and Hearing Research*, *53*(6), 1742-1756.

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *45*(5), 998-1014.

Oetting, J. B. (2018). Prologue: Toward accurate identification of developmental language disorder within linguistically diverse schools. *Language, Speech, and Hearing Services in Schools*, *49*(2), 213–7.

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *36*(3), 172-187.

Paradis, J., Genesee, F., & Crago, M. B. (2011). Dual Language Development & Disorder. Baltimore: Paul H. Brookes Publishing Co.

Pérez, A. M., Peña, E. D., & Bedore, L. M. (2010). Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services (San Diego, Calif.)*, *4*(1), 55.

Rice, M. L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. In Y. Levy & J. Schaeffer (Eds.), Language competence across populations: Toward a definition of specific language impairment (pp. 63–95). Mahwah, NJ: Lawrence Erlbaum Associates.

Rice, M. L. (2004). Growth models of developmental language disorders. *Developmental language disorders: From phenotypes to etiologies*, 207-240.

Rice, M. L. (2010). Evaluating maturational parallels in second language children and children with specific language impairment. *Applied Psycholinguistics*, *31*(2), 320-327.

Rosselli, M., Ardila, A., Jurado, M. B., & Salvatierra, J. L. (2014). Cognate facilitation effect in balanced and non-balanced Spanish–English bilinguals using the Boston Naming Test. *International Journal of Bilingualism*, *18*(6), 649-662.

Spaulding, T. J., Plante, E., & Farinella, K. A. (2006). Eligibility Criteria for Language Impairment: Is the Low End of Normal Always Appropriate? *Language, Speech, and Hearing Services in Schools*, *37*(1), 61-72.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1284-1294.

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech, Language, and Hearing Research*, *38*(6), 1349-1355.