Title
Continuity in language development: Predictions from decontextualized vocabulary and lexical access

Author
Smolak, Erin L

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Continuity in language development:
Predictions from decontextualized vocabulary and lexical access

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

in

Language and Communicative Disorders

by

Erin Laurel Smolak

Committee in charge:

University of California San Diego

Professor Sarah Creel
Professor Victor Ferreira

San Diego State University

Professor Margaret Friend, Chair
Professor Alyson Abel-Mills
Professor Sonja Pruitt-Lord

2019

The Dissertation of Erin Laurel Smolak is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

San Diego State University

2019

DEDICATION

For Jake, my love and my rock,

and for my parents, all my friends and family.

Thank you for all you do.

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Table of Contents v

List of Figures viii

List of Tables ix

Acknowledgements.. xii

Vita xv

Abstract of the Dissertation xix

Introduction 1

Introduction 2

References 5

Chapter 1: Word Knowledge and Processing: A Domain General Developmental Account 8

Introduction 9

Vocabulary, Processing, and Working Memory 10

Models of Early Word Learning 19

Vocabulary and Speed of Processing Predict Downstream Outcomes 25

Summary 33

Conclusion 34

References 36

Chapter 2: A Cross-Language Study of Decontextualized Vocabulary Comprehension in Toddlerhood and Kindergarten Readiness.. 47

Kindergarten Readiness and Emergent Literacy 49

Early Vocabulary and Literacy 49

Decontextualized Vocabulary 50

The Present Research 50

Study 1 51
Method 51
Results 52
Discussion 54
Study 2 55
Method 55
Results 56
Discussion 57
Study 3 58
Method 58
Results 58
Discussion 60
General Discussion 60
Early Vocabulary and Kindergarten Readiness 60
The Role of Decontextualized Vocabulary 61
Limitations 61
References 62
Chapter 3: Language Status at Age 3: Group and Individual Prediction From Vocabulary Comprehension in the Second Year 69
The Present Research 71
Study 1 71
Method 71
Results 73
Discussion 74
Study 2 75
Method 75
Results 76
Discussion 76
Study 3 77
Method 77
Results 79
Discussion 80

General Discussion 80
Predicting Preschool Language 80
Limitations and Directions for Future Research 81
References 81
Language Status at Age 3: Group and Individual Prediction from Vocabulary Comprehension in the Second Year Supplemental Material 84
Chapter 4: Predicting Preschool Oral Language: Convergence and Divergence in Vocabulary and Speed of Lexical Access 89
Introduction 90
The Relation Between Vocabulary Size and Lexical Access 92
Long-Term Prediction 94
Current Study 96
Method 97
Results 107
Discussion 114
References 125
Chapter 5: General Discussion 140
References 153

LIST OF FIGURES

Figure 3-1.	Measures of sensitivity, specificity, and positive and negative likelihood ratio for each of the potential cutoff determinations for low language skill. Whereas the parameters appear stronger at higher cutoffs, in fact these cutoffs eliminate many children with language skills sufficiently low to be markers of impairment. Further, the shape of the curve under these cutoffs suggests that these parameter estimates are unstable	79
Figure 3-2.	Receiver operating characteristic (ROC) plot for empirical data and binormal-smoothed ROC curves from Study 3 using the Computerized Comprehension Task (CCT) predictor (*N* = 104)	79
Figure 3-3.	Receiver operating characteristic (ROC) plot for empirical data and binormal-smoothed ROC curves from Study 3 using the MacArthur-Bates Communicative Development Inventories (MCDI) predictor (*N* = 104)	80
Figure 3-S1.	CCT comprehension (16 and 22 months) and Oral Language Factor (36 months) z-scores for Low Language (*N* = 13) and Higher Language (*N* = 91) groups	85
Figure 3-S2.	MCDI/IFDC production (16 and 22 months) and Oral Language Factor (36 months) z-scores for Low Language (*N* = 13) and Higher Language (*N* = 91) groups	86
Figure 4-1.	Illustration of the interaction between age and vocabulary on sentence repetition	138

LIST OF TABLES

Table 2-1. Distribution of selected demographic characteristics of participants in Study 1 51

Table 2-2. Relative importance analysis predicting fourth-year kindergarten readiness for Study 1 53

Table 2-3. Relative importance analysis predicting fourth-year vocabulary for Study 1 54

Table 2-4. Regression parameters for models predicting fourth-year kindergarten readiness from second-year vocabulary for Study 1 (*N* = 48) 54

Table 2-5. Regression parameters for models predicting fourth-year vocabulary from second-year vocabulary for Study 1 (*N* = 48) 54

Table 2-6. Regression parameters for models predicting fourth-year kindergarten readiness from second- and fourth-year vocabulary for Study 1 (*N* = 48) 54

Table 2-7. Distribution of selected demographic characteristics of participants in Study 2 55

Table 2-8. Relative importance analysis predicting fourth year kindergarten readiness for Study 2 56

Table 2-9. Relative importance analysis predicting fourth year vocabulary for Study 2 57

Table 2-10. Regression parameters for models predicting fourth year kindergarten readiness from second year vocabulary for Study 1 (*N* = 58) 57

Table 2-11. Regression parameters for models predicting fourth-year vocabulary from second-year vocabulary for Study 2 (*N* = 60).......... 57

Table 2-12. Regression parameters for models predicting fourth-year kindergarten readiness from second- and fourth-year vocabulary for Study 1 (*N* = 60) 57

Table 2-13. Distribution of selected demographic characteristics of participants in Study 3 58

Table 2-14 Unstandardized B values for all regression models presented in Chapter 2, Study 1 67

Table 2-15 Unstandardized B values for all regression models presented in Chapter 2, Study 2 68

Table 3-1.	Distribution of selected demographic characteristics of participants in Study 1	72
Table 3-2.	Bivariate correlations for all predictor and 36-month measures (z scores) in Study 1 (*N* = 44)	74
Table 3-3.	Component matrix of the language factor extracted after the factorial analysis with language measures at 36 months in Study 1	74
Table 3-4.	Hierarchical regression parameters for 16- and 22-month models, Study 1 (*N* = 44)	75
Table 3-5.	AIC values for models containing combinations of the significant predictors from regression analyses in Study 1	75
Table 3-6.	Distribution of selected demographic characteristics of participants in Study 1	76
Table 3-7.	Bivariate correlations for all predictor and 36-month measures in Study 2 (z scores; *N* = 60)	77
Table 3-8.	Component matrix of the language factor extracted after the factorial analysis with language measures at 36 months in Study 2	77
Table 3-9.	Hierarchical regression parameters for 16- and 22-month models, Study 2 (*N* = 60)	78
Table 3-10.	AIC values for models containing combinations of the significant predictors from regression analyses in Study 1	78
Table 3-S1.	Scores on 36-month measures for Higher Language (*N* = 91) and Low Language (*N* = 13) groups	84
Table 3-11	Unstandardized B values for all regression models presented in Chapter 3	88
Table 4-1.	Distribution of selected demographic characteristics of participants from English- and French-speaking samples	131
Table 4-2.	Descriptive information of all measures of interest by sample. *T*-tests conducted across samples	132
Table 4-3.	Correlation matrix of predictive and longitudinal relations in the English sample and French sample	133
Table 4-4.	Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable conceptual knowledge	134
Table 4-5.	Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable PPVT Vocabulary	135

Table 4-6.	Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable sentence repetition	136
Table 4-7.	Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable nonword repetition	137

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge and thank my mentor, Margaret Friend, for her support as chair of my committee and for all the advice she has given me over the years. It goes without saying that this dissertation would not be possible without her, and I will forever be grateful for her encouragement in my professional and personal development. Through my doctoral studies and with her mentorship I have grown and matured as a scientist in so many ways. Maggie has shared with me her wisdom and her dedication to the production of high quality scientific work. She has always worked hard with me on the many drafts of the chapters contained in this dissertation, through late nights and long discussions. Maggie, thank you for your investment in my development as a scientist. I am so grateful to you. I would also like to acknowledge all students and members of the Infant and Child Development Lab with whom I have worked throughout the years. To the many research assistants who contributed countless hours of coding: you have made this work possible. To all the graduate students who work tirelessly on your own work but also find time to discuss my work, analyses, and interpretations: thank you. Science is a collaborative endeavor and I have benefitted from each of you.

I have also benefited from the support and counsel of the other members of my committee: Alyson Abel-Mills, Sonja Pruitt-Lord, Sarah Creel, and Vic Ferreira. I am fortunate to have benefited from your expertise. Thank you for your time, contributions, and questions. Finally, I must acknowledge Jeff Elman, with whom I worked during a lab rotation during the second year of my doctoral program and who was a member of my dissertation committee until his death in June, 2018. Jeff was an extraordinary and brilliant scientist who was consistently generous with his time and expertise. Our discussions during my lab rotation impacted my work immensely and his advice has strengthened my thinking in the work that appears in this dissertation. Thank you, Jeff, for your contributions.

I also have to thank everyone in my personal life for their tremendous support. First, to my parents, Lisa Hunter and Rick Smolak: from the time I was born you instilled in me curiosity and a thirst for knowledge. I even participated in a scientific study of leopard frogs when I was seven years old! And to this day your support and encouragement has made all the difference. Mom, thank you for being a role model for women in science – I've looked up to you my entire life, and it's no wonder I've followed in your footsteps. Dad, thank you for encouraging my love of mathematics and my quizzical mind. You have always been there for me with love and understanding. Thank you both. To my husband, Jake Shannon: I could not have done this without you. You came with me all the way from Ohio to stand by me and support me while I earned my Ph.D. and I can never thank you enough for everything you've done for me over the years. I love you, and I am so excited for our future. Finally, thank you to all my other family members and friends for your support.

Chapter 2, in full, is a reprint of the material as it appears in Developmental Psychology. Friend, M., Smolak, E., Liu, Y. Poulin-Dubois, D., & Zesiger, P. (2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. *Developmental Psychology, 54*, 1317-1333. doi: 10.1037/dev0000514. The dissertation author was the co-investigator and co-author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in Developmental Psychology. Friend, M., **Smolak, E.,** Patrucco-Nanchen, T., & Zesiger, P. (2019). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year.

Developmental Psychology, 55, 9-22. doi: 10.1037/dev0000617. The dissertation author was the co-investigator and co-author of this paper.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Smolak, E., Helm, J., Patrucco-Nanchen, T., Poulin-Dubois, D., Zesiger, P., & Friend, M. (in preparation). Convergence and divergence in prediction from vocabulary and speed of lexical access. The dissertation author was the primary investigator and author of this material.

VITA

EDUCATION

Ph.D	San Diego State University/University of California San Diego Joint Doctoral Program in Language and Communicative Disorders	2014 – 2019
B.A.	Ohio University Bachelor of Arts, Psychology Graduated *cum laude*	2010 – 2014

TEACHING EXPERIENCE

Teaching Assistant — Fall 2015, Fall 2016
SDSU, SLHS 513
Undergraduate course: Foundations of Speech-Language Development
Course Instructor: Dr. Alyson Abel-Mills

Teaching Assistant — Spring 2017
UCSD, PSYC 176
Undergraduate course: Psychology of Creativity
Course Instructor: Dr. Craig McKenzie

PUBLICATIONS (PEER REVIEWED)

Smolak, E., DeAnda, S., Enriquez, B., Poulin-Dubois, D., & Friend, M. (in press, *Bilingualism: Language and Cognition)*. Code-Switching in Young Bilingual Toddlers: A Longitudinal, Cross-Language Investigation.

Friend, M., Smolak, E., Patrucco-Nanchen, T., & Zesiger, P. (2019). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year. *Developmental Psychology, 55,* 9-22. doi: 10.1037/dev0000617

Friend, M., Smolak, E., Liu, Y. Poulin-Dubois, D., & Zesiger, P. (2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. *Developmental Psychology, 54*, 1317-1333. doi: 10.1037/dev0000514

Woody, A., Smolak, E. L., Rabideau, E. M., Figueroa, W. S., & Zoccola, P. M. (2015). Trait rumination moderates the effect of mentation type on heart rate responses to stressor recall. *Stress*, *18*, 554-560. doi: 10.3109/10253890.2015.1055726

PRESENTATIONS

Oral presentations

Smolak, E. (March 2019). *Continuity in Vocabulary Development: Predictions from Decontextualized Vocabulary and Lexical Access.* Invited talk presented at Purdue University, West Lafayette, IN.

Smolak, E., Hendrickson, K., Zesiger, P., & Friend, M. (March 2019). *Continuity in Vocabulary Development: Predictions from Decontextualized Vocabulary and Lexical Access.* Presentation to be given at the Society for Research in Child Development, Baltimore, MD.

Friend, M., Smolak, E., Zesiger, P., Poulin-Dubois, D., Hojen, A., Jensen, P., Bleses, D., Dale, P., & Barlow, J. (March 2019). *The Computerized Comprehension Task: Cross-language evaluation of psychometric properties and prediction to developmental outcomes.* Presentation to be given at the Society for Research in Child Development, Baltimore, MD.

Smolak, E., Villanueva, A., Liu, Y., Campos, A., Poulin-Dubois, D., Zesiger, P., & Friend, M. (June 2018). *Executive Function in Young Spanish-English Bilingual Children: Are They Really Special?* Flash talk presented at the International Congress on Infant Studies, Philadelphia, PA.

Smolak, E., Enriquez, B., Poulin-Dubois, D., Zesiger, P., & Friend, M. (July 2017). *Code-Switching in Young Bilingual Toddlers and Their Parents.* Presentation given at the International Association for the Study of Child Language, Lyon, France.

Smolak, E. (April 2017). *Out of the Mouths of Babes: Child Language and the Economy.* Presentation given at the UCSD grad SLAM competition, San Diego, CA.

Smolak, E., & Friend, M. (March 2017). *Convergence and Divergence in the Predictions of Looking vs. Haptic Measures.* Presentation given at the SDSU Student Research Symposium, San Diego, CA.

Smolak, E. (September 2016). *Convergence and Divergence in Prediction from Looking vs. Haptic Measures.* Presentation given at the SDSU/UCSD Joint Doctoral Program in Language and Communicative Disorders Doctoral Student Colloquium, San Diego, CA.

Smolak, E. (September 2015). *Age, Vocabulary Size, and Individual Trajectory Influence Early Predictions of Language Outcomes.* Presentation given at the SDSU/UCSD Joint Doctoral Program in Language and Communicative Disorders Doctoral Student Colloquium, San Diego, CA.

Smolak, E., & Friend, M. (March 2015). *Age and Individual Trajectory Influence Early Predictions of Language Outcomes.* Presentation given at the SDSU Student Research Symposium, San Diego, CA.

Poster Presentations

Villanueva, A., Smolak, E., & Friend, M. (March 2018). *Executive Function and Translation Equivalents in Young Spanish-English Bilingual Children.* Poster presented at the SDSU Student Research Symposium, San Diego, CA.

Smolak, E., Villanueva, A., Liu, Y., Campos, A., Poulin-Dubois, D., Zesiger, P., & Friend, M. (October 2017). *Executive Function and Translation Equivalents in Young Spanish-English Bilingual Children.* Poster presented at the Bilingualism Matters research conference, Riverside, CA.

Smolak, E., Patrucco-Nanchen, T., Zesiger, P., Poulin-Dubois, D., & Friend, M. (April 2017). *Language Screening Using Early Decontextualized Vocabulary.* Poster presented at the Society for Research in Child Development, Austin, TX.

Villanueva, A., Smolak, E., & Friend, M. (March 2017). *The Relationship of Code-Switching and Translation Equivalents, and How Does It Relate to Executive Functioning in Young Children?* Poster presented at the SDSU Student Research Symposium, San Diego, CA.

Smolak, E., Barlam, D., Zesiger, P., Poulin-Dubois, D., & Friend, M. (May 2016). *It's Comprehension and Production! The Efficacy of Canonical Analysis for Predicting Language Outcomes.* Poster presented at the International Conference on Infant Studies, New Orleans, LA.

Smolak, E., Poulin-Dubois, D., Zesiger, P., & Friend, M. (October 2015). *Using Comprehension to Predict Production: Variability in Individual Trajectories.* Poster presented at the Cognitive Development Society, Columbus, Ohio.

Smolak, E., Zesiger, P., Poulin-Dubois, D., & Friend, M. (June 2015). *Predicting Language Outcomes from Early Comprehension: Contributions of Age, Vocabulary Size, and Trajectory.* Poster presented at the Workshop on Infant Language Development, Stockholm, Sweden.

AWARDS AND HONORS

SDSU Graduate Student Travel Fund – Student Success Fee, 2019, $1,000

SDSU University Graduate Fellowship, Fall 2018 – Spring 2019

Program of support at San Diego State University designed to provide funding for graduate students enrolled in a PhD program. Thirty fellowships are awarded through a merit-based, competitive process. Role: Research Fellow.

IASCL travel bursary award, 2017, $1250

Travel bursary award to attend the 14th International Congress for the Study of Child Language, Lyon, France, July 2017.

UCSD Grad SLAM competition, 2017, 3rd place award, $1500

Annual competition in the University of California system. Participants aim to present accessible research to the general public in three minutes or less. Earned the 3rd place prize at UC San Diego out of approximately 50 competitors.

SDSU Center for Clinical and Cognitive Neuroscience travel award, 2017, $500

Travel award received from SDSU CCN to attend the biennial meeting of the Society for Research in Child Development, Austin, Texas, April 2017.

NIH T32, NIDCD 5T32DC007361-09, Fall 2014 – Spring 2016; Fall 2017 – Spring 2018

Neurocognitive Approaches to Communication Disorders. This training grant under the JDP-LCD at San Diego State University supports training of doctoral students. PI: Dr. Lew Shapiro. Role: Doctoral Trainee.

Phi Beta Kappa honor society member, 2014 – present

Ohio University Gateway Excellence Scholarship, 2010 – 2014

Merit-based academic scholarship covering full tuition and fees.

PROFESSIONAL ACTIVITIES AND DEVELOPMENT

Graduate Certificate Workshop: Introduction to Evidence-Based Teaching, 2019

Service

Doctoral Student Representative, Fall 2015 – Present

Responsibilities: revise and update a handbook for current and future students in the doctoral program, solicit feedback from students and discuss the program and potential changes with program directors especially as a part of the program review during Spring 2017.

Ad Hoc Reviewer

Developmental Psychology, co-reviewer with Dr. Margaret Friend 2018

Child Development, co-reviewer with Dr. Margaret Friend 2019

ABSTRACT OF THE DISSERTATION

Continuity in language development:
Predictions from decontextualized vocabulary and lexical access

by

Erin Laurel Smolak

Doctor of Philosophy in Language and Communicative Disorders

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

Professor Margaret Friend, Chair

Vocabulary knowledge and speed of lexical access are early components of linguistic skill that co-develop rapidly in the first two years of life. Evidence suggests that these components are foundational for downstream linguistic and cognitive skills and that early individual differences set the stage for a developmental cascade throughout the lifespan. The three studies in this dissertation seek to clarify how early, decontextualized vocabulary predicts

later development, whether it has utility in individual prediction, and how vocabulary and speed of lexical access support language development together and separately.

Chapter 1 reviews the current literature on vocabulary development and speed of lexical access, exploring how these skills are interrelated and their relation to other domain-general learning mechanisms, closing with extant evidence on predictive utility. Chapter 2 follows from recent findings that early vocabulary predicts later literacy but that it accounts for a modest amount of variance in outcomes. Chapter 2 uses an assessment of decontextualized vocabulary and shows that early, decontextualized vocabulary predicts later vocabulary and kindergarten readiness at age four, explaining additional variance above and beyond a parent report assessment of early vocabulary. In a similar vein, results from Chapter 3 reveal decontextualized vocabulary to be a robust predictor of language abilities at age three and present preliminary evidence for its utility as a screening measure. Finally, Chapter 4 presents preliminary evidence for the hypothesis that although vocabulary knowledge and lexical access are related, they are somewhat distinct in the extent to which they depend on long-term learning vs. processing efficiency. Results tentatively support this conclusion by revealing dissociation in prediction between these two components of early language. All three chapters examine these effects across groups differing linguistically and geographically and explore whether patterns across samples suggest generalization of effects across languages or reveal linguistic differences. Practically, the goal of this research is to elucidate the processes of early language development, improve identification of children at risk for language delay and inform intervention strategies.

INTRODUCTION

Introduction

Vocabulary knowledge and speed of word processing (lexical access) develop rapidly in the first two years of life (Fenson et al., 1994; Fernald, Perfors, & Marchman, 2006; Samuelson & McMurray, 2017). Recent research suggests that early word knowledge exists on a continuum from fragile, context-bound associations to more stable, decontextualized word representations (Bion et al., 2013; Hendrickson et al., 2015; 2017; Gershkoff-Stowe & Hahn; 2013). Of particular importance to the current research is the predictive power of decontextualized word knowledge that we hypothesize is built in an iterative process relying in part on domain general mechanisms such as cross-situational learning and hypothesis testing (Fazly et al., 2010; Yu & Smith, 2012; Vlach & DeBrock, 2017).

Extant research investigating prediction from early vocabulary to later linguistic and cognitive skills has relied on measures that likely assess the full continuum of children's word knowledge. This research has found that early language experience predicts preschool and school-age language (Bleses, Makransky, Dale, Højen, & Ari, 2016; Duff et al., 2015; Fernald & Marchman, 2012; Friend, et al., 2012; Henrichs et al., 2011; Reilly et al., 2010) but explains a modest amount of variance and does not scale down to prediction at the individual level (Ghassabian et al., 2013; Heilmann et al., 2005; Law et al., 2000; Wallace et al., 2015): previous research has failed to appropriately differentiate higher- from lower-performing children. Instead of assessing the full continuum of children's vocabulary knowledge, studies included in this dissertation seek to tap into decontextualized vocabulary specifically. We further seek to assess children's lexical access in order to inform our understanding of the continuity of language development and the role of early language as a building block for downstream linguistic and cognitive skills and conceptual knowledge. Since vocabulary knowledge and lexical access are related but separable, a study in this dissertation will also explore whether these constructs differentially predict downstream skills along a continuum from long-term knowledge to processing efficiency to. To this end, the research will address the following aims:

1) Assess the relation between decontextualized vocabulary and downstream language at the group and individual levels (Chapter 2 and 3)**.** Research investigating prediction from early vocabulary size has found that it predicts a modest proportion of outcome variance and inadequately differentiates children who will ultimately evince low language performance from those who will evince average to high language (Ghassabian et al., 2013; Heilmann et al., 2005; Law et al., 2000; Wallace et al., 2015; Westerlund et al., 2006). The perspective that guides the current research is that there is continuity in language skills across early childhood (e.g., Bornstein, Hahn, Putnick, & Suwalsky, 2014) but that previous research may have underestimated vocabulary's predictive power by utilizing parent reports that evaluate fragile, in conjunction with stable, word knowledge. Thus the first aim of the current research is to investigate prediction from decontextualized vocabulary size in the second year to language abilities one year later. Since early language is malleable and children benefit from targeted intervention (Cates, Weisleder, & Mendelsohn, 2016; Hassinger-Das et al., 2016; Reese, Sparks, & Leyva, 2010), we further seek to evaluate whether decontextualized vocabulary at age two can differentiate children with low language performance from children with average to high language performance at age three using a signal detection approach. We expect it will significantly predict later language abilities at the group level and will better differentiate high from lower performing children than prior approaches.

2) Investigate differential prediction from vocabulary and lexical access (Chapter 4)**.** Vocabulary size and lexical access are negatively correlated, such that children with larger vocabularies process words more efficiently (DeAnda, Hendrickson, Zesiger, Poulin-Dubois, & Friend, 2018; Fernald & Marchman, 2012; Fernald et al., 2006; Hendrickson et al., 2015; 2017; Legacy, Zesiger, Friend, & Poulin-Dubois, 2016). Both measures predict later language (e.g., Bleses et al., 2016; Bornstein & Haynes, 1998; Duff et al., 2015; Fernald & Marchman, 2012; Friend, Schmitt, & Simpson, 2012; Ghassabian et al., 2013; Henrichs et al., 2011; Reilly et al., 2010), and cognitive skills (Duff et al., 2015; Marchman & Fernald, 2008; Morgan et al., 2015). Indeed, one extant study compared these measures prospectively and found that both vocabulary size and lexical access predicted significant unique variance in working memory. Vocabulary size significantly predicted IQ at age 8 and this prediction was mediated by working memory (Marchman & Fernald, 2008). However, the shared variance between vocabulary size and lexical access is only about 20 percent, suggesting these are related, but separable, components of early language competency. We hypothesize that they may diverge in the extent to which they predict later skills along a continuum from working memory to conceptual development. The second aim of this work therefore extends the focus on prediction in Aim 1 by investigating differential prediction from vocabulary knowledge and lexical access to later language skills along this continuum.

3) Assess generalization of findings in French- and English-speaking children (Chapters 2, 3, and 4)**.** A final aim of this dissertation was to evaluate generalization of our effects across samples of children from differing linguistic backgrounds. In all three chapters, we will study effects in English- and French-speaking children. In addition, in Chapter 2 we also include a study of bilingual development in French-English speakers. We include multiple samples to determine if our findings generalize across languages or if linguistic differences also play a role.

References

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*, 39-53.

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*, 1461-1476.

Bornstein, M. H., Hahn, C. S., Bell, C., Haynes, O. M., Slater, A., Golding, J., ... & ALSPAC Study Team. (2006). Stability in cognition across early childhood: A developmental cascade. *Psychological Science*, *17*, 151-158.

Cates, C. B., Weisleder, A., & Mendelsohn, A. L. (2016). Mitigating the effects of family poverty on early child development through parenting interventions in primary care. *Academic Pediatrics*, *16*, S112-S120.

DeAnda, S., Hendrickson, K., Zesiger, P., Poulin-Dubois, D., & Friend, M. (2018). Lexical access in the second year: a study of monolingual and bilingual vocabulary development. *Bilingualism: Language and Cognition*, *21*, 314-327.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*, 848-856.

Fazly, A., Ahmadi-Fakhr, F., Alishahi, A., & Stevenson, S. (2010, January). Cross-situational learning of low frequency words: The role of context familiarity and age of exposure. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 32, No. 32).

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, i-185.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*, 203-222.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*, 98-116.

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the Computerized Comprehension Task: Comprehension predicts production. *Developmental Psychology*, *48*, 136.

Ghassabian, A., Rescorla, L., Henrichs, J., Jaddoe, V. W., Verhulst, F. C., & Tiemeier, H. (2014). Early lexical development and risk of verbal and nonverbal cognitive delay at school age. *Acta Paediatrica*, *103*, 70-80.

Gershkoff-Stowe, L., & Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. *Journal of Experimental Child Psychology*, *114*, 489-509.

Hassinger-Das, B., Ridge, K., Parker, A., Golinkoff, R. M., Hirsh-Pasek, K., & Dickinson, D. K. (2016). Building vocabulary knowledge in preschoolers through shared book reading and gameplay. *Mind, Brain, and Education*, *10*, 71-80.

Heilmann, J., Weismer, S. E., Evans, J., & Hollar, C. (2005). Utility of the MacArthur-Bates Communicative Development Inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology*, *14*, 40-51.

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*, 723-735.

Hendrickson, K., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2017). Assessing a continuum of lexical–semantic knowledge in the second year of life: A multimodal approach. *Journal of Experimental Child Psychology*, *158*, 95-111.

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W., Hofman, A., ... & Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The Generation R study. *Journal of Speech, Language, and Hearing Research*, *54*, 854-869.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). The feasibility of universal screening for primary speech and language delay: Findings from a systematic review of the literature. *Developmental Medicine and Child Neurology*, *42*, 190-200.

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2016). Vocabulary size and speed of word recognition in very young French–English bilinguals: A longitudinal study. *Bilingualism: Language and Cognition*, 1-13.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, F9-16.

Reese, E., Sparks, A., & Leyva, D. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, *10*, 97-117.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., ... & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: findings from Early Language in Victoria Study. *Pediatrics*, *126*, e1530-e1537.

Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word? *Wiley Interdisciplinary Reviews: Cognitive Science*, *8,* 1-16.

Wallace, I. F., Berkman, N. D., Watson, L. R., Coyne-Beasley, T., Wood, C. T., Cullen, K., & Lohr, K. N. (2015). Screening for speech and language delay in children 5 years old and younger: A systematic review. *Pediatrics*, 1-15.

Westerlund, M., Berglund, E., & Eriksson, M. (2006). Can severely language delayed 3-year-olds be identified at 18 months? Evaluation of a screening version of the MacArthur–Bates Communicative Development Inventories. *Journal of Speech, Language, and Hearing Research*, *49*, 237-247.

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word–referent learning: Prior questions. *Psychological Review*, *119*, 21.

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217-230.

CHAPTER 1

Word Knowledge and Processing:

A Domain General Developmental Account

Introduction

Vocabulary knowledge develops rapidly in the first two years of life (Reznick & Goldfield, 1992; Fenson et al., 1994; Dapretto & Bjork, 2000; Mayor & Plunkett, 2010; Samuelson & McMurray, 2017), as does the efficiency with which children are able to access their word knowledge (Fernald, Perfors, & Marchman, 2006). It is hypothesized that early word knowledge ranges along a continuum from fragile, context-bound knowledge to more robust, mature word representations (operationalized as those word-referent relations that are decontextualized and generalized to other exemplars, e.g., Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Of interest are the domain general mechanisms underlying the development of mature word knowledge (hereafter referred to as decontextualized word knowledge) and the interplay between decontextualized word knowledge, speed of processing, and working memory in laying a foundation for the development of linguistic and cognitive skills throughout the lifespan.

In reviewing the data on this topic, I begin with a brief review of vocabulary knowledge during the second year of life and describe several domain general mechanisms that have been proposed to contribute to the development of decontextualized word knowledge: speed of word processing and working memory. In the second section, I review current models of early word learning that focus on domain general processes, compare these to similar models of word learning in adulthood, and consider the empirical support for these models and how they can reconcile competing interpretations of the empirical literature. Finally, I highlight the predictions that can be made from early language to later linguistic and cognitive skills.

Vocabulary, Processing, and Working Memory

Vocabulary. Infants first begin to understand word-world pairings around the age of 6 to 8 months (Bergelson & Swingley, 2012; 2015; Fenson et al., 1994; Tincoff & Jusczyk, 1999; 2012), and produce their first words at approximately 10 to 15 months (Fenson et al., 1994). By the age of two most children are able to produce over 250 words (Fenson et al., 1994). However, there is considerable variability in the number of words children are able to comprehend and produce at any given age (Bates et al., 1994; Fenson et al., 1994). Additionally, although it has been shown that comprehension of word meanings precedes production of those words and increases at a more rapid pace, at least within the first year (Bornstein & Hendricks, 2012; Goldin-Meadow, Seligman, & Gelman, 1976) there is also variability both in the number of words comprehended at the start of production and in the time lag between comprehension and production (Harris, Yeeles, Chasin, & Oakley, 1995; Snyder, Bates, & Bretherton, 1980). Specifically, some children understand many words before they begin to produce any whereas others have a relatively equal number of words in comprehension and production across development (Goldin-Meadow, Seligman, & Gelma, 1976; Snyder et al., 1980).

Although the source of this variability in early vocabulary is not fully understood, it has been shown that genetic heritability contributes substantially to individual differences in language achievement (e.g., Bishop, Price, Dale, & Plomin, 2003; Dale et al., 1998). Nevertheless, there is empirical support suggesting that the process of word learning is guided by domain-general learning mechanisms such as sensitivity to statistical regularities and associative pattern learning (Akhtar & Montague, 1999; Blythe, Smith, & Smith, 2010; Fazly, Ahmadi-Fakhr, Alishani, & Stevenson, 2010; Siskind, 1996; Smith & Yu, 2008; Yu & Smith, 2011; Vlach & DeBrock, 2017; Vlach & Johnson, 2013), speed of processing (Bornstein et al., 2006;

Fernald & Marchman, 2008) and working memory/short term memory (Baddeley, 2003; Gathercole, 2006; Marini, Ruffino, Sali, & Molteni, 2017; Parra, Hoff, & Core, 2011; Stokes & Klee, 2009). These learning mechanisms also support cognitive development more generally (e.g., Bornstein et al., 2006; Rose, Feldman, Jankowski, & Van Rossem, 2008; Rose, Feldman, & Jankowski, 2012; Thiessen, Kronstein, & Hufnagle, 2013). The perspective that guides the present paper is that early vocabulary knowledge is foundational to later linguistic and cognitive skills that are similarly subserved by domain general learning mechanisms and that early individual differences set the stage for a developmental cascade of language and cognitive abilities throughout childhood and adolescence (e.g., Bornstein et al., 2006; Rose et al., 2008; van der Maas et al., 2006). That is, small individual differences in the application of general learning mechanisms early on produce variability in vocabulary acquisition, altogether culminating in important differences in outcomes. I begin by reviewing the process of word acquisition during the first two years of life, focusing on how children develop mature word-referent representations, and how this process is subserved by domain general learning mechanisms.

After a period of slow learning (about 8 to 11 words per month), researchers have observed an increase in the rate of lexical growth in some children in both comprehension and production, culminating in a rate of 22 to 37 words learned per month (Goldfield & Reznick, 1990; Reznick & Goldfield, 1992; Samuelson & McMurray, 2017). Most children show an increased learning rate however the absolute change in the slope of word acquisition varies from child to child (Bloom, 2000; Goldfield & Reznick, 1990). Researchers originally suggested that this increased efficiency was a result of a qualitative shift in the way children learn words (e.g., Nazzi & Bertoncini, 2003; Reznick & Goldfield, 1992) however more recent evidence suggests that it is due to a steady increase in the efficiency of learning mechanisms such as associative learning (Fazly et al., 2010; Plunkett, Sinha, Møller, & Strandsby, 1992; McMurray, 2007) or the maturation of general memory and attentional processes (e.g., Samuelson & McMurray, 2017; Vlach & DeBrock, 2017). Children's existing word knowledge may also play a role by constraining the possible meanings of novel words (Fazly et al., 2010; Plunkett et al., 1992; Li, Zhao, & MacWhinney, 2007; Yu, 2008). For example, after a period of slow, effortful learning children extract regularities across the words they know and use this domain-specific knowledge to interpret new words (Namy, 2012; Plunkett et al., 1992). In a computational model, McMurray showed that an increase in word-learning efficiency is a mathematically expected result of two facts: 1) children are learning multiple word meanings simultaneously and 2) the words children are learning vary in difficulty (McMurray, 2007). Indeed, McMurray showed that this type of increasing efficiency results from any type of learning that displays those two properties.

Together, this research suggests that word learning is supported by increases in the efficiency of basic learning mechanisms across the second year of life. Nevertheless, there has been some debate on the extent to which early word-referent representations are mature and adult-like. For example, some research has suggested that early-acquired words are context-specific: they are comprehended or produced only in the context in which they were learned (e.g., Harris, Barret, Jones, & Brookes, 1988; Harris, Yeeles, Chasin & Oakley, 1995). A child may use the word "shoe" only when her parents are helping her put her shoes on. In this case, the child does not have a mature representation of the meaning of the word "shoe" and is using it only in the context of an established routine. Increased experience with word-referent pairings is likely necessary to form more mature, decontextualized word knowledge (Bion, Borovsky &

Fernald, 2013; Horst & Samuelson, 2008), defined by Bates, Bretherton, Carlson, Carpen, and Rosser (1979, p. 273) as “occurring in a broader range of situations with decreasing perceptual support”. At this point, the child knows that “shoe” refers to the broader referent class *shoe* and is able to extend the referent beyond the context in which it was learned.

In addition to being potentially context-bound, words acquired after only brief exposures are difficult for toddlers to retain (Bion et al., 2013; Friederich & Friederici, 2011; Horst & Samuelson, 2008) and the representations of those word meanings may be immature, fragile, and subject to interference (Gershkoff-Stowe and Hahn; 2013). For example, children are more likely to make production errors on newly learned words by substituting a familiar word that they have produced recently (Gershkoff-Stowe, 2002). Further, Gershkoff-Stowe and Hahn (2013) found, in teaching new words to adults and two-year-old children, that these words were unstable and moved in and out of comprehension and production across three testing periods spaced one week apart. They conclude that word learning is not a gradual progression from initial, fragile understanding to mature word knowledge with consistent, defined stages of acquisition; rather it is a dynamic process with contextual dependencies that yields graded understanding of word meanings. This interpretation is consistent with recent research examining behavioral dissociations associated with partial word knowledge early in development.

Behavioral dissociations, in which children display knowledge in one modality but not another, help to illustrate the distinction between fragile, context-bound knowledge and decontextualized knowledge. In two recent studies comparing vocabulary size (on a forced-choice haptic response measure) and speed of word processing (indexed by a gaze shift to the target referent), processing speed was a more sensitive index of toddlers’ early understanding of word meaning (Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015; Hendrickson, Poulin-Dubois, Zesiger, & Friend, 2017). Specifically, children's haptic responses and visual fixations converged when they correctly touched the target referent and when they did not touch either image. However their responses diverged when they touched the distractor—they were as quick to fixate the target as when they executed a correct haptic response even though in this case they responded incorrectly. Similarly, Bannard and Tomasello (2012) taught children novel words in a socially-supported or nonsocial context and tested their knowledge of the words using pointing or looking to the target object as evidence of learning. They found that in the social context children evinced knowledge of the words in both measures however in the nonsocial context children evinced knowledge by looking at, but not pointing to, the target object. Whereas Bannard and Tomasello interpreted this dissociation between modalities as evidence that learning was superior in the social context, Hendrickson and colleagues have interpreted similar dissociations as evidence for graded word knowledge, with early word-referent relations existing along a continuum of strength. When the associative relation between a word and its potential referents is weak, children may orient to the target referent upon hearing the familiar word but nevertheless be prone to errors in executing an accurate haptic response. Indeed this type of dissociation is a hallmark of partial or graded knowledge (e.g., Goldin-Meadow & Wagner, 2005; Munakata, 2001). Together these studies suggest that the strength of word-referent associations depends in part on learning conditions, that high task demands preferentially select for strong, rather than weak, associations, and that robust learning may require referential, rather than mere associative, knowledge.

In sum, this evidence suggests that word knowledge is not all-or-none but exists on a continuum from fragile, context-bound to decontextualized knowledge. Further, stable word-referent associations require additional experience (Bion et al., 2013), developing incrementally over time from multiple presentations in different contexts and with several exemplars (Bates et al., 1979; Bunce & Scott, 2017; Scott & Fisher, 2012; Smith & Yu, 2008; Twomey, Ranson, & Horst, 2014; Vlach & Johnson, 2013; Vouloumanos & Werker, 2009). Early word knowledge is complemented, in the first two years of life, by improvements in the ability to access that knowledge.

Processing efficiency. Like vocabulary knowledge, efficiency of word recognition and language processing develop rapidly over the first two years. For example, infants are able to recognize familiarized words and segment them from fluent speech as early as 7.5 months of age (Jusczyk, 1999). This skill matures with age, as does children's efficiency in recognizing familiar words such that by 24 months of age they recognize words with a speed and accuracy similar to that of adults. Using the looking-while-listening task (LWL; Fernald, Pinto Swingley, Weinberg, & McRoberts, 1998; Swingley, Pinto, & Fernald, 1999) in which toddlers view two images and hear a prompt in which one is named, Fernald and colleagues found an increase in the speed with which familiar words are processed from 15 to 24 months of age (Fernald, Perfors, & Marchman, 2006; Hurtado, Marchman, & Fernald, 2007). Efficiency at 18 months was significantly correlated with efficiency at 21 months however stability was not observed across other ages. Nevertheless this improvement in processing over the second year has been replicated with children from diverse socioeconomic backgrounds (Fernald, Marchman, & Weisleder, 2013) and with children learning Spanish in the United States (Hurtado et al., 2010).

Moreover, toddlers, like adults, appear to process verbal information incrementally as the input unfolds such that at 15 months, children who initially fixate a distractor image do not initiate a look to the target image until after the offset of the target word. However, by 24 months of age children reliably initiate a look to the target image before the entire target word is uttered

(Fernald et al., 1998; Swingley et al., 1999). Indeed, children as young as 18 months of age are able to initiate a visual shift to the correct image using partial phonetic information (Fernald, Swingley, & Pinto, 2001). That is, toddlers reliably fixate the target image *baby* upon hearing the partial word /*bei*/. Not only do children recognize familiar words with partial phonetic information at this age, they recognize familiar words with some level of acoustic degradation (Zangl, Klarman, Thal, Fernald, & Bates, 2005).

Similar results have been obtained using a haptic measure of word processing (DeAnda, Hendrickson, Zesiger, Poulin-Dubois, & Friend, 2018; Legacy, Zesiger, Friend, & Poulin-Dubois, 2016). For example, Legacy and colleagues (2016) examined speed of word processing longitudinally in French monolingual and French-English bilingual children at 16 and 22 months of age and found that efficiency significantly improved over this time period in both groups. However, speed of processing at 16 months was not correlated with speed of processing at 22 months. Similarly, DeAnda and colleagues (2018), examined speed of processing using the same measure in English monolingual, Spanish monolingual, and English-Spanish bilingual children at 16 and 22 months of age. The authors found significant increases of a comparable magnitude in speed of processing over time and no correlations between efficiency at 16 and 22 months. However, children who made greater gains in speed of processing across time were faster at both 16 and 22 months of age. This latter finding suggests that there may be reliable individual differences in speed of word processing in the first two years of life.

Empirical evidence has shown that vocabulary size and speed of word processing are negatively correlated, such that children with larger vocabularies process words more efficiently (DeAnda et al., 2018; Fernald & Marchman, 2012; Fernald et al., 2006; Hendrickson et al., 2015; 2017; Hurtado, Grüter, Marchman & Fernald, 2014; Legacy et al., 2016). This is generally considered to be a bidirectional relation, such that increases in vocabulary size lead to more connected and efficient semantic networks, driving increases in speed of word processing. Complementarily, as children begin to process familiar words more quickly, they have additional attentional resources to devote to unknown word meanings within the speech stream, thus increasing the efficiency with which they can learn new words. Two studies have indirectly investigated the direction of effect between vocabulary size and speed of word processing: one found that children who were exposed to higher quality input were quicker to process familiar words, an effect explained by children's vocabulary size. This suggests a primary effect of vocabulary on processing speed (Hurtado et al., 2014). However, another study found that the positive effect of maternal input on vocabulary size was mediated by how quickly children processed familiar words (Weisleder & Fernald, 2013). These studies leave the directionality of the relation between speed of word processing and vocabulary size unclear. More recently, however, Borovosky and colleagues (2016) directly tested this relationship by showing that semantic density (operationalized as the number of words parents reported that children knew in a given semantic category) predicts speed of processing of words and sentences. Children were faster to orient to the referent after hearing a target word when the semantic network for that category was relatively dense. This suggests that semantic organization, rather than vocabulary size, per se, influences processing efficiency. However as both vocabulary and speed of word processing are theoretically undergirded by working memory, this review of domain general processes in vocabulary acquisition would be incomplete without a consideration of this contribution.

Working memory. Working memory is responsible for maintaining information and the processing of that information over short intervals of time, and is one component of a memory system that is interrelated with other types of short term, working, and long term memory processes. Vocabulary knowledge, speed of word processing, and the way in which they interact may also depend on working memory. In fact, Baddeley and Hitch (1974), in describing working memory, posited that it plays an important role in a variety of aspects of cognition, of which language is one factor. Not only does it seem likely that working memory is critical to evaluating the continuum of possible word-referent associations, there is a large body of evidence that working memory, especially the phonological loop component associated with auditory information, is involved in both native spoken language acquisition by children and adult second language learning (Archibald, 2017; Baddeley, 2003; Baddeley, Gathercole, and Papagno, 1998; Stokes & Klee, 2009). Logically, the acquisition of language relies on the ability to retain and manipulate verbal information over short time periods. Phonological working memory/short term memory span specifically appears to exhibit a causal effect on vocabulary acquisition early in development, with the relation becoming more bidirectional later in childhood (Bowey, 2001; Gathercole & Baddeley, 1993, but cf. Melby-Lervåg et al., 2012). Overall this suggests a tight link between phonological memory and vocabulary such that phonological working memory supports language acquisition (Gathercole, 2006; Marini et al., 2017; Stokes & Klee, 2009) and long-term vocabulary knowledge influences performance on tasks of phonological working memory (Gupta & Tisdale, 2009; Jones, Gobet, & Pine, 2007; Pierce, Genesee, Delcenserie, & Morgan, 2017).

Working memory capacity and speed of information processing are also tightly linked, such that increases in working memory capacity are associated with more efficient processing across childhood (Fry & Hale, 1996; Kail & Salthouse, 1994). Indeed, speed of word processing in infancy is related to working memory at age eight (Marchman & Fernald, 2008). Similar to the relation between vocabulary size and phonological working memory, the relation between working memory and processing speed is likely to be bidirectional, such that increased working memory capacity leads to more efficient processing, and speeded processing frees up additional working memory capacity.

In summary, vocabulary knowledge, speed of word processing, semantic organization, and working memory develop rapidly and are related over the first few years of life. I turn now to models of early word learning: specifically, I review evidence that toddlers may be sensitive to the statistical regularities across word naming events, using cross-situational associative learning and hypothesis testing mechanisms to build word representations by strengthening those associations that consistently map onto observed word-world relations and weakening those that do not (e.g., Akhtar & Montague, 1999; Blythe, Smith, & Smith, 2010; Siskind, 1996; Smith & Yu, 2008; Smith, Suanda & Yu, 2014; Yu, Ballard, & Aslin, 2005).

Models of Early Word Learning

Cross-situational learning model. Cross-situational learning was originally proposed to explain how children could solve the referential ambiguity problem. That is, given the sheer number of potential referents in a visual scene, how could a child infer the correct word meaning? Recent research however suggests that there is less ambiguity in word learning that originally assumed. For example, data from head-mounted cameras suggest that infants and toddlers have a more constrained view of the visual scene with fewer potential referents available relative to adults and that they may focus attention on one object at a time (Pereira, Smith, & Yu, 2014; Yu & Smith, 2011). Further, parents often create unambiguous naming instances for their children (Frank, Tenenbaum, & Fernald, 2013). This being the case, statistical regularities may be less important for referent selection (especially given evidence that children's moment-by-moment attention is vitally important to referent selection; Yu & Smith, 2011; Smith & Yu, 2013) but more important for decontextualized word learning. That is, with increasing exposure to different exemplars of a word (Twomey et al., 2014) in varying contexts (Bunce & Scott, 2017; Hills, Maouene, Riordan, & Smith, 2010; Vlach & DeBrock, 2017; Vouloumanos & Werker, 2009), the child may be able to extract the regularities among the many exemplars of the word to discover those properties that are relevant to its category (e.g., a dog has four legs, a tail, and barks). It should additionally be noted that mere co-occurrence of words and objects alone are not sufficient for word learning: toddlers must connect words and their referents in some way to store those co-occurrences (Smith & Yu, 2013).

Computational models have shown cross-situational associative learning to be consistent with available behavioral data, and robust in the face of realistic and noisy input (Fazly et al., 2010, Siskind, 1996; Yu & Smith, 2011). For example, models of probabilistic associative learning have replicated findings from experimental studies on the increased efficiency in word learning and comprehension of word meanings preceding production (Fazly et al., 2010). Indeed, young children can learn both nouns (Smith & Yu, 2008; Vlach & Johnson, 2013) and verbs (Scott & Fisher, 2012) from cross-situational information even when the label and the object do not occur together 100 percent of the time and when there are competitor referents that co-occur with the target object half of the time (but not more frequently; Bunce & Scott, 2017; Vouloumanos & Werker, 2009). Contextual diversity in the input may also play a role in associative word learning such that words that occur in many different contexts are learned more easily than words occurring in similar contexts (Hills et al., 2010; see Roembke & McMurray, 2016 for a similar argument from empirical work with adults). Children younger than two years of age appear to learn novel words more effectively with naming instances presented in succession as opposed to in alternation (Schwab & Lew-Williams, 2016; Vlach & Johnson, 2013) however by 30 months of age distributed naming events across spaced exposures yields the most robust word learning (Childers & Tomasello, 2002). Finally, performance in these tasks is related not to age but rather to general language and memory abilities (Vlach & DeBrock, 2017), adding additional support to the hypothesis that word learning is at least in part subserved by domain-general learning mechanisms.

Propose-but-verify model. Another model of word learning argues that children do not keep track of all word-referent co-occurrences but instead utilize hypothesis testing/elimination strategies (Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013; Woodard, Gleitman, & Trueswell, 2016). During hypothesis elimination, also referred to as propose but verify models, children start out with one hypothesis as to the meaning of a new word (i.e., a partial understanding of the word-referent relation that may be context bound). With each presentation of the word children either retain the hypothesis if it is confirmed, or if it is disconfirmed they reject the hypothesis and choose a new referent. Associative learning and hypothesis elimination models are similar however the major difference is associative learning models propose children attend to and build up information on the co-occurrences of words and objects whereas in hypothesis elimination children have a hypothesis as to the meaning of a word and that hypothesis is either confirmed or rejected. Empirical evidence for these models comes from findings that participants (adults and children) who initially made an incorrect word-referent mapping were at chance on the next trial in which the chosen referent was absent, though the unselected referent was present on both trials (Aravind et al., 2018; Trueswell et al., 2013; Woodward et al., 2016). The authors conclude that participants did not retain co-occurrence information from the referent that was present, but not chosen, during learning. However there is evidence from other research that, following an incorrect haptic response, adults' performance significantly improves on subsequent trials (Yurovsky, Fricker, Yu and Smith, 2014). This seems to suggest that, at least for adults, some combination of hypothesis testing and associative learning may be at play.

Hybrid model. A final proposal suggests that children use both inferential hypothesis testing and associative mechanisms (Roembke & McMurray, 2016; Stevens, Gleitman, Trueswell, & Yang, 2016; Tenenbaum & Xu, 2000; Xu & Tenenbaum, 2005; 2007; Yu & Smith, 2008; Yurovksy & Frank, 2015): these may be thought of as hybrid models of word learning. One form of this combined approach takes advantage of Bayesian learning statistics (Tenenbaum & Xu, 2000; Xu & Tenenbaum, 2005; 2007; Yurovksy & Frank, 2016). Specifically, these models suggest that children make rational hypotheses as to the most likely referent of a given word based on the word-learning context (e.g., infer the referent the object that is primarily in view). However, the models take advantage of the graded output of associative learning by suggesting that children entertain multiple hypotheses at a time and these hypotheses are not either correct or incorrect but rather have a certain probability of being correct. In other words, upon the presentation of a word children form a hypothesis space of word meanings, with prior probabilities of the likelihood of each word meaning given by the word-learning context and domain-specific knowledge. With each occurrence of the word, the likelihood of each hypothesis changes according to Bayesian learning rules. That is, the strength changes in proportion to the probability of the hypothesized word meaning being correct given the available data and with the probability of the available data given the hypothesis. Further, one model suggests that individuals may rely more on single-referent tracking when memory and attention demands are high (e.g., in situations of high ambiguity), but retain more graded referent hypotheses when demands are low (Yurovsky & Frank, 2016). Other, non-Bayesian combinatorial models similarly suggest that children and adults propose hypotheses to word meanings (possibly only one hypothesis with each word presentation) but also keep track of the co-occurrence statistics of other possible word-referent pairs (Stevens et al., 2016), and that learning is built gradually through many co-occurrences (Roembke & McMurray, 2016). These models better account for the empirical data than do purely associative- or hypothesis-testing-driven models (Stevens et al., 2016; Yurovksy & Frank, 2015) and show that the debate between the prior two models may have been due to difference in memory demands across tasks used in different empirical studies (e.g., differences in the number of foils used; Yurovsky & Frank, 2015).

Recent empirical findings support this hybrid model. Aravind et al. (2018) found that, when three, four, and five year old monolingual and bilingual children initially failed to identify the referent of a newly learned word, they did not immediately benefit from this mistake. Rather, on a second trial they remained at chance in identifying the correct referent, consistent with hypothesis testing/propose but verify accounts. However, this study presented only two trials to children. Due to this, as the authors acknowledge, the findings could also be consistent with hybrid models: if hypothesis testing and associative learning work together, it may be the case that many additional word-referent pairings are required to strengthen the appropriate association before hypothesis testing can yield an accurate result. This is the case in adults: although after an initially incorrect hypothesis adults remain at chance at identifying the referent in the next trial, there is evidence of gradual learning across a large number of trials (Roembke & McMurray, 2016).

This combinatorial approach has the advantage that word meanings can be acquired through hypothesis testing and word-referent pairing combined with the graded knowledge built through associative learning (Roembke & McMurray, 2016; Smith & Yu, 2013; Stevens et al., 2016; Tenenbaum & Xu, 2000; Xu & Tenenbaum, 2005; 2007; Yurovksy & Frank, 2016; Yu & Smith 2012). One possibility is that associative learning and hypothesis testing support the development of the mature decontextualized word meanings discussed in the beginning of this review. That is, although associative learning may yield word-referent relations that vary along a continuum of strength, over time the iterative application of associative learning combined with hypothesis testing across situations should contribute to the development of strong, decontextualized representations of word meaning.

Further, as previously reported research suggests that children's existing word knowledge scaffolds the acquisition of new words (Borovsky & Elman, 2006; Plunkett et al., 1992; Li et al., 2007; Namy, 2012; Thom & Sandhofer, 2009; Yu, 2008). It is not only the absolute number of words known that can facilitate new lexical acquisition, however. As children learn new words they begin to organize these words into semantic networks (e.g., animals; Beckage, Smith, & Hills, 2011). Indeed a recent study finds that that infants are sensitive to semantic similarities between words even in the very earliest stages of word learning: in one study, six-month-olds' familiar word recognition was suppressed due to interference when it was paired with a semantically related object relative to when it was paired with a semantically unrelated object (Bergelson & Aslin, 2017), although these results may be due more to fragile early word knowledge than semantic category interference *per se*. By 24 months of age, semantic categories are more robust: at this point, upon hearing a word that is a member of a given semantic category (e.g., "dog") toddlers will activate other semantically related words (e.g., "cat") (Arias-Trejo & Plunkett, 2009; 2010). With development, these semantic networks appear to support the acquisition of new words such that the more words that a child acquires within a given semantic category, the easier it is for them to learn new words within that category (Borovsky, Ellis, Evans, & Elman, 2016; Hills, Maouene, Maouene, Sheya, & Smith, 2009; Thom & Sandhofer, 2009). Further, semantic density (i.e., the number of words that share semantic meaning) in a given semantic network is associated with increased efficiency in processing words in that network (Borovsky et al., 2016). Together, these studies suggest that across the latter half of the second year of life, the lexicon becomes more stable and organized and continues to support vocabulary development.

In summary, evidence that vocabulary knowledge in toddlerhood exists along a continuum from fragile, context-bound to more stable, decontextualized knowledge and that it is scaffolded by existing knowledge, suggests that mature vocabulary is formed via an iterative process. As knowledge of words develops and becomes decontextualized, words are organized into semantic categories in the lexicon and this accrued experience with word-referent pairings and relations between words supports new word learning. It is this stable decontextualized knowledge that may provide the best support for downstream learning. Of interest in the next section of this paper are the skills that emerge in early childhood that are dependent upon these early foundations.

Vocabulary and Speed of Processing Predict Downstream Outcomes

Appropriate language and cognitive development is largely dependent upon the earliest language skills established in infancy and toddlerhood. As summarized above, the second year of life is a critical time during which children's vocabularies and speed of word processing are rapidly developing (Fernald et al., 2006; Reznick & Goldfield, 1992; Fenson et al., 1994) and this early language experience has been shown to predict language (e.g., Bleses, Makransky, Dale, Højen, & Ari, 2016; Bornstein & Haynes, 1998; Duff, Reen, Plunkett, & Nation, 2015;

Fernald & Marchman, 2012; Friend, Schmitt, & Simpson, 2012; Ghassabian et al., 2014; Henrichs et al., 2011; Reilly et al., 2010), nonverbal intelligence (Bornstein & Haynes, 1998; Marchman & Fernald, 2008), reading accuracy and comprehension (Duff et al., 2015), school readiness and achievement (Christensen, Zubrick, Lawrence, Mitrou, & Taylor, 2014; Friend et al., 2018; Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015) and behavioral functioning (Morgan et al., 2015) in the preschool and school-age periods. In this section I summarize previous work investigating the predictive power of vocabulary and speed of word processing and highlight areas for further research.

Developmental predictions from early vocabulary. Much of the research investigating the types of downstream outcomes that early vocabulary predicts has focused on assessing vocabulary at three years of age or older (Christensen et al., 2014; Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Forget-Dubois et al., 2009; Storch & Whitehurst, 2002; Zubrick, Taylor, & Christensen, 2015). In general, this research has shown that vocabulary competence in the preschool years is an important building block for later language, reading, literacy, and school achievement. For example, vocabulary at age four predicts later language and literacy at ages eight and ten (Christensen et al., 2014; Zubrick et al., 2014) and low receptive vocabulary is a risk factor for continued vocabulary delay (Christensen et al., 2014). Similarly, the pace of vocabulary acquisition across the preschool years is associated with language skill at age five (Rowe, Raudenbush, & Goldin-Meadow, 2012). Vocabulary knowledge prior to school entry (between three and five years of age) contributes to reading development in the elementary school years both directly and indirectly by supporting phonological awareness skills that contribute to print decoding (Dickinson et al., 2010; Dickinson et al., 2005; Storch & Whitehurst, 2002). Longer-term prediction has also been established, with research showing that vocabulary at kindergarten entry significantly predicts literacy into adulthood when controlling for a broad set of risk factors (e.g., maternal education and home literacy experience; Law, Rush, Parsons & Schoon, 2013).

Together this research reveals the importance of preschool vocabulary knowledge for a variety of school-age outcomes. However, until more recently, less was known about whether vocabulary skills as early as the second year of life predicted long-term linguistic and cognitive skills. In fact, due to the substantial variability evidenced in children's early vocabulary knowledge (e.g., Fenson et al., 2000), some researchers have argued that there is not sufficient stability in vocabulary across the first few years of life to yield adequate prediction (Feldman et al., 2000; Law, Boyle, Harris, Harkness, & Nye, 2000), while other researchers suggest this variability is meaningful and stable across at least across short time intervals (Fenson et al., 2000). Since early language is malleable and children benefit from targeted intervention (Cates, Weisleder, & Mendelsohn, 2016; Hassinger-Das et al., 2016; Mol, Bus, & de Jong, 2009; Wasik, Bond, & Hindman, 2006; Reese, Sparks, & Leyva, 2010), it is critical to more thoroughly examine the predictive power of vocabulary as assessed as early as age two.

More recent research investigating this very question has yielded somewhat mixed results. Although vocabulary production as early as 16 months of age predicts later language (Bleses et al., 2016; Duff et al., 2015; Ghassabian et al., 2014; Henrichs et al., 2011; Reilly et al., 2010), literacy and reading (Duff et al., 2015; Lee, 2011), and academic and behavioral performance (Morgan et al., 2015), some studies employing large, demographically representative samples of children have found that it predicts a relatively small amount of variance in these outcomes (Bleses et al., 2016; Henrichs et al., 2011; Reilly et al., 2010). Specifically, a study of more than 3000 children in the Netherlands found that expressive and receptive vocabulary at 18 months predicted 11.5 percent of the variance in vocabulary at 30 months (Henrichs et al., 2011). Similarly, Lee (2011), studying 1000 children in the United States, found that a binary variable of high or low vocabulary at age two predicted 11 to 19 percent of the variance in expressive language at age three and 12 percent of the variance in language at age five after controlling for gender, birth order, socioeconomic status, and ethnicity. These studies predict across fairly narrow age ranges, however. Extending this prediction, late-talker status (having relatively low expressive vocabulary) at age 2 in a sample of 1910 Australian children predicted between 5 and 10 percent of unique variance in language outcomes at age four above and beyond other measurable risk factors (e.g., child sex, family history of language impairment; Reilly et al., 2010). Finally, in a sample of more than 2000 Danish children, vocabulary production assessed between 16 and 30 months of age predicted a significant but small amount of variance in sixth-grade reading and math outcomes (between 3 and 4 percent of the variance in language and literacy, and only 2 percent of the variance in math achievement). The authors found significant prediction from as early as 16 months of age however from this age range the variance accounted for was generally less than 3 percent (Bleses et al., 2016). This research shows that the variance in outcomes accounted for generally decreases as the age at outcomes assessment increases. Indeed, the proportion of variance accounted for in the U.S. study decreased with the age of outcome measurement such that early vocabulary predicted only about 7 percent of the variance in language and reading comprehension at the first and third grade levels (Lee, 2011).

More promisingly, a study of 300 British children found that vocabulary between 16 and 24 months of age significantly predicted language and reading skills between 4 and 9 years of age (Duff et al., 2015). The authors created a latent predictor from vocabulary comprehension and production between 16 and 24 months of age (with age regressed out) and used this factor to model longitudinal relations between early vocabulary and several measures of language and literacy. Early vocabulary accounted for 16% of the variance in school-age vocabulary, 11% of the variance in reading accuracy, and 18% of the variance in reading comprehension. Family risk of language and reading difficulties did not predict additional variance in school-age vocabulary, but did predict an additional 10 percent of the variance in reading accuracy and 12 percent of the variance in reading comprehension, which is similar to that explained by early vocabulary alone. Finally, a recent study of 8650 children found that vocabulary production at age two predicted academic reading and math achievement as well as behavioral performance at age five (Morgan, et al., 2015). The standardized effect sizes for early vocabulary's prediction to later reading and math were .22 and .27, respectively.

Importantly, the findings reviewed above confirm that early vocabulary knowledge provides a strong foundation for the development of later language, math and reading achievements that are important for school success. However, in general, vocabulary measured as early as the second year of life predicted a modest proportion of outcome variance (Bleses et al., 2016; Reilly et al., 2010) and did not add much unique prediction above and beyond other demographic risk factors. Further, the practical significance of these findings for individual children is uncertain.

Although early vocabulary size significantly predicts language and cognitive outcomes at the group level in these larger-scale studies, research suggests that prediction at the individual level is inadequate especially from young ages and across longer time periods (Law et al., 2000; Law & Roy, 2008). For example, a longitudinal study aimed at predicting children's risk for language delay at 30 months of age based on their vocabulary size at 24 months of age found that many children with low language scores at 30 months were not properly identified (Heilmann, Weismer, Evans, & Hollar, 2005). Specifically, classifying children at or below the 11th percentile in vocabulary production at 24 months as at risk for language delay yielded specificity and sensitivity estimates of .98 and .68, respectively. In practical terms, this means that although vocabulary correctly identified 98 percent of children who had 30-month language scores within the typical range it captured only 68 percent of children who underperformed at 30 months of age. A similar pattern has been found in many other studies of individual prediction from early vocabulary. For example, within a group of twins found to have language difficulties at three and four years, only about half were identified by a measure of vocabulary taken two years prior (Dale, Price, Bishop, & Plomin, 2003), yielding a sensitivity of 51.5 percent and a specificity of 80.5 percent. Lexical development at 18 and 30 months of age in a large sample of children in the Netherlands modestly predicted language comprehension at age six, and again only about half of toddlers with low vocabulary early on displayed persistent delay (Ghassabian et al., 2014). Finally, using a vocabulary screener 18 months, Westerlund and colleagues that although specificity was high at 90 percent, sensitivity measured only 50 percent (Westerlund, Berglund, & Eriksson, 2006). If the criterion for language delay on the screener was raised to yield the best balance of sensitivity and specificity, both were low to moderate, with a sensitivity of 66 percent and a specificity of 73 percent, both of which are inadequate. In summing up all these findings, a recent review of measures of toddler vocabulary suggests that there is a moderately higher likelihood of language delay for a positive screening on a parent report, but individual prediction is not sufficient to warrant screening for language delay as early as age two (Wallace et al., 2015).

In summary, vocabulary assessed in the second year of life predicts a modest percentage of variance in language and cognitive outcomes at the group level. Further, when vocabulary is used distinguish children with average to high language abilities from children with low language abilities or language delay, specificity (or the percent of typically-developing children who are appropriately identified), is at an acceptable level. However, sensitivity is inadequate, meaning many children who later develop language delay are not identified. One explanation for these findings is that researchers have not focused on children's early, decontextualized word understanding. Decontextualized comprehension and production are the products of multiple iterations of domain general associative learning and hypothesis testing and therefore represent a child's most mature understanding of word-referent relations. Arguably, it is this mature knowledge that is foundational to the cognitive and linguistic skills that early vocabulary size predicts. Therefore, extending this work using measures that tap into children's decontextualized word knowledge is crucial (Law & Roy, 2008).

Developmental predictions from early speed of word processing. The efficiency with which children are able to process incoming information and their ability to manipulate this information in short term storage is associated with a wide array of linguistic and cognitive skills (Alloway, Gathercole, Adams, Willis, Eaglen, & Lamont, 2005; Baddeley, 2003; Marchman & Fernald, 2008). Together, these studies provide evidence for the continuity of language and cognitive skills over time, such that the children's earliest facility with working memory and language processing seem to lay the substrate for the development of future skills.

In fact, previous studies have found that speed of word processing is associated with a variety of linguistic and cognitive outcomes. Specifically, in an early study, Fernald and colleagues assessed word processing efficiency longitudinally in conjunction with parent-reported vocabulary size (Fernald, Perfors, & Marchman, 2006). The authors found that speed of word processing was negatively correlated with parent reported vocabulary comprehension at 15 months and production at 18, 21, and 25 months. In addition, speed of processing was significantly associated with vocabulary comprehension at 25 months. Finally, speed of word processing at 25 months predicted the speed and acceleration of expressive vocabulary growth over the second year of life in a growth curve model. More recent research has found that speed of word processing assessed at 18 months of age is associated with concurrent vocabulary size: children classified as late talkers exhibited significantly slower speed of word processing than their typically developing peers (Fernald & Marchman, 2012). In addition, speed of processing significantly predicted the rate and acceleration of vocabulary growth over the second year of life, explaining an additional 8 percent of the variance in rate of growth, 4 percent of the variance in acceleration, and 12 percent of the variance in the intercept (vocabulary size at the final period of measurement at 30 months) above and beyond parent-reported vocabulary alone.

Speed of processing also significantly predicts expressive vocabulary and general IQ at age 8 (Marchman & Fernald, 2008). Speed of word processing and parent-reported expressive vocabulary size both significantly predicted later expressive vocabulary size when controlling for the other variable but only parent-reported vocabulary size predicted significant unique variance in IQ. Interesting effects were also found when children were organized into four groups based on median splits of vocabulary and speed of processing: children with large vocabularies and fast speed of word processing performed significantly better than other groups on expressive vocabulary and IQ and children who scored low on both measures performed significantly worse, with children who were high on one measure and low in the other both performing at intermediary levels. Finally, although both vocabulary and speed of word processing did predict

IQ, this effect appeared to be mediated by working memory, such that processing speed and vocabulary both no longer predict IQ when controlling for working memory (Marchman & Fernald, 2008).

Although early vocabulary size alone has been shown to be inadequate at differentiating children with low language from typically developing children (Dale et al., 2003; Heilmann et al., 2005; Law et al., 2000; Law & Roy, 2008; Wallace et al., 2015; Westerlund et al., 2006), there is evidence that speed of word processing can add to this prediction (Fernald & Marchman, 2012). Specifically, this study found that vocabulary size alone at 18 months yielded sensitivity of 78 percent and specificity of 66 percent when predicting vocabulary at 30 months of age. Children who were classified as late talkers evinced significantly reduced speed of processing compared to their typically developing peers. Within this group of late talkers, those with persistent low language at 30 months evinced significantly reduced speed of processing compared those who were typically developing at 30 months. Although speed of word processing in combination with vocabulary size lowers sensitivity to 61 percent, it provides better discrimination than vocabulary size alone by reducing the number of false positives (i.e., the number of children classified as low language in the second year who are typically developing at 30 months of age). However, given the lower sensitivity there is still progress to be made in individual prediction from language measures in the second year of life.

Summary

This section has reviewed evidence from extant research suggesting that, as early acquired language skills, both vocabulary size (Bleses et al., 2016; Duff et al., 2015; Friend et al., 2012; Henrichs et al., 2011; Morgan et al., 2015; Reilly et al., 2010) and speed of word processing (Fernald et al., 2006; Fernald & Marchman, 2008; Marchman & Fernald, 2012) predict later linguistic and cognitive outcomes including literacy, working memory, and IQ. However, prediction is moderate at the group level and inadequate at the individual level. Furthermore, although vocabulary size and speed of processing are somewhat distinct assessments of early language ability, only one study to our knowledge has directly compared these constructs prospectively (Marchman and Fernald, 2008). Thus, extending the existing body of research necessitates a more thorough consideration of early vocabulary size as a predictor of language outcomes and examination of how vocabulary size and speed of processing converge and diverge in prediction.

Conclusion

Accounts of early language acquisition suggest 1) that early acquired word knowledge exists along a continuum of strength (Bion, Borovsky & Fernald, 2013; Bretherton et al., 1979; Hendrickson et al., 2015; 2017; Horst & Samuelson, 2008), 2) that young children acquire word meanings over time and across contexts using, at least in part, general learning mechanisms (Akhtar & Montague, 1999; Blythe et al., 2010; Fazly et al., 2010; Siskind, 1996; Smith & Yu, 2008; Yu & Smith, 2011; Vlach & DeBrock, 2017; Vlach & Johnson, 2013), and 3) vocabulary size and speed of word processing are related but not completely overlapping skills (e.g., DeAnda et al., 2018; Marchman & Fernald, 2008).

These early language measures predict a small to moderate amount of variance in later language and school readiness/achievement and have so far been shown to be inadequate predictors of individual language outcomes such that a substantial proportion of children at risk for low language abilities are not accurately identified. One possible reason for the modest prediction from vocabulary is that most research to date has utilized assessments that capture the full continuum of children's vocabulary knowledge. That is, these assessments may have collapsed vocabulary estimates across words for which children have robust, decontextualized representations and words for which representations are more fragile. As discussed, word knowledge built up through experience and domain-general learning mechanisms might provide the best foundation for the acquisition of downstream conceptual learning. It is therefore unclear from this research whether strength of prediction can be improved by focusing on children's more robust word knowledge. Moreover, it remains unclear the exact way in which vocabulary and processing speed provide the foundation for later learning: that is, whether vocabulary and speed of processing differentially predict linguistic and cognitive outcomes. These constructs may be seen as varying along a continuum from long-term semantic knowledge (vocabulary) to short-term memory and processing of linguistic information (speed of word processing), and may differentially predict language outcomes that lie along a similar continuum. The goal of this dissertation is to address these remaining questions by investigating the utility of decontextualized vocabulary in the second year of life for predicting language and cognitive outcomes downstream (Chapters 1 and 2) and for differentiating high performing from lower-performing children (Chapter 2), and to explore whether these constructs differentially predict downstream outcomes existing along a continuum from processing and working memory constructs to long-term knowledge (Chapter 3).

References

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*, 347-358.

Alloway, T. P., Gathercole, S. E., Adams, A. M., Willis, C., Eaglen, R., & Lamont, E. (2005). Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. *British Journal of Developmental Psychology*, *23*, 417-426.

Aravind, A., de Villiers, J., Pace, A., Valentine, H., Golinkoff, R., Hirsch-Pasek, K., ... & Wilson, M. S. (2018). Fast mapping word meanings across trials: Young children forget all but their first guess. *Cognition*, *177*, 177-188.

Archibald, L. M. (2017). Considering the (separable) influences of phonological sensitivity and working memory on language learning outcomes. *Applied Psycholinguistics*, *38*, 1301-1304.

Arias-Trejo, N., & Plunkett, K. (2009). Lexical–semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 3633-3647.

Arias-Trejo, N., & Plunkett, K. (2010). The effects of perceptual similarity and category membership on early word-referent identification. *Journal of Experimental Child Psychology*, *105*, 63-80.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189-208.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). San Diego, CA: Academic Press.

Bannard, C., & Tomasello, M. (2012). Can we dissociate contingency learning from social learning in word acquisition by 24-month-olds? *PloS ONE*, *7*, e49881.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The Emergence of Symbols*. New York: Academic Press.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., ... & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, *21*, 85-123.

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PloS ONE*, *6*(5), e19348.

Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences*, *114*, 12916-12921.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*, 3253-3258.

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, *11*, 369-380.

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*, 39-53.

Bishop, D. V., Price, T. S., Dale, P. S., & Plomin, R. (2003). Outcomes of early language delay: II. Etiology of transient and persistent language difficulties. *Journal of Speech, Language, and Hearing Research*, *46*, 561-575.

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*, 1461-1476.

Blythe, R. A., Smith, K., & Smith, A. D. (2010). Learning Times for Large Lexicons Through Cross-Situational Learning. *Cognitive Science*, *34*, 620-642.

Bornstein, M. H., Hahn, C. S., Bell, C., Haynes, O. M., Slater, A., Golding, J., ... & ALSPAC Study Team. (2006). Stability in cognition across early childhood: A developmental cascade. *Psychological Science*, *17*, 151-158.

Bornstein, M. H., & Haynes, O. M. (1998). Vocabulary competence in early childhood: Measurement, latent construct, and predictive validity. *Child Development*, *69*, 654-671.

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Lexical leverage: Category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science*, *19*, 918-932.

Borovsky, A., & Elman, J. (2006). Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language*, *33*, 759-790.

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics*, *22*, 441-469.

Bunce, J. P., & Scott, R. M. (2017). Finding meaning in a noisy world: exploring the effects of referential ambiguity and competition on 2· 5-year-olds' cross-situational word learning. *Journal of Child Language*, *44*, 650-676.

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, *44*, 431-443.

Cates, C. B., Weisleder, A., & Mendelsohn, A. L. (2016). Mitigating the effects of family poverty on early child development through parenting interventions in primary care. *Academic Pediatrics*, *16*, S112-S120.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*, 967.

Christensen, D., Zubrick, S. R., Lawrence, D., Mitrou, F., & Taylor, C. L. (2014). Risk factors for low receptive vocabulary abilities in the preschool and early school years in the Longitudinal Study of Australian Children. *PloS ONE*, *9*, e101476.

Dale, P. S., Price, T. S., Bishop, D. V., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research*, *46*, 544-560.

Dale, P., Simonoff, E., Bishop, D., Eley, T., Oliver, B., Price, T., ... & Plomin, R. (1998). Genetic influence on language delay in two-year-old children. *Nature Neuroscience*, *1*, 324-328.

DeAnda, S., Hendrickson, K., Zesiger, P., Poulin-Dubois, D., & Friend, M. (2018). Lexical access in the second year: a study of monolingual and bilingual vocabulary development. *Bilingualism: Language and Cognition*, *21*, 314-327.

Dell, G.S, & Chang, F. (2014). The P-Chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 369,* 20120394.

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, *39*, 305-310.

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, *95*, 465.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*, 848-856.

Fazly, A., Ahmadi-Fakhr, F., Alishahi, A., & Stevenson, S. (2010, January). Cross-situational learning of low frequency words: The role of context familiarity and age of exposure. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 32, No. 32).

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development*, *71*, 310-322.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J. S., & Thal, D. (2000). Reply: Measuring variability in early child language: Don't shoot the messenger. *Child Development*, *71*, 323-328.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, i-185.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*, 203-222.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*, 234-248.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*, 98-116.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*, 228-231.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*, 1003-1015.

Forget-Dubois, N., Dionne, G., Lemelin, J. P., Pérusse, D., Tremblay, R. E., & Boivin, M. (2009). Early child language mediates the relation between home environment and school readiness. *Child Development*, *80*, 736-749.

Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: fast encoding–weak retention. *Journal of Cognitive Neuroscience*, *23*, 3228-3240.

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the Computerized Comprehension Task: Comprehension predicts production. *Developmental Psychology*, *48*, 136.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237-241.

Ghassabian, A., Rescorla, L., Henrichs, J., Jaddoe, V. W., Verhulst, F. C., & Tiemeier, H. (2014). Early lexical development and risk of verbal and nonverbal cognitive delay at school age. *Acta Paediatrica*, *103*, 70-80.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*, 513-543.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*, 259.

Gershkoff-Stowe, L. (2002). Object naming, vocabulary growth, and the development of word retrieval abilities. *Journal of Memory and Language*, *46*, 665-687.

Gershkoff-Stowe, L., & Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. *Journal of Experimental Child Psychology*, *114*, 489-509.

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*, 171-183.

Goldin-Meadow, S., & Wagner, S. M. (2005). How our hands help us learn. *TRENDS in Cognitive Sciences*, *9*, 234-241.

Goodman, J.C, & Bates, E. (2013). On the emergence of grammar from the lexicon. In *The emergence of language* (pp. 47-98). Psychology Press.

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language*, *61*, 481-502.

Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language*, *15*, 77-94.

Harris, M., Yeeles, C., Chasin, J., & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension and production. *Journal of Child Language*, *22*, 1-18.

Hassinger-Das, B., Ridge, K., Parker, A., Golinkoff, R. M., Hirsh-Pasek, K., & Dickinson, D. K. (2016). Building vocabulary knowledge in preschoolers through shared book reading and gameplay. *Mind, Brain, and Education*, *10*, 71-80.

Heilmann, J., Weismer, S. E., Evans, J., & Hollar, C. (2005). Utility of the MacArthur-Bates Communicative Development Inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology*, *14*, 40-51.

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*, 723-735.

Hendrickson, K., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2017). Assessing a continuum of lexical–semantic knowledge in the second year of life: A multimodal approach. *Journal of Experimental Child Psychology*, *158*, 95-111.

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W., Hofman, A., ... & Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The Generation R study. *Journal of Speech, Language, and Hearing Research*, *54*, 854-869.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? *Psychological Science*, *20*, 729-739.

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language*, *63*, 259-273.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*, 128-157.

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish–English bilingual toddlers. *Bilingualism: Language and Cognition*, *17*, 189-202.

Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language*, *34*, 227-249.

Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: a computational model of the learning of new words. *Developmental Science*, *10*, 853-873.

Jusczyk, P. W. (1999). How infants begin to extract words from speech. *TRENDS in Cognitive Sciences*, *3*, 323-328.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199-225.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). The feasibility of universal screening for primary speech and language delay: Findings from a systematic review of the literature. *Developmental Medicine and Child Neurology*, *42*, 190-200.

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health*, *13*, 198-206.

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, *32*, 69-92.

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2016). Vocabulary size and speed of word recognition in very young French–English bilinguals: A longitudinal study. *Bilingualism: Language and Cognition*, 1-13.

Li, P., Zhao, X., & MacWhinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive Science*, *31*(4), 581-612.

March, V.A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339-366.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, F9-16.

Marini, A., Ruffino, M., Sali, M. E., & Molteni, M. (2017). The role of phonological working memory and environmental factors in lexical development in Italian-speaking late talkers: A one-year follow-up study. *Journal of Speech, Language, and Hearing Research*, *60*, 3462-3473.

Mayor, J., & Plunkett, K. (2010, January). Vocabulary Spurt; Are Infants full of Zipf? In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 32, No. 32).

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*, 631-631.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*, 9014-9019.

Melby-Lervåg, M., Lervåg, A., Lyster, S. A. H., Klem, M., Hagtvet, B., & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science*, *23*, 1092-1098.

Mol, S. E., Bus, A. G., & De Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, *79*, 979-1007.

Montgomery, J.W. (2006). Real-time language processing in school-age children with specific language impairment. *International Journal of Language & Communication Disorders, 41*, 275–291.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, *86*, 1351-1370.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *TRENDS in Cognitive Sciences*, *5*, 309-315.

Namy, L. L. (2012). Getting specific: Early general mechanisms give rise to domain-specific expertise in word learning. *Language Learning and Development*, *8*, 47-60.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, *6*, 136-142.

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, *21*, 178-185.

Pierce, L. J., Genesee, F., Delcenserie, A., & Morgan, G. (2017). Variations in phonological working memory: Linking early language experiences and language learning outcomes. *Applied Psycholinguistics*, *38*, 1265-1300.

Plunkett, K., Sinha, C., Møller, M. F., & Strandsby, O. (1992). Symbol grounding or the emergence of symbols? Vocabulary growth in children and a connectionist net. *Connection Science*, *4*, 293-312.

Reese, E., Sparks, A., & Leyva, D. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, *10*, 97-117.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., ... & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: findings from Early Language in Victoria Study. *Pediatrics*, *126*, e1530-e1537.

Reznick, J. S., & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology*, *28*, 406-413.

Roembke, T. C., & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*, *87*, 105-127.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382-439.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2008). A cognitive cascade in infancy: Pathways from prematurity to later mental development. *Intelligence*, *36*, 367-378.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2012). Information processing from infancy to 11 years: Continuities and prediction of IQ. *Intelligence*, *40*, 445-457.

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development*, *83*, 508-525.

Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word? *Wiley Interdisciplinary Reviews: Cognitive Science*, *8,* 1-16.

Schafer, G. (2005). Infants can learn decontextualized words before their first birthday. *Child Development*, *76*, 87-96.

Schwab, J. F., & Lew-Williams, C. (2016). Repetition across successive sentences facilitates young children's word learning. *Developmental psychology*, *52*(6), 879.

Scott, R. M., & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, *122*, 163-180.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*, 39-91.

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word–referent learning. *TRENDS in cognitive sciences*, *18*(5), 251-258.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558-1568.

Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development*, *9*, 25-49.

Stevens, J. S., Gleitman, L. R., Trueswell, J. C., & Yang, C. (2017). The pursuit of word meanings. *Cognitive Science*, *41*, 638-676.

Stokes, S. F., & Klee, T. (2009). Factors that influence vocabulary development in two-year-old children. *Journal of Child Psychology and Psychiatry*, *50*, 498-505.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, *38*, 934-947.

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, *71*, 73-108.

Tenenbaum, J. B., & Xu, F. (2000, January). Word learning as Bayesian inference. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 22, No. 22).

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin*, *139*, 792-814.

Thom, E. E., & Sandhofer, C. M. (2009). More is more: The relationship between vocabulary size and word extension. *Journal of Experimental Child Psychology*, *104*, 466-473.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*, 172-175.

Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*, 432-444.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *TRENDS in ognitive Sciences, 4,* 156-163.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*, 126-156.

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's more like it: Multiple exemplars facilitate word learning. *Infant and Child Development*, *23*, 105-122.

van der Maas, H. L., Dolan, C. V., Grasman, R. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842-861.

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217-230.

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*, 375-382.

Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, *45*, 1611.

Wallace, I. F., Berkman, N. D., Watson, L. R., Coyne-Beasley, T., Wood, C. T., Cullen, K., & Lohr, K. N. (2015). Screening for speech and language delay in children 5 years old and younger: A systematic review. *Pediatrics*, 1-15.

Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology*, *98*, 63-74.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*, 2143-2152.

Westerlund, M., Berglund, E., & Eriksson, M. (2006). Can severely language delayed 3-year-olds be identified at 18 months? Evaluation of a screening version of the MacArthur–Bates Communicative Development Inventories. *Journal of Speech, Language, and Hearing Research*, *49*, 237-247.

Woodard, K., Gleitman, L. R., & Trueswell, J. C. (2016). Two-and three-year-olds track a single meaning during word learning: Evidence for Propose-but-verify. *Language Learning and Development*, *12*(3), 252-261.

Xu, F., & Tenenbaum, J. B. (2005). Word learning as Bayesian inference: Evidence from preschoolers. In *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (Vol. 23).

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*, 245.

Yu, C. (2008). A statistical associative account of vocabulary growth in early word learning. *Language Learning and Development*, *4*, 32-62.

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*, 165-180.

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word–referent learning: Prior questions. *Psychological Review*, *119*, 21.

Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53-62.

Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. B. (2014). The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review*, *21*, 1-22.

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, *6*, 179-208.

Zubrick, S. R., Taylor, C. L., & Christensen, D. (2015). Patterns and predictors of language and literacy abilities 4-10 years in the longitudinal study of Australian children. *PloS ONE*, *10*, e0135612.

CHAPTER 2:

A Cross-Language Study of Decontextualized Vocabulary Comprehension in Toddlerhood and Kindergarten Readiness

American Psychological Association

Developmental Psychology

0012-1649/18/$12.00

2018, Vol. 54, No. 7, 1317–1333
http://dx.doi.org/10.1037/dev0000514

A Cross-Language Study of Decontextualized Vocabulary Comprehension in Toddlerhood and Kindergarten Readiness

Margaret Friend
San Diego State University

Erin Smolak
San Diego State University and University of California, San Diego

Yushuang Liu
San Diego State University

Diane Poulin-Dubois
Concordia University

Pascal Zesiger
University of Geneva

Recent studies demonstrate that emerging literacy depends on earlier language achievement. Importantly, most extant work focuses on *parent-reported production* prior to 30 months of age. Of interest is whether and how *directly assessed vocabulary comprehension* in the 2nd year of life supports vocabulary and kindergarten readiness in the 4th year. We first contrasted orthogonal indices of parent-reported production and directly assessed vocabulary comprehension and found that comprehension was a stronger predictor of child outcomes. We then assessed prediction from vocabulary comprehension controlling for maternal education, preschool attendance, and child sex. In 3 studies early, decontextualized vocabulary comprehension emerged as a significant predictor of 4th year language and kindergarten readiness accounting for unique variance above demographic control variables. Further we found that the effect of early vocabulary on 4th year kindergarten readiness was not mediated by 4th year vocabulary. This pattern of results emerged in English monolingual children ($N = 48$) and replicated in French monolingual ($N = 58$) and French–English bilingual children ($N = 34$). Our findings suggest that early, decontextualized vocabulary may provide a platform for the establishment of a conceptual system that supports both later vocabulary and kindergarten readiness, including the acquisition of a wide range of concepts including print and number. Differences between parent-reported and directly assessed vocabulary and the mechanisms by which decontextualized vocabulary may contribute to conceptual development are discussed.

Keywords: decontextualized vocabulary, vocabulary size, kindergarten readiness, toddler, preschool

Understanding the origins of early literacy is a priority in psychological and educational research. The perspective that guides the present article is that emerging literacy is dependent upon earlier developing language achievement (e.g., Dickinson, Golinkoff, & Hirsh-Pasek, 2010; NICHD Early Child Care Research Network, 2005; Storch & Whitehurst, 2002). Of interest in the present article is the efficacy of *decontextualized vocabulary comprehension* in the second year of life for predicting kindergarten readiness. Decontextualized vocabulary consists of those word–referent relations that the child recognizes across contexts, contributing to a more stable lexicon and setting the stage for subsequent learning (e.g., Suanda, Mugwanya, & Namy, 2014). The vast majority of the work associating vocabulary with school readiness and reading focuses on children 3 years of age and older (Cristofaro & Tamis-LeMonda, 2011; Forget-Dubois et al., 2009; Kim, Im, & Kwon, 2015; NICHD Early Child Care Research Network, 2005) and the limited research conducted prior to age 3 is based primarily on *parent report of vocabulary production*

This article was published Online First April 5, 2018.

Margaret Friend, Department of Psychology, San Diego State University; Erin Smolak, School of Speech, Language, and Hearing Sciences, San Diego State University, and Doctoral Program in Language and Communicative Disorders, University of California, San Diego; Yushuang Liu, Department of Psychology, San Diego State University; Diane Poulin-Dubois, Department of Psychology, Concordia University; Pascal Zesiger, Department of Psychology and Psycholinguistics, University of Geneva.

This research was supported by NICHD HD468058, NIDCD DC007361, and the NSERC and does not necessarily represent the views of the National Institutes of Health or the Natural Sciences and Engineering Research Council of Canada. We gratefully acknowledge Bianka Enriquez, Tamara Patrucco, Laura Alaria, Monika Rodrigues, Katherine Gittins, and Olivia Kuzyk for assistance in data collection and coding and all of our participant families.

Correspondence concerning this article should be addressed to Margaret Friend, Department of Psychology, San Diego State University, 6505 Alvarado Road, Suite 101, San Diego, CA 92120. E-mail: mfriend@mail.sdsu.edu

(Duff, Reen, Plunkett, & Nation, 2015; Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015).

In Study 1, we contrast decontextualized vocabulary comprehension with parent-reported vocabulary production in monolingual English-speaking children to determine whether decontextualized vocabulary accounts for unique variance in kindergarten readiness and assess the extent to which the effect of early vocabulary is mediated by concurrent vocabulary at preschool age. In Study 2, we assess the generalizability of our findings in an independent sample of monolingual French-speaking children who differ geographically, culturally, and in first language acquisition. In Study 3, we evaluate whether these findings hold in children who are simultaneously acquiring French and English in Montreal, Quebec. This extension to bilingualism is important as the population of children who learn more than one language from birth (bilingual first language acquisition or BFLA) is growing. In Quebec, Canada, 31.7% of the population 5 years or older can hold a conversation in either English or French (Statistics Canada, 2016). Similarly, in California, 43.6% of the population 5 years or older speak a language other than English at home (United States Census Bureau, 2015). Although there are some differences in monolingual and bilingual acquisition, in many ways the mechanisms are quite similar (De Houwer, 2009). In Study 3, we hold language constant with Studies 1 and 2 to assess generalizability across monolingual with bilingual acquisition.

We begin by considering how emergent literacy and kindergarten readiness are related. We then proceed to the current literature on the relation between early vocabulary and literacy. Finally, we discuss decontextualized vocabulary as precursor to deep word knowledge, setting the stage for conceptual development.

Kindergarten Readiness and Emergent Literacy

Emergent literacy is an approach that views literacy as existing on a continuum with preliteracy such that skills that develop early in life are the precursors on which literacy is based (Crain-Thoreson & Dale, 1992; Sulzby & Teale, 1985; Whitehurst & Lonigan, 1998). The focus of the present article is on vocabulary comprehension as a predictor of kindergarten readiness. Kindergarten readiness taps into components on the developmental continuum of emergent literacy (e.g., narrative and recognition of letters and numbers in print). Specifically, narrative skills mediate the relation between early language and emergent literacy in kindergarten (Gardner-Neblett & Iruka, 2015), letter naming contributes to preschool writing (Milburn et al., 2016), and number naming is a strong predictor of later numeracy (Koponen, Aunola, Ahonen, & Nurmi, 2007). The focus of kindergarten readiness is on a specific set of skills that children possess at the point in the continuum that they enter school.

Children embark on this continuum as early as the first 2 years of life. By age 2, children develop phonetic inventories that largely match adult forms (Stoel-Gammon, 1987) and phonological awareness is key to making the link later on between phonological representations and words (Whitehurst & Lonigan, 1998) such that early phonological awareness predicts reading ability (Lonigan, Burgess, Anthony, & Barker, 1998). In typically developing children, parent report of vocabulary production at 24 months predicts language and nonverbal intelligence at 4 years (Marchman & Fernald, 2008; Reilly et al., 2010). Vocabulary production across the second and third year (Forget-Dubois et al., 2009), and comprehension alone at 36 months (Cristofaro & Tamis-LeMonda, 2011), are associated with school readiness. Finally, there is evidence for a direct link between oral language skills at 3 years of age and first grade reading (NICHD Early Child Care Research Network, 2005). Recent findings suggest further that composite vocabulary (comprehension and production) at preschool age is concurrently associated with an array of decoding skills related to reading (Kim et al., 2015). This, in conjunction with previous research showing that reading comprehension is critically dependent on vocabulary comprehension (Cain & Oakhill, 2011; Ouellette, 2006; Quinn, Wagner, Petscher, & Lopez, 2015), suggests that early vocabulary may set the stage for later reading.

Such a general relation between early language and later cognitive development should hold across languages and monolingual and bilingual acquisition. What is less clear is how early in life vocabulary size predicts preschool outcomes. However, few instruments assess vocabulary size prior to third year of life. Two primary approaches, parent-reported vocabulary and directly assessed vocabulary, are discussed in the remainder of this review. Parent report is easily administered and provides a reliable assessment of child vocabulary size relative to their peers. For comparison, directly assessed vocabulary is portable, can be administered with minimal training in a variety of settings (e.g., preschools and well-baby visits) in about 10 min, and provides unique information about children's knowledge of word–world relations outside of the context(s) in which they were learned and are used in daily life. Further, direct assessment can overcome the need for multiple reporters in the bilingual case, in which it is important to obtain caregiver-reported vocabulary from the interlocutor most familiar with the child's use of each of their languages (De Houwer, Bornstein, & Putnick, 2014).

Early Vocabulary and Literacy

Recently, in a study of 300 British infants, Duff, Reen, Plunkett, and Nation (2015) found that vocabulary between 16 and 24 months of age significantly predicted language and reading skills between 4 and 9 years of age. Both parent-reported *comprehension and production* on the Oxford Communicative Development Inventory (OCDI: Hamilton, Plunkett, & Schafer, 2000) contributed to a single latent predictor. From this, Duff et al. (2015) modeled longitudinal relations between early vocabulary and several measures of language and literacy. Early vocabulary accounted for 16% of the variance in school-age vocabulary, 11% of the variance in reading accuracy, and 18% of the variance in reading comprehension. Increases in vocabulary size are thought to lead to more efficient written word identification (Perfetti & Hart, 2001; Perfetti & Stafura, 2014), supporting reading through the ability to infer spelling–sound relations (Mitchell & Brady, 2013). These findings suggest that vocabulary as early as the toddler period predicts language and reading outcomes into the early school years.

Population studies support this interpretation. A recent study of 8,650 children found that parent-reported vocabulary *production* on a modified version of the MacArthur Communicative Development Inventories (Fenson et al., 2007) at age 2 predicted academic achievement (in reading and math) and behavioral performance at kindergarten entry (Morgan et al., 2015). Standardized effect sizes for reading and math were .22 and .27, respectively.

Similarly, in a sample of 6,941 children followed from 5 to 34 years of age, vocabulary *comprehension* at age 5 was a significant predictor of adult literacy when controlling for a broad set of risk factors (including, e.g., maternal education, preschool attendance, and parent–child reading; Law, Rush, Parsons, & Schoon, 2013).

Critically, the Morgan et al. (2015) and Duff et al. (2015) findings support the view that early vocabulary contributes to a positive manifold that increases educational and social opportunities. However, in each case, parent-reported vocabulary predicted a modest proportion of outcome variance (Duff et al., 2015; Morgan et al., 2015). As Law, Rush, Parsons, and Schoon (2013) point out, this leaves two issues unaddressed. First, the practical significance of these findings is uncertain. Relatedly, whereas parent report has dominated much of the work on early vocabulary (with the notable exception of work on speed of word processing; e.g., Marchman & Fernald, 2008), other measures are crucial to replicating and extending these findings (Law & Roy, 2008). For the purposes of this article, we begin to tackle these issues by asking whether decontextualized vocabulary as early as the second year accounts for unique variance in children's kindergarten readiness.

Decontextualized Vocabulary

When parents report on early vocabulary, it is not clear that they discriminate between words that are context-dependent and words that are decontextualized. When parents tell us that their child produces the word "milk," they may mean that the child spontaneously says "milk" in breakfast or snack rituals that are richly contextualized and in which many potential referents are present. They may have associated "milk" with the rituals themselves rather than with the referent in adult usage. Alternatively, the parent may be able to elicit the work "milk" from the child when comprehension of the word–referent relationship is not strong enough to guide spontaneous production. In contrast, Bates, Bretherton, Carlson, Carpen, and Rosser (1979, p. 273) described decontextualized language as "occurring in a broader range of situations with decreasing perceptual support."

Words are decontextualized when the word–referent relation is recognized in the *absence of* the supportive context in which it was learned (e.g., correctly mapping "milk" to a glass of milk rather than to a cookie in an unfamiliar context). The only extant measure for assessing children's decontextualized vocabulary *size* prior to the third year of life is the computerized comprehension task (CCT; Friend & Keplinger, 2008). Other measures designed for this age range (e.g., Fernald, Perfors, & Marchman, 2006; Zimmerman, Steiner, & Pond, 2011), assess processing efficiency for familiar words or attention to sound, respectively. Vocabulary size on the CCT is operationalized as the number of discrete haptic responses to a referent (Friend & Keplinger, 2003, 2008). The task captures the size of the lexicon that is decontextualized and stable: haptic responses are nonrandom and tend to be elicited only when children's understanding of the word–world relation in any particular trial is stable (Friend & Keplinger, 2008; Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015).

Recent research suggests that children may make use of statistical regularities in the repeated pairing of the word and its referent(s) to reach this level of word knowledge (e.g., cross-situational learning; McMurray, Horst, & Samuelson, 2012; Smith & Yu, 2008). Decontextualized word knowledge may reflect domain-general learning processes that support vocabulary acquisition and other kinds of learning (Suanda et al., 2014). For example, both associative learning and hypothesis testing are thought to play a role in the acquisition of new words from infancy through adulthood (Yu & Smith, 2012). Decontextualized vocabulary reflects well-established concepts that are the product of these processes across many pairings of words and referents. This level of word comprehension is essential before children can begin to develop the semantic networks that support what Hadley, Dickinson, Hirsh-Pasek, Golinkoff, and Nesbitt (2015) refer to as "deep word knowledge" (p. 182). Hadley, Dickinson et al. (2015) argue, based on Perfetti's (2007) lexical quality hypothesis, that high-quality word knowledge is crucial to both reading speed and comprehension. High-quality or deep word knowledge can be difficult to define and measure: It can refer to mastery in production, or mastery of word meaning, and most importantly, the extent to which the word is part of a larger network of semantic associations (Schmitt, 2014). In the present article, we consider children's word recognition in the *absence of* the supportive context in which the relation was learned, as measured on the CCT, as evidence of decontextualized word knowledge, an early step on the road to deep word knowledge.

To summarize, extant evidence suggests a link between early language and later language and academic performance, raising the possibility that early vocabulary size is an indicator of learning mechanisms that are crucial to subsequent success in school. Only two large-scale studies connect vocabulary as early as the second year of life to these outcomes (Duff et al., 2015; Morgan et al., 2015) and the practical significance of these findings is unknown. Further, there has been no work on the role of *decontextualized vocabulary* as early as the second year in supporting subsequent development. The broad implications of academic achievement for development across the life course make exploring early prediction of kindergarten readiness imperative since remedying early deficits could have longstanding implications for success in school and beyond.

The Present Research

In three studies, the present article assesses the relation between directly assessed and parent-reported vocabulary in the second-year and kindergarten readiness in the fourth year. In addition, we assess whether vocabulary size in the fourth year mediates this effect. The article expands upon extant literature on the relation between early vocabulary and early literacy by assessing the relation between a direct measure of vocabulary size in the second year of life and subsequent kindergarten readiness. By doing so, it clarifies the contribution of *decontextualized* vocabulary for developmental outcomes.

In Study 1, we assessed the contribution of early vocabulary in the second year to kindergarten readiness using both parent-reported production and directly assessed comprehension. Our choice of measures was predicated on the fact that there are no other measures of vocabulary size for children in the second year and because parent-reported production has been used broadly in the previous literature. Evaluating the relative importance of parent-reported and directly assessed vocabulary allows us to extend the literature by evaluating the relative contribution of these approaches to predicting preschool outcomes.

We expected directly assessed (or *decontextualized*) vocabulary in the second year to yield stronger prediction than parent-reported vocabulary for both kindergarten readiness and concurrent vocabulary comprehension in the fourth year. However, it is important to consider why early vocabulary should predict kindergarten readiness. One possibility is that decontextualized vocabulary is foundational to the establishment of a conceptual system and, as such, it provides the scaffolding that directly supports both later vocabulary and kindergarten readiness. An alternate view is that early vocabulary directly contributes to later vocabulary and only indirectly to the concepts and skills that underlie kindergarten readiness. From this second view, it is later vocabulary that should best reflect the lexicon of concepts related to kindergarten readiness. To test these competing hypotheses, we conducted mediation analyses to determine whether decontextualized vocabulary in the second year predicts kindergarten readiness in the fourth year and whether this effect is mediated by concurrent vocabulary.

This research was conducted under the project, The Path from Language to Literacy. Authors were supported by the NICHD and NIDCD, in the United States, and the Natural Sciences and Engineering Research Council, in Canada. The project was approved by the Institutional Review Boards at San Diego State University under protocol #603057, Concordia University under protocol #UH2003-058 – 6, and the University of Geneva.

Study 1

Method

Participants. Participants were part of a larger longitudinal study and were recruited through parenting magazines, community-based Internet resources, newspapers, daycare, nutrition centers, and state birth records in a large city in the United States. Sixty-eight monolingual English children (35 girls) and their primary caregivers participated when children were 22 months of age (M = 22;28 months, range = 21;6 to 25;12). An additional 10 children were tested but were excluded due to fussiness (N = 1), being a twin (N = 1), or for missing data at one of the waves (N = 9). Forty-nine children returned for a second wave of testing at 48 months of age. Attrition over the 2-year period from first to second testing was 28% (N = 19; 10 of whom had moved out of state). To assess whether the children who did not return were different in any way from the final sample, we calculated mean scores for all demographic variables and predictors. In every case, mean differences between groups were small with overlapping confidence intervals indicating no differences on these measures. The final longitudinal sample was comprised of 48 children (29 girls; M age = 49;15, range = 47;0 to 53;0 months). Of these, 29 had begun attending preschool. Full demographic information on the final sample is presented in Table 1.

Measures.

Language exposure assessment tool (LEAT). Participants' relative exposure to language was assessed via the Language Exposure Assessment Tool (DeAnda, Bosch, Poulin-Dubois, Zesiger, & Friend, 2016). The LEAT is a parent-report measure that takes the form of a systematic interview. The LEAT gathers information on each of the individuals who regularly interact with the child, the languages they speak, whether they are a native speaker, and the number of hours spent talking to/being overheard by the child in each language per day. The LEAT yields the following four quantitative measures of relative exposure for each language to which the child is exposed: hours per day, hours per week, percent exposure, and a parent estimate. Percent exposure to each language is determined by weighting hours of exposure by duration of exposure across the child's life for each interlocutor. As a check, parents provide an independent estimate the percent exposure to each language. Internal consistency for the four estimates of relative exposure on the LEAT is excellent (Cronbach's alpha = .96). Further, LEAT percent relative exposure significantly predicts concurrent vocabulary size in each of bilingual toddlers' languages as measured by parent report (R^2 = .36) and direct experimental measures (R^2 = .22) above and beyond maternal education, age, and parent estimates (DeAnda et al., 2016). Participants were included in the English monolingual sample if their relative exposure to English was 80% or greater at 22 months (M = .98, range = .80 to 1).

Table 1
Distribution of Selected Demographic Characteristics of Participants in Study 1

Number (%) of participants	Female	Male	Total
Maternal education			
High school or less	3 (6.3)	5 (10.4)	8 (16.7)
Some college	9 (18.8)	2 (4.2)	11 (22.9)
College graduate	7 (14.6)	5 (10.4)	12 (25.0)
Post-baccalaureate	10 (20.8)	7 (14.6)	17 (35.4)
Approximate income			
\$18,000–\$40,000	5 (10.4)	2 (4.2)	7 (14.6)
\$41,000–\$60,000	1 (2.1)	5 (10.4)	6 (12.5)
\$61,000–\$80,000	5 (10.4)	0 (.0)	5 (10.4)
\$81,000–\$100,000	11 (22.9)	7 (14.6)	18 (37.5)
>\$100,000	7 (14.6)	5 (10.4)	12 (25.0)
Ethnicity			
Asian	0 (.0)	1 (2.1)	1 (2.1)
Black/not Hispanic	2 (4.2)	0 (.0)	2 (4.2)
Hispanic	8 (16.7)	1 (2.1)	9 (18.8)
White/not Hispanic	14 (29.2)	16 (33.3)	30 (62.4)
Mixed race	5 (10.4)	1 (2.1)	6 (12.5)

Note. Income reported in U.S. dollars.

MacArthur-Bates communicative development inventory: Words and sentences (MCDI). The MCDI (Fenson et al., 2007) is a well-established checklist of 680 items that allows parents to indicate the words that their child currently produces. Vocabulary production on this measure exhibits excellent short-term test–retest reliability (r = .95) and is highly correlated with sentence complexity (r = .80) and grammar (r = .91; Fenson et al., 1994). Vocabulary production was estimated from the MCDI: Words and Sentences Form.

Computerized comprehension task (CCT, Friend & Keplinger, 2003; Friend, Schmitt, & Simpson, 2012). The CCT consists of four training trials, 41 test trials, and 13 reliability trials in opposite left-right orientation. All trials are forced-choice and there are two forms of the task such that each target on one form serves as a foil on the other and vice versa. The assessment is experimenter-controlled with a maximum duration of 7 s per trial and assesses comprehension of nouns, verbs, and adjectives. There are equal numbers of easy (comprehension ≥66%), moderately difficult (comprehension = 33%–66%), and difficult words (comprehension <33%) based on normative data at 16 months of age (Dale & Fenson, 1996). All image pairs presented during training, testing, and reliability were matched for word difficulty (easy, moderately difficult, difficult), part of speech (noun, adjective, verb), category (animal, human, artifact, activity, color, or size), and visual salience (color, size, luminance). For each trial, exemplars are prototypical photographic images of the referent and foil. Finally, inclusion of difficult items from the MCDI results in a high ceiling enabling the CCT to be productively extended to 24 months of age.

The CCT yields test–retest reliability ($r = .70$) and convergent validity ($r = .64$) with parent report and responses on the CCT are nonrandom (Friend & Keplinger, 2008; Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015). This finding has been replicated across languages (Friend & Zesiger, 2011) and in bilinguals (Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2013).

The task was administered in a dimly lit testing room with the toddler seated in a low, cushioned chair or on the parent's lap. Static images were presented in a forced-choice format on a touch sensitive screen positioned 30 cm from the child and 16 cm above the floor at a 60-degree angle. Infants were prompted to touch the named image (e.g., Where's the dog?, Who is running?, Which one is blue?). The member of the forced-choice pair that constituted the target was counterbalanced across participants resulting in two forms of the assessment. The primary experimenter sat next to the child, delivered the prompts, and advanced the assessment. A coder observed the task through a one-way glass and coded responses as correct if the first touch/point was to the target image. A second, reliability coder then coded the videotape for each participant independently. Interrater agreement on CCT responses was 95%. Decontextualized vocabulary comprehension at 22 months was estimated from the CCT.

Peabody picture vocabulary test-III (PPVT; Dunn & Dunn, 1997). The PPVT is a direct measure of vocabulary appropriate for individuals from 30 months of age through adulthood. The test is standardized on a large and representative sample and widely used in the field with reliability and validity coefficients in the .90s. Like the CCT, participants are prompted to identify the pictorial referent associated with a target word. Decontextualized vocabulary comprehension at 48 months was estimated from the PPVT.

The lollipop test: A diagnostic test of school readiness III (Chew, 2007). The lollipop test is a brief, validated measure of kindergarten readiness that assesses knowledge of shapes, numbers, letters, and spatial concepts (Chew & Morris, 1984, 1987; Eno & Woehlke, 1995) and has been used in previous research on predictors of school readiness (Forget-Dubois et al., 2009; Lemelin et al., 2007). Prekindergarten scores are correlated with the developmental indicator for the assessment of learning in prekindergarten children ($r = .71$, $p < .001$; Chew & Morris, 1987) and with the metropolitan readiness test in kindergarten ($r = .76$, $p < .001$; Chew & Morris, 1984). Kindergarten scores explain significant variance in word reading, sentence reading, and math in the first grade (R^2s = .75, .63, .72, *p*s < .001) as well as reading and math in the third grade (R^2s = .58 and .55, *p*s < .001; Chew, 2007). Scores for those children entering kindergarten predict performance on the California Achievement Test at the end of kindergarten and successfully identify those children not prepared to proceed to a regular first-grade program (Eno & Woehlke, 1995).

Procedure. At the 22-month visit, participants completed the CCT while seated in their parents' laps. Parents wore blackout glasses and headphones playing music to reduce the possibility of interference and were instructed to allow their children to respond without assistance. During the same visit, parents completed the MCDI:WS. At the 48-month visit, participants completed the PPVT and the lollipop test as part of a battery of tests that included assessment of executive functioning, narrative ability, and general language abilities.

Data analytic strategy. We first conduct a relative weight analysis to determine the relative importance of parent-reported and decontextualized vocabulary to predicting language and kindergarten readiness. It is possible that the correlation between these two approaches obscures their orthogonal importance to prediction. To assess the hypothesis that decontextualized vocabulary in the second year is a stronger predictor of subsequent kindergarten readiness and concurrent vocabulary comprehension in the fourth year, we conduct two relative importance analyses (Johnson, 2000). In relative importance analysis, orthogonal variables are created to determine how much unique variance in the dependent measure is explained by each predictor. Multiple regression weights may fail to accurately capture this when predictors are correlated.

Next, in order to determine whether the relation between second-year vocabulary and fourth-year kindergarten readiness is mediated by concurrent vocabulary, we test each of the causal steps to establish mediation (Judd & Kenny, 1981; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). These steps constitute three regression models that test the three necessary conditions for mediation: (a) the independent variable (second-year vocabulary) affects the outcome variable (kindergarten readiness); (b) the independent variable affects the mediator variable (fourth-year vocabulary) and the mediator variable affects the outcome variable controlling for the effect of the independent variable; and (c) the independent variable does not affect the outcome variable when controlling for the mediator variable. As a further check on mediation, we employ Sobel's test, which is a *t* test of the significance of the indirect effect (Sobel, 1982).

Results

22 months. Average relative exposure to English according to the LEAT was .99 (range = .89 to 1.0). MCDI production vocabulary ranged from five to 633 words ($M = 256.94$, $SD = 168.18$), corresponding to the first to the 99th percentile. Sample-specific internal consistency was $\alpha = .99$. Two children had expressive vocabulary scores at the first percentile whereas their CCT comprehension scores corresponded to the 20th to 30th deciles for the sample suggesting that they knew more words than they were producing. Visual inspection of the scatterplot of MCDI and CCT scores suggested that, although they performed more poorly than their peers, they were not outliers. CCT comprehension vocabulary ranged from 10 to 39 words ($M = 27.44$, $SD = 6.95$). Sample-specific internal consistency was excellent across forms (Cronbach's alpha = .86 and .93, respectively). Immediate test–retest reliability was good, $r(42) = .65$, $p < .001$, and CCT scores correlated significantly with concurrent MCDI production, $r(47) = .44$, $p = .002$.

48 months. Comprehension vocabulary on the PPVT ranged from 31 to 106 words ($M = 72.23$, $SD = 15.48$), reflecting a range from the 13th to the 99th percentile. Sample-specific internal consistency was $\alpha = .94$. Lollipop test scores ranged from 14 to 65 ($M = 50.54$, $SD = 10.87$), corresponding to roughly the seventh to the 99th percentile for children entering kindergarten (Chew &

Morris, 1987) and sample-specific internal consistency was moderate (α = .69).

Zero-order correlations revealed significant longitudinal relations between vocabulary at 22 months and later vocabulary and kindergarten readiness scores ($r(46)$ = .41, p = .004) and $r(46)$ = .51, $p < .001$, respectively). Parent-reported vocabulary was correlated with kindergarten readiness at 48 months ($r(46)$ = .37, p = .011) but not with vocabulary at 48 months (p = .253). Maternal education was significantly associated with decontextualized vocabulary at 22 months, $r(46)$ = .37, p = .010, and marginally associated with kindergarten readiness at 48 months, $r(46)$ = .27, p = .063, but not with any other variable (all ps > .18). Sex was marginally correlated with decontextualized vocabulary at 22 months, $r(46)$ = .25, p = .087, and kindergarten readiness at 48 months, $r(46)$ = .27, p = .064, but not with any other variable (all ps > .15). Finally, number of hours in preschool was not associated with either predictor or outcome variables (all ps > .6).

Relative weights analysis. Each relative importance analysis included the decontextualized and parent-reported vocabulary as predictors and child sex and maternal education as control variables. All relative importance analyses were conducted in R (R Development Core Team, 2008) using the script provided by Tonidandel and LeBreton (2014). First, orthogonal raw correlations and 95% confidence intervals were computed for sex, maternal education, decontextualized vocabulary, and parent-reported vocabulary. These reflect the raw correlations between a set of orthogonal predictors (that are maximally similar to the original predictors) and the outcome. Following this, rescaled weights representing the proportion of total variance accounted for by each predictor were calculated. Each weight was then statistically tested using a bootstrap procedure of 10,000 iterations with replacement (Tonidandel, LeBreton, & Johnson, 2009). Finally, we tested whether the weights for decontextualized and parent-reported vocabulary were significantly different. Nominal alpha was set at .05, however, it is important to note that the relative weights procedure is overly conservative with regard to Type I error rates (Tonidandel et al., 2009). For this reason, application of a family wise correction would be inappropriate in conjunction with this procedure.

The first analysis examined the relative contribution of each independent variable in predicting kindergarten readiness scores at 48 months. Table 2 displays the raw and rescaled relative weights for each independent variable in predicting kindergarten readiness with 95% confidence intervals. In total, the predictors and control variables accounted for 19% of the variance in kindergarten readiness. The largest (and only statistically significant) weight was for decontextualized vocabulary, which explained 13% of the variance in kindergarten readiness and 69% of the total variance explained by all four predictors. However, the confidence intervals for decontextualized and parent-reported vocabulary overlap, indicating that their relative weights are not significantly different.

This analysis was repeated using vocabulary at 48 months as the outcome. Table 3 displays raw and rescaled weights for each independent variable with 95% confidence intervals. In total, the predictors and control variables accounted for 22% of the variance in vocabulary at 48 months. Similarly to the results for kindergarten readiness, the largest (and only statistically significant) weight was for the decontextualized vocabulary at 22 months, which explained 16% of the variance in vocabulary at 48 months and 71% of the total variance explained by all four predictors. As in the previous analysis, the confidence intervals for the predictors overlap, indicating that their relative weights are not significantly different. In summary, decontextualized vocabulary at 22 months was the largest and the only statistically significant relative weight in predicting both vocabulary and kindergarten readiness at 48 months, however, due to the large confidence intervals for both predictors, these weights were not significantly different.

These relative importance analyses identify decontextualized vocabulary as a stronger predictor, relative to parent-reported vocabulary, of both kindergarten readiness and subsequent vocabulary size at 48 months. We now turn our attention to how decontextualized vocabulary is related over time and to kindergarten readiness. Specifically, does decontextualized vocabulary as early as the second year predict decontextualized vocabulary in the fourth year, and how does vocabulary, in turn, support kindergarten readiness?

Early vocabulary and kindergarten readiness. To test the hypothesis that decontextualized vocabulary comprehension in the second year would predict kindergarten readiness in the fourth year, we conducted a hierarchical regression with the effects of sex and maternal education controlled on the first step with decontextualized vocabulary at 22 months entered on the second step and kindergarten readiness scores as the dependent measure. Nominal alpha was set at .05 and a sequential Bonferroni procedure was applied to control family wise false discovery rate (Benjamini & Hochberg, 1995). The first model revealed that maternal education and sex were not significant predictors. The inclusion of decontextualized vocabulary led to a significant increment in R^2 and accounted for significant unique variance in children's kindergarten readiness scores (η^2 = .29). As expected, decontextualized vocabulary predicted subsequent kindergarten readiness above variance accounted for by control variables. Regression parameters for the models described above are presented in Table 4.

To test the hypothesis that decontextualized vocabulary in the second year predicts vocabulary in the fourth year, we conducted a second hierarchical regression in which sex and maternal education were entered as control variables on the first step and second-year decontextualized vocabulary was entered on the second step with fourth-year vocabulary as the dependent measure. Only the second model was significant indicating the expected prediction from the second- to the fourth-year vocabulary (η^2 = .22; see Table 5 for full regression parameters).

Table 2
Relative Importance Analysis Predicting Fourth Year Kindergarten Readiness for Study 1

Variables	Relative weights		95% Confidence interval	
	Raw	Rescaled	Lower bound	Upper bound
Child sex	.029	15.284	<.001	.156
Maternal education	.003	1.333	<.001	.017
Second-year CCT comprehension	.132*	69.360	.018	.323
Second-year MCDI production	.027	14.023	.002	.121

* Significant at $\alpha < .05$ (95% CI).

Table 3
Relative Importance Analysis Predicting Fourth Year Vocabulary for Study 1

Variables	Relative weights: Raw	Relative weights: Rescaled	95% Confidence interval: Lower bound	95% Confidence interval: Upper bound
Child sex	.036	15.835	.002	.156
Maternal education	.006	2.642	<.001	.015
Second-year CCT comprehension	.161*	71.306	.037	.374
Second-year MCDI production	.023	10.216	.002	.121

* Significant at $\alpha < .05$ (95% CI).

To complete testing of the causal requirements for mediation, we conducted a final hierarchical regression. On the first step, we controlled for maternal education and sex. We entered fourth-year vocabulary scores on the second step and second-year vocabulary scores on the third step with kindergarten readiness scores as the dependent measure. If the relation between the second-year vocabulary and kindergarten readiness is mediated by fourth-year vocabulary, we would expect fourth-year, but not second-year, vocabulary to be significant in the final model. The first model was significant and neither maternal education nor sex emerged as significant predictors. The inclusion of fourth-year vocabulary scores on the second step led to a significant increment in R^2 ($\eta^2 = .31$); however, when CCT scores were included on the third step, no single predictor reached significance although the model itself was significant ($\eta^2 = .36$). Observed power for the final model was estimated using G* Power (Faul, Erdfelder, Buchner, & Lang, 2009) at .90. In all models, collinearity was assessed using SPSS diagnostics. Tolerance was consistently high across predictors indicating low, acceptable levels of shared variance. Regression parameters are presented in Table 6.

As a final test for mediation, we entered the parameters in the second (second-year vocabulary predicting fourth-year vocabulary) and third (fourth-year vocabulary predicting kindergarten readiness, controlling for second-year vocabulary) regressions into a Sobel test (Sobel, 1982). The result was not significant ($z' = 1.87$, $SE = .13$, $p = .062$), consistent with the prior regression test for mediation: fourth-year vocabulary did not significantly predict kindergarten readiness when controlling for demographic variables and second-year vocabulary. These findings suggest that the effect of early, decontextualized vocabulary on kindergarten readiness is not mediated by concurrent vocabulary.

In sum, these models indicate that decontextualized vocabulary in the second year is a significant predictor of both vocabulary and kindergarten readiness in the fourth year. Further, the effect of early vocabulary on kindergarten readiness was not mediated by concurrent vocabulary comprehension.

Table 4
Regression Parameters for Models Predicting Fourth Year Kindergarten Readiness From Second Year Vocabulary for Study 1 (N = 48)

Full model	Model 1: R^2	Model 1: *SE*	Model 1: β	Model 1: *p*	Model 2: R^2	Model 2: *SE*	Model 2: β	Model 2: *p*
	.10			.035	**.24**			**.002**
Measure								
Maternal education		.69	.26	.073		.68	.10	.460
Child sex		3.05	.26	.071		2.89	.16	.238
Second-year CCT comprehension						.22	.43	**.004**

Note. R^2 is adjusted. Significant *p* values after family-wise error correction are bolded.

Table 5
Regression Parameters for Models Predicting Fourth Year Vocabulary From Second Year Vocabulary for Study 1 (N = 48)

Full model	Model 1: R^2	Model 1: *SE*	Model 1: β	Model 1: *p*	Model 2: R^2	Model 2: *SE*	Model 2: β	Model 2: *p*
	−.02			.567	.17			**.010**
Measure								
Maternal education		1.05	.10	.496		1.01	−.07	.608
Child sex		4.62	−.13	.393		4.30	−.24	.087
Second-year CCT comprehension						.33	.49	**.002**

Note. R^2 is adjusted. Significant *p* values after family-wise error correction are bolded.

Discussion

Parent-reported vocabulary in the second year was significantly correlated with kindergarten readiness consistent with Duff et al. (2015) and Morgan, Farkas, Hillemeier, Hammer, and Maczuga (2015). Importantly, however, decontextualized vocabulary emerged as the only significant predictor of both kindergarten readiness and concurrent vocabulary in the fourth year in the relative importance analyses. These results support previous findings indicating that vocabulary knowledge in the toddler period is associated with later school-related outcomes but extending Duff et al. (2015) we find in the relative importance analyses that decontextualized vocabulary is a stronger predictor of kindergarten readiness than is parent-reported vocabulary.

Table 6
Regression Parameters for Models Predicting Fourth-Year Kindergarten Readiness From Second- and Fourth-Year Vocabulary for Study 1 (N = 48)

Full model	Model 2: R^2	Model 2: *SE*	Model 2: β	Model 2: *p*	Model 3: R^2	Model 3: *SE*	Model 3: β	Model 3: *p*
	.28			**.001**	.30			.001
Measure								
Maternal education		.62	.21	.099		.65	.13	.347
Child sex		2.77	.31	.018		2.86	.23	.082
Fourth-year PPVT comprehension		.09	.42	**.002**		.10	.31	.030
Second-year CCT comprehension						.24	.28	.073

Note. R^2 is adjusted. Significant *p* values after family-wise error correction are bolded.

The CCT and MCDI capture both unique and overlapping aspects of early vocabulary given shared variance but differential prediction to subsequent vocabulary and kindergarten readiness scores. Indeed, the measures differ in two important ways in the present study. First, we assessed vocabulary comprehension on the CCT whereas we assessed production on the MCDI. In the Morgan et al. (2015) population study, MCDI production was associated with subsequent reading and math achievement. Thus, we contrasted directly assessed vocabulary comprehension with parent-reported production, a known predictor of developmental outcomes. Nevertheless, only decontextualized vocabulary accounted for significant variance, which brings us to the second difference between the CCT and the MCDI.

The CCT assesses decontextualized word knowledge: Children must respond to a word–referent relation without the context in which the referent is usually encountered. One can think of children's decontextualized word knowledge as essential to building the "deep" knowledge that is crucial to conceptual development generally and to reading skill specifically (Hadley et al., 2015). As we argued earlier, parents likely report this decontextualized knowledge on the MCDI as well but also report words for which children show evidence of context-dependent, or partial knowledge. From this perspective, the CCT provides an estimate of early word knowledge based on those words for which children have a sufficiently robust representation to execute a response in a decontextualized setting whereas MCDI scores inform us about where children are, relative to same-age peers, in making a wide variety of word–world associations. Our findings suggest that early vocabulary is foundational to the establishment of a conceptual system and, as such, it provides the scaffolding that directly supports both later vocabulary and kindergarten readiness, including the acquisition of a wide range of concepts including print and number. To assess the generalizability of these findings, we tested two additional samples that differed geographically and linguistically from the sample in Study 1.

Study 2

Method

Participants. Participants were part of a larger longitudinal study and were recruited via birth records in a large French-speaking city in Switzerland. Sixty-three monolingual French children (31 girls) and their primary caregivers participated in data collection when children were 22 months of age (M = 21;29 months, range = 21 to 22;6 months). An additional three children were tested but excluded due to missing data at one of the waves. Fifty-eight children returned for a third wave of testing at 48 months of age (M = 47;25 months, range = 47;3 to 49;3 months). Seven of these children had begun attending preschool. Attrition between testing occasions was 8%. Additional demographic information is presented in Table 7.

Table 7
Distribution of Selected Demographic Characteristics of Participants in Study 2

Number (%) of participants	Male	Female	Total
Maternal education			
High school or less	6 (10.3)	8 (13.8)	14 (24.1)
Some college	9 (15.5)	4 (6.9)	13 (22.4)
College graduate	2 (3.4)	3 (5.2)	5 (8.6)
Post-baccalaureate	12 (20.7)	14 (24.1)	26 (44.8)
Approximate income			
$18,000–$40,000	0 (.0)	2 (5.9)	2 (5.9)
$41,000–$60,000	1 (3.0)	1 (3.0)	2 (5.9)
$61,000–$80,000	3 (8.8)	2 (5.9)	5 (14.7)
$81,000–$100,000	2 (5.9)	2 (5.9)	4 (11.8)
>$100,000	10 (29.4)	11 (32.3)	21 (61.7)
Ethnicity			
Black/not Hispanic	1 (1.7)	1 (1.7)	2 (3.4)
Hispanic	0 (0)	1 (1.7)	1 (1.7)
White/not Hispanic	28 (48.3)	27 (46.6)	55 (94.8)

Note. Income reported in Swiss Francs (CHF). Median family income at time of data collection was approximately 126,000 CHF. 24 participants declined to provide income information. Some percentages may not sum to 100 due to rounding error.

Measures.

Language exposure assessment tool (LEAT). Identical to Study 1. Participants were included in the French monolingual sample if their relative exposure to French was 80% or greater at 22 months (M = .96, range = .80 to 1).

L'inventaire français du développement communicatif (IFDC). The IFDC (Kern, 2007; Kern & Gayraud, 2010) is the European French adaptation of the MCDI. The IFDC has been shown to be sensitive to vocabulary development over time (Kern, 2010) and to have strong short-term test–retest reliability (r > .90; Kern & Gayraud, 2010). The IFDC: Mots et Phrases (Words and Sentences) consists of 698 vocabulary items and can be administered from 16 to 30 months of age. Vocabulary production was estimated from the IFDC Mots et Phrases Form.

CCT. The French CCT is an adaptation of the English CCT described in Study 1. The design and administration are the same with items on the French adaptation based on comprehension norms from the IFDC. The French adaptation evinces significant test–retest reliability (r = .42) and convergent validity (r = .54) with parent report on the IFDC (Friend & Zesiger, 2011). As many translation equivalents (words across languages with the same referential meaning) as possible were maintained in the adaptation from English to French with the restriction that the French, like the English, contained one third each easy, moderately difficult, and difficult items (based on parent-report norms from the IFDC). Similarly, the proportion of nouns, verbs, and adjectives was roughly equivalent. Images were prototypical exemplars in the region in which children were tested.

Échelle de vocabulaire en images Peabody (EVIP). The EVIP (Dunn, Dunn, & Thériault-Whalen, 1993) is the French adaptation of the English PPVT. Like the PPVT, it is a measure of receptive vocabulary from 30 months of age to adulthood. The EVIP was normed on a large representative sample of French speakers. It has satisfactory test–retest reliability (r = .72), as well as internal consistency (α = .81). Validity of the French version has also been established through high correlations with other vocabulary tests (r = .86) and with IQ measures (r = .62-.72; Dunn et al., 1993).

Lollipop test. The Lollipop test is identical to Study 1 and was administered in French.

Procedure and data analytic strategy. Identical to Study 1.

Results

22 months. Average relative exposure to French according to the LEAT was .95 (range = .80 to 1.0). IFDC production vocabulary ranged from 13 to 523 words (M = 201.66, SD = 133.61), corresponding to the fifth to the 90th percentile and sample-specific internal consistency was excellent (α = .99). CCT comprehension vocabulary ranged from 14 to 40 words (M = 29.67, SD = 5.58). Sample-specific internal consistency was excellent across forms (α = .92 and .91, respectively). Immediate test–retest reliability was significant, $r(51) = .67$, $p < .001$, and CCT scores correlated with concurrent IFDC production, $r(56) = .41$, p = .001.

48 months. Comprehension vocabulary on the EVIP ranged from 27 to 92 words (M = 55.17, SD = 14.83), reflecting a range from the 16th to the 99th percentile and sample-specific internal consistency was excellent (α = .97). Lollipop test scores ranged from 14 to 67 (M = 38.79, SD = 13.78). This range is almost identical to that observed in the English sample in Study 1. Sample-specific internal consistency was good (α = .78).

Early vocabulary and kindergarten readiness. Zero-order correlations revealed significant longitudinal relations between vocabulary at 22 months and later vocabulary and kindergarten readiness scores and kindergarten readiness scores at 48 months, $r(56) = .54$, $p < .001$, and $r(56) = .40$, $p = .002$, respectively. For Study 2, we used a dichotomous measure of preschool attendance because, unlike the English sample, all children who attended preschool attended for the same number of hours. Maternal education was associated with attendance at preschool, $r(58) = .31$, $p = .020$, but not with any other measure (all ps > .27). Sex was marginally correlated with attendance at preschool, $r(56) = .27$, $p = .045$, but not with any other measure (all ps > .32). In large part, these correlations mirror the pattern observed in the English sample in Study 1.

Relative weights analysis. Following the logic of Study 1, we first examined the relative contribution of decontextualized and parent-reported vocabulary in predicting fourth-year kindergarten readiness scores. Table 8 displays the raw and rescaled relative weights for each independent variable along with the 95% confidence intervals. In total, the predictors and control variables accounted for 30% of the variance in kindergarten readiness scores. The largest (and only statistically significant) weight was for the decontextualized vocabulary, which explained 13% of the variance in kindergarten readiness scores and 44% of the total variance explained by all four predictors. However, the confidence intervals for the predictors overlap, indicating that their relative weights are not significantly different.

Table 8
Relative Importance Analysis Predicting Fourth-Year Kindergarten Readiness for Study 2

	Relative weights		95% Confidence interval	
Variables	Raw	Rescaled	Lower bound	Upper bound
Child sex	.045	15.175	.002	.178
Maternal education	.040	13.384	.002	.157
Second-year CCT comprehension	.130*	44.012	.030	.278
Second-year MCDI production	.081	27.428	.009	.206

* Significant at α < .05 (95% CI)

This analysis was repeated using fourth-year vocabulary as the outcome variable. Table 9 displays the raw and rescaled weights for each independent variable, along with the 95% confidence intervals. In total, the predictors and control variables accounted for 35% of the variance in fourth-year vocabulary. The largest (and only statistically significant) weight was for decontextualized vocabulary, which explained 31% of the variance in fourth-year vocabulary and 88% of the total variance explained by all four predictors. In contrast to the previous analysis, the confidence intervals for the predictors did not overlap, indicating that the relative weight of decontextualized vocabulary was significantly greater than the relative weight of the parent-reported vocabulary. In summary, decontextualized vocabulary was the largest and the only statistically significant relative weight in predicting both vocabulary and kindergarten readiness scores in the fourth year, however, due to the large confidence intervals for both predictors, these weights were only significantly different when predicting fourth-year vocabulary.

To test the mediation model put forward in Study 1, we conducted parallel hierarchical regressions on the French data. Nominal alpha was set at .05 and a sequential Bonferroni procedure was applied to control family wise false discovery rate (Benjamini & Hochberg, 1995). In all models, collinearity was assessed using SPSS diagnostics. Tolerance was consistently high across predictors indicating low, acceptable levels of shared variance. In the first regression, the effects of sex, maternal education, and attendance at preschool were controlled on the first step. Decontextualized vocabulary in the second year was entered on the second step with kindergarten readiness scores as the dependent measure. The first model was significant with attendance at preschool the only factor approaching significance and the inclusion of vocabulary in second year in the second step accounted for significant unique variance in the model (η^2 = .26). Observed power for the final model was estimated using G* Power (Faul et al., 2009) at .87. This provides further support for the hypothesis that decontextualized vocabulary predicts subsequent kindergarten readiness. Regression parameters for the models described above are presented in Table 10.

As in Study 1, we conducted a second hierarchical regression in which maternal education, sex, and preschool attendance were entered as control variables on the first step, Decontextualized vocabulary scores in the second year were entered on the second step, and vocabulary scores in the fourth year served as the dependent measure. Only the second model was significant (η^2 = .32) indicating the expected relation between vocabulary in the second and fourth years (see Table 11 for full regression parameters). Consistent with the English data, second-year vocabulary predicted both kindergarten readiness and fourth-year vocabulary.

To determine whether the relation between second-year vocabulary and fourth-year kindergarten readiness was mediated by concurrent vocabulary in the fourth year, we conducted a final hierarchical regression. On the first step, we controlled for mater-

Table 9
Relative Importance Analysis Predicting Fourth-Year Vocabulary for Study 2

Variables	Relative weights: Raw	Relative weights: Rescaled	95% Confidence interval: Lower bound	95% Confidence interval: Upper bound
Child sex	.002	.446	<.001	.004
Maternal education	.024	6.814	<.001	.131
Second-year CCT comprehension	.306*[a]	88.207	.159	.477
Second-year MCDI production	.016[a]	4.533	.001	.041

[a] Indicates significant weight difference at the $\alpha < .05$ level.
* Significant at $\alpha < .05$ (95% CI).

nal education, sex, and preschool attendance. We entered fourth-year vocabulary scores on the second step and second-year vocabulary scores on the third step with kindergarten readiness scores as the dependent measure. If the relation between second-year vocabulary and kindergarten readiness is mediated by fourth-year vocabulary, we would expect the fourth-year, but not second-year, vocabulary to be significant in the final model. The first model was significant and preschool attendance was the only predictor to approach significance. The inclusion of fourth-year vocabulary on the second step lead to a marginally significant increment in R^2 ($\eta^2 = .23$) but when second-year vocabulary scores were included on the third step, neither fourth-year, nor second-year, vocabulary scores reached significance even though the model was significant ($\eta^2 = .28$). Regression parameters are presented in Table 12. Interestingly, preschool attendance contributed modestly to the models although it was not significant once the sequential Bonferroni correction was applied.

Finally, to further test our mediation model, we entered the appropriate parameters into a Sobel test. As in Study 1, the test was not significant ($z' = 1.14$, $SE = .20$, $p = .255$), suggesting the effect of decontextualized vocabulary on kindergarten readiness was not mediated by concurrent vocabulary.

Discussion

In contrast to Study 1, Duff et al. (2015), and Morgan et al. (2015), we did not find evidence that parent-reported vocabulary production was significantly correlated with kindergarten readiness. In addition, as in Study 1, decontextualized vocabulary was a stronger predictor, relative to parent-reported vocabulary, of both kindergarten readiness and subsequent vocabulary in our relative importance analyses. Thus, the general finding that decontextualized vocabulary is a stronger predictor of kindergarten readiness than is parent-reported vocabulary was replicated. It is not clear why parent-reported production in the French-speaking sample failed to predict kindergarten readiness but it is possible that parents in the U.S. and Switzerland have different expectations for what constitutes production and, consequently, complete the MCDI and IFDC forms differently. This illustrates the fact that language adaptations of instruments cannot necessarily be expected to capture vocabulary estimates in the same way due to the potential for cultural differences in reporting practices. Although these differences in prediction could result from differences in kindergarten readiness, the fact that directly assessed vocabulary was a strong predictor of readiness across languages and cultures argues against this interpretation.

Consistent across both English and French monolingual children is the fact that the relation between decontextualized vocabulary in the second year and children's kindergarten readiness in the fourth year is not mediated by fourth-year vocabulary comprehension.

Table 10
Regression Parameters for Models Predicting Fourth-Year Kindergarten Readiness From Second-Year Vocabulary for Study 2 (N = 58)

Full model	Model 1 R^2	Model 1 *SE*	Model 1 β	Model 1 *p*	Model 2 R^2	Model 2 *SE*	Model 2 β	Model 2 *p*
	.10			**.038**	.20			**.003**
Measure								
Maternal education		.65	−.04	.797		.62	−.01	.923
Child sex		3.60	.13	.335		3.40	.10	.444
Preschool		5.80	.33	.020		5.49	.28	.036
Second-year CCT comprehension						.30	.34	**.006**

Note. R^2 is adjusted. Significant *p* values after family-wise error correction are bolded.

Table 11
Regression Parameters for Models Predicting Fourth-Year Vocabulary From Second-Year Vocabulary for Study 2 (N = 60)

Full model	Model 1 R^2	Model 1 *SE*	Model 1 β	Model 1 *p*	Model 2 R^2	Model 2 *SE*	Model 2 β	Model 2 *p*
	−.02			.628	.27			**<.001**
Measure								
Maternal education		.75	.12	.401		.63	.16	.200
Child sex		4.12	.02	.876		3.50	−.03	.803
Preschool		6.63	.09	.527		5.65	.02	.898
Second-year CCT comprehension						.31	.55	**<.001**

Note. R^2 is adjusted. Significant *p* values after family-wise error correction are bolded.

Table 12
Regression Parameters for Models Predicting Fourth-Year Kindergarten Readiness From Second- and Fourth-Year Vocabulary for Study 2 (N = 60)

Full model	Model 1 R^2	Model 1 *SE*	Model 1 β	Model 1 *p*	Model 2 R^2	Model 2 *SE*	Model 2 β	Model 2 *p*
	.18			**.006**	.21			**.004**
Measure								
Maternal education		.63	−.07	.581		.62	−.04	.765
Child sex		3.44	.12	.338		3.39	.10	.420
Preschool		5.55	.30	.026		5.48	.28	.037
Fourth-year EVIP comprehension		.11	.31	.016		.13	.17	.247
Second-year CCT comprehension						.35	.25	.085

Note. R^2 is adjusted. Significant *p* values after family-wise error correction are bolded.

Thus, as we argued in Study 1, it appears that early, decontextualized vocabulary may scaffold both later vocabulary and kindergarten readiness. However, it should be noted that these findings are restricted to children learning only one language from birth when the majority of children in the world learn two or more languages. To better understand whether the observed relation between early vocabulary and subsequent kindergarten readiness holds for children exposed to more than one language, we conducted a study with children who acquired two languages from birth. To follow as directly as possible from Studies 1 and 2, we recruited children who acquired both French and English as their first languages.

Because we were interested in extending our findings from Studies 1 and 2 on the relation between early vocabulary and kindergarten readiness to BFLA children, we were particularly interested in how children's vocabulary in their dominant language in toddlerhood was related to their language and kindergarten readiness skills in the fourth year.

Study 3

Method

Participants. Participants were part of a larger longitudinal study and were recruited via birth records in a large bilingual French-English city in Québec, Canada. Forty-eight French-English bilingual children (19 girls) and their primary caregivers participated when children were 22 months of age (M = 23;24 months, range = 20;24 to 26;6 months for the English visit and M = 23;26 range = 20;27 to 26;6 for the French visit). An additional 24 children were tested but were excluded due to fussiness (N = 7) or missing data at one of the waves (N = 17). Attrition between testing occasions was 29% (N = 14, eight of whom had moved out of the area). Thirty-four children returned for a third wave of testing at 48 months of age (M = 48;26, range = 46;15 to 51;12 months for the English visit and M = 49, range = 47;15 to 52;3 for the French visit). Of these 34 children, all were enrolled in daycare. Additional demographic information is presented in Table 13.

Measures.

Language exposure assessment tool (LEAT). Identical to Studies 1 and 2. Participants who were included in the bilingual sample were exposed to French and English from birth and their exposure to each language was 80% or less and exposure to the least-exposed language was at least 20%.

Inventaire MacArthur du développement de la communication (IMDC). The IMDC (Frank, Poulin-Dubois, & Trudeau, 1997) is the Quebec French adaptation of the MCDI. In contrast to the IFDC, the items were normed on children acquiring Québécois French. Thus, there is overlap between the IFDC and the IMDC but the IMDC is more appropriately targeted to the present sample. The Mots et Énoncés (Words and Sentences) Form is designed for use for children from 16 to 30 months of age. Vocabulary production on IMDC exhibits excellent short-term test–retest reliability (r = .90) and is highly correlated with sentence complexity (r = .78) and grammar (r = .70; Trudeau, Aktouf, Boudreault, & Breault, 2008).

Results have confirmed the validity of the IMDC and a significant correlation between the IMDC and chronological age suggests the IMDC is sensitive to language growth (Boudreault, Cabirol, Trudeau, Poulin-Dubois, & Sutton, 2007).

MCDI. The MCDI is identical to Study 1.

CCT. The English CCT and the French adaptation were utilized. The instruments are identical to those described in Studies 1 and 2.

PPVT. The PPVT is identical to Study 1.

EVIP. The EVIP is identical to Study 2.

Lollipop test. The lollipop test is identical to Studies 1 and 2 and was administered in the child's dominant language.

Procedure. All children were assessed on two different days, once in English and once in French. The language of first and second visit was counterbalanced across children. During visits, the experimenter instructed the parent to speak only the language of that visit (i.e., English *or* French) and assessments were completed in that language to elicit optimal performance. The two appointments were scheduled approximately 2 weeks apart (M interval = 16 days, range = 6 to 43 days). Language of testing was counterbalanced across participants such that half of the children completed the first visit in their dominant language and the second in their non-dominant language. Otherwise the procedure was identical to Studies 1 and 2.

Data analytic strategy. The strategy is similar to Study 1 although a reduction in sample size required us to use a correlational approach instead of performing regressions. Due to this, in Study 3 we test mediation using only the causal steps tests (Judd & Kenny, 1981; MacKinnon et al., 2002).

Results

Language dominance. In BLFA, both languages are acquired simultaneously from birth. Here we refer to L1 and L2 as a function of relative exposure. Therefore, L1 refers to the language

Table 13
Distribution of Selected Demographic Characteristics of Participants in Study 3

Number (%) of participants	Male	Female	Total
Maternal education			
High school or less	1 (3.0)	2 (6.1)	3 (9.1)
Some college	4 (12.1)	2 (6.1)	6 (18.2)
College graduate	8 (24.2)	2 (6.1)	10 (30.3)
Post-baccalaureate	7 (21.2)	7 (21.2)	14 (42.4)
Approximate income			
\$18,000–\$40,000	2 (6.5)	4 (12.9)	6 (19.4)
\$41,000–\$60,000	3 (9.7)	0 (0)	3 (9.7)
\$61,000–\$80,000	5 (16.1)	2 (6.5)	7 (22.6)
\$81,000–\$100,000	2 (6.5)	3 (9.7)	5 (16.1)
>\$100,000	7 (22.6)	3 (9.7)	10 (32.3)
Ethnicity			
Asian	1 (3.0)	0 (0)	1 (3.0)
Black/not Hispanic	1 (3.0)	3 (9.1)	4 (12.1)
Hispanic	0 (0)	1 (3.0)	1 (3.0)
White/not Hispanic	15 (45.5)	9 (27.3)	24 (72.8)
Mixed race	1 (3.0)	0 (0)	1 (3.0)
Other	1 (3.0)	1 (3.0)	2 (6.1)

Note. Income reported in Canadian dollars (CAN). Median income at the time of the study was approx. 74,000 CAN. One participant declined to provide demographic information and two participants declined to provide income information. Some percentages may not sum to 100 due to rounding error.

with the greatest relative exposure and L2 refers to the language with less exposure rather than to which language was acquired first or second. According to the LEAT, at 22 months of age, 24 children had greater exposure to English, and 24 had greater exposure to French. Average relative exposure to L1 was 65% (range = 42% to 79%), whereas relative exposure to L2 was 34% (range = 21% to 61%). Four children were also exposed to a third language (*M* exposure = .12, range = .02 to .26). In these cases, we collapsed across the two nondominant languages to obtain an estimate of nondominant exposure. The average L1/L2 exposure ratio was 2.04 (*SD* = .89), reflecting that children received roughly twice as much exposure to the dominant, relative to nondominant, languages. Parent-reported L1/L2 exposure was higher at 48 months (*M* L1/L2 = 3.08, *SD* = 5.43).

There was evidence of a change in relative exposure in some participants from 22 to 48 months: Five children who received greater French exposure at 22 months had greater exposure to English at 48 months, and six children who had greater English exposure at 22 months had greater exposure to French at 48 months. There was no change in relative language exposure for 23 children. Children whose relative exposure to French was higher than to English at 22 months also had a larger French vocabulary at 48 months. The same was true for children whose relative exposure was higher to English than to French. Of the 34 children who participated at 48 months, a majority ($N = 19$) had larger vocabularies in English than in French and the rest ($N = 15$) had larger vocabularies in French. Whereas exposure is a good predictor of vocabulary size in the toddler period, this relation did not hold at 48 months of age. Language exposure estimates at 48 months were uncorrelated with vocabulary size ($ps \geq .171$) consistent with the literature on dominance as children approach school age (Bedore et al., 2012; Lust et al., 2016; Sheng, Lu, & Gollan, 2014, but cf. Brebner, McCormack, & Liow, 2016). All children were included in the analyses.

22 months. Parent-reported vocabulary production ranged from three to 446 words ($M = 186.6$, $SD = 120.5$) in French and from five to 643 words ($M = 171.0$, $SD = 162.1$) in English. Sample-specific internal consistency was $\alpha = .99$ for both French and English parent-reported vocabulary. CCT vocabulary comprehension ranged from four to 36 words ($M = 23.1$, $SD = 8.3$) in French and from two to 37 words ($M = 23.3$, $SD = 8.3$) in English. Sample-specific internal consistency was excellent across forms (Cronbach's alpha = .89 and .90, respectively). Immediate test–retest reliability was significant, $r(29) = .610$, $p < .001$ and $r(23) = .508$, $p = .013$ in French and English, respectively; and CCT scores correlated with concurrent parent-reported production, $r(44) = .337$, $p = .024$ and $r(45) = .457$, $p = .001$, in French and English, respectively.

48 months. Comprehension vocabulary ranged from four to 72 words ($M = 34.66$, $SD = 17.19$) in French and 11 to 87 words ($M = 52.47$, $SD = 17.70$) in English. Vocabulary size in the dominant language ranged from 16 to 87 words ($M = 52.9$, $SD = 17.85$). Sample-specific internal consistency was $\alpha = .92$ and .95, on the EVIP on the PPVT, respectively. Lollipop test scores ranged from 12 to 56 ($M = 38.18$, $SD = 10$). This range of scores was comparable with the monolingual samples in Studies 1 and 2. Sample-specific internal consistency was moderate ($\alpha = .68$).

Below we present findings for the dominant language at 22 and 48 months. Zero-order correlations revealed significant longitudinal relations between decontextualized vocabulary at 22 months and L1 vocabulary and kindergarten readiness at 48 months, $r(33) = .35$ $p = .045$) and $r(33) = .47$, $p = .006$, respectively. In Study 3, all children attended preschool and varied only in how many days they attended per week so we used number of days per week in preschool as a control variable ($M = 4.73$, range = 3 to 5 days/week). Maternal education was associated with days/week in preschool, $r(33) = .43$, $p = .014$, but not with any other measure (all $ps > .11$). Sex was significantly correlated with parent-reported language production and marginally correlated with CCT scores at 22 months, $r(30) = .37$, $p = .048$ and $r(33) = .31$, $p = .083$, respectively; and with kindergarten readiness scores at 48 months, $r(33) = .44$, $p = .01$, but not with other variables ($ps > .272$). Finally, attendance at preschool was associated with maternal education (as described above) and with L1 vocabulary at 48 months, $r(32) = .39$, $p = .027$, but not with any other variables (all $ps > .275$). Importantly, parent-reported L1 vocabulary production was related only to sex, $r(30) = .37$, $p = .048$, and not to any outcome variable (all $ps > .25$).

To test whether the mediation model in Studies 1 and 2 applies to BFLA, we conducted a set of partial correlations to parallel the hierarchical regressions on the English and French samples. This change in analytic procedure was undertaken due to the smaller sample in Study 3. The first partial correlation assessed the relation between CCT comprehension at 22 months and kindergarten readiness at 48 months controlling for sex, maternal education, and days/week at preschool. CCT comprehension was significantly associated with subsequent kindergarten readiness scores, $r(27) = .43$, $p = .02$.

We then tested the relation between CCT at 22 months and L1 vocabulary comprehension at 48 months again controlling for the effects of sex, maternal education, and days/week at preschool. This relation was marginal, $r(27) = .39$, $p = .039$, when controlling for false discovery rate (Bonferroni $\alpha = .033$). Early L1 vocabulary predicted kindergarten readiness scores and marginally predicted subsequent vocabulary at 48 months.

Finally, to assess whether the relation between CCT scores and kindergarten readiness on the kindergarten readiness were mediated by fourth-year vocabulary, we tested the relation between the CCT and kindergarten readiness controlling for sex, maternal education, days/week at preschool, and fourth-year L1 vocabulary comprehension. This correlation was not significant ($p = .094$). Similarly, the relation between L1 vocabulary at 48 months and kindergarten readiness was not significant ($p = .133$) when controlling for sex, maternal education, days/week at preschool, and second-year vocabulary. Together, these findings suggest, consistent with Studies 1 and 2, that vocabulary in the fourth year does not mediate the relation between vocabulary in the second year and kindergarten readiness. Instead, the size of early, decontextualized vocabulary in the second year of life predicts both subsequent vocabulary (although this was marginal when controlling for false detection rate) and kindergarten readiness. Post hoc analyses to assess prediction from L2 revealed no association between early L2 vocabulary comprehension and fourth-year vocabulary in L1 ($p = .30$) or fourth-year L2 ($r(27) = .34$, $p = .08$) and no relation between early L2 vocabulary and kindergarten readiness ($p = .40$). This suggests that early L1 vocabulary provides stronger support for subsequent language acquisition and kindergarten readiness in bilingual children.

Discussion

The purpose of Study 3 was to extend a model in which early vocabulary was associated with subsequent vocabulary and kindergarten readiness to BFLA children. Consistent with Studies 1 and 2, children's decontextualized vocabulary comprehension in the second year predicted kindergarten readiness outcomes and vocabulary comprehension in their dominant language at 48 months of age. Given the small sample size, these findings are preliminary but are important for guiding future research.

Interestingly, the dominant language of exposure at 22 months marginally predicted the language with the largest vocabulary at 48 months even though children's relative exposure changed over time. Some children who received greater exposure to French at 48 months of age had larger vocabularies in English and vice versa. This sheds light on two important facts. First, language exposure is fluid and depends upon environmental influences. Importantly, all of these children had begun preschool, which likely contributed to changing language exposure. Second, children's vocabulary in their dominant language at 22 months marginally predicts vocabulary in their dominant, but not their nondominant, language at 48 months even when relative exposure to these languages has changed. The present research suggests that, in monolingual children and their BFLA peers, early vocabulary is a key longitudinal predictor of language and kindergarten readiness.

General Discussion

Early Vocabulary and Kindergarten Readiness

In Study 1, second-year decontextualized vocabulary explained unique variance in both fourth-year vocabulary and kindergarten readiness after controlling for demographic variables. Second-year decontextualized vocabulary (measured by the CCT) and parent-reported vocabulary (measured by the MCDI) were both related to children's kindergarten readiness although a relative importance analysis that allowed us to assess the independent contributions of these instruments revealed that only CCT comprehension accounted for significant unique variance.

It is important to note that the CCT and MCDI were moderately correlated. This, in conjunction with differential prediction to developmental outcomes, reveals that these measures capture unique as well as overlapping aspects of early vocabulary. Whereas the MCDI yields a reliable estimate of the lexicon inclusive of context-dependent and decontextualized items, the CCT yields a more circumscribed estimate focused on word-world relations that are sufficiently stable to elicit an accurate haptic response. Thus, the two instruments may serve different but complementary purposes. In the present study directly assessed vocabulary yielded effect sizes that were similar to or larger than those obtained in large-scale studies of parent-reported vocabulary. We discuss this point in more detail at the end of this section.

Consistent with previous work finding a relation between comprehension at 18 months and a language sample at 30 months (Friend, Schmitt, & Simpson, 2012), early decontextualized vocabulary was also a predictor of fourth-year vocabulary. These findings, consistent with larger scale studies (Duff et al., 2015; Morgan et al., 2015), suggest that kindergarten readiness and, by extension, emergent literacy is subserved in part by vocabulary size as early as the second year of life.

Study 2 confirmed and extended this finding in a sample of French-speaking children in Switzerland. Decontextualized vocabulary at 22 months of age was the only significant predictor of kindergarten readiness at 48 months. There was a notable difference in parent-reported vocabulary size across our monolingual samples. Parents of English-speaking monolingual children reported that children produced, on average, 259 words whereas parents of French-speaking monolingual children reported that their children produced 198 words on average. This finding is consistent with previous cross-language comparisons using the MCDI (Bleses et al., 2008) and an earlier version of the MCDI (Bornstein et al., 2004). In both reports, there is variability across languages in parent-reported vocabulary for children of the same age and reported vocabulary in French is lower than in English. The same pattern emerged for parent-reported production in the present French and English monolingual samples. However, children's directly assessed vocabulary *comprehension* was similar across monolingual samples: English monolingual children correctly identified the referent for an average of 27 words whereas French children did so for 29 words.

These findings are consistent with Bornstein et al.'s (2004) previous report specific to French and with reports by Bleses et al. (2008) and Eriksson et al. (2012) revealing language group differences in parent-reported vocabulary. Whether these differences reflect real differences in acquisition across languages, differences in reporting practices, or differences in predictive validity is indeterminate. To the last concern, it is important to note three sources of evidence that speak to the reliability and validity of the IFDC. First, Kern and Gayraud (2010) report reliability estimates for the IFDC similar to those reported for the MCDI. Second, sample-specific internal consistency on this measure was excellent and similar to the estimates for the other parent-report instruments in the study and finally, the IFDC converges with directly assessed vocabulary estimates.

In both Studies 1 and 2, decontextualized vocabulary in the second year predicted both vocabulary and kindergarten readiness in the fourth year. In Study 3, we demonstrated that this model holds for children who acquire two languages from birth and importantly, the effect is robust to changes in language status suggesting that general language ability drives later language achievement and readiness for school. This is an important, if preliminary, extension of the mechanisms underlying cognitive development from the toddler to the preschool period to the majority of children in the world who grow up learning more than one language.

In the present study, we calculate eta-squared to estimate effect size to contrast our effects with those in the extant literature. This estimate can be interpreted similarly to R^2. The present study yielded similar or larger effect sizes to those obtained by Duff et al. (2015) and Morgan et al. (2015) for both monolingual samples. For example, whereas Duff et al. (2015) found that parent-reported second-year vocabulary accounted for about 16% of the variance in school age vocabulary, in the present study, directly assessed second-year vocabulary accounted for between 22% and 32% of the variance. Similarly, whereas Morgan et al. (2015) found that parent-reported second-year vocabulary accounted for 22% of the variance in school age reading and 27% of the variance in math, directly assessed second-year vocabulary accounts for between 26% and 29% of the variance in kindergarten readiness. The estimates for the bilingual sample, which are somewhat smaller, should be interpreted with caution given the small sample size and preliminary nature of this investigation. Nevertheless, these findings extend previous research by showing that decontextualized vocabulary may be effective for predicting downstream developmental outcomes across language (English and French) and language status (monolingual and bilingual).

The Role of Decontextualized Vocabulary

One of the difficulties that novice word learners have in generalizing across referents for a word is that they must dissociate, to some extent, the word–referent relation from the context in which it was learned. Otherwise, children's early vocabulary is "context-bound:" children can only show recognition of the word–referent relation in its original supporting context. We emphasize decontextualized vocabulary because this level of knowledge reflects stable word–referent relations across contexts, the ability to generalize relations to other exemplars, and the ability to extract the relevant properties of the referent class (McMurray et al., 2012; Smith & Yu, 2008). This is an essential precursor to what Hadley et al. (2015) refer to as "deep word knowledge:" lexical knowledge that sets the stage for acquiring new concepts and for later literacy.

From this view, early, decontextualized knowledge supports the development of stable conceptual networks. Children must first form word–referent relations that are consistent across multiple contexts. These stable word–referent relations can, in turn, support stable conceptual networks. Finally, these networks become a platform from which the concepts and skills necessary for beginning kindergarten are built. For example, knowing number words facilitates the development of numeric skills (Barner & Bachrach, 2010; Mix, 2009) and knowing letter names facilitates recognition of letters in print (Justice & Ezell, 2004). Although our findings indicate that early, decontextualized vocabulary is related to kindergarten readiness, this is not to say that it is *causal* in and of itself. It is possible that children who have large, stable lexicons at the end of the second year may be "good learners" relative to their peers such that, when they reach preschool age, they are also better prepared for kindergarten. That is, the same general learning mechanisms that support the acquisition of a stable lexicon may also support conceptual development and kindergarten readiness. For example, working memory may scaffold both early vocabulary and a broad range of nonlinguistic skills (e.g., understanding of numbers and shapes).

Previous work focused on processing *efficiency* illustrates this point (Marchman & Fernald, 2008). Processing efficiency and parent-reported production at 25 months each account for unique variance (16% and 17%, respectively) in linguistic and cognitive skills at age 8 and these predictors are correlated. In fact, processing efficiency is related to *both* parent-reported (Hurtado, Marchman, & Fernald, 2007; Marchman, Fernald, & Hurtado, 2010) and directly assessed vocabulary (Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015; Legacy, Zesiger, Friend, & Poulin-Dubois, 2018). Further, vocabulary size and processing efficiency are associated with later working memory. Marchman and Fernald (2008) argue that vocabulary growth is driven by the same learning principles that support the development of an array of skills and concepts. Evaluating a causal model that takes into account the relative contributions of vocabulary, processing efficiency, and general learning to children's kindergarten readiness is beyond the scope of the present article and an important direction for future research.

Limitations

The samples in the present research are somewhat small relative to some of the large-scale studies that we have cited (e.g., Duff et al., 2015; Forget-Dubois et al., 2009; Morgan et al., 2015). Nevertheless, the samples are within range of several related studies (e.g., Cristofaro & Tamis-LeMonda, 2011; Fernald & Marchman, 2012; Marchman & Fernald, 2008) and, in particular, of previous work using the CCT, an individually administered, lab-based assessment (Friend & Keplinger, 2003, 2008; Hendrickson et al., 2015; Legacy et al., 2016; Legacy, Zesiger, Friend, & Poulin-Dubois, 2018; Poulin-Dubois et al., 2013). Our findings are in line with, and extend, previous work with both large and small samples. An important focus of the present article was to determine whether decontextualized vocabulary in the second year was sufficiently robust to account for unique variance in preschool outcomes. The fact that the same pattern of effects emerged across language and geographic/cultural differences with effect sizes in the range of previous reports or larger, suggests that these effects are robust. We are involved currently in collaborative research scaling this assessment to environments outside the laboratory to facilitate testing of larger samples.

In addition, more work is required to test monolingual findings in bilingual samples. The pattern of effects reported here must be tested in larger samples and across languages. For example, it is not clear whether these effects would hold for children who are learning languages from different language families (e.g., Chinese and English). Finally, that we find the same pattern of effects in bilinguals as in monolinguals tells us little with regard to how bilingual children make the transition to proficiency in the language of schooling that is critical for academic success. That is, when a child's dominant language does not match the language of schooling, it is possible that whereas dominant language vocabulary size predicts kindergarten readiness, it may not predict academic achievement.

Finally, our findings differ from previous research in that, whereas zero order correlations revealed maternal education was associated with some measures in Study 1, it was not associated with parent-reported vocabulary production in the second year or with vocabulary in the fourth year and it was not a significant predictor in our regression models. Further, maternal education did not emerge as significantly associated with either predictors or outcome measures in Studies 2 and 3. In contrast, Forget-Dubois et al. (2009) reported a significant effect of socioeconomic status (SES; i.e., parent education and household income) with an effect size of $R^2 = .10$. Our models in Study 1 yielded an identical effect size for control variables (child sex and maternal education) suggesting that larger samples may be required to detect this effect. Because maternal education has been shown to exert an effect in previous work (Hoff, 2013) and in the interest of providing better-adjusted estimates of the relation of early vocabulary to later outcomes, we controlled for it in all of our analyses. Scaling up testing to larger samples would facilitate more comprehensive models of the role of decontextualized vocabulary vis-à-vis variables with smaller effect sizes (e.g., demographic factors) in contributing to children's preparedness for school.

To conclude, this research contributes to the literatures on language development and emergent literacy by providing evidence that vocabulary comprehension as early as 22 months of age is associated with subsequent vocabulary and kindergarten readiness. In conjunction with recent evidence that early vocabulary predicts later literacy (Duff et al., 2015; Law et al., 2013; Morgan et al., 2015), this article emphasizes the significance of early language comprehension for broad developmental outcomes and, of particular importance, we have extended these findings across two languages and across monolingual and bilingual children. Of interest for future research are the mechanisms that underlie this continuity. We have suggested that early, decontextualized vocabulary supports later vocabulary and broader conceptual learning. However, an alternative is that vocabulary knowledge is the product of an efficient cognitive system and it is this cognitive efficiency that underlies both vocabulary acquisition and school readiness. Both language-based and general cognitive accounts are consonant with these findings. Distinguishing between these accounts is a direction for future research that may have implications for preschool interventions.

References

Barner, D., & Bachrach, A. (2010). Inference and exact numerical representation in early language development. *Cognitive Psychology, 60,* 40–62. http://dx.doi.org/10.1016/j.cogpsych.2009.06.002

Bates, E., Bretherton, I., Carlson, V., Carpen, K., & Rosser, M. (1979). Next steps: A follow-up study and some pilot research. In E. Bates, L. Benigni, I. Bretherton, L. Camaioni, & V. Volterra (Eds.), *The emergence of symbols: Cognition and communication in infancy* (pp. 271–311). Waltham, MA: Academic Press. http://dx.doi.org/10.1016/B978-0-12-081540-1.50012-0

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., . . . Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism: Language and Cognition, 15,* 616–629. http://dx.doi.org/10.1017/S1366728912000090

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, Methodological, 57,* 289–300.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language, 35,* 619–650. http://dx.doi.org/10.1017/S0305000908008714

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., . . . Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75,* 1115–1139. http://dx.doi.org/10.1111/j.1467-8624.2004.00729.x

Boudreault, M.-C., Cabirol, É.-A., Trudeau, N., Poulin-Dubois, D., & Sutton, A. (2007). Les Inventaires Macarthur du Développement de la Communication: Validité et données normatives préliminaires [The MacArthur Communicative Development Inventories: Preliminary validity and normative data]. *Journal of Speech-Language Pathology and Audiology, 31,* 27–37.

Brebner, C., McCormack, P., & Liow, S. R. (2016). Marking of verb tense in the English of preschool English–Mandarin bilingual children: Evidence from language development profiles within subgroups on the Singapore English Action Picture Test. *International Journal of Language & Communication Disorders, 51,* 31–43. http://dx.doi.org/10.1111/1460-6984.12181

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities, 44,* 431–443. http://dx.doi.org/10.1177/0022219411410042

Chew, A. L. (2007). *Developmental and interpretive manual for the Lollipop Test: A Diagnostic Screening Test of School Readiness-III*. Atlanta, GA: Humanics Learning.

Chew, A. L., & Morris, J. D. (1984). Validation of the lollipop test: A diagnostic screening test of school readiness. *Educational and Psychological Measurement, 44,* 987–991. http://dx.doi.org/10.1177/0013164484444022

Chew, A. L., & Morris, J. D. (1987). Investigation of the lollipop test as a pre-kindergarten screening instrument. *Educational and Psychological Measurement, 47,* 467–471. http://dx.doi.org/10.1177/0013164487472019

Crain-Thoreson, C., & Dale, P. S. (1992). Do early talkers become early readers? Linguistic precocity, preschool language, and emergent literacy. *Developmental Psychology, 28,* 421–429. http://dx.doi.org/10.1037/0012-1649.28.3.421

Cristofaro, T. N., & Tamis-LeMonda, C. S. (2011). Mother–child conversations at 36 months and at pre-kindergarten: Relations to children's school readiness. *Journal of Early Childhood Literacy, 12,* 68–97. http://dx.doi.org/10.1177/1468798411416879

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28,* 125–127. http://dx.doi.org/10.3758/BF03203646

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). The language exposure assessment tool: Quantifying language exposure in infants and children. *Journal of Speech, Language, and Hearing Research, 59,* 1346–1356. http://dx.doi.org/10.1044/2016_JSLHR-L-15-0234

De Houwer, A. (2009). *An introduction to bilingual development*. Bristol, UK: Multilingual Matters. Retrieved from https://books.google.com/books?hl=en&lr=&id=R0-SY0ryj0oC&oi

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics, 35,* 1189–1211. http://dx.doi.org/10.1017/S0142716412000744

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher, 39,* 305–310. http://dx.doi.org/10.3102/0013189X10370204

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56,* 848–856. http://dx.doi.org/10.1111/jcpp.12378

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. Circle pines, MN: American Guidance Service.

Dunn, L. M., Dunn, L. M., & Thériault-Whalen, C. M. (1993). *Échelle de Vocabulaire en Images Peabody: Série de Planches* [Peabody Picture Vocabulary Scale: Series of Stages]. Toronto, Ontario, Canada: Psychcorp.

Eno, L., & Woehlke, P. (1995). Use of the Lollipop Test as a predictor of California achievement test scores in kindergarten and transitional first-grade status. *Psychological Reports, 76,* 145–146. http://dx.doi.org/10.2466/pr0.1995.76.1.145

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., . . . Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology, 30,* 326–343. http://dx.doi.org/10.1111/j.2044-835X.2011.02042.x

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149–1160. http://dx.doi.org/10.3758/BRM.41.4.1149

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59,* 1–185. http://dx.doi.org/10.2307/1166093

Fenson, L., & Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Reilly, J. (2007). *MacArthur communicative development inventories: User's guide and technical manual*. San Diego, CA: Singular Publishing Group.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83,* 203–222. http://dx.doi.org/10.1111/j.1467-8624.2011.01692.x

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42,* 98–116. http://dx.doi.org/10.1037/0012-1649.42.1.98

Forget-Dubois, N., Dionne, G., Lemelin, J. P., Pérusse, D., Tremblay, R. E., & Boivin, M. (2009). Early child language mediates the relation between home environment and school readiness. *Child Development, 80,* 736–749. http://dx.doi.org/10.1111/j.1467-8624.2009.01294.x

Frank, I., Poulin-Dubois, D., & Trudeau, N. (1997). *Inventaires MacArthur du Développement de la Communication: Mots et Énoncés*. Montreal, Quebec, Canada: Concordia University.

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers, 35,* 302–309. http://dx.doi.org/10.3758/BF03202556

Friend, M., & Keplinger, M. (2008). Reliability and validity of the computerized comprehension task (CCT): Data from American English and Mexican Spanish infants. *Journal of Child Language, 35,* 77–98. http://dx.doi.org/10.1017/S0305000907008264

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the computerized comprehension task: Comprehension predicts production. *Developmental Psychology, 48,* 136–148. http://dx.doi.org/10.1037/a0025511

Friend, M., & Zesiger, P. (2011). A systematic replication of the psychometric properties of the CCT in three languages: English, Spanish, and French. *Enfance, 3,* 329–344. http://dx.doi.org/10.4074/S0013754511003053

Gardner-Neblett, N., & Iruka, I. U. (2015). Oral narrative skills: Explaining the language-emergent literacy link by race/ethnicity and SES. *Developmental Psychology, 51,* 889–904. http://dx.doi.org/10.1037/a0039274

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2015). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly, 51,* 181–198. http://dx.doi.org/10.1002/rrq.130

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language, 27,* 689–705. http://dx.doi.org/10.1017/S0305000900004414

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,* 723–735. http://dx.doi.org/10.1111/desc.12250

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology, 49,* 4–14. http://dx.doi.org/10.1037/a0027238

Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language, 34,* 227–249. http://dx.doi.org/10.1017/S0305000906007896

Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35,* 1–19. http://dx.doi.org/10.1207/S15327906MBR3501_1

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5,* 602–619. http://dx.doi.org/10.1177/0193841X8100500502

Justice, L. M., & Ezell, H. K. (2004). Print referencing: An emergent literacy enhancement strategy and its clinical applications. *Language, Speech, and Hearing Services in Schools, 35,* 185–193. http://dx.doi.org/10.1044/0161-1461(2004/018)

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language, 27,* 227–250. http://dx.doi.org/10.1177/0142723706075789

Kern, S. (2010). Les premiers mots du jeune enfant français: Analyse quantitative et qualitative du vocabulaire réceptif et productif des deux premières années de vie [The first words of the young French child: Quantitative and qualitative analysis of the receptive and productive vocabulary of the first two years of life]. *Rééducation orthophonique, 244,* 149–165.

Kern, S., & Gayraud, F. (2010). *L'inventaire français du développement communicatif* [French inventories of communicative development]. Grenoble, France: Editions La Cigale.

Kim, S., Im, H., & Kwon, K. A. (2015). The role of home literacy environment in toddlerhood in development of vocabulary and decoding skills. *Child and Youth Care Forum, 44,* 835–852. http://dx.doi.org/10.1007/s10566-015-9309-y

Koponen, T., Aunola, K., Ahonen, T., & Nurmi, J. E. (2007). Cognitive predictors of single-digit and procedural calculation skills and their covariation with reading skill. *Journal of Experimental Child Psychology, 97,* 220–241. http://dx.doi.org/10.1016/j.jecp.2007.03.001

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health, 13,* 198–206. http://dx.doi.org/10.1111/j.1475-3588.2008.00503.x

Law, J., Rush, R., Parsons, S., & Schoon, I. (2013). The relationship between gender, receptive vocabulary, and literacy from school entry through to adulthood. *International Journal of Speech-Language Pathology, 15,* 407–415. http://dx.doi.org/10.3109/17549507.2012.721897

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2016). Vocabulary size, translation equivalents, and efficiency in word recognition in very young bilinguals. *Journal of Child Language, 43,* 760–783. http://dx.doi.org/10.1017/S0305000915000252

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2018). Vocabulary size and speed of word recognition in very young French-English bilinguals: A longitudinal study. *Bilingualism: Language and Cognition, 21,* 137–149. http://dx.doi.org/10.1017/S1366728916000833

Lemelin, J.-P., Boivin, M., Forget-Dubois, N., Dionne, G., Séguin, J. R., Brendgen, M., . . . Pérusse, D. (2007). The genetic-environmental etiology of cognitive school readiness and later academic achievement in early childhood. *Child Development, 78,* 1855–1869. http://dx.doi.org/10.1111/j.1467-8624.2007.01103.x

Lonigan, C. J., Burgess, S. R., Anthony, J. L., & Barker, T. A. (1998). Development of phonological sensitivity in 2-to 5-year-old children. *Journal of Educational Psychology, 90,* 294–311. http://dx.doi.org/10.1037/0022-0663.90.2.294

Lust, B., Flynn, S., Blume, M., Park, S. W., Kang, C., Yang, S., & Kim, A. Y. (2016). Assessing child bilingualism: Direct assessment of bilingual syntax amends caretaker report. *The International Journal of Bilingualism, 20,* 153–172. http://dx.doi.org/10.1177/1367006914547661

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83–104. http://dx.doi.org/10.1037/1082-989X.7.1.83

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11,* F9–F16. http://dx.doi.org/10.1111/j.1467-7687.2008.00671.x

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language, 37,* 817–840. http://dx.doi.org/10.1017/S0305000909990055

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review, 119,* 831–877. http://dx.doi.org/10.1037/a0029872

Milburn, T. F., Hipfner-Boucher, K., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2016). Cognitive, linguistic and print-related predictors of preschool children's word spelling and name writing. *Journal of Early Childhood Literacy, 17,* 111–136. http://dx.doi.org/10.1177/1468798415624482

Mitchell, A. M., & Brady, S. A. (2013). The effect of vocabulary knowledge on novel word identification. *Annals of Dyslexia, 63,* 201–216. http://dx.doi.org/10.1007/s11881-013-0080-1

Mix, K. S. (2009). How Spencer made number: First uses of the number words. *Journal of Experimental Child Psychology, 102,* 427–444. http://dx.doi.org/10.1016/j.jecp.2008.11.003

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development, 86,* 1351–1370. http://dx.doi.org/10.1111/cdev.12398

NICHD Early Child Care Research Network. (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41,* 428–442. http://dx.doi.org/10.1037/0012-1649.41.2.428

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98,* 554–556. http://dx.doi.org/10.1037/0022-0663.98.3.554

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 357–383. http://dx.doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. S. Gorfien (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10459-004

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18,* 22–37. http://dx.doi.org/10.1080/10888438.2013.827687

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *The International Journal of Bilingualism, 17,* 57–70. http://dx.doi.org/10.1177/1367006911431198

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86,* 159–175. http://dx.doi.org/10.1111/cdev.12292

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., . . . Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from early language in Victoria Study. *Pediatrics, 126,* e1530–e1537. http://dx.doi.org/10.1542/peds.2010-0254

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64,* 913–951. http://dx.doi.org/10.1111/lang.12077

Sheng, L., Lu, Y., & Gollan, T. H. (2014). Assessing language dominance in Mandarin–English bilinguals: Convergence and divergence between subjective and objective measures. *Bilingualism: Language and Cognition, 17,* 364–383.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106,* 1558–1568. http://dx.doi.org/10.1016/j.cognition.2007.06.010

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology, 13,* 290–312. http://dx.doi.org/10.2307/270723

Statistics Canada. (2016). *Language highlights tables, 2016 Census*. Retrieved from http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/hlt-fst/lang/Table.cfm?Lang=E&T=21&Geo=00&SP=1&view=1&age=4

Stoel-Gammon, C. (1987). Phonological skills of 2-year-olds. *Language, Speech, and Hearing Services in Schools, 18,* 323–329. http://dx.doi.org/10.1044/0161-1461.1804.323

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934–947. http://dx.doi.org/10.1037/0012-1649.38.6.934

Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology, 126,* 395–411. http://dx.doi.org/10.1016/j.jecp.2014.06.003

Sulzby, E., & Teale, W. H. (1985). Writing development in early childhood. *Educational Horizons, 64,* 8–12.

Tonidandel, S., & LeBreton, J. M. (2014). RWA-Web—A free, comprehensive, web-based, and user-friendly tool for relative weight analysis. *Journal of Business and Psychology, 30,* 207–216. http://dx.doi.org/10.1007/s10869-014-9351-z

Tonidandel, S., Lebreton, J. M., & Johnson, J. W. (2009). Determining the statistical significance of relative weights. *Psychological Methods, 14,* 387–399. http://dx.doi.org/10.1037/a0017735

Trudeau, N., Aktouf, K., Boudreault, M., & Breault, C. (2008). *Les Inventaires MacArthur du développement de la communication: Mots et gestes* [MacArthur Inventories of the Development of communication: words and gestures]. Unpublished Manual, School of Speech Therapy and Audiology, University of Montreal, Montréal, Québec, Canada.

United States Census Bureau. (2015). *Detailed languages spoken at home and ability to speak English for the population 5 years and over: 2009–2013*. Retrieved from http://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848–872. http://dx.doi.org/10.1111/j.1467-8624.1998.tb06247.x

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word-referent learning: Prior questions. *Psychological Review, 119,* 21–39. http://dx.doi.org/10.1037/a0026182

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *PLS-5: Preschool Language Scale-5*. San Antonio, TX: Psychological Corporation.

Received February 6, 2017
Revision received October 30, 2017
Accepted December 31, 2017 ■

Correction to Friend et al. (2018)

In the article "A Cross-Language Study of Decontextualized Vocabulary Comprehension in Toddlerhood and Kindergarten Readiness," by Margaret Friend, Erin Smolak, Yushuang Liu, Diane Poulin-Duboisand, and Pascal Zesiger (*Developmental Psychology*, Advance online publication. April 5, 2018. http://dx.doi.org/10.1037/dev0000514), the reference for Legacy, Zesiger, Friend, & Poulin-Dubois (2016) should be Legacy, Zesiger, Friend, & Poulin-Dubois (2018). The correct reference for the article is listed below:

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2018). Vocabulary size and speed of word recognition in very young French-English bilinguals: A longitudinal study. *Bilingualism: Language and Cognition, 21*, 137–149. https://doi.org/10.1017/S1366728916000833.

All versions of this article have been corrected.

http://dx.doi.org/10.1037/dev0000590

Chapter 2, in full, is a reprint of the material as it appears in Developmental Psychology. Friend, M., Smolak, E., Liu, Y. Poulin-Dubois, D., & Zesiger, P. (2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. *Developmental Psychology, 54*, 1317-1333. doi: 10.1037/dev0000514. The dissertation author was the co-investigator and co-author of this paper.

Table 14. Unstandardized B values for all regression models presented in Chapter 2, Study 1.

Predicting Kindergarten Readiness	Model 1		Model 2	
	SE	*B*	*SE*	*B*
Measure				
Maternal education	.69	1.27	.68	.51
Child sex	3.05	5.63	2.89	3.45
2nd year CCT comprehension			.22	.67

Predicting Vocabulary	Model 1		Model 2	
	SE	*B*	*SE*	*B*
Measure				
Maternal education	1.05	.72	1.01	-.53
Child sex	4.62	-3.98	4.30	-7.54
2nd year CCT comprehension			.33	1.10

Predicting Kindergarten Readiness	Model 2		Model 3	
	SE	*B*	*SE*	*B*
Measure				
Maternal education	.62	1.06	.65	.62
Child sex	2.77	6.82	2.86	5.09
4th year PPVT comprehension	.09	.30	.10	.22
2nd year CCT comprehension			.24	.44

Table 15. Unstandardized B values for all regression models presented in Chapter 2, Study 2.

Predicting Kindergarten Readiness	Model		Model 2	
	SE	*B*	*SE*	*B*
Measure				
Maternal education	.65	-.17	.62	-.06
Child sex	3.60	3.50	3.40	2.62
Preschool	5.80	13.88	5.49	11.84
2nd year CCT comprehension			.30	.85
Predicting Vocabulary	**Model 1**		**Model 2**	
	SE	*B*	*SE*	*B*
Measure				
Maternal education	.75	.64	.63	.82
Child sex	4.12	.65	3.50	-.87
Preschool	6.63	4.22	5.65	.73
2nd year CCT comprehension			.31	1.46
Predicting Kindergarten Readiness	**Model 2**		**Model 3**	
	SE	*B*	*SE*	*B*
Measure				
Maternal education	.63	-.35	.62	-.19
Child sex	3.44	3.32	3.39	2.76
Preschool	5.55	12.68	5.48	11.73
4th year EVIP comprehension	.11	.28	.13	.16
2nd year CCT comprehension			.35	.62

CHAPTER 3:

Language Status at Age 3: Group and Individual Prediction From Vocabulary Comprehension in the Second Year

American Psychological Association

Developmental Psychology

0012-1649/19/$12.00

2019, Vol. 55, No. 1, 9–22
http://dx.doi.org/10.1037/dev0000617

Language Status at Age 3: Group and Individual Prediction From Vocabulary Comprehension in the Second Year

Margaret Friend
San Diego State University

Erin Smolak
San Diego State University/University of California, San Diego

Tamara Patrucco-Nanchen
University of Geneva

Diane Poulin-Dubois
Concordia University

Pascal Zesiger
University of Geneva

The present research extends recent work on the prediction of preschool language skills by exploring prediction from decontextualized vocabulary comprehension. Vocabulary comprehension was a stronger predictor than parent-reported production, yielding a quadrupling of variance accounted for relative to prior studies. Parallel studies (Studies 1 and 2) are reported for two linguistically and geographically distinct samples. In both samples, decontextualized vocabulary comprehension late in the second year provided the best balance between model fit and parsimony in predicting language skills at age three. In Study 3, vocabulary comprehension prospectively identified children with low language status 2 years earlier than other prospective studies but with similar sensitivity and specificity. The present paper provides evidence on three questions of practical and theoretical significance: the relation between decontextualized vocabulary prior to 30 months of age and language outcomes, how prediction from decontextualized vocabulary compares with parent-reported vocabulary, and finally how early stable predictions to language outcomes can be made.

Keywords: comprehension, Cross-Language, practical significance, prediction

Supplemental materials: http://dx.doi.org/10.1037/dev0000617.supp

Word learning develops rapidly in the first two years (Dapretto & Bjork, 2000; Fenson et al., 1994; Mayor & Plunkett, 2010; Reznick & Goldfield, 1992; Samuelson & McMurray, 2017) and vocabulary production (Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015), comprehension (Friend, Schmitt, & Simpson, 2012), and speed of word processing (Fernald & Marchman, 2012; Marchman & Fernald, 2008) are building blocks for later linguistic and cognitive development. The present paper evaluates the role of early vocabulary in predicting language skill at age three in geographically and linguistically distinct samples of monolingual children at the group and individual levels. In so doing we attempt to reconcile evidence that word learning emerges from domain-general processes that are expected to be stable with evidence that, in general, early vocabulary accounts for only a small proportion of variance at the group level and is unstable at the individual level.

Recent research reveals that domain general mechanisms can account for the pattern of vocabulary acquisition with age (Samuelson & McMurray, 2017; Vlach & DeBrock, 2017; Vlach & Johnson, 2013; Vlach & Sandhofer, 2011; Yu & Smith, 2012). For instance, young children's attentional biases reduce referential ambiguity (Samuelson & McMurray, 2017; Yurovsky, Smith, & Yu, 2013) and both vocabulary size and memory contribute to cross-situational word learning (Smith & Yu, 2013), supporting the development of stable word-referent relations (Vlach & DeBrock, 2017; Vlach & Johnson, 2013; Vlach & Sandhofer, 2011). Initially weak word-referent relations may be strengthened over time through the iterative application of domain general learning (Bion, Borovsky, & Fernald, 2013; Gershkoff-Stowe & Hahn, 2013; Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend,

This article was published Online First October 25, 2018.

Margaret Friend, Department of Psychology, San Diego State University; Erin Smolak, Joint Doctoral Program in Language and Communicative Disorders, San Diego State University/University of California, San Diego; Tamara Patrucco-Nanchen, Psychologie et Sciences de l'Education, University of Geneva; Diane Poulin-Dubois, Department of Psychology, Concordia University; Pascal Zesiger, Psychologie et Sciences de l'Education, University of Geneva.

We gratefully acknowledge Laura Alaria, Kristi Hendrickson, and Danielle Rosen for assistance in data collection and coding and all of our participant families. This research was supported by NICHD R01HD468058 and NIDCD T32DC00736 and does not necessarily represent the views of the National Institutes of Health.

Correspondence concerning this article should be addressed to Margaret Friend, Department of Psychology, San Diego State University, 6505 Alvarado Road, Suite 101, San Diego, CA 92120. E-mail: mfriend@sdsu.edu

2015; Hendrickson, Poulin-Dubois, Zesiger, & Friend, 2017; Yu & Smith, 2012). From this view, early vocabulary should evince stability with later abilities that build on these mechanisms (e.g., language, school readiness, and achievement; Duff, Reen, Plunkett, & Nation, 2015; Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018; Morgan et al., 2015). Because these mechanisms are presumed to be universal, this expectation also applies across languages.

In support of this idea, by 24 months of age, parent-reported vocabulary predicts later language (Duff et al., 2015; Ghassabian et al., 2014; Henrichs et al., 2011; Kemp et al., 2017; Reilly et al., 2010), literacy and reading (Bleses, Makransky, Dale, Højen, & Ari, 2016; Duff et al., 2015), and academic and behavioral functioning (Morgan et al., 2015) at the group level in English-, Dutch-, and Danish-speaking children. However, it accounts for a small to modest proportion of variance (Duff et al., 2015; Morgan et al., 2015; Reilly et al., 2010). At the individual level, prediction is inadequate (Law et al., 2000; Law & Roy, 2008). Across studies, parent report prospectively identifies only roughly one half of children who develop language problems (e.g., Dale, Price, Bishop, & Plomin, 2003; Heilmann, Weismer, Evans, & Hollar, 2005; Westerlund, Berglund, & Eriksson, 2006). From these findings, Dale et al. (2003) concluded that supplemental assessment is necessary to identify developmental risk.

How can we explain the weak prediction from vocabulary prior to the third year to subsequent language and literacy at the group level and to language problems at the individual level? Imagine a child who produces the words "dog" and "milk." This child may have a strong association between the word "dog" and its referents and use it appropriately across contexts but a relatively weak association between "milk" and its referent, using it only in the context of breakfast. Indirect assessments such as parent report may assess this full continuum of word-referent associations from weak to strong. In both cases, parents should report these as words their child produces. In contrast, direct assessments that require active lexical retrieval and hypothesis testing ("Is this a duck or is that a duck?") should preferentially tap strong, rather than weak, associations (Yu & Smith, 2012). Weak associations are considered fragile, subject to interference, and context-bound (Bion et al., 2013; Gershkoff-Stowe & Hahn, 2013), whereas strong associations, formed through iterative domain general processes, are stable across situations. We refer to these stable associations as decontextualized.

Decontextualized, in contrast to context-bound, associations may provide the substrate for subsequent word knowledge and conceptual development (Schmitt, 2014) and predict downstream language and cognitive ability. Friend et al. (2018) found that decontextualized vocabulary in the second year predicted vocabulary comprehension and kindergarten readiness at age four in monolingual English and French-speaking children and in bilingual children acquiring French and English simultaneously. Decontextualized vocabulary was a stronger predictor than parent report with effect sizes comparable to or greater than those in prior studies.

The Present Research

This research builds on recent research on predicting language and literacy from early vocabulary. First, we assess prediction from decontextualized vocabulary at 16 and 22 months of age to preschool language in two groups (American English and Swiss French) that differ geographically, culturally, and in native language. Because these languages have distinct prosody, syntax, and grammar and differ both in age-related vocabulary and MLU (Bleses et al., 2008; Thordardottir, 2005), we conduct parallel analyses in Studies 1 and 2 to assess generalizability. We utilize spontaneous and elicited measures to capture the breadth of preschool language (vocabulary comprehension, expressive language complexity, sentence comprehension, grammar, and syntax) and extract a single factor to estimate core language ability and eliminate method variance (Bornstein, Hahn, & Putnick, 2016). We anticipate prediction from early vocabulary to core language ability that is statistically and practically significant at the group and individual levels.

Second, we evaluate the relative contributions of decontextualized vocabulary and parent report. We anticipate decontextualized vocabulary to account for unique variance beyond parent report, resulting in lower error and better fit. Third, consistent with guidelines from the Individuals with Disabilities Act (Wallace et al., 2015), we contrast models across time points to find the age of earliest prediction (Bornstein, Putnick, & Esposito, 2017) and expect models to become more stable with age (Bornstein et al., 2016). Finally, in Study 3 we assess the sensitivity and specificity of decontextualized vocabulary to prospectively identify individual children with low language.

This research was conducted under the project "The Path from Language to Literacy," supported by the NICHD and approved by the Institutional Review Boards at San Diego State University (protocol #603057) and at the University of Geneva.

Study 1

Method

Participants. Seventy-nine English-speaking monolingual children (41 girls) were recruited as part of a larger, multiinstitutional, longitudinal project investigating children's path to literacy. Participants were recruited from Women, Infant, and Children Centers, the YMCA, local churches, parenting groups, swap meets, child-oriented festivals, and birth records in a large city in the Southwestern U.S. Thirty-five children were excluded for failure to complete at least one task across waves (n = 11), becoming bilingual (n = 1), or attrition (n = 23 or 30%). Roughly one-half of the attrition was attributable to parents moving out of state (n = 10) with the rest attributable to lost contact (n = 12) or a change in lab location (n = 1). The final sample consisted of 44 children (26 girls). To test whether this sample size was appropriate to our aims, we conducted an a priori power analysis in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) using the observed correlations between the CCT, MCDI, and language sample MLU found in Friend et al. (2012). The analysis confirmed that the present sample size was sufficient to detect similar effect sizes with power = .8.

Participants visited the lab on three occasions: Wave 1 at 16 months of age (M = 16;21, range 15;15 – 18;3), Wave 2 at 23 months (M = 23;0, range 21;6 – 25;12), and Wave 3 at 36 months (M = 37;23; range 35;9 – 41;24). All infants were full term, had no diagnosed hearing or vision impairments, and were exposed at least 80% of the time to their native language. A $25 gift card to a major retailer and a small toy were provided as incentives at each Wave. See Table 1 for demographic data on the final sample.

Measures.

Language Exposure Assessment Tool (LEAT). Language exposure was estimated on the LEAT (DeAnda, Bosch, Poulin-Dubois, Zesiger, & Friend, 2016) prior to the initial visit to insure monolingual status. This estimate derived from parent reports of the number of hours of language input by all interlocutors over the course of the child's life. Internal consistency is excellent (Cronbach's alpha = .96).

MacArthur-Bates Communicative Development Inventories (MCDI). The MCDI is a parent report measure of early comprehension and production (Fenson et al., 1994, 2007). It consists of two forms: Words and Gestures (WG), for children from 8 to 18 months of age, and Words and Sentences (WS), for children from 16 to 30 months. The WG measure contains a checklist of 396 words, on which caregivers indicate words children understand and words they understand and say. At 16 months, receptive and expressive vocabulary were estimated from the WG form. The WS form contains a vocabulary checklist of 680 words and assesses only words that children say. At 23 months, expressive vocabulary was estimated from the WS form. The MCDI: WG and WS evince moderate to high internal consistency and good test–retest reliability. Six-month stability is moderate for both forms (Fenson et al., 1994).

Computerized Comprehension Task (CCT). The CCT is a forced-choice measure of vocabulary comprehension administered on a touchscreen (Friend & Keplinger, 2003, 2008; available at https://childes.talkbank.org). Paired images appear on the screen as an experimenter delivers a prompt in which the target word is embedded (e.g., "Where is the *dog*?" "Who is *swimming*?" "Which one is *old*?"). It is expected that retrieving word-referent associations upon hearing the prompt and selecting an association via haptic response correspond to processes of lexical retrieval and hypothesis testing. Correct responses are thought to reflect children's decontextualized word-referent associations. Each trial has a maximum duration of 7 seconds (sufficient to execute a haptic response), and trials are interleaved with a blank blue screen. The pace is experimenter-controlled to ensure that trials are presented only when the child is quiet, alert, and looking at the screen. Administration followed Friend et al. (2012) with the following additional criteria for repeating or terminating trials. Repetitions of trials were allowed under the following conditions: (a) the child attempts a response but does not complete the touch before the end of the trial, (b) the child becomes distracted and misses the trial, (c) the child accidentally touches the screen, or (d) the child has not made any attempts for the last 3 consecutive trials. In the last case, the experimenter attempts to reengage the child by moving the child's hand to the target, or by touching the target to elicit the rewarding stimulus. If the child becomes fussy during the procedure, and three attempts to reengage have failed, the experimenter terminates the procedure. If this is necessary and the child has completed one-third of the trials or less, they are excluded from analyses. Three children met this criterion ($n = 2$ at Wave 1 and $n = 1$ at Wave 2).

Table 1
Distribution of Selected Demographic Characteristics of Participants in Study 1

Characteristic	Number (%) of participants		
	Female	Male	Total
Maternal education			
High school or less	2 (4.5)	4 (9.1)	6 (13.6)
Some college	7 (15.9)	1 (2.3)	8 (18.2)
College graduate	7 (15.9)	6 (13.6)	13 (29.5)
Post-baccalaureate	10 (22.7)	7 (15.9)	17 (38.6)
Family income			
18,000–40,000	4 (9.1)	2 (4.5)	6 (13.6)
41,000–60,000	1 (2.3)	3 (6.8)	4 (9.1)
61,000–80,000	5 (11.4)	0 (0.0)	5 (11.4)
81,000–100,000	10 (22.7)	8 (18.2)	18 (40.9)
>100,000	6 (13.6)	5 (11.4)	11 (25.0)
Ethnicity			
Asian	0 (0.0)	2 (4.5)	2 (4.5)
Black/not Hispanic	1 (2.3)	0 (0.0)	1 (2.3)
Hispanic	6 (13.6)	1 (2.3)	7 (15.9)
White/not Hispanic	14 (31.8)	14 (31.8)	28 (63.6)
Mixed Race	5 (11.4)	1 (2.3)	6 (13.6)

Note. Income reported in U.S. dollars. Some values may not sum to 100 because of rounding error.

Words on the CCT are derived from the MCDI vocabulary checklist (Fenson et al., 2007). Referents are high-quality, colorful digital images that are prototypical exemplars. Pairs are matched on color, size, saliency, word class, and difficulty (see Table S1 for full item set). Word difficulty was based on 16-month norming data (Frank, Braginsky, Yurovsky, & Marchman, 2017). There are approximately equal numbers of easy (comprehension ≥66%), moderate (comprehension = 33% to 66%), and difficult (comprehension ≤33%) words randomly distributed throughout the test. The inclusion of more difficult items allows the CCT to be extended up to 2 years of age. Two forms are counterbalanced across participants, such that each word serves as both target and distractor. Finally, target side is randomized, with the restriction that it not appear on the same side on more than two consecutive trials following Hirsh-Pasek and Golinkoff (1996).

The CCT has strong immediate test–retest reliability and moderate short-term stability over a 4-month period (Friend & Keplinger, 2008) and correlates concurrently with parent-report and predictively with a language sample (Friend et al., 2012). Internal consistency is strong across forms (Cronbach's alpha = .86 and .93, respectively; Friend et al., 2018). It is the only measure of decontextualized vocabulary size prior to 30 months of age.

Peabody Picture Vocabulary Test (PPVT-III). The PPVT is an adaptive measure of vocabulary comprehension appropriate from 30 months of age through adulthood (Dunn & Dunn, 1997). Vocabulary comprehension is associated with subsequent language, literacy, and academic success (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Friend et al., 2018; Oakhill & Cain, 2012; Silva & Cain, 2015) and is therefore an important measure of 36-month language skill.

An experimenter displays four pictures and asks the child to point to the one that corresponds to a target word. Difficulty increases with age, and the final score is the number of items to reach ceiling minus errors. Like the CCT, the PPVT yields a direct estimate of vocabulary comprehension. It was standardized on a sample representative of the U.S. population and has generally strong reliability for all age ranges tested and strong internal consistency (Dunn & Dunn, 1997).

Free play language sample. Children and caregivers played with an extended Fisher-Price farm play set, which included several structures, vehicles, toy people, and animals, for 15 min. The full session was recorded with a Zoom H2n Handy Recorder microphone. This allowed us to assess spontaneous, as opposed to elicited, language usage at 36 months. Child language samples were transcribed, coded for grammatical morphemes, and analyzed for mean length of utterance (MLU) using the Systematic Analysis of Language Transcripts software (SALT; Miller & Iglesias, 2012). MLU reflects general language ability that is correlated with grammatical and semantic development (Dethorne, Johnson, & Loeb, 2005) and is lower in children with language impairment than in typically developing peers (Hewitt, Hammer, Yont, & Tomblin, 2005).

To ensure stability across transcripts varying in child talkativeness, we restricted MLU analysis to the first 100 complete and intelligible utterances. Thirty-five children (80%) met this criterion, leaving nine children whose transcripts included fewer than 100 utterances (M = 85.78, range = 58 to 97). Case-by-case review indicated no systematic difference in language skills from the larger sample. We retained these cases and calculated MLU in morphemes (MLU) over the entire transcript.

Five trained assistants transcribed 13 language samples each using Express Scribe Transcription Software (available at: http://www.nch.com.au/scribe/) and coded them for lexical units, plurals, articles, tense markers, possessives, and contractions. The assistants were trained using the SALT video training module to an interrater agreement of .90. Reliability checks were performed by four trained transcribers/coders for two to four transcripts from each transcriber for a total of 12 transcripts or approximately 20% of the full sample. Morpheme-level interrater agreement was .90.

Sentence repetition (SR). This task was based on Devescovi and Caselli (2007). The test included 27 sentences of varying complexity and length (see Table S2) accompanied by images depicting sentence-level meaning. SR improves significantly with age and correlates with concurrent spontaneous production. This elicited measure complements our measure of spontaneous production by tapping into diverse skills in language processing, sentence comprehension, production, and syntax (Klem et al., 2015). It is also a cross-linguistic marker for language impairment (Armon-Lotem & Meir, 2016; Conti-Ramsden, Botting, & Faragher, 2001).

The experimenter told the child to repeat after her. With the image covered, she modeled the sentence then revealed the image. Sentences were repeated up to three times. The child's first attempt at repetition was scored. The second author coded all SR data. The number of sentences repeated correctly ranged from 0 to 27. A correct repetition included all words and morphemes in the correct order with no extraneous words. Twenty-seven percent of the data were reliability-coded by one additional coder. Sentence-level interrater agreement was .92.

Procedure. Before each Wave, caregivers completed the LEAT interview over the phone. At the lab, following a brief warm-up period, caregivers and children were escorted to a playroom for testing. At Waves 1 and 2, this warm-up included a game to familiarize children with the touch-sensitive screen. Next, toddlers completed the CCT seated on their caregivers' laps approximately 30 cm from the screen; parents wore opaque sunglasses and listened to masking music over noise-cancelling headphones. Following the CCT, caregivers completed the MCDI. The WG form was completed at Wave 1, and the WS form was completed at Wave 2.

At Wave 3, dyads participated in three assessments: free play, PPVT, and SR. During free play, caregivers were instructed to play with their children as they would at home. Next, children were administered the PPVT and SR task. Free play always occurred first to facilitate optimal performance across tasks and the order of the other tasks was counterbalanced across participants.

Results

Descriptive data are reported below on the raw scores for all measures. English language exposure ranged from .87 to 1.00 (M = .99, SD = .03). At 16 months, expressive vocabulary on the MCDI ranged from 0 to 233 words (M = 43.80), corresponding to the 1st to the 99th percentile. Receptive vocabulary on the MCDI ranged from 63 to 355 words (M = 184.16), corresponding to the 1st to the 99th percentile. CCT receptive vocabulary ranged from 0 to 29 words (M = 12.11) and internal consistency was excellent (Cronbach's alpha = .92 and .95 for forms A and B, respectively). Twenty-eight children completed reliability trials, and immediate test–retest reliability was high ($r(26) = .79$, $p < .001$). Measures were approximately normally distributed with the exception of MCDI expressive vocabulary (*skewness* = 2.25, *kurtosis* = 5.51), indicating floor effects. A log transformation yielded equivalent results to the raw data in all analyses therefore raw data are presented.

At 23 months, MCDI expressive vocabulary ranged from 5 to 614 (M = 259.34), corresponding to the 1st to the 98th percentile. CCT receptive vocabulary ranged from 10 to 37 words (M = 27.93). Internal consistency was strong (Cronbach's alpha = .86 and .97 for forms A and B, respectively). Forty-one children completed reliability trials, and immediate test–retest reliability was moderate ($r(39) = .55$, $p < .001$). Both measures were approximately normally distributed.

At 36 months, PPVT receptive vocabulary ranged from 10 to 93 (M = 51.43), corresponding to the 2nd to the 99th percentile. MLU in morphemes ranged from 1.23 to 5.04 (M = 3.40), within the expected range at this age (Miller & Chapman, 1981). SR scores ranged from 0 to 26 (M = 15.48, SD = 7.51) and mirrored Devescovi and Caselli's (2007) findings. All measures were approximately normally distributed.

Control variables and zero-order correlations. We first evaluated the role of control variables (age, sex, and maternal education) on language skills. Maternal education served as a proxy for SES due to the relation between maternal education and early vocabulary (Hoff, 2013). Age was not significantly correlated with any predictors or dependent measures (all ps $> .18$). There was a negative correlation between sex and SR scores ($r(42) = -.32$, $p = .03$) and between sex and MLU ($r(42) = -37$, $p = .013$): boys performed slightly more poorly than girls. Maternal education correlated with CCT receptive vocabulary at 23 months ($r(42) = .31$, $p = .04$) and with SR at 36 months ($r(42) = .35$, $p = .02$). Both maternal education and child sex were included as control variables in subsequent analyses.

We evaluated zero-order relations between predictor and 36-month language skill variables to provide context for our predictive analyses (see Table 2). At 16 months, MCDI comprehension and production and the CCT were significantly correlated with SR at 36 months but no 16-month measure correlated with MLU or PPVT ($ps > .90$). At 23 months, MCDI production was significantly correlated with SR and MLU but not with PPVT at 36 months ($p = .09$) whereas the CCT correlated significantly with the PPVT, SR, and MLU.

Prediction to 36-month language skills. We transformed the PPVT, MLU, and SR to sample-specific z-scores and entered these into an exploratory factor analysis (see Table 3 for the component matrix). All participants contributed data for each indicator. A single factor explained 60.54% of the variance. This Language Factor was significantly correlated with 16- and 23-month MCDI production, $r(42) = .39$, $p = .009$ and $r(42) = .51$, $p < .001$, respectively, and 23-month CCT, $r(42) = .62$, $p < .001$. This composite is derived by removing unshared variance of the indicators to arrive at a more robust representation of broad language skill at 36 months (e.g., Bornstein et al., 2016).

We conducted two stepwise, hierarchical linear regressions to independently assess prediction at 16 and 23 months of age. We took this approach because the more proximal measures may suppress the predictive power of the earlier measures (Bornstein et al., 2016). We used backward selection to remove nonsignificant predictors sequentially. This permits the unique contribution of the remaining variables to be more accurately estimated. The criterion for the removal was $p > .10$. In the first model, Language Factor was entered as the dependent measure with child sex, maternal education, 16-month CCT comprehension, and 16-month MCDI comprehension and production entered as predictors. The final model, $F(2, 41) = 7.09$, $p = .002$ included child sex and MCDI production. Tolerance was excellent at .999. Observed power = .92 (G*Power; Faul et al., 2007). A follow-up t test of the effect of sex on Language Factor scores was not significant ($p = .05$). In the second model, child sex, maternal education, 23-month CCT comprehension, and 23-month MCDI production were entered using the stepwise method with the Language Factor as the dependent measure. Tolerance was good at .798. The final model, $F(1, 42) = 16.83$, $p < .001$ included CCT comprehension and MCDI production. Observed power = .99. See Table 4 for parameter estimates for the final model and excluded variables.

Table 3
Component Matrix of the Language Factor Extracted After the Factorial Analysis With Language Measures at 36 Months in Study 1

Measure	Component 1
PPVT comprehension 36 months	.711
SR sentences correct 36 months	.854
MLU morphemes 36 months	.763

Note. SR = Sentence repetition; PPVT = Peabody Picture Vocabulary Test; MLU = mean length of utterance. Extraction method: Principal component analysis. One component extracted.

To identify the most parsimonious model with the best fit, we contrasted all possible models from significant predictors identified in the stepwise regressions using the Akaike Information Criterion (AIC; Posada & Buckley, 2004, see Table 5). The AIC evaluates the loss of information in each model in approximating the data using maximum likelihood estimation and imposes a penalty for model complexity. Lower AIC scores are associated with higher quality. The lowest AIC values were obtained for 23-month CCT Comprehension alone (AIC = 108.49), 23-month MCDI production and CCT comprehension (AIC = 105.48), and the full complement of predictors at both ages (AIC = 107.03). For the current sample size, a 10-point spread in AIC scores would indicate a meaningful difference in the fit of the candidate models (Hilbe, 2011). Whereas these models cannot be distinguished in terms of fit, we can conclude that, of the candidate models, the one with 23-month CCT as the sole predictor provides the best balance of fit and parsimony.

Discussion

As predicted, decontextualized vocabulary comprehension at 23 months uniquely predicted language skills at 36 months. Parent-

Table 2
Bivariate Correlations for All Predictor and 36-Month Measures (z Scores) in Study 1 (N = 44)

Measure	1	2	3	4	5	6	7	8	9	10
1. Maternal education										
2. Sex	−.05									
3. MCDI comprehension 16 months	.09	−.06								
4. MCDI production 16 months	.21	−.03	.66**							
5. CCT comprehension 16 months	.18	.05	.38*	.28						
6. MCDI production 23 months	.17	−.23	.63**	.65**	.06					
7. CCT comprehension 23 months	.31*	−.24	.47**	.35*	.24[a]	.45*				
8. PPVT comprehension 36 months	.04	−.07	.12	.15	.24	.26	.45**			
9. SR sentences correct 36 months	.35*	−.32*	.34*	.48**	.35*	.50**	.67**	.44**		
10. MLU morphemes 36 months	.10	−.37*	−.07	.26	−.04	.41**	.31*	.27	.51**	

Note. MCDI = MacArthur-Bates Communicative Development Inventories; CCT = Computerized Comprehension Task; SR = Sentence repetition; PPVT = Peabody Picture Vocabulary Test; MLU = mean length of utterance.
[a] Three children improved dramatically on the CCT from Wave 1 to Wave 2; at Wave 1 they gave between zero and two correct responses, and at Wave 2 they gave between 31 and 35 correct responses. These children are not outliers, however they are visually distinct from the rest of the sample and, when they are removed from the data, the test–retest correlation for the CCT at Wave 1 and Wave 2 is significant consistent with previous reports, $r(41) = .360$, $p = .021$.
* $p < .05$. ** $p < .01$.

Table 4
Hierarchical Regression Parameters for 16- and 22-Month Models, Study 1 (N = 44)

Final model	Model 1 R^2	Model 1 *SE*	Model 1 β	Model 1 *p*	Model 2 R^2	Model 2 *SE*	Model 2 β	Model 2 *p*
Measure	.22			**.002**	.42			**<.001**
Sex		.27	−.32	**.022**				
MCDI production 16 months		.14	.38	**.007**				
CCT comprehension 23 months						.13	.49	**<.001**
MCDI production 23 months						.13	.29	**.032**

Excluded variables	Model 1 B^a	Model 1 *t*	Model 1 *p*	Model 2 *B*	Model 2 *t*	Model 2 *p*
Measure						
Maternal Education	.13	.93	.358	.02	.19	.848
Sex				−.16	−1.37	.177
CCT comprehension 16 months	.17	1.19	.241			
MCDI comprehension 16 months	−.16	−.89	.377			

Note. MCDI = MacArthur-Bates Communicative Development Inventories; CCT = Computerized Comprehension Task. Follow-up test on the effect of sex at 16 months was not significant ($p = .05$). *p*-values < .05 are bolded.
[a] Unstandardized *B* values.

reported vocabulary comprehension at 16 months of age was the earliest predictor of later language skills, however, by 23 months, decontextualized vocabulary offered the best balance of fit and parsimony consistent with our expectation of increased stability with age. Before assessing the practical significance of these findings, we first assess the generalizability of our findings in a sample of French-speaking children in Switzerland.

Study 2

Method

Participants. Sixty-six Swiss-French-speaking monolingual children (33 girls) were recruited through birth lists in a large city in Switzerland. Six children were excluded for not completing a task at one visit ($n = 4$) or attrition ($n = 2$; 3%). The final sample consisted of 60 toddlers (30 girls), all of whom had been carried to term and had normal hearing and vision. Participants made three visits to the lab: Wave 1 at 16 months ($M = 16;0$, range = 15;6–17;1), Wave 2 at 22 months ($M = 21;28$, range = 21;0–22;6), and Wave 3 at 36 months ($M = 35;25$, range = 34;8–37;2). See Table 6 for demographic data on the final sample.

Measures.

Language Exposure Assessment Tool (LEAT). Administration of the LEAT was identical to Study 1.

L'Inventaire Français du Développement Communicatif (IFDC). The IFDC (Kern, 2003, 2007; Kern & Gayraud, 2010) is the European-French adaptation of the MCDI. The IFDC: Mots et Gestes (MG) corresponds to the MCDI: WG and the IFDC: Mots et Phrases (MP) corresponds to the MCDI: WS. Vocabulary comprehension and production were estimated from the IFDC: MG at 16 months and production was estimated from the IFDC: MP at 22 months.

Computerized Comprehension Task (CCT). The French CCT was adapted from the English CCT. The design and administration were the same as Study 1. Translation equivalents across languages were included whenever possible while maintaining the same distribution of word class and difficulty (see Table S3). Images were prototypical exemplars in the region where children were tested. The French CCT has moderate test–retest reliability and convergent validity with the IFDC (Friend & Zesiger, 2011).

Table 5
AIC Values for Models Containing Combinations of the Significant Predictors From Regression Analyses in Study 1

Model	Model ID	AIC
Child Sex + 16-month MCDI Production	1	118.79
23-month MCDI production	2	116.68
23-month CCT comprehension	3	108.49
Child Sex + 16-month MCDI Production + 23-month MCDI Production	4	116.75
23-month MCDI Production + 23-month CCT Comprehension	5	105.48
Child Sex + 16-month MCDI Production + 23-month MCDI Production + 23-month CCT Comprehension	6	107.03

Note. MCDI = MacArthur-Bates Communicative Development Inventories; CCT = Computerized Comprehension Task; AIC = Akaike Information Criterion. Child sex was included only in models that contain 16-month variables because it did not reach significance in the 23-month model.

Table 6
Distribution of Selected Demographic Characteristics of Participants in Study 2

Characteristic	Number (%) of participants		
	Female	Male	Total
Maternal education			
High school or less	9 (15.0)	6 (10.0)	15 (25.0)
Some college	4 (6.7)	9 (15.0)	13 (21.7)
College graduate	3 (5.0)	3 (5.0)	6 (10.0)
Post-baccalaureate	14 (23.3)	12 (20.0)	26 (43.3)
Approximate income			
18,000–40,000	2 (3.3)	0 (0.0)	2 (3.3)
41,000–60,000	1 (1.7)	1 (1.7)	2 (3.3)
61,000–80,000	2 (3.3)	2 (3.3)	4 (6.7)
81,000–100,000	2 (3.3)	3 (5.0)	5 (8.3)
>100,000	10 (16.7)	10 (16.7)	20 (33.3)
Ethnicity			
Asian	0 (0.0)	0 (0.0)	0 (0.0)
Black/not Hispanic	1 (1.7)	2 (3.3)	3 (5.0)
Hispanic	1 (1.7)	0 (0.0)	1 (1.7)
White/not Hispanic	28 (46.7)	28 (46.7)	56 (93.3)
Mixed Race	0 (0.0)	0 (0.0)	0 (0.0)

Note. Twenty-seven participants declined to provide income information. Income reported in Swiss Francs. Some values may not sum to 100 because of rounding error.

Internal consistency is strong across forms (Cronbach's alpha = .92 and .91, respectively; Friend et al., 2018).

Échelle de Vocabulaire en Images Peabody (EVIP). The EVIP (Dunn, Thériault-Whalen, & Dunn, 1993) is the French adaptation of the PPVT normed on a large representative sample of French speakers in Canada.

Free play language sample. Administration of the free play language sample was identical to Study 1. Fifty-nine children (98%) met the criterion of 100 complete and intelligible utterances, leaving one child whose transcript contained only 89 utterances. We retained this case and calculated MLU over the entire transcript. Five trained assistants transcribed and coded child language samples. Morpheme-level interrater agreement for 25% of the total sample was .89.

Sentence repetition (SR). The SR was identical to Study 1, adapted to French. All SR data were reliability-coded. Sentence-level interrater agreement was .99. See Table S4.

Procedure. The procedure for Study 2 was identical to Study 1.

Results

Exposure to French ranged from .80 to 1.00 (M = .96 SD = .06). At 16 months, IFDC expressive vocabulary ranged from 0 to 185 words (M = 26.26) corresponding to the 5th to the 90th percentile. IFDC receptive vocabulary ranged from 52 to 387 words (M = 200.45), corresponding to the 5th to the 90th percentile. CCT receptive vocabulary ranged from 2 to 32 words (M = 15.95). The internal consistency of the CCT was excellent across forms (Cronbach's alpha = .92 and .90, respectively). Thirty-three children completed reliability trials and immediate test-retest stability was moderate, $r(31) = .54$, $p = .001$. Measures were approximately normally distributed with the exception of expressive vocabulary on the IFDC (*skewness* = 2.82, *kurtosis* = 9.46). A log transformation yielded equivalent results so subsequent analyses are reported on the raw data.

At 22 months, IFDC expressive vocabulary ranged from 13 to 523 (M = 196.40), corresponding to the 5th to the 90th percentile; CCT receptive vocabulary ranged from 12 to 40 words (M = 28.71). Internal consistency was strong across forms (Cronbach's alpha = .92 and .87, respectively). Fifty-seven children completed the reliability trials and immediate test–retest stability was moderate, $r(55) = .49$, $p < .001$. All measures were approximately normally distributed.

At 36 months, MLU from the language sample ranged from 2.18 to 5.82 (M = 4.03), EVIP receptive vocabulary ranged from 7 to 61 (M = 27.71), corresponding to the 2nd to the 99th percentile, and sentences correct on the SR task ranged from 0 to 25 (M = 10.96). Measures were approximately normally distributed.

Control variables and zero-order correlations. There was no relation between child sex or maternal education and language measures at any wave. However, to parallel Study 1, maternal education and child sex were included as control variables in the analyses. At 16 months, IFDC comprehension was significantly correlated with EVIP and MLU at 36 months, but not with SR (p = .26). IFDC production was not correlated with any 36 month variable ($ps > .53$) and CCT comprehension was significantly correlated with EVIP and SR, but not with MLU (p = .70). At 22 months, IFDC production was significantly correlated with SR but not with EVIP or MLU ($ps > .07$). CCT comprehension was significantly correlated with EVIP and SR, but not with MLU (p = .06; see Table 7).

Prediction to 36-month language skills. Following the procedure in Study 1, we computed a composite language score at 36 months (see Table 8 for the component matrix) using sample-specific z scores. A single factor explained 54.01% of the variance. This Language Factor was significantly correlated with 16-month IFDC, $r(58) = .36$, $p < .01$, and CCT comprehension, $r(58) = .27$, $p = .04$, and 22-month IFDC production, $r(58) = .32$, $p = .01$, and CCT comprehension, $r(58) = .56$, $p < .001$.

We conducted two stepwise, hierarchical linear regressions to independently assess prediction at 16 and 22 months of age. At 16 months, the final model included only MCDI comprehension, $F(1, 58) = 8.78$, $p = .004$ and at 22 months, the final model included only CCT comprehension, $F(1, 58) = 26.63$, $p < .001$ (see Table 9). Observed power was .84 and .99 for the final models in regressions 1 and 2, respectively. We contrasted all possible models using the significant predictors at 16 and 22 months by calculating the AIC values for each model (see Table 10). The lowest AIC values obtained for 22-month CCT comprehension (AIC = 152.60) and 16-month MCDI production and 22-month CCT comprehension (AIC = 147.18). As in Study 1, these models cannot be distinguished in terms of fit, but we can conclude that, of the candidate models, the one with 23-month CCT as the sole predictor provides the best balance of fit and parsimony.

Discussion

Results for the French sample largely paralleled those for English: parent-reported vocabulary was a stronger predictor of 36-month language skill than was decontextualized comprehension at 16 months, and decontextualized comprehension at 22 months was

Table 7
Bivariate Correlations for All Predictor and 36-Month Measures in Study 2 (z Scores; N = 60)

Measure	1	2	3	4	5	6	7	8	9	10
1. Maternal education										
2. Sex	−.012									
3. IFDC comprehension 16 months	−.147	−.004								
4. IFDC production 16 months	−.095	.048	.369**							
5. CCT comprehension 16 months	.023	−.061	.236[a]	.058						
6. IFDC production 22 months	.006	.089	.136	.390**	.122					
7. CCT comprehension 22 months	−.042	.185	.149	−.088	.360*	.250[a]				
8. EVIP comprehension 36 months	−.039	.087	.402**	.041	.273*	.194	.509**			
9. SR sentences correct 36 months	−.038	.043	.147	.082	.261*	.276*	.484**	.325*		
10. MLU morphemes 36 months	.022	−.088	.264*	.044	.052	.272	.243	.254	.349**	

Note. CCT = Computerized Comprehension Task; SR = Sentence repetition; MLU = mean length of utterance; EVIP = Échelle de Vocabulaire en Images Peabody; IFDC = L'Inventaire Français du Développement Communicatif.
[a] One outlier was discovered at 3 standardized residuals from the regression line describing the relation between the CCT and IFDC comprehension at 16 months. This parent had reported that the child knew all of the words on the IFDC, but this was not consistent with the child's performance on the CCT. With this outlier removed, the correlation between the CCT and the IFDC at each wave is significant consistent with previous reports, $r(59) = .355$, $p = .006$, and $r(59) = .405$, $p = .001$, respectively.
* $p < .05$. ** $p < .01$.

a stronger predictor than parent report. The only difference was that, in the English sample at 16 months of age, parent-reported vocabulary *production* predicted language skill whereas in the French sample, parent-reported *comprehension* was predictive. Consistent with our expectation, the balance of model fit and parsimony was superior for the model including only decontextualized vocabulary at 22 months.

Next we evaluate the practical significance of these findings for identifying children with low language skills at the individual level. Studies 1 and 2 each yielded a one-factor solution for the 36-month language variables with factor loadings that were remarkably similar suggesting that this factor has a similar underlying structure across samples. Therefore, we combine samples to take advantage of the increase in sample size.

Study 3

Method

Participants. All English-speaking monolingual children from Study 1 and French-speaking monolingual children from Study 2 were included in Study 3, resulting in a final sample of 104 children. Although a typically developing sample was recruited at 16 months, demographic data from the 22- and 36-month visits indicated that some children had received speech/language services or a diagnosis of language delay as a primary feature after the first visit. In the English sample, five children received services prior to 36 months of age for diagnosed phonological disorder ($n = 1$), expressive language delay, sensory processing disorder, and mild autism spectrum disorder (ASD; $n = 1$), and low expressive language with no other identified deficits ($n = 3$). In the French sample, three children received services prior to 36 months of age for universal dyslalia ($n = 1$), dysphasia ($n = 1$), and low expressive language ($n = 1$).

Table 8
Component Matrix of the Language Factor Extracted After the Factorial Analysis With Language Measures at 36 Months in Study 2

Measure	Component 1
EVIP comprehension 36 months	.703
SR sentences correct 36 months	.776
MLU morphemes 36 months	.724

Note. SR = Sentence repetition; MLU = mean length of utterance; EVIP = Échelle de Vocabulaire en Images Peabody. Extraction method: Principal component analysis. One component extracted.

Choice of Gold Standard. Consistent with our interest in discriminating children with low language skills at 36 months from their average-to-high language peers, we chose the Language Factor score as our "gold standard" for signal detection analysis to assess the practical significance of Studies 1 and 2 for individual children. Choosing a single gold standard is difficult (Eriksson, Westerlund, & Miniscalco, 2010; Heilmann et al., 2005; Westerlund et al., 2006). For example, spontaneous language is subject to contextual variation: length of session, time of day, and the conditions under which it was recorded can all influence the quality of the sample. On the other hand, standardized assessments may not capture the richness of child language and history of speech-language services can conflate children with primary language impairment with those who "catch up." We chose the Language Factor score as the gold standard because it takes into account vocabulary, grammar, and general language ability derived from both spontaneous language and standardized assessments. Vocabulary is associated with later language and literacy, and SR and MLU are recognized markers of language impairment. Finally, the Language Factor is more robust than any individual measure since the process of factor construction removes unshared variance.

Procedure. We used signal detection analysis to investigate the practical significance of decontextualized vocabulary in the second year for prospectively discriminating individual language skills. In this approach, a receiver operating characteristic (ROC) curve plots accuracy in identifying low language children against accuracy in identifying their average-to-high language peers. The ROC analysis was conducted in R (R Core Team, 2016) with the

Table 9
Hierarchical Regression Parameters for 16- and 22-Month Models, Study 2 (N = 60)

Final model	Model 1 R^2	Model 1 *SE*	Model 1 β	Model 1 *p*	Model 2 R^2	Model 2 *SE*	Model 2 β	Model 2 *p*
Measure	.12			**.004**	.30			**<.001**
MCDI comprehension 16 months		.13	.37	**.004**				
CCT comprehension 23 months						.11	.56	**<.001**

Excluded variables	Model 1 B^a	Model 1 *t*	Model 1 *p*	Model 2 *B*	Model 2 *t*	Model 2 *p*
Measure						
Maternal education	.03	.23	.816	−.001	−.01	.992
Sex	.02	.16	.871	−.09	−.79	.431
CCT comprehension 16 months	.19	1.54	.129			
MCDI production 16 months	−.07	−.50	.623			
MCDI production 23 months				.19	1.74	.087

Note. CCT = Computerized Comprehension Task; MCDI = MacArthur-Bates Communicative Development Inventories.
[a] Unstandardized *B* values are reported. *p*-values < .05 are bolded.

pROC script (Robin et al., 2011). The CCT at 22–23 months was entered as the predictor and score on the Language Factor at 36 months was the dependent measure. We evaluated the empirical ROC curve and a binormal-smoothed curve. Binormal smoothing serves to normally distribute scores separately for the low and average-to-high vocabulary groups (Robin et al., 2011). This transformation is robust, provides a good fit to the empirical data, and provides a better estimate of the area under the curve (AUC) than raw data, especially when there are few positive cases (i.e., low language). Estimated AUC and confidence intervals were performed on the smoothed curve.

The signal detection analysis yields several measures of discriminability. The measures of interest include AUC, sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR–). The AUC can be interpreted as the average sensitivity over all points on the curve. An AUC of .5 indicates no discrimination, whereas an AUC of 1 indicates perfect discrimination. Sensitivity refers to the ability of the predictor to accurately detect children who will have low language skills, whereas specificity refers to the ability to discriminate these children from their average-to-high language peers. If a test has a sensitivity of .70 and a specificity of .80, it accurately captures 70% of children with low language skills and correctly rejects 80% of children with average-to-high language skills.

Table 10
AIC Values for Models Containing Combinations of the Significant Predictors From Regression Analyses in Study 1

Model	Model ID	AIC
16-month MCDI comprehension	1	166.78
23-month CCT Comprehension	2	152.60
16-month MCDI Comprehension + 23-month CCT Comprehension	3	147.18

Note. CCT = Computerized Comprehension Task; MCDI = MacArthur-Bates Communicative Development Inventories; AIC = Akaike Information Criterion.

The positive likelihood ratio (LR+) is the likelihood that a child identified as low language by the predictor is also identified as low language at 36 months. For example, an LR+ of 5 indicates that a child identified as low language on the predictor is 5 times more likely to have low, than average-to-high, language skills at 36 months. Complementarily, LR– is the likelihood that a child identified as average-to-high language on the predictor has low language at 36 months. For example, an LR– of .2 indicates that a child identified as average-to-high language on the predictor variable is 1/5 times as likely to have a low, rather than average-to-high, language skills.

Children were classified as Low Language (LL) if their Language Factor score at 36 months was more than 1 *SD* below the mean and Average-to-High Language (HL) if their score was at or above 1 *SD* below the mean. This approach resulted in 13 LL children (eight French-speaking children and five English-speaking children; 12.5% of the total sample), and 91 HL children. This is within the expected incidence in the population (American Speech and Hearing Association, 2017). The LL sample had an average score of −1.55 *SD* below the mean (range = −2.81 to −1.03), whereas as the HL children had an average score of .22 *SD* above the mean (range = −0.99 to +2.36) relative to their sample-specific Language Factor.

Other classification cutoffs (e.g., 1.25 *SD* below the mean, 1.5 *SD* below the mean) were considered: first, we visually inspected the resulting ROC curve for smoothness and symmetry and considered the number of children who were identified as LL, which affected the ROC curve and the reliability of the discriminability measures. We also examined multiple indices of discriminability (e.g., sensitivity, specificity, LR+, and LR−) to find the cutoff with the best balance across indices. See Figure 1 for a graphical representation of the discriminability measures across cutoff points. Based on these considerations, we concluded that 1 *SD* below the mean provided the best balance of smoothness and symmetry with maximum discrimination across indices and a sufficient number of children classified as LL. With more stringent cutoffs, although estimates of sensitivity, specificity, and likelihood ratios appear better, they are less likely to be reliable: the

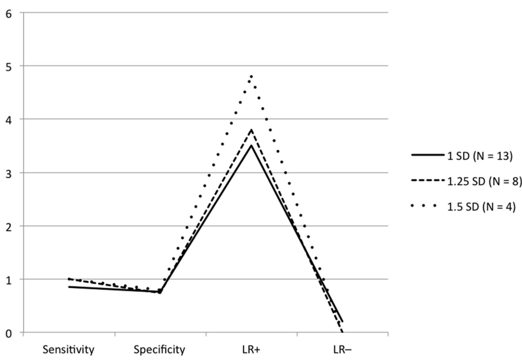


Figure 1. Measures of sensitivity, specificity, and positive and negative likelihood ratio for each of the potential cutoff determinations for low language skill. Whereas the parameters appear stronger at higher cutoffs, in fact these cutoffs eliminate many children with language skills sufficiently low to be markers of impairment. Further, the shape of the curve under these cutoffs suggests that these parameter estimates are unstable.

number of children identified is reduced and the curve becomes less smooth and symmetric.

Results

The AUC (see Figure 2) was .83 (95% confidence interval [.72, .94]), indicating good discriminability. Visual inspection suggested high specificity for low scores on the CCT. For example, a cutoff on the CCT predictor at approximately the 10th percentile (less than 21 words correct) yielded a specificity of .95 and a LR+ of 5.6, indicating that a child who knew fewer than 21 words on the CCT was 5.6 times more likely to be LL at 36 months than HL. However, sensitivity and LR– at this cutoff were inadequate at .31 and .73, respectively. On the other side of the curve, higher scores on the CCT evinced excellent sensitivity. For example, a cutoff on the CCT predictor at approximately the 50th percentile (fewer than 29 words correct) yielded a sensitivity of .92 and a LR– of .12. At this cutoff, all but one LL child at 36 months were captured. However, specificity and LR+ were insufficient at .62 and 2.4, respectively. Because there is no prior evidence to suggest an optimal cutoff, we examined two statistics to maximize the balance of sensitivity and specificity.

First, we determined the cutoff on the CCT that produces the optimal balance of sensitivity and specificity using Youden's J (Youden, 1950). The cutpoint was 26.5 words (of 41), which yielded a sensitivity of .85 and specificity of .76. The LR+ was 3.5, indicating that children who knew fewer than 26.5 words were 3.5 times as likely to be LL at 36 months than to be HL. Complementarily, the LR– was .20, indicating that children who knew more than 26.5 words were 1/5 times as likely to be LL at 36 months than to be HL. Second, we evaluated the point closest to the top left corner of the ROC plot (perfect sensitivity and specificity). This yields an optimal cutoff of 24.5 words, with a sensitivity of .77 and a specificity of .84. The LR+ was 4.67, and the LR– was .28. These statistics varied primarily in the relative balance of sensitivity and specificity.

Using the first criterion, CCT scores at 22 months correctly identified four of five children referred for services in the English sample and one of three children in the French sample. Using the second criterion, CCT scores at 22 months correctly identified three of five children referred for services in the English sample and one of three children in the French sample. The one child who was consistently not identified in the English sample received services prior to 24 months of age but had language skills that were above average at 36 months. Of the two children not identified in the French sample, one received services for articulation difficulties and had language skills at the classification borderline (1 *SD* below the mean). The other was diagnosed with expressive, but not receptive, delay. However, his score on the Language Factor at 36 months was 2 standard deviations below the mean, indicating low language skills relative to his peers. This child was a true "miss" in signal detection terms.

To situate these results within the literature, we also conducted a signal detection analysis with MCDI production as the predictor. We evaluated both the empirical ROC curve and a binormal-smoothed curve. The AUC was .77 (95% confidence interval = .63 to .88), indicating fair discriminability. Next, we determined the cutoff on the MCDI that produces the optimal balance of sensitivity and specificity using Youden's J (Youden, 1950). The cutpoint was 119.5 words (between the 30th and 35th percentiles on the MCDI and approximately the 40th percentile on the IFDC), with a sensitivity of .74 and specificity of .77. The LR+ was 2.92 and the LR– was .31. The point closest to the top left corner of the ROC plot yielded the same cutoff (see Figure 3). It is noteworthy that this cutoff is considerably higher than cutoffs based on normative considerations, resulting in improved sensitivity but poorer specificity relative to previous reports (e.g., Heilmann et al., 2005). Because the AIC model contrasts suggested that the 22-month model including both the MCDI and the CCT was equivalent in strength at the group level to the model containing only the CCT,

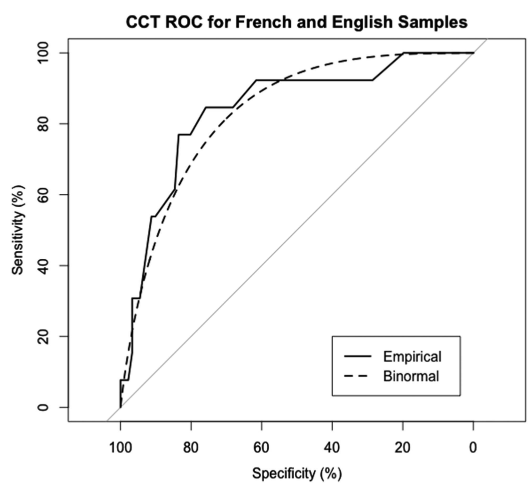


Figure 2. Receiver operating characteristic (ROC) plot for empirical data and binormal-smoothed ROC curves from Study 3 using the Computerized Comprehension Task (CCT) predictor (N = 104).

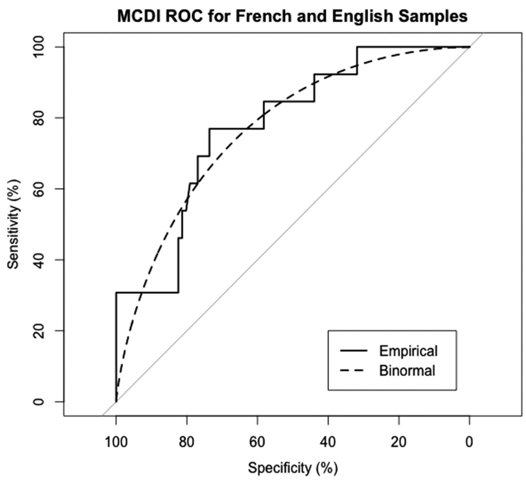


Figure 3. Receiver operating characteristic (ROC) plot for empirical data and binormal-smoothed ROC curves from Study 3 using the MacArthur-Bates Communicative Development Inventories (MCDI) predictor (N = 104).

we repeated this analysis using a composite MCDI/CCT approach. Children were classified based on whether they fell below the statistically determined cutoff on both measures. This approach yielded an improvement in specificity (90% to 93%, depending on the CCT criterion), but a reduction in sensitivity (61% to 69%).

Discussion

This study represents the first attempt to estimate the sensitivity and specificity of an assessment of decontextualized vocabulary for prospectively identifying children whose language at 36 months is a full standard deviation below their peers. These were typically developing samples at the time of recruitment with no known risks of language impairment. Therefore, our findings should be interpreted as a preliminary indication of the practical significance of this approach.

Sensitivity and specificity were moderately strong, although their relative strength depended on the cutpoint. In general, a cutoff between about 24 and 26 words comprehended (of 41) yielded the best balance of performance. This corresponded well to both LR+ and LR− estimates. Sensitivity ranged between .76 and .85 and specificity ranged between .77 and .84. Vocabulary size on the CCT at 22 months correctly identified 77% to 85% of the "true" cases of low language and correctly rejected 76% to 84% of cases of average-to-high performance at 36 months of age. Compared with the MCDI/IFDC, the CCT yielded either higher sensitivity or specificity depending on the cutoff. Finally, sensitivity for the CCT was generally better than in prospective studies with other measures (Frisk et al., 2009; Klee et al., 1998; Klee, Pearce, & Carson, 2000; McIntyre et al., 2017; McKean et al., 2016, 2017; Stott, Merricks, Bolton, & Goodyer, 2002; Westerlund et al., 2006). Our findings are most comparable with McIntyre et al. (2017) and McKean et al. (2017), with the notable exception that we obtain comparable sensitivity and specificity a full two years earlier. Thus a significant empirical contribution is a potential screen for language difficulties that can be used as early as the second year.

General Discussion

This research follows from recent efforts to predict developmental achievements from early vocabulary (Duff et al., 2015; Kemp et al., 2017; Morgan et al., 2015; Reilly et al., 2010). In prior work, effect sizes are generally modest and prediction at the individual level is weak (e.g., Westerlund et al., 2006). Thus, an overarching goal was to overcome these limitations. With this in mind, this paper addresses three primary aims: to predict preschool language skills in typically developing children from a measure of decontextualized vocabulary, to contrast this measure with parent report of comprehension at 16 months and production at 22 months, and to determine how early predictions to preschool language can be made.

Predicting Preschool Language

Both parent-reported and decontextualized vocabulary were associated with preschool language. At 16 months of age parent-reported production was the strongest predictor of 36-month skills in English, whereas parent-reported comprehension was the strongest predictor in French. This finding reflects variation across samples in the underlying pattern of correlation between parent-reported comprehension and production in the second year and diverse metrics of language ability (vocabulary, MLU, sentence processing) in the third year.

By 22 months, decontextualized vocabulary accounted for more variance in preschool language skills and provided a better balance of fit and parsimony than parent report in both French and English. The unique variance in language skills accounted for at the group level was 20% to 25%: four to five times that reported for large-scale studies using parent-reported vocabulary (Duff et al., 2015; Morgan et al., 2015; Reilly et al., 2010).

This observed stability may lie in underlying domain general processes rather than in language itself. This follows from the idea that a characteristic (e.g., language) might appear stable because some other characteristic mediates the relation between measures of language at different points in time (Bornstein et al., 2016). Simple associative processes, which themselves rely on attention and memory, can account for cross-situational word learning (Yu & Smith, 2012) and, hypothetically, yield graded word-referent associations that are strengthened over time and situations. When faced with a choice between a target referent and a perceptually similar distractor (i.e., color, size, saliency) from the same word class and conceptual category, weak, context-bound associations are not sufficient to elicit a correct response. By estimating the number of generalized word-referent relations, decontextualized vocabulary reflects the efficiency of domain general learning. By 22 months of age, the word–world relations that children recognize beyond the context in which they were acquired predict downstream language skills.

Stronger prediction at 22 relative to 16 months is in line with previous work: prediction to language samples has been reported by 18–24 months of age but not earlier (Friend et al., 2012;

Westerlund et al., 2006), paralleling an acceleration in word learning late in the second year (Samuelson & McMurray, 2017). This may reflect richer semantic organization with stronger word–world relations that facilitate retrieval of word meaning late in the second year. Alternatively, better prediction at 22 relative to 16 months may be a function of proximity to the 36-month measures (Bornstein et al., 2016) although previous research suggests stability in decontextualized vocabulary from the second year through at least the beginning of the fourth year (Friend et al., 2018).

Finally, decontextualized vocabulary evinced practical significance for prospectively identifying children with low language skills. Sensitivity and specificity compared favorably with other prospective studies (Frisk et al., 2009; Klee et al., 1998, 2000; Stott et al., 2002; Westerlund et al., 2006; Wetherby, Goldstein, Cleary, Allen, & Kublin, 2003), overcoming the longstanding difficulty of harnessing early vocabulary prior to 30 months of age to predict development at the group and individual levels. This is the first paper to show strong prospective sensitivity as early as two years of age with implications for the early assessment of developmental risk.

Limitations and Directions for Future Research

First, whereas a strength of this paper is replication across two distinct samples, in Study 3, we collapsed across samples to obtain stable estimates of sensitivity, specificity, and likelihood ratios. We expect these estimates to be consistent across languages, but this remains to be empirically tested. Second, we focused on prediction in a sample with no known risk factors thus we identified only a small number of children with low language skills at age three. Whereas the proportion of children identified is consistent with the population incidence, future work is needed to establish norms and cut points to estimate sensitivity and specificity in high-risk samples (e.g., children from families with low-income/a history of language difficulties). Third, although one could question the component structure of the Language Factor, average scores on the component measures suggest that it adequately classifies low- relative to higher-language children (please see Table S5 and Figures S1 and S2). Finally, although we cannot know how well our research generalizes to other languages or language families, the fact that our findings parallel each other so well in English and in French is encouraging.

References

American Speech and Hearing Association. (2017). *Spoken language disorders*. Practice Portal [On-line]. Retrieved from https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589935327&section=Overview

Armon-Lotem, S., & Meir, N. (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: Evidence from Russian and Hebrew. *International Journal of Language & Communication Disorders, 51,* 715–731. http://dx.doi.org/10.1111/1460-6984.12242

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word -object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition, 126,* 39–53. http://dx.doi.org/10.1016/j.cognition.2012.08.008

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics, 37,* 1461–1476. http://dx.doi.org/10.1017/S0142716416000060

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of Child Language, 35,* 619–650. http://dx.doi.org/10.1017/S0305000908008714

Bornstein, M. H., Hahn, C. S., & Putnick, D. L. (2016). Long-term stability of core language skill in children with contrasting language skills. *Developmental Psychology, 52,* 704–716. http://dx.doi.org/10.1037/dev0000111

Bornstein, M. H., Putnick, D. L., & Esposito, G. (2017). Continuity and stability in development. *Child Development Perspectives, 11,* 113–119. http://dx.doi.org/10.1111/cdep.12221

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42,* 741–748. http://dx.doi.org/10.1111/1469-7610.00770

Dale, P. S., Price, T. S., Bishop, D. V., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research, 46,* 544–560. http://dx.doi.org/10.1044/1092-4388(2003/044)

Dapretto, M., & Bjork, E. L. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development, 71,* 635–648. http://www.jstor.org/stable/1132382. http://dx.doi.org/10.1111/1467-8624.00172

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). The language exposure assessment tool: Quantifying language exposure in infants and children. *Journal of Speech, Language, and Hearing Research, 59,* 1346–1356. http://dx.doi.org/10.1044/2016_JSLHR-L-15-0234

Dethorne, L. S., Johnson, B. W., & Loeb, J. W. (2005). A closer look at MLU: What does it really measure? *Clinical Linguistics & Phonetics, 19,* 635–648. http://dx.doi.org/10.1080/02699200410001716165

Devescovi, A., & Caselli, M. C. (2007). Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language & Communication Disorders, 42,* 187–208. http://dx.doi.org/10.1080/13682820601030686

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher, 39,* 305–310. http://dx.doi.org/10.3102/0013189X10370204

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56,* 848–856. http://dx.doi.org/10.1111/jcpp.12378

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.

Dunn, L. M., Thériault-Whalen, C. M., & Dunn, L. M. (1993). *Échelle de vocabulaire en images Peabody: Série de planches* [Peabody Picture Vocabulary Scale: Series of Boards]. Toronto, Ontario, Canada: Psycan.

Eriksson, M., Westerlund, M., & Miniscalco, C. (2010). Problems and limitations in studies on screening for language delay. *Research in Developmental Disabilities, 31,* 943–950. http://dx.doi.org/10.1016/j.ridd.2010.04.019

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur-Bates communicative development inventories*. Baltimore, MD: Brookes.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59,* 1–173. http://dx.doi.org/10.2307/1166093

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83,* 203–222. http://dx.doi.org/10.1111/j.1467-8624.2011.01692.x

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language, 44,* 677–694. http://dx.doi.org/10.1017/S0305000916000209

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers, 35,* 302–309. http://dx.doi.org/10.3758/BF03202556

Friend, M., & Keplinger, M. (2008). Reliability and validity of the computerized comprehension task (CCT): Data from American English and Mexican Spanish infants. *Journal of Child Language, 35,* 77–98. http://dx.doi.org/10.1017/S0305000907008264

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the computerized comprehension task: Comprehension predicts production. *Developmental Psychology, 48,* 136–148. http://dx.doi.org/10.1037/a0025511

Friend, M., Smolak, E., Liu, Y., Poulin-Dubois, D., & Zesiger, P. (2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. *Developmental Psychology, 54,* 1317–1333. http://dx.doi.org/10.1037/dev0000514

Friend, M., & Zesiger, P. (2011). A systematic replication of the psychometric properties of the CCT in three languages: English, Spanish, and French. *Enfance, 3,* 329–344. http://dx.doi.org/10.4074/S0013754511003053

Frisk, V., Montgomery, L., Boychyn, E., Young, R., McLachlan, D., & Neufeld, J. (2009). Why screening Canadian preschoolers for language delays is more difficult than it should be. *Infants & Young Children, 22,* 290–308. http://dx.doi.org/10.1097/IYC.0b013e3181bc4db6

Gershkoff-Stowe, L., & Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. *Journal of Experimental Child Psychology, 114,* 489–509. http://dx.doi.org/10.1016/j.jecp.2012.11.005

Ghassabian, A., Rescorla, L., Henrichs, J., Jaddoe, V. W., Verhulst, F. C., & Tiemeier, H. (2014). Early lexical development and risk of verbal and nonverbal cognitive delay at school age. *Acta Paediatrica, 103,* 70–80. http://dx.doi.org/10.1111/apa.12449

Heilmann, J., Weismer, S. E., Evans, J., & Hollar, C. (2005). Utility of the MacArthur-Bates communicative development inventory in identifying language abilities of late-talking and typically developing toddlers. *American Journal of Speech-Language Pathology, 14,* 40–51. http://dx.doi.org/10.1044/1058-0360(2005/006)

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,* 723–735. http://dx.doi.org/10.1111/desc.12250

Hendrickson, K., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2017). Assessing a continuum of lexical-semantic knowledge in the second year of life: A multimodal approach. *Journal of Experimental Child Psychology, 158,* 95–111. http://dx.doi.org/10.1016/j.jecp.2017.01.003

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W., Hofman, A., . . . Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The generation R study. *Journal of Speech, Language, and Hearing Research, 54,* 854–869. http://dx.doi.org/10.1044/1092-4388(2010/09-0255)

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,* 197–213. http://dx.doi.org/10.1016/j.jcomdis.2004.10.002

Hilbe, J. M. (2011). *Negative binomial regression*. New York, NY: Cambridge University Press.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 105–124). Cambridge, MA: MIT Press.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology, 49,* 4–14. http://dx.doi.org/10.1037/a0027238

Kemp, N., Scott, J., Bernhardt, B. M., Johnson, C. E., Siegel, L. S., & Werker, J. F. (2017). Minimal pair word learning and vocabulary size: Links with later language skills. *Applied Psycholinguistics, 38,* 289–314. http://dx.doi.org/10.1017/S0142716416000199

Kern, S. (2003). Le compte-rendu parental au service de l'évaluation de la production lexicale des enfants français entre 16 et 30 mois [Parent report in service of the evaluation of the lexical production of French children between 16 and 30 months]. *Glossa, 85,* 48–62.

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language, 27,* 227–250. http://dx.doi.org/10.1177/0142723706075789

Kern, S., & Gayraud, F. (2010). *L'inventaire français du développement communicatif* [The French Communicative Development Inventory]. Grenoble, France: La Cigale.

Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research, 41,* 627–641. http://dx.doi.org/10.1044/jslhr.4103.627

Klee, T., Pearce, K., & Carson, D. K. (2000). Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research, 43,* 821–833. http://dx.doi.org/10.1044/jslhr.4304.821

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science, 18,* 146–154. http://dx.doi.org/10.1111/desc.12202

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). The feasibility of universal screening for primary speech and language delay: Findings from a systematic review of the literature. *Developmental Medicine & Child Neurology, 42,* 190–200. http://dx.doi.org/10.1017/S0012162200000335

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the communicative development inventories. *Child and Adolescent Mental Health, 13,* 198–206. http://dx.doi.org/10.1111/j.1475-3588.2008.00503.x

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11,* F9–F16. http://dx.doi.org/10.1111/j.1467-7687.2008.00671.x

Mayor, J., & Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning. *Psychological Review, 117,* 1–31. http://dx.doi.org/10.1037/a0018130

McIntyre, L. L., Pelham, W. E., III, Kim, M. H., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2017). A brief measure of language skills at 3 years of age and special education use in middle childhood. *The Journal of Pediatrics, 181,* 189–194. http://dx.doi.org/10.1016/j.jpeds.2016.10.035

McKean, C., Law, J., Mensah, F., Cini, E., Eadie, P., Frazer, K., & Reilly, S. (2016). Predicting meaningful differences in school-entry language skills from child and family factors measured at 12 months of age. *International Journal of Early Childhood, 48,* 329–351. http://dx.doi.org/10.1007/s13158-016-0174-0

McKean, C., Reilly, S., Bavin, E. L., Bretherton, L., Cini, E., Conway, L., . . . Mensah, F. (2017). Language outcomes at 7 years: Early predictors and co-occurring difficulties. *Pediatrics, 139,* e20161684. http://dx.doi.org/10.1542/peds.2016-1684

Miller, J. F., & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech & Hearing Research, 24,* 154–161. http://dx.doi.org/10.1044/jshr.2402.154

Miller, J., & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT), Research (Version 2012) [Computer Software]. Middleton, WI: SALT Software, LLC.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development, 86,* 1351–1370. http://dx.doi.org/10.1111/cdev.12398

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16,* 91–121. http://dx.doi.org/10.1080/10888438.2010.529219

Posada, D., & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology, 53,* 793–808. http://dx.doi.org/10.1080/10635150490522304

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., . . . Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from Early Language in Victoria Study. *Pediatrics, 126,* e1530–e1537. http://dx.doi.org/10.1542/peds.2010-0254

Reznick, J. S., & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology, 28,* 406–413. http://dx.doi.org/10.1037/0012-1649.28.3.406

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12,* 77. http://dx.doi.org/10.1186/1471-2105-12-77

Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word? *WIREs Cognitive Science, 8,* e1421. http://dx.doi.org/10.1002/wcs.1421

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64,* 913–951. http://dx.doi.org/10.1111/lang.12077

Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology, 107,* 321–331. http://dx.doi.org/10.1037/a0037769

Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development, 9,* 25–49. http://dx.doi.org/10.1080/15475441.2012.707104

Stott, C. M., Merricks, M. J., Bolton, P. F., & Goodyer, I. M. (2002). Screening for speech and language disorders: The reliability, validity and accuracy of the General Language Screen. *International Journal of Language & Communication Disorders, 37,* 133–151. http://dx.doi.org/10.1080/13682820110116785

Thordardottir, E. T. (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. *International Journal of Language & Communication Disorders, 40,* 243–278. http://dx.doi.org/10.1080/13682820410001729655

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of memory and language, 93,* 217–230.

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition, 127,* 375–382. http://dx.doi.org/10.1016/j.cognition.2013.02.015

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology, 108,* 394–401.

Wallace, I. F., Berkman, N. D., Watson, L. R., Coyne-Beasley, T., Wood, C. T., Cullen, K., & Lohr, K. N. (2015). Screening for speech and language delay in children 5 years old and younger: A systematic review. *Pediatrics, 136,* e448–e462. http://dx.doi.org/10.1542/peds.2014-3889

Westerlund, M., Berglund, E., & Eriksson, M. (2006). Can severely language delayed 3-year-olds be identified at 18 months? Evaluation of a screening version of the MacArthur-Bates Communicative Development Inventories. *Journal of Speech, Language, and Hearing Research, 49,* 237–247. http://dx.doi.org/10.1044/1092-4388(2006/020)

Wetherby, A. M., Goldstein, H., Cleary, J., Allen, L., & Kublin, K. (2003). Early identification of children with communication disorders: Concurrent and predictive validity of the CSBS Developmental Profile. *Infants and Young Children, 16,* 161–174. http://dx.doi.org/10.1097/00001163-200304000-00008

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3,* 32–35. http://dx.doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word–referent learning: Prior questions. *Psychological Review, 119,* 21–39. http://dx.doi.org/10.1037/a0026182

Yurovsky, D., Smith, L. B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science, 16,* 959–966. http://dx.doi.org/10.1111/desc.12036

Received June 12, 2017
Revision received May 21, 2018
Accepted July 23, 2018 ■

Language Status at Age 3: Group and Individual Prediction from Vocabulary Comprehension in the Second Year

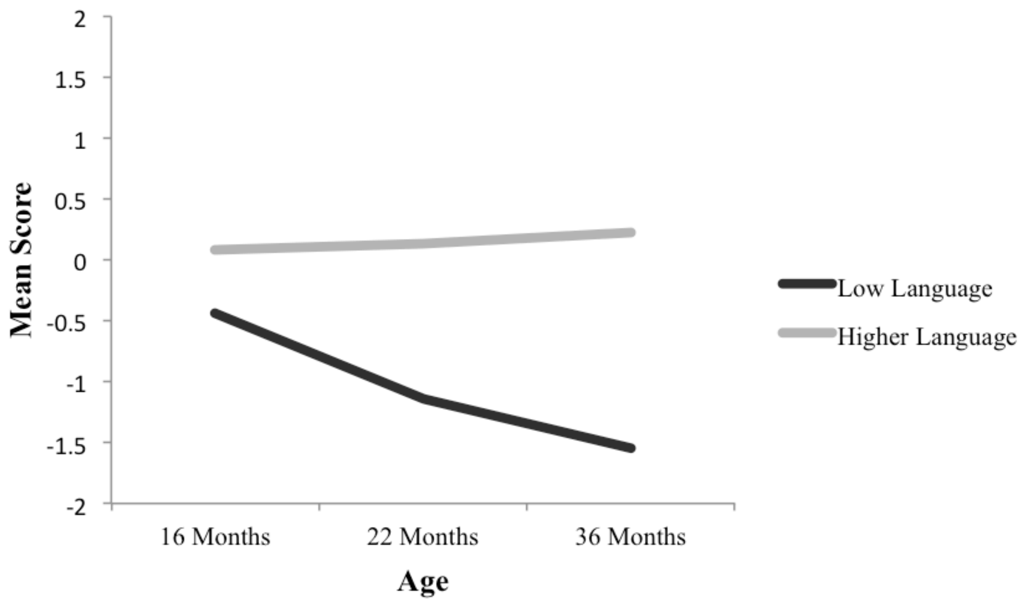
Supplemental Material

The following table and figures provide additional detail on performance across measures of the low- and average-to-higher language groups defined in Study 3.

Table S1. Scores on 36-month measures for Higher Language (*N* = 91) and Low Language (*N* = 13) groups.

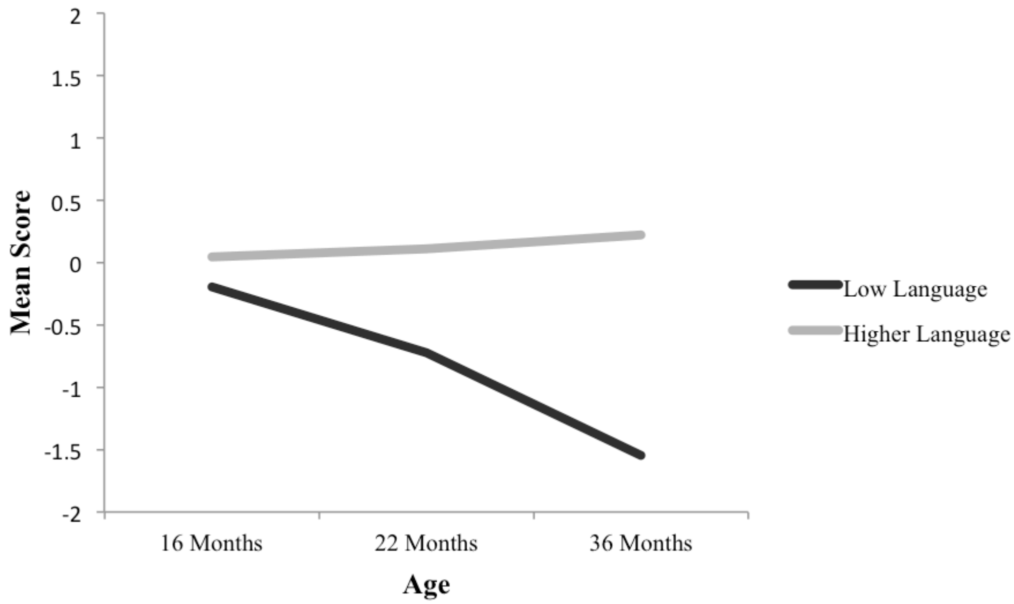
	PPVT Vocabulary Comprehension	Sentence Repetition	Mean Length of Utterance - Morphemes
Higher Language	39.88	14.22	3.92
Low Language	22.85	3.46	2.69

Figure S1. CCT Comprehension (16 and 22 months) and Oral Language Factor (36 months) z-scores for Low Language ($N = 13$) and Higher Language ($N = 91$) groups.



Note: The low-language group falls further behind their peers over time, even on the same measure at 16 and 22 months of age. This gap is wider for the CCT than for the MCDI/IFDC.

Figure S2. MCDI/IFDC Production (16 and 22 months) and Oral Language Factor (36 months) z-scores for Low Language ($N = 13$) and Higher Language ($N = 91$) groups.



Note: The low-language group falls further behind their peers over time, even on the same measure at 16 and 22 months of age. This gap is wider for the CCT than for the MCDI/IFDC.

Chapter 3, in full, is a reprint of the material as it appears in Developmental Psychology. Friend, M., **Smolak, E.,** Patrucco-Nanchen, T., & Zesiger, P. (2019). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year. *Developmental Psychology, 55,* 9-22. doi: 10.1037/dev0000617. The dissertation author was the co-investigator and co-author of this paper.

Note. Because all analyses in Study 1 and 2 were conducted on standard scores, unstandardized-beta and standardized-beta are the same.

CHAPTER 4

Predicting Preschool Oral Language: Convergence and Divergence in Vocabulary and Speed of Lexical Access

Introduction

Empirical evidence has shown that both vocabulary size and speed of word processing (hereafter speed of lexical access) develop rapidly over the second year of life (Fenson et al., 1994; Fernald, Perfors, & Marchman, 2006; Fernald, Pinto, & Swingley, 1998; Samuelson & McMurray, 2017). Vocabulary refers to the number of word-referent relations that children know, whereas speed of lexical access refers to how quickly children access those word-referent relations. Vocabulary and speed of lexical access are correlated such that children with larger vocabularies access the meanings of words more efficiently (DeAnda, Hendrickson, Zesiger, Poulin-Dubois, & Friend, 2018; Fernald & Marchman, 2012; Fernald et al., 2006; Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015; Hendrickson, Poulin-Dubois, Zesiger, & Friend, 2017; Legacy, Zesiger, Friend, & Poulin-Dubois, 2018; Marchman, Fernald, & Hurtado, 2010). The directionality of this effect remains unclear: one possibility is that growth in vocabulary size leads to more connected and efficient semantic networks, increasing the speed of lexical access (Borovsky, Ellis, Evans, & Elman, 2016). Alternatively, as children process familiar words more quickly, they have increased resources to devote to unknown word meanings within the speech stream, thus increasing the efficiency with which they can learn new words.

Both vocabulary size (Bleses, Makransky, Dale, Højen, & Ari, 2016; Duff, Reen, Plunkett, & Nation, 2015; Friend, Schmitt, & Simpson, 2012; Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018; Friend, Smolak, Patrucco-Nanchen & Zesiger, 2019; Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015) and speed of lexical access (Fernald & Marchman, 2012; Fernald et al., 2006; Marchman & Fernald, 2008) assessed within the first three years of life predict language and cognitive development later in childhood. However, despite the fact that vocabulary size and lexical access are correlated, we hypothesize that they also partially diverge in the extent to which they rely on long-term knowledge compared to processing/working memory systems. That is, the development of word-referent knowledge is hypothesized to depend on the iterative application of learning mechanisms such as cross-situational associative learning and hypothesis testing (Smith & Yu, 2008; Reuter, Borovsky, & Lew-Williams, 2019; Vlach & DeBrock, 2017; Yu & Smith 2012). Increasing experience with a word and its referent builds a robust representation of the meaning of the word that is long-term and resistant to decay or interference (Bion, Borovsky & Fernald, 2013; Horst & Samuelson, 2008). In contrast, speed of processing is highly associated with other processing skills and working memory (Fry & Hale, 1996; Kail & Salthouse, 1994). Of course, it is also the case that vocabulary, processing, and working memory are all linked (Baddeley & Hitch, 1974; Marchman & Fernald, 2008; Stokes & Klee, 2009; Vlach & DeBrock, 2017) however we hypothesize that links are stronger between speed of lexical access and processing/working memory components and between vocabulary and long-term knowledge than between processing components and long-term knowledge components. If this is the case, then models of prediction from to downstream abilities that are differentially dependent on processing and long-term memory abilities are likely to differ.

Only one extant study to our knowledge has prospectively compared these components of language development (Marchman & Fernald, 2008) and in that case, vocabulary and speed of word processing each explained unique variance in working memory and language development at age eight but only vocabulary size explained unique variance in IQ at age eight. Further, the prediction from early vocabulary and speed of word processing to age-eight language outcomes was fully mediated by concurrent age-eight working memory. The current study seeks, in contrast, to evaluate the potential for divergent longitudinal models of prediction from early vocabulary size and speed of lexical access to childhood language skills. Additionally, we assess generalizability of these models by estimating prediction across two linguistically and geographically distinct languages.

The Relation Between Vocabulary Size and Lexical Access

Although the relation between vocabulary size and lexical access is generally considered to be bidirectional, few studies have directly investigated this relation. Indirectly, one study of English-Spanish bilingual language development found that children exposed to a larger relative quantity of parent language input in English were quicker to process familiar words and knew more words in English. The same pattern of results was obtained for Spanish input and Spanish vocabulary and word processing. Further, lexical access predicted vocabulary development at 30 and 36 months of age above and beyond the prediction of language exposure (Hurtado, Grüter, Marchman, & Fernald, 2014). This result was paralleled in a study of monolingual children, in which it was found through mediation analyses that the effect of parental input on vocabulary development was mediated by processing efficiency (Weisleder & Fernald, 2013). Thus these studies suggest a primary directional relation such that input supports lexical access, which in turn supports the building of vocabulary knowledge. More recently however, Borovsky and colleagues (2016) found that semantic density (operationalized as the number of words parents reported that children knew in a given semantic category) predicted lexical access. Children were faster to orient to the referent after hearing a target word when the density of words known within the target word category was relatively high and slower to orient when semantic density was sparse. For instance, if a child knew many food words but few clothing words, this child would process the meanings of known words in the food category faster than the meanings of known words in the clothing category. The results of this study therefore suggest that semantic organization, rather than vocabulary size *per se*, influences lexical access but points to a different directional relation such that vocabulary size facilitates lexical access rather than vice versa. Directionality of the effect notwithstanding, it is clear from this research that vocabulary acquisition and lexical access are related throughout the first two years of life.

However, despite this, the shared variance between vocabulary and speed of lexical access is only about 20 percent (e.g., DeAnda et al., 2018). Even with the consideration of method variance, it is unlikely that these abilities are completely redundant. Therefore, they may be considered related, but separable, components of early language ability. We hypothesize that these components represent two ends of a continuum from short-term memory and speed of symbolic processing (lexical access) to long-term semantic knowledge (vocabulary). Consider that the incoming speech stream is processed and manipulated in working memory prior to and during access to word meaning. Indeed, processing speed is highly related to working memory capacity (Kail & Salthouse, 1994). An early report on this topic found that most of the developmental changes in spatial and verbal working memory throughout childhood and adolescence can be explained by increases in information processing speed (Fry & Hale, 1996). In the language domain, speed of lexical access in toddlers predicts childhood working memory at age eight (Marchman & Fernald, 2008).

In order to develop stable word-referent representations, in contrast, it is necessary for children to encounter many word-referent pairings (e.g., Yu & Smith, 2012; Vlack & DeBrock, 2017). In this way, word meanings are acquired and their representations become more robust with time and experience (Yu & Smith, 2012). Although the development of word knowledge also depends on processing and working memory, recent evidence suggests that it is robust, decontextualized knowledge built through multiple learning experiences that best supports language and cognitive development (Friend et al., 2018; Friend et al., 2019). Thus, one way in which vocabulary and lexical access may be separated is to consider their dependence on different components of the memory system. This has yet to be fully explored, leaving open the possibility that vocabulary size and lexical access differentially predict later linguistic and cognitive skills.

Long-Term Prediction

Both vocabulary size (Bleses et al., 2016; Duff et al., 2015; Friend et al., 2012, 2018; Henrichs et al., 2011; Lee, 2011; Morgan et al., 2015; Reilly et al., 2010) and lexical access (Fernald & Marchman, 2012; Fernald et al., 2006; Marchman & Fernald, 2008) predict childhood language, literacy, working memory, and IQ. Specifically, research with fairly large, demographically diverse samples has found that vocabulary assessed between 16 and 24 months of age significantly predicts language and reading skills between 4 and 9 years of age (Duff et al., 2015) and, similarly, that expressive and receptive vocabulary at 18 months of age predict vocabulary size at 30 months (Henrichs et al., 2011). Moreover, a binary variable of high or low vocabulary at age two predicts expressive language at ages three and five after controlling for gender, birth order, socioeconomic status, and ethnicity (Lee, 2011). Late-talker status (falling below the 10^{th} percentile in productive vocabulary) at age two predicts unique variance in language outcomes at age four above and beyond other measurable risk factors (e.g., child sex, family history of language impairment; Reilly et al., 2010). Finally, vocabulary production at age two predicts academic reading and math achievement as well as behavioral performance at age five (Morgan, et al., 2015).

Similar patterns are observed for lexical access. It is concurrently related to parent-reported vocabulary comprehension at 15 months of age and with vocabulary production, grammatical complexity, and directly-assessed receptive vocabulary at 25 months (Fernald et al., 2006). Further, children who had greater growth and acceleration in vocabulary size across 12, 15, 18, 21, and 25 months also had more efficient processing at 25 months. Recently this research has been extended to include typically developing children and children who are late talkers (falling at or below the 20th percentile in vocabulary scores at 18 months of age). Children who were classified as late talkers exhibited significantly slower processing than typically developing peers (Fernald & Marchman, 2012). Additionally, those late talkers who remained delayed in language (at or below the 20th percentile) at 30 months of age had significantly slower processing than late talkers who scored in the typically developing range (above the 20th percentile) at 30 months. Similar to previous reports, this study found that lexical access at 18 months significantly predicted the rate and acceleration of vocabulary growth over the second year of life, explaining additional unique variance above and beyond parent-reported vocabulary production alone. Finally, lexical access predicts expressive vocabulary and working memory at age 8 (Marchman & Fernald, 2008).

Only one study to our knowledge has directly compared vocabulary size and speed of lexical access prospectively. Marchman and Fernald (2008) utilized the Looking-While-Listening paradigm (Fernald et al., 1998; Fernald, Zangl, Portillo & Marchman, 2008), an online, visual measure of word processing efficiency (similar to lexical access) in addition to a parent-report measure of expressive vocabulary size (Fenson et al., 2007) at 25 months of age to investigate prediction to linguistic and cognitive outcomes at age eight. The authors found that vocabulary size and lexical access each predicted significant unique variance in age-eight language and working memory. However, only early vocabulary size accounted for significant unique variance in IQ and this effect was fully mediated by concurrent working memory. The current research seeks to extend this work by examining differential prediction from speed of lexical access and vocabulary size to outcomes that lie along the continuum from processing/working memory to long-term knowledge.

Current Study

The current research investigates whether vocabulary size and lexical access differentially predict linguistic and cognitive skills at three, four, and five years of age. As discussed, we hypothesize that vocabulary and lexical access are related but separable components of language skill and that one way in which they may differ is the extent to which they rely on long-term knowledge vs. processing/working memory resources. As such, we are interested in the models of prediction from vocabulary and lexical access to conceptual vocabulary specific to a repeated context (conceptual development), vocabulary size (semantic memory), sentence repetition (sentence processing/production), and nonword repetition (phonological short-term/working memory).

We hypothesize that decontextualized vocabulary will more strongly predict vocabulary and conceptual knowledge in early childhood than will lexical access due to the fact that vocabulary and conceptual development both involve long-term knowledge. In contrast, we hypothesize that lexical access will more strongly predict verbal short-term memory than will vocabulary size. Finally, we hypothesize that both vocabulary and lexical access should equally predict sentence repetition, a measure that taps into both long-term semantic/syntactic knowledge and language processing. We test these hypotheses in a sample of English-speaking monolingual children in the United States and a sample of French-speaking monolingual children in

Switzerland in order to evaluate the generalizability of our findings and explore whether models of prediction differ across children from distinct language backgrounds. English and French differ linguistically in ways that may be relevant to the present research. For example, French is more complex in its inflectional morphology than is English, having gender suffixes for nouns and verbs, and a greater number of inflectional verb endings. Indeed, between two and four years of age, French-speaking children have longer (but not statistically different) mean lengths of utterances in morphemes (reflecting more utterance complexity) but also smaller vocabulary sizes than English-speaking children (Thordardottir, 2005). Further, there are differences in the phonological systems of the two languages. For example, in French, final unstressed vowels are unpronounced when they precede a word beginning with a vowel and most final consonants are not pronounced. However, there are also similarities observed between development of English and French: in both cases, development of vocabulary is positively associated with utterance complexity (Thordardottir, 2005). For these reasons, we are interested in evaluating the generalizability of models of the prediction from early vocabulary and processing measures to conceptual development, semantic memory, sentence processing/production, and phonological working memory.

Method

Participants

Seventy-nine English-speaking monolingual children (41 girls) and sixty-six Swiss-French-speaking monolingual children (33 girls) were recruited to a large, multi-site longitudinal study investigating children's language and literacy development. English-speaking participants were recruited from Women, Infant, and Children Centers, parenting groups, swap meets, and birth records in a large city in the United States whereas French-speaking participants were recruited through birth lists in a large city in Switzerland. All children were carried to term (at least 37 weeks), had normal hearing and vision, and had no reported diagnoses or cognitive delay at the start of the study. In the English-speaking sample, 30 children were excluded from the current study for the following reasons: attrition and/or missing data on either the predictor variables or all three outcome time points (*N* = 28), or falling below 80% exposure to English (*N* = 2). Children with data for at least one outcome at one time point were included due to an interest in retaining as large a sample as possible. The final sample consisted of 49 children (29 girls) at ages two and three, 46 children at age four, and 39 children at age five. In the French-speaking sample, 7 children were excluded due to attrition and/or missing data. The final sample consisted of 59 children (31 girls) at ages two and three, 57 children at age four, and 54 children at age five. We conducted an a priori power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) for a linear multiple regression fixed model. It is difficult to estimate the models that will be run because we plan on evaluating only models that fit the data well. However, consideration of all possible models revealed that we have sufficient sample size to detect a medium effect with a power of .80 in all cases except when evaluating a model with *both predictors by group*. Participants came into the lab for six visits, of which four are reported here: two years of age (*M* = 22;17, range = 21;6 – 25;12), three years of age (*M* = 36;27, range = 34;27 – 41;24), four years of age (*M* = 49;2, range = 47;9 – 55;12), and five years of age (*M* = 61;13; range = 59;15 – 68;13). A $25 gift card to a major retailer and a small toy were provided as incentives at each visit. See Table 1 for demographic data on the final sample.

Measures

Language Exposure Assessment Tool (LEAT). The LEAT (DeAnda et al., 2016) was used to estimate relative language exposure to insure that participants were primarily exposed (at least 80 percent of the time) to a single, native language (either English or French). The LEAT is a systematic parent interview that gathers information on each of the individuals who regularly interact with the child, the languages they speak, whether they are a native speaker, and the number of hours spent talking to/being overheard by the child in each language per day. The LEAT yields the following estimates of relative exposure: hours per day, hours per week, percent exposure, and a parent estimate. Percent exposure to each language is determined by the hours of exposure to each language across the child's life. Internal consistency on the LEAT is excellent. Further, LEAT percent relative exposure significantly predicts concurrent vocabulary size above and beyond maternal education, age, and parent estimates (DeAnda et al., 2016) even at minimal levels (20%) of exposure to the language. All children included in the present study were exposed to their primary language at least 80 percent of the time.

Computerized Comprehension Task (CCT). The CCT is a forced-choice, direct measure of vocabulary comprehension administered on a touch screen (Friend & Keplinger, 2003; 2008). Two images are presented to the left and right center of the screen while an experimenter delivers a prompt naming a target image (e.g., Where is the dog? Who is swimming? Which one is old?). Trials are experimenter-controlled such that the experimenter advances to the next trial or is able to repeat trials, but each trial has a maximum duration of 7 seconds. Touches to the screen automatically end the trial. In the current study the administration of the CCT followed Friend, et al. (2012) and repetitions of trials were when: 1) the child initiates a response before the end of the trial but does not touch the screen prior to image offset, 2) the child is distracted, not looking at the screen, and misses the trial, or 3) in the case of an accidental screen touch. If the child became fussy or disengaged at any point during administration, the experimenter attempted to re-engage the child. If three attempts to re-engage failed, the experimenter terminated the procedure. Only children who were engaged for at least one-third of trials were retained in analyses.

Words on the CCT were obtained from the MacArthur-Bates Communicative Development Inventories: Words and Gestures form (MCDI; Fenson et al., 2007). Paired images are high-quality, colorful digital images that are prototypical exemplars of the word. Target and distractor items are matched on color, size, saliency, word class, and word difficulty. Difficulty is defined based on 16-month normative data from the MCDI (Frank, Braginsky, Yurovsky, & Marchman, 2016). There are approximately equal numbers of easy (comprehension ≥ 66 percent), moderate (comprehension between 33 and 66 percent) and difficult (comprehension ≤ 33 percent) words, which are randomly distributed. Additionally, the test includes nouns, verbs, and adjectives. Two forms of the CCT exist such that all word-image pairs are presented as both target and distractor. Forms were counterbalanced across participants. The side of the target image (left or right) is randomized across trials but does not appear on the same side more than twice consecutively following Hirsh-Pasek and Golinkoff (1996).

The French adaptation of the CCT is the same in design and administration. Items on the French adaptation were selected based on comprehension norms for the French adaptation of the MCDI, Les Inventaires Français du Développement Communicative (IFDC; Frank et al., 2016; Kern 2007). Translation equivalents (synonyms across languages) were maintained in the adaptation to the extent that they yielded equivalent distributions of word class and word difficulty as the English adaptation of the CCT. Images were selected to be prototypical for children in Switzerland.

Both adaptations of the CCT have strong test-retest reliability and internal consistency, and moderate short-term stability over a 4-month period (Friend & Keplinger, 2008; Friend &

Zesiger, 2011). The CCT correlates with parent-reported vocabulary and predicts the number of unique words and the mean length of utterance in subsequent child language samples (Friend, et al., 2012). Finally, the CCT is the only direct measure of receptive vocabulary size prior to 30 months of age.

Coding vocabulary size. Decontextualized vocabulary size was operationalized as the number of correct target image touches. Coding was completed offline, and inter-rater reliability (coded for approximately 20 percent of participants) was excellent (.93 or above).

Coding lexical access. Lexical access was operationalized as the latency to touch the correct target image as in DeAnda et al. (2018). For the French-speaking sample, the CCT program computed reaction time automatically from the appearance of the images on the screen to the child's touch to the screen. For the English-speaking sample, coding of reaction time was completed offline in Eudico Linguistics Annotator (ELAN) (http://tla.mpi.nl/tools/tla-tools/elan/), Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands; Lausberg & Sloetjes, 2009) using a video recording of the child's touch responses. Onset of coding began time-locked with the appearance of the images on the screen, which occurred on average 238 ms after the target word onset and prior to target word offset in the first carrier phrase (e.g., "Where is the DOG? Touch dog."; "Which one is SWIMMING? Touch swimming."). Latency to touch was coded in the frame during which the child touched the screen (i.e., when forward momentum ceased and the child's finger was slightly bent, indicating contact with the screen). Inter-rater reliability was established for approximately 20 percent of participants. Reliability exceeded .90 within three frames for image and touch onset. Following DeAnda et al. (2018), only trials during which the child touched the correct image were included in the analysis of reaction time. Further, trials with short latencies likely reflect a child's decision to touch prior to accessing the meaning of the word: therefore, trials with latencies less than 400 ms were excluded from analysis, corresponding to longest cutoffs typically used in eye-tracking studies to allow for the time necessary to execute haptic responses (Fernald et al., 2008; DeAnda et al., 2018). Coding of latency to touch began at the onset of the images on the screen.

Free play language sample (LS). Caregivers were instructed to play with their children just as they would at home during three separate visits to the laboratory. Children and caregivers played with a Fisher-Price farm play set including a rich assortment of farm structures (barn, silo, etc.), toy animals and people, and vehicles. At the 3-year visit children and parents played for 15 minutes whereas at the 4- and 5-year visits children and parents played for 10 minutes with an added complexity toy. The session was recorded in full using a Zoom H2n Handy Recorder microphone. Subsequently, the audio files were transferred to Express Scribe Transcription Software (available at http://www.nch.com.au/scribe/), and were fully transcribed and coded for grammatical morphemes (lexical units, plurals, tense markers, and contractions) using the Systematic Analysis of Language Transcripts software (SALT; Miller & Iglesias, 2012). All assistants who transcribed the audio files were trained using the online SALT training platform and were trained on practice transcripts until they obtained reliability of at least 90 percent with a completed sample transcript. Assistants were fluent in the language of the transcript (either English or French). Finally, approximately 20 percent of the transcripts were transcribed by a second assistant in order to ensure reliability. Reliability was calculated for the analysis set as the number of agreements divided by the number of agreements plus the number of disagreements. For the English-speaking sample, word- and morpheme-level reliability was as follows: .91 and .92 at the 3-year visit, .94 and .94 at the 4-year visit, and .91 and .92 at the 5-

year visit. For the French-speaking sample, reliability was .90 and .89 at the 3-year visit, .91 and .89 at the 4-year visit, and .96 and .95 at the 5-year visit.

In order to assess conceptual development, we coded transcripts for the number of context-related vocabulary items that children and parents produced as well as the number of utterances containing a farm-related vocabulary item. First, we generated a list of specific farm-related vocabulary items. This list was created by taking farm-related words from the MCDI and its French adaptation (IFDC), searching online farm word lists in English and French (https://relatedwords.org/relatedto/farm; https://www.frenchlearner.com/vocabulary/farm/) and taking commonly used words from the English and French transcripts. Once this list was generated, the first author and a coder fluent in French confirmed the appropriateness of words included as well as adding potential synonyms from both languages (e.g., tractor/truck, étable/écurie). This vocabulary list was used to identify utterances and words related to the farm context. The final list contained 303 English, and 264 French, farm-related nouns (e.g., horse, barn, fence) and verbs (feed, plow, grow).

Two coders (the first author and an assistant fluent in French and English) first coded each utterance of the transcripts for words that were farm-related (included in the vocabulary list), coding only content words (i.e., nouns and verbs) but excluding function words (articles, conjunctions, prepositions, etc.). Then, each utterance that contained a farm-related word was coded as a farm-related utterance. Approximately 20 percent of the transcripts were coded for reliability, calculated as the number of agreements divided by the number of agreements plus the number of disagreements. For the English sample, utterance-level and word-level agreement was as follows: .93 and .93 for the 3-year visit .95 and .96 for the 4-year visit, and .98 and .98 for the 5-year visit. For the French sample, reliability calculations were performed on English translations of the French transcripts. Agreement was as follows: 98 and .97 for the 3-year visit, .96 and .97 for the 4-year visit, and .98 and .98 for the 5-year visit.

Our primary variable of interest was the number of different (unique) word forms children produced that were related to the farm context. In the interest of holding constant the number of utterances that children produced (and thereby controlling for child talkativeness) we divided the number of different farm words by the total number of utterances. Thus the final variable was the average number of different farm words produced per utterance, which will hereafter be referred to as conceptual knowledge. We anticipate that these measures index children's knowledge of the farm context (e.g., Farrar, Friend, & Forbes, 1993), such that children with greater conceptual knowledge will produce more words and utterances consistent with that context. We expect farm-related vocabulary to increase over time, indicating a more mature farm concept. Since these estimates are taken over successive free play sessions with the farm set, we expect that they reflect children's long-term conceptual knowledge.

Peabody Picture Vocabulary Test (PPVT) and Échelle de Vocabulaire en Images Peabody (EVIP). The PPVT is a direct measure of vocabulary comprehension for use for individuals from 30 months of age through adulthood (Dunn & Dunn, 1997) and the EVIP is its French adaptation (Dunn, Dunn, & Thériault-Whalen, 1993). During administration of the PPVT/EVIP an experimenter displays four pictures and asks the child to point to the one that corresponds to a target word. The test is adaptive such that its difficulty increases with age, and the test ends once the ceiling criterion is met. The child's final score is the number of items to reach ceiling minus the number of errors. Thus, like the CCT, the PPVT yields a direct estimate of receptive vocabulary size. The PPVT and EVIP were standardized on representative samples, and have generally strong reliability for all age ranges tested, as well as strong internal consistency (Dunn & Dunn, 1997; Dunn et al., 1993). The measure of interest from the PPVT are the raw scores corresponding to the number of words correct and we expect this measure to capture long-term semantic knowledge.

Sentence repetition (SR). The sentence repetition task used in the current study is a modified version of the Italian task created by Devescovi and Caselli (2007). For the purposes of the current study, new sentences in each category (e.g., simple sentences with copula, simple sentences with one argument) were created in English and French in order to equate, as much as possible, the number of words and morphemes in each sentence across languages. Devescovi and Caselli (2007) found that performance on the sentence repetition task significantly increased with age and was concurrently significantly correlated with spontaneous production measures. Recent research suggests that SR abilities are indicative of underlying language competence, requiring diverse skills in language processing, comprehension, and production (Klem et al., 2015). Clinically, SR tasks have often been used as a marker for Specific Language Impairment (SLI; Conti-Ramsden, Botting & Faragher, 2001). At the year-3 visit, the test included 27 sentences of varying complexity and length. An additional nine sentences at were added at the year-4 and year-5 visits to increase complexity and reduce the chances of a ceiling effect. During test, the experimenter told the child to repeat after her, modeled the sentence, and then revealed an image depicting sentence-level meaning to the child. The sentence could be repeated up to three times to the child however the child's first attempt at repetition was scored. The primary variable of interest is proportion of morphemes repeated correctly on the task. Approximately 20 percent of the sentence repetition data were coded for reliability. For the English-speaking sample, sentence-level inter-rater agreement was .92 at the year-3 visit, .91 at the year-4 visit, and .92 at the year-5 visit. For the French-speaking sample, sentence-level inter-rater agreement was .92 at the 3-year visit, .99 at the year-4 visit, and .99 at the year-5 visit. Since repeating sentences requires skills in language processing and comprehension, syntax knowledge, vocabulary knowledge, and language production, we anticipate that this measure lies midway on the continuum between long-term knowledge and processing/working memory.

Nonword repetition (NW). The nonword repetition task in the current research was created to equate syllable and word length across languages: pseudowords were created using the MCDI adaptation in each language and shuffling syllables to create phonotactically plausible words for each language separately. In general, nonword repetition tasks are a measure of phonological working/short term memory: in practice, the capacity to remember novel sound sequences (Gathercole, Willis, Baddeley, & Emslie, 1994). Recent evidence suggests that repetition of shorter nonwords involves short-term memory (short-term storage of information) whereas repetition of longer nonwords involves working memory (short term storage and concurrent manipulation of information; Unsworth & Engle, 2007). In the current study, the test is adaptive such that words increase in length and complexity (i.e., inclusion of consonant clusters) with age. At 36 months, the test included four 1-syllable words, four 2-syllable words, four 3-syllable words, and four 4-syllable words. At 48 months the task was similar but added four 5-syllable words. At 60 months, the one-syllable items were removed and another four 5-syllable words were added. During test, the experimenter asked the child to repeat after her. She presented a toy figure and says, “This one’s name is ___. Can you say ___?” The child’s first attempt at repetition was scored. The measure of interest was percent of consonants repeated correctly (PCC). We anticipate that the adaptive nature of this task will result in an equal number of short nonwords and long nonwords at each age, giving us a similar assessment of short-term and working memory across time. For the English-speaking sample, phoneme-level inter-rater agreement was .88 at the 3-year visit, .91 at the year-4 visit, and .94 at the year-5 visit. For the French-speaking sample, phoneme-level inter-rater agreement was .97 at the 3-year visit, .99 at the year-4 visit, and .96 at the year-5 visit. Although there is evidence that correct repetition of nonwords involves long-term word knowledge (Jones, Gobet, & Pine, 2007), it is primarily considered a measure of phonological working/short term memory. For this reason, we expect that this measure taps more into the working memory/processing end of the continuum.

Results

Analytic Approach

The primary goal of this study is to investigate differential prediction from early vocabulary and speed of lexical access to performance in conceptual development, PPVT vocabulary, sentence repetition, and nonword repetition across 3, 4, and 5 years of age. We will first present descriptive results and zero-order correlations from our predictors of interest to outcomes at each time point. These analyses will be conducted separately by language in order to determine whether there are differences in development across groups for which we need to account. Next, we will conduct a series of mixed effects models investigating prediction from vocabulary size and lexical access to performance on each outcome of interest over time. These analyses will include both the English and the French samples and a binary language group control variable as well as predictor by language group interactions to account for differences across samples. Note that the pattern of missing data on outcome measures and attrition may result in varying degrees of freedom in the results presented. In the interest of retaining as large a sample as possible, and because mixed effects models in R can handle missing data, participants were not excluded list-wise.

Descriptive Data

See Table 2 for descriptive data for the English and French samples. We performed *t*-tests across groups to determine if group differences in variables of interest existed. A family-wise correction for multiple tests was applied using a sequential Bonferroni procedure (Benjamini & Hochberg, 1995) for 20 comparisons. Nominal alpha was set at .05. Conceptual knowledge and PPVT vocabulary were approximately normally distributed for each group at each time point. However, for sentence repetition, we observed a negative skew in the English sample at age four (*skewness* = -3.39, *kurtosis* = 15.61) and age five (*skewness* = -3.24, *kurtosis* = 12.98) and in the French sample at age four (*skewness* = -2.45, *kurtosis* = 6.35) and age five (*skewness* = -2.13, *kurtosis* = 5.76). This indicates a slight to moderate ceiling effect for sentence repetition at later ages. For nonword repetition, there were four outliers. In the English sample at age three, one child scored .16 and another scored .33. In the French sample one child scored .14. These scores were > 4SD below the mean for the group and were consequently excluded as outliers. In the English sample at age five, one child scored .55: a score > 4 SD below the mean. This score was also excluded as an outlier. No skewness was observed for nonword repetition in the English sample. However, in the French sample, we observed a negative skew at age four only (*skewness* = -2.28, *kurtosis* = 4.41) which might indicate a ceiling effect.

Due to the fact that we observed negative skewness, indicating ceiling effects, at some ages in sentence repetition (both samples) and in nonword repetition (French sample only), we transformed all sentence repetition and nonword repetition variables using a logit transformation. We present results from these transformed variables.

Prediction to Outcomes

For zero-order correlation information please see Table 3. Family-wise error corrections were applied to the correlation results for each outcome variable using a sequential Bonferroni procedure (Benjamini & Hochberg, 1995) for six correlation assessments per family for predictive relations and three correlation assessments per family for longitudinal outcome assessments. Nominal alpha was set at .05.

Mixed Effects Analyses

In order to investigate prediction to the outcomes of interest, we conducted a mixed effects model in R (R Core Team, 2019) using the lmer function of the lme4 package (Bates, Maechler, Boker, & Walker, 2015). The dependent variable is the repeated measure of the outcome of interest. For the purposes of these analyses, we will first evaluate the role of relevant control variables (child sex and maternal education as a proxy for SES) and will only include those with significant or marginal effects as control variables in subsequent models. Given that there are differences across language groups in the development of the predictors and outcome measures of interest, as well as some differences in zero-order correlations, all analyses will include a binary language group control variable as well as language group by predictor interactions. Therefore, the models included fixed effects of age, language group, the predictor of interest, an age by predictor interaction, a language group by predictor interaction, and a three-way interaction between age, language group, and predictor. Finally, we will include a by-subjects random intercept. We first assessed the fit of three models per dependent variable: a model with age-two vocabulary as a predictor by itself, a model with age-two speed of lexical access by itself, and a model with both predictors together. We present these models as a way of determining whether early vocabulary and lexical access predict outcomes independently as well as the unique variance in outcomes they predict when controlling for the other variable. We compared these models using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), which are measures of model fit that penalize for model complexity (i.e., the addition of additional parameters): BIC penalizes for model complexity to a greater extent then does AIC. The AIC and BIC values alone are not informative however when comparing models, lower values indicate better fit. For our purposes and sample size, we can conclude a model is a better fit if its AIC and BIC values are 6 less than the competing model (Hilbe, 2011). We also include Log Likelihood estimations for all models: for this measure, higher numbers indicate better fit. We evaluate estimates of parameter significance only for models that provide the best fit to the data. In the absence of model differences, we report estimates for all models. In order to assess significance of added parameters we will conduct likelihood ratio tests on nested models with the addition of each parameter. Finally, we assess the proportion of variance accounted for in the outcome by all significant predictors using the MuMIn package in R (Bartón, 2019) and assess the unique variance accounted for by the addition of our predictors of interest.

Conceptual knowledge. Assessment of control variables revealed that the addition of maternal education to the model was significant ($\chi^2 = 9.81$, $p = .002$) but the addition of child sex was not ($\chi^2 = .24$, $p = .62$). Maternal education was therefore controlled in subsequent analyses. Overall, evaluation of model fit revealed that the best fitting, most parsimonious model was the model with vocabulary as the sole predictor. This model was a better fit than the model with lexical access only. In addition, it was a better fit than the full model according to the BIC (but not AIC) value. We therefore present results from the vocabulary only model. In this model, the addition of age to the model was significant, revealing that children increased in their conceptual knowledge across time. Maternal education and age-two vocabulary were also significant predictors however no other effects were significant. See Table 4 for results. The model with significant parameters (age, maternal education, and age-two vocabulary) accounted for 20.76 percent of the variance in conceptual knowledge across age. The addition of vocabulary accounted for 3.23 percent above age and maternal education. Consistent with our hypothesis, age-two vocabulary predicted conceptual knowledge, however age and maternal education accounted for the bulk of the variance, 13.1 and 4.43 percent, respectively, suggesting that children's repeated exposure to the farm context was the strongest predictor of farm-related conceptual knowledge. In contrast, the model with speed of lexical access was not a good fit to the data.

PPVT vocabulary. Assessment of control variables revealed that neither the addition of maternal education (χ^2 = 9.81, p = .002) nor child sex (χ^2 = .24, p = .62) to the model was significant so these were not retained as control variables. Overall, evaluation of model fit revealed that the best fitting, most parsimonious model was the model with age-two vocabulary alone. This model was a better fit than the model with lexical access alone. In addition, it was a better fit than the full model according to the BIC (but not AIC) value. We therefore present results from the vocabulary only model. In this model, the addition of age to the model was significant, accounting for 51 percent of the variance, revealing growth in vocabulary over time. The addition of language group was also significant (13.47 percent of the variance), reflecting somewhat lower vocabularies at each age in the French sample relative to the English sample. Finally, age-two vocabulary (9.57 percent of the variance) was also a significant predictor of PPVT vocabulary across ages three, four, and five after taking into account the effects of age and language group. No other effects were significant. See Table 5 for results. The final model (age, language group, and age-two vocabulary) accounted for 74.04 percent of the variance in vocabulary across ages. In contrast, models with lexical access as a predictor were not a good fit to the data. Thus, we conclude, consistent with out hypothesis, that vocabulary size at age two, but not speed of lexical access, predicted downstream vocabulary development over and above age and language group.

Sentence repetition. Assessment of control variables revealed that the addition of child sex to the model was significant ($\chi^2 = 4.59$, $p = .03$) but the addition of maternal education was not ($\chi^2 = .12$, $p = .73$). Child sex was therefore controlled in subsequent analyses. Overall, evaluation of model fit revealed that the best fitting, most parsimonious model was the model with vocabulary as the sole predictor. This model was a better fit than the model with lexical access only. In addition, it was a better fit than the full model according to the BIC (but not AIC) value. However, due to our a prior hypothesis that lexical access should predict sentence repetition, and because the model fit with vocabulary alone was not a better fit than the model with both predictors according to the AIC value, we evaluated parameter significance for both models. In the vocabulary only model, the addition of age to the model was significant, accounting for 13.10 percent of the variance, revealing improved sentence repetition scores over time. Child sex (2.76 percent of the variance) and age-two vocabulary (11.09 percent of the variance) were also significant. Finally, there was a significant age by vocabulary interaction (4.28 percent of the variance), suggesting the effect of vocabulary on sentence repetition varies with age. Adding lexical access to the model yielded no significant effects. See Table 6 for results. The final model (age, child sex, age-two vocabulary, and age X age-two vocabulary) accounted for 31.22 percent of the variance in sentence repetition across ages.

We turn now to regressions conducted at each age point to evaluate the changing effect of vocabulary on sentence repetition with age. After controlling for the effect of child sex, vocabulary significantly predicted sentence repetition at age three ($\beta = .52$, $SE = .02$, $p < .001$), age four ($\beta = .33$, $SE = .01$, $p < .001$), and age five ($\beta = .28$, $SE = .01$, $p = .008$). However, the strength of this effect weakens with age. This might be a result of a ceiling effect on the sentence repetition task, which remained even after the logistic transformation. See Figure 1 for scatterplots of vocabulary X sentence repetition at each age. In contrast to our hypothesis, lexical access did not predict sentence repetition.

Nonword repetition. Assessment of control variables revealed that the addition of child sex to the model was marginally significant ($\chi^2 = 3.22$, $p = .07$) but the addition of maternal education was not ($\chi^2 = 1.40$, $p = .24$). Child sex was therefore controlled in subsequent analyses. Overall, evaluation of model fit revealed that the best fitting, most parsimonious model was the model with vocabulary as the sole predictor. This model was a better fit than the model with lexical access only. In addition, it was a better fit than the full model according to the BIC (but not AIC) value. However, due to our a priori hypothesis that lexical access should predict nonword repetition we evaluated the fit of both the vocabulary only model and the full model. In the vocabulary only model, the addition of age to the model was significant, accounting for 13.79 percent of the variance, revealing improvement in nonword repetition over time. The addition of language group (17.46 percent of the variance) and child sex (2.66 percent) were also significant: girls performed better on nonword repetition than boys and children in the French group performed better than children in the English group. The addition of age-two vocabulary was also significant, accounting for 1.93 percent of unique variance over and above age, sex, and language group. No other effects were significant and adding lexical access to the model did not yield any significant effects. See Table 7 for results. The final model (age, language group, child sex, and age-two vocabulary) accounted for 35.84 percent of the variance in nonword repetition from age three to five. In contrast to our hypothesis, there were no effects of lexical access.

Discussion

The primary goal of the current research was to investigate longitudinal prediction from vocabulary size compared to speed of lexical access at age two. The impetus for this work is twofold. First, there is a body of evidence indicating that both vocabulary and lexical access significantly predict downstream language and cognition and can therefore be considered foundational skills for later learning (Bleses et al., 2016; Duff et al., 2015; Fernald & Marchman, 2012; Fernald et al., 2006; Friend et al., 2012; 2018; 2019; Henrichs et al., 2011; Lee, 2011; Marchman & Fernald, 2008; Morgan et al., 2015; Reilly et al., 2010). Second, vocabulary size and lexical access are correlated but their shared variance is only about 20 percent (DeAnda et al., 2018; Fernald et al., 2006; Legacy et al., 2016). Thus, we consider these components of language ability to be related, but dissociable. Given this, the current study adds to extant work on the stability of early language abilities by specifically exploring whether there is differential prediction from vocabulary and lexical access to downstream linguistic skills. We conducted longitudinal analyses predicting linguistic outcomes across ages three, four, and five in two linguistically and geographically distinct groups of children: English- and French-speaking monolingual children.

The linguistic outcomes of interest were conceptual knowledge, vocabulary size, sentence repetition, and nonword repetition. We chose these outcomes because we expect that they vary along a continuum from long-term concepts (conceptual knowledge) to verbal working memory (nonword repetition). Specifically, the conceptual knowledge variable was composed of children’s acquisition of context-related words and phrases in a free-play context over time. In order to have high scores on this task, children needed to develop broad knowledge about the context in which they were playing. Similarly, vocabulary size is an index of the breadth of long-term semantic knowledge. Sentence repetition requires both syntactic knowledge and efficient language processing. Finally, nonword repetition is thought to index phonological short-term and working memory. We hypothesized that one of the ways in which early vocabulary and speed of lexical access might be dissociated is the extent to which they depend on long-term storage versus processing/working memory. Since long-term storage of knowledge and processing are related and interdependent we do not expect these components can be completely dissociated in this way. Nevertheless, the distinction between storage and processing may be one source of the non-overlapping variance between vocabulary size and lexical access. Overall, this hypothesis was not supported in the present study. We turn first to a discussion of the results of prediction to each outcome measure of interest followed by an examination of possible explanations for the pattern of effects.

Differential Prediction

First, results revealed that vocabulary size at age two significantly predicted conceptual knowledge across age and language groups. Although the pattern of zero-order correlations suggested that the effect of vocabulary on conceptual knowledge decreased with age such that vocabulary was significantly correlated with conceptual knowledge at age three but not at age four or age five, the results from the mixed effects models, after controlling for age and maternal education, revealed that the effect of age-two vocabulary was consistent. Thus we conclude that vocabulary knowledge significantly predicted conceptual development consistent with our hypothesis that early long-term semantic knowledge should load onto long-term concepts. Further, vocabulary at age two significantly predicted conceptual knowledge at age three, which was associated with conceptual knowledge at age four, which was in turn associated with conceptual knowledge at age five. In contrast, speed of lexical access did not predict conceptual development.

Similar patterns emerged in the prediction of PPVT vocabulary size across ages. Vocabulary size at age two significantly predicted PPVT vocabulary across ages and language groups. In contrast, lexical access did not predict PPVT vocabulary. Although it is perhaps not surprising that vocabulary at age two would more strongly predict later vocabulary in the preschool years, this result in combination with the results from the conceptual knowledge analyses suggests that early vocabulary predicts later long-term knowledge, as expected.

Performance on the sentence repetition task, as mentioned previously, requires diverse language skills including syntactic knowledge, language processing, and syntactic production. Due to this, we anticipated that sentence repetition lies midway on the continuum from long-term knowledge to processing/working memory. If this hypothesis were correct, we would expect prediction from both vocabulary size and speed of lexical access at age two. Vocabulary size at age two did significantly predict sentence repetition. However, this analysis also revealed an age by sentence repetition interaction. This result may be partially due to a ceiling effect on the measure, which we observed at ages four and five. Follow-up tests revealed that vocabulary significantly predicted sentence repetition at all ages but that there was a weakening of this effect with age. Despite the fact that we subjected the sentence repetition data to a logit transformation, it is possible that remaining ceiling effects restricted the variance to be accounted for by predictors. In contrast to our hypothesis, speed of lexical access did not predict sentence repetition any age.

Finally, nonword repetition is considered an index of phonological working/short-term memory (Gathercole et al., 1992; Unsworth & Engle, 2007). We hypothesized that performance on this task would be more strongly predicted by speed of lexical access and less strongly predicted by vocabulary size. However, results revealed that vocabulary size predicted nonword repetition across age. Although the zero-order correlations suggest a weakening effect of vocabulary size on nonword repetition with age, similar to the results for conceptual vocabulary, after controlling for age, sex, and language group, the effect of age-two vocabulary size was small but consistent over time. It is not surprising that this is a relatively small effect given the observation that the predictive links between vocabulary and nonword repetition shift in the preschool years (Gathercole et al., 1992). Specifically, between 4 and 6 years of age vocabulary exerts a weakening influence on phonological memory. In contrast to our hypothesis, speed of lexical access did not predict nonword repetition at any age.

Overall, our hypothesis that early vocabulary and speed of lexical access can be partially disassociated according to their links to skills in long-term memory versus working memory and processing was not supported. We anticipated that vocabulary size would more strongly predict conceptual knowledge and PPVT vocabulary but would more weakly predict sentence repetition and nonword repetition. This was partially supported: the strongest unique variance accounted for by vocabulary size at age two was for PPVT vocabulary (9.57%) and sentence repetition (11.09%), followed by conceptual knowledge (3.23%) and then nonword repetition (1.93%). The predictions to PPVT vocabulary and sentence repetition are consistent with our hypothesis of a relation between early vocabulary and long-term knowledge (semantic knowledge for PPVT vocabulary and general language/syntactic knowledge for sentence repetition). Similarly, the weaker prediction to nonword repetition is also consistent with our hypothesis that vocabulary would predict a working memory measure more weakly than a long-term memory measure. However, the prediction to conceptual knowledge was weaker than to PPVT vocabulary, perhaps reflecting the much stronger effect of repeated exposure to the farm context indexed by the age variable, which accounted for 13.1 percent of the variance in the use of farm-related words. However, it is possible that our assessment of conceptual vocabulary did not fully assess long-term conceptual knowledge for this age range: we return to this in the limitations and directions for future research section.

Importantly, the lack of prediction from lexical access to any outcome measure is inconsistent with our predictions. There are a few explanations for this finding. First, regarding the prediction from sentence repetition, if it is true that the age-two measures can be dissociated by the extent to which they index long-term knowledge compared to processing, it is possible that sentence repetition better indexes syntactic knowledge and production than working memory. This would also be consistent with the finding of relatively stronger prediction from age-two vocabulary to sentence repetition in terms of the variance accounted for by vocabulary. Indeed although there is evidence that short term memory and language processing are important to performance on sentence repetition tasks (Riches, 2012), other researchers stress that sentence repetition measures general language ability more so than working memory or the episodic buffer (Klem et al., 2015). The view consistent with this hypothesis is that children do not process the sentence to be repeated in working memory but rather create a representation of the meaning of the sentence that is then processed through semantic and syntactic knowledge to be reproduced. Our results lend support to this idea by suggesting early vocabulary plays a greater role in later sentence repetition than does lexical access.

However, this does not explain the lack of an effect of lexical access on nonword repetition. Nonword repetition is an assessment of phonological working memory and since working memory and processing speed are known to be related (Fry & Hale, 1996; Kail & Salthouse, 1994) we would expect our measure of lexical access to predict later development of phonological working memory. An explanation for this finding (and the finding that speed of lexical access did not predict sentence repetition) is that our assessment of speed of lexical access is not solely a measure of processing speed but that it also involves other cognitive and motor abilities. Our measure of lexical access is based on a haptic response to a two-alternative forced-choice vocabulary task administered at two years of age. Previous work on this task suggests that children execute a relatively quick visual saccade from the distractor to the target image whenever they touch the screen, regardless of whether the touch is to the target image. (Hendrickson et al., 2015; Hendrickson et al., 2017). The authors concluded that a touch to the target image requires overriding the prepotent response to touch the image upon which children originally fixated (the distractor image). This may index a difference in response strategies for computationally expensive touch responses relative to visual responses. Classic theories on action selection suggest that a certain amount of evidence or confidence is necessary in order to reach a decision threshold to execute a response while maximizing accuracy (Gold & Shadlen, 2007; Luce, 1986; Shadlen & Gold, 2004). This varies among individuals and we might also expect that computationally expensive haptic responses might have a more stringent decision threshold for execution. Thus our measure of lexical access likely involves not only word processing but also executive control and levels of decision confidence.

Relatedly, research also suggests that stable representations of word meaning are built through repeated experiences with word-referent associations and the application of learning mechanisms such as hypothesis testing and cross-situational learning (Aravind et al., 2018; Yu & Smith, 2012, Vlach & DeBrock, 2017). This being the case, it is possible that when children choose to touch an image in a vocabulary assessment, it is only after considerable deliberation and hypothesis testing of the two potential referents. Thus, *speed* to touch may reflect deliberative processes, and/or recruitment of executive control mechanisms in addition to speed of lexical access. Indeed, prospective motor control in toddlerhood is related to inhibition and working memory components of executive function (Gottwald, Achermann, Marciszko, Lindskog, & Gredebäck, 2016). In contrast, a visual fixation to a target image is relatively less computationally expensive and may therefore be a more pure assessment of lexical access. We return to these points in the discussion of directions for future research.

Language Group Differences.

A secondary goal of the present work was to investigate models of prediction across English- and French- speaking monolingual children for the purpose of assessing generalization. We anticipated that, given similarities of the relations between linguistic components of each language (Thordardottir, 2005), we would not find many differences across groups in the development of the skills of interest or in relations between skills. On the whole, it was the case that correlations and models were highly similar across groups. However, some differences between samples were observed. First, when evaluating descriptive data on our variables of interest it was revealed that French-speaking children responded more slowly to vocabulary items on the CCT. Additionally, they evinced smaller vocabularies at ages three, four, and five (although this may be due to scale differences on the PPVT compared to the EVIP) but performed significantly better on the nonword repetition task. Stronger performance on nonword repetition and weaker vocabulary in French relative to English may reflect the importance of complex French phonology and morphology as children are moving into early word combinations. In this light however, it is somewhat surprising that we did not see language group effects on sentence repetition. Nevertheless, there were no significant language group interactions in any of the mixed effects models and we conclude that our results largely generalized across languages.

Limitations and Directions for Future Research

Although results from this project provide important evidence on the pattern of early language development, the project is not without its limitations. For example, there are some measurement issues that need to be addressed. First, all of the tasks used in this project were adaptive. This was a necessary component of our research since we were evaluating the development of skills from age three through age five. In order to keep the task the same at each age it would have been necessary either to have items that were far too complex for children as young as age three (which would likely lead to fussiness and frustration), or to risk ceiling level performance at ages four and five. Nevertheless, we did observe a ceiling effect in sentence repetition at age four and five (in both samples) and in nonword repetition at age four (in the French sample only). To address this limitation, we transformed these variables using a logit transformation useful for proportional variables and bounded effects. However, in the mixed model results we still observed a ceiling effect for sentence repetition such that the effect age-two vocabulary weakened across age.

Secondly, our measure of speed of lexical access (operationalized as the speed to touch a referent image in a two-alternative forced-choice task) did not predict any outcome measures in the mixed model analyses. This is in contrast to work by Marchman & Fernald (2008) indicating that visual speed of processing in the second year predicts expressive language and working memory as late as eight years of age. One possibility is that differences in the timing window for haptic and visual responses may have contributed to our failure to find similar prediction at three, four, and five years of age in the present study. However, evidence from prior work utilizing this haptic response measure argues against this interpretation. Specifically, in three published reports, haptic speed of processing, like visual speed of processing, decreases with age and is correlated with vocabulary size in French, English, and Spanish monolingual children and in Spanish-English and French-English bilingual children (Legacy et al., 2016; Legacy et al., 2018; DeAnda et al., 2018). That is, concurrently, haptic and visual speed of processing perform similarly suggesting the issue is not one of timing. Rather, we hypothesize that the lack of longitudinal prediction observed in the present study is due to confounds between speed and other cognitive processes.

That is, as discussed previously, this measure may index executive control, decision confidence, and hypothesis-testing mechanisms in addition to speed of processing. Therefore, we propose some directions for future research. First, evidence from dissociations in modalities on the CCT suggests that visual assessments are not as subject to deliberation effects as are haptic responses such that quick visual fixations to the target image are observed in the absence of a target touch (the child touched the distractor image instead, which is the object upon which they were originally fixating). Thus, latency of visual fixations may be a relatively more pure assessment of the speed of processing of word meaning without the intrusion of deliberation and executive control components. Future research should investigate other assessments of speed of word processing to determine if the original hypothesis of this study can be supported. Additionally, future research should explore whether speed of haptic response in a two-alternative forced-choice assessment is related to the development of executive functions. Finally, it is important to evaluate in greater detail how variability in speed of haptic response is related to concurrent vocabulary size. In the current study and in previous reports (DeAnda et al., 2018; Legacy et al, 2016), vocabulary size is concurrently associated with speed of haptic response such that children who execute more correct touches do so more quickly. What remains unclear is the relative accuracy of these children. It is possible that children who respond more quickly and get more words correct also make relatively many mistakes and that children who respond more slowly are more deliberate and therefore have a lower proportion of mistakes. One future avenue for research is to describe children's proportions of correct vs. erroneous responses as a function of haptic speed. Results may help to clarify why we find a concurrent (but not longitudinal) association between the haptic assessment of lexical access and vocabulary size.

Finally, although we expected age-two vocabulary size to predict conceptual development across ages three, four, and five, its unique contribution to conceptual development was weaker than its prediction to later vocabulary and sentence repetition. This is due in part to the number of control variables, which varied across models. However, this may be due in part to the nature of the task. Our choice of this measure was predicated on prior research showing that children's event knowledge is evinced in a familiar play context (e.g., Farrar et al., 1993). However, it is important to note that, in the Farrar et al. study, children's event knowledge was assessed over a matter of weeks whereas in the present research, repeated exposure to the farm context occurred on an annual basis. This more limited frequency of exposure may have influenced the development of the farm concept. Nevertheless, a metric of conceptual development would be beneficial in evaluating the influence of early vocabulary on long-term knowledge and establishing best practices for such metrics would be a useful direction for future research.

In conclusion, it is critically important to detail the early developmental trajectories of different types of language skills in toddlerhood and examine the predictive utility of each. Specifically, early language is malleable and children benefit from targeted intervention (Cates, Weisleder, & Mendelsohn, 2016; Reese, Sparks, & Leyva, 2010). Pinpointing the best predictors of later language development and how prediction varies across different types of skills will inform strategies for early identification and intervention. In practical terms, combining assessment of multiple early foundational skills has the potential to increase the efficacy of early identification of children at risk and inform development of early intervention strategies catered to a child's specific pattern of deficits. Results from this work are some of the first steps in clarifying these early developmental patterns.

References

Aravind, A., de Villiers, J., Pace, A., Valentine, H., Golinkoff, R., Hirsch-Pasek, K., ... & Wilson, M. S. (2018). Fast mapping word meanings across trials: Young children forget all but their first guess. *Cognition*, *177*, 177-188. doi: 10.1016/j.cognition.2018.04.008

Archibald, L. M. D. (2017). Working memory and language learning: A review. *Child Language Teaching and Therapy*, *33*(1), 5–17. doi: 10.1177/0265659016654206

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189–208. doi: 10.1016/S0021-9924(03)00019-4

Baddeley, A. & Hitch, G. J. (1974) Working memory. In: *Recent advances in learning and motivation, vol. 8*, ed. G. Bower. Academic Press.

Bartoń (2019). MuMIn: Multi-Model Inference. R package version 1.43.6. https://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1-48. doi:10.18637/jss.v067.i01.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B,* 289-300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*, 39-53. doi: 10.1016/j.cognition.2012.08.008

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*(6), 1461–1476. doi: 10.1017/S0142716416000060

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Semantic Structure in Vocabulary Knowledge Interacts With Lexical and Sentence Processing in Infancy. *Child Development*, *87*(6), 1893–1908. doi: 10.1111/cdev.12554

Cates, C. B., Weisleder, A., & Mendelsohn, A. L. (2016). Mitigating the Effects of Family Poverty on Early Child Development through Parenting Interventions in Primary Care. *Academic Pediatrics*, *16*, S112–S120. doi: 10.1016/j.acap.2015.12.015

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*, 741-748. doi: https://doi.org/10.1111/1469-7610.00770

DeAnda, S., Bosch, L., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). The language exposure assessment tool: Quantifying language exposure in infants and children. *Journal of Speech, Language, and Hearing Research*, *59*, 1346-1356. doi: 10.1044/2016_JSLHR-L-15-0234

DeAnda, S., Hendrickson, K., Zesiger, P., Poulin-Dubois, D., & Friend, M. (2018). Lexical access in the second year: a study of monolingual and bilingual vocabulary development. *Bilingualism: Language and Cognition*, *21*(2), 314–327. doi: 10.1017/S1366728917000220

Devescovi, A., & Cristina Caselli, M. (2007). Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language & Communication Disorders*, *42*, 187-208. doi: 10.1080/13682820601030686

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*(8), 848–856. doi: 10.1111/jcpp.12378

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody Picture Vocabulary Test.* Circle pines, MN: American Guidance Service.

Dunn, L. M., Dunn, L. M., & Thériault-Whalen, C. M. (1993). Échelle de Vocabulaire en Images Peabody: Série de Planches: Psycan.

ELAN (Version 5.2) [Computer software]. (2018, April 04). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

Farrar, M. J., Friend, M. J., & Forbes, J. N. (1993). Event knowledge and early language acquisition. *Journal of Child Language*, *20*, 591-606. doi: 10.1017/S0305000900008497

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*, 1149-1160. doi: 10.3758/BRM.41.4.1149

Fenson, L., Dale, P. S., Reznick, S. J., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*, i-185. doi: 10.2307/1166093

Fenson, L., Marchman, V. A., Thal, D., Dale, P., Reznick, J. S. & Bates, E. (2007). MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual. 2nd Edition. Baltimore, MD: Brookes Publishing Co.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*(1), 203–222. doi: 10.1111/j.1467-8624.2011.01692.x

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116. doi: 10.1037/0012-1649.42.1.98

Fernald, A., Pinto, J. P., & Swingley, D. (1998). Rapid gains in the speed and efficiency of word recognition by infants in the second year. *Psychological Science*, *9*(3), 228–231. doi: 10.1111/1467-9280.00044

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E.M. Fernández & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: J. Benjamins.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*, 677-694. doi: 10.1017/S0305000916000209.

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers, 35, 302–309.* doi: 10.3758/BF03202556

Friend, M., & Keplinger, M. (2008). Reliability and validity of the Computerized Comprehension Task (CCT): Data from American English and Mexican Spanish infants. *Journal of Child Language, 35, 77–98.* doi: 10.1017/S0305000907008264

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the Computerized Comprehension Task: Comprehension predicts production. *Developmental Psychology*, *48*(1), 136–148. doi: 10.1037/a0025511

Friend, M., Smolak, E., Liu, Y., Poulin-Dubois, D., & Zesiger, P. (2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. *Developmental Psychology*, *54*(7), 1317–1333. doi: 10.1037/dev0000514

Friend, M., Smolak, E., Patrucco-Nanchen, T., Poulin-Dubois, D., & Zesiger, P. (2019). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year. *Developmental Psychology*, *55*, 9–22. doi: 10.1037/dev0000617

Friend, M., & Zesiger, P. (2011). A systematic replication of the psychometric properties of the CCT in three languages: English, Spanish, and French. *Enfance*, *3*, 329-344.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237-241. doi: 10.1111/j.1467-9280.1996.tb00366.x

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology, 28*, 887. doi: 10.1037/0012-1649.28.5.887

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535-574. doi: 10.1146/annurev.neuro.29.051605.113038

Gottwald, J. M., Achermann, S., Marciszko, C., Lindskog, M., & Gredebäck, G. (2016). An embodied account of early executive-function development: prospective motor control in infancy is related to inhibition and working memory. *Psychological Science*, *27*, 1600-1610. doi: 10.1177/0956797616667447

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. doi: 10.1111/desc.12250

Hendrickson, K., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2017). Assessing a continuum of lexical-semantic knowledge in the second year of life: A multimodal approach. *Journal of Experimental Child Psychology*, *158*, 95–111. doi: 10.1016/j.jecp.2017.01.003

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W. V., Hofman, A., … Tiemeier, H. (2011). Examining Continuity of Early Expressive Vocabulary Development: The Generation R Study. *Journal of Speech Language and Hearing Research*, *54*(3), 854–869. doi: 10.1044/1092-4388(2010/09-0255)

Hilbe, J. M. (2011). *Negative Binomial Regression*. Cambridge University Press.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (Eds.), *Language, Speech, and Communication. Methods for Assessing Children's Syntax* (pp. 105-124). Cambridge, MA, US: The MIT Press.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*, 128-157. doi: 10.1080/15250000701795598

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish-English bilingual toddlers. *Bilingualism: Language and Cognition*, *17*(1), 189–202. doi: 10.1017/S136672891300014X

Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: a computational model of the learning of new words. *Developmental Science*, *10*, 853-873. doi: https://doi.org/10.1111/j.1467-7687.2007.00638.x

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199–225. doi: 10.1016/0001-6918(94)90003-5

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, *27*(3), 227-250. doi: 10.1177/0142723706075789

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S.A.H., Gustafsson, J.E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science, 18,* 146-154. doi: 10.1111/desc.12202

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Resarch Methods*, *41*(3), 841–849. doi: 10.3758/BRM.41.3.841

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, *32*, 69–92. doi: 10.1017/S0142716410000299

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2018). Vocabulary size and speed of word recognition in very young French-English bilinguals: A longitudinal study. *Bilingualism: Language and Cognition*, *21*(1), 137–149. doi: 10.1017/S1366728916000833

Luce, R.D. (1986). *Response times: Their role in inferring elementary mental organization.* Oxford: Oxford University Press.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16. doi: 10.1111/j.1467-7687.2008.00671.x

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language*, *37*(4), 817–840. doi: 10.1017/S0305000909990055

Miller, J., & Iglesias, A. (2012). Systematic analysis of language transcripts (SALT), research version 2012 [computer software]. *Middleton, WI: SALT Software, LLC*.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, *86*(5), 1351–1370. doi: 10.1111/cdev.12398

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reese, E., Sparks, A., & Leyva, D. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, *10*, 97–117. doi: 10.1177/1468798409356987

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., … Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from Early Language in Victoria Study. *Pediatrics*, *126*(6), E1530–E1537. doi: 10.1542/peds.2010-0254

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Developmental Psychology*. doi: 10.1037/dev0000754

Riches, N. G. (2012). Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language & Communication Disorders*, *47*, 499-510. doi: 10.1111/j.1460-6984.2012.00158.x

Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word? *WIREs Cognitive Science*, *8*, e1421. doi: 10.1002/wcs.1421

Shadlen, M. N., & Gold, J. I. (2004). The neurophysiology of decision-making as a window on cognition. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1229–1241). Boston, Mass: MIT Press

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568. doi: 10.1016/j.cognition.2007.06.010

Stokes, S. F., & Klee, T. (2009). Factors that influence vocabulary development in two-year-old children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *50*(4), 498–505. doi: 10.1111/j.1469-7610.2008.01991.x

Thordardottir, E. T. (2005). Early lexical and syntactic development in Quebec French and English: implications for cross-linguistic and bilingual assessment. *International Journal of Language & Communication Disorders*, *40*, 243-278. doi: 0.1080/13682820410001729655

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104. doi: 10.1037/0033-295X.114.1.104

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children’s cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217-230. doi: 10.1016/j.jml.2016.10.001

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152. doi: 10.1177/0956797613488145

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262. doi: 10.1016/j.cognition.2012.06.016

Table 1. Distribution of Selected Demographic Characteristics of Participants from English- and French-speaking samples

	Number (%) of participants	
	English	French
Maternal education		
High School or Less	8 (16.3)	13 (22.0)
Some College	11 (22.4)	13 (22.0)
College Graduate	15 (30.6)	6 (10.2)
Post-Baccalaureate	15 (30.6)	27 (45.8)
Approximate Income		
15,000-34,999	6 (12.2)	1 (3.2)
35,000-49,999	6 (12.2)	1 (3.2)
50,000-74,999	4 (8.2)	5 (16.1)
75,000-99,999	12 (24.5)	2 (6.4)
>100,000	21 (42.9)	22 (71.0)
Mother's Ethnicity		
Asian	2 (4.1)	0 (0)
Black/ not Hispanic	2 (4.1)	3 (5.2)
Hispanic	8 (16.3)	1 (1.7)
White/not Hispanic	28 (57.1)	54 (91.5)
Multicultural/Other	8 (16.3)	0 (0)

Note. For the English sample, income is reported in U.S. Dollars (USD). For the French sample, income is reported in Swiss Francs (CHF). 2 participants declined to provide ethnicity information. 28 participants declined to provide income information. Some percentages may not sum to 100 due to rounding error.

Table 2. Descriptive information of measures of interest by sample with *t*-tests across samples.

	English Sample				French Sample					
Variables	N	Mean	SD	Range	N	Mean	SD	Range	*t*-test	Sig.
CCT Vocabulary	49	27.53	7.39	9.00–39.00	59	29.29	5.76	12–40	-1.36	.18
Lexical Access	49	2633.11	602.69	1253.77–4096.75	59	3133.43	707.65	1901.82–4739.50	-3.97	<.001*
Number of Different Farm Words										
Age Three	46	25.37	10.12	1–44	58	31.21	12.80	5–71	-2.60	.01*
Age Four	46	26.91	12.82	9–58	57	30.82	12.95	9–67	-1.53	.13
Age Five	39	29.23	11.45	8–58	53	31.62	10.33	12–51	-1.03	.31
Number of Utterances										
Age Three	46	127.02	38.59	50–240	58	148.86	42.92	38–270	-2.73	.008*
Age Four	46	95.87	33.03	36–204	57	98.35	27.65	40–156	-.41	.68
Age Five	39	101.15	26.96	50–180	53	101.91	31.85	36–176	-.12	.90
Concept										
Age Three	46	.21	.09	.02–.42	58	.21	.08	.06–.41	-.33	.74
Age Four	46	.29	.11	.10–.52	57	.32	.13	.10–.77	-1.53	.13
Age Five	39	.30	.11	.14–.60	53	.32	.10	.14–.60	-1.18	.24
PPVT Vocabulary										
Age Three	49	49.75	18.58	11–93	59	28.42	11.33	7–61	7.02	<.001*
Age Four	46	71.80	16.09	31–106	55	55.91	14.68	27–92	5.14	<.001*
Age Five	39	89.44	13.01	54–124	54	73.63	13.15	39–101	5.76	<.001*
Sentence Repetition										
Age Three	43	.84	.14	.36–.99	46	.86	.15	.42–.99	-.58	.57
Age Four	45	.93	.08	.54–1.00	54	.93	.09	.55–1.00	.49	.63
Age Five	39	.96	.05	.73–1.00	49	.97	.03	.87–1.00	-1.21	.23
Nonword Repetition										
Age Three	41	.80	.09	.61–.93	55	.88	.11	.49–1.00	-4.03	<.001*
Age Four	45	.89	.08	.62–.97	57	.94	.08	.68–1.00	-3.30	.001*
Age Five	38	.91	.05	.81–1.00	53	.96	.04	.85–1.00	-5.07	<.001*

Note. *significant after family-wise FDR correction.

PPVT and EVIP values are raw scores.

Table 3. Correlation matrix of predictive and longitudinal relations in the English sample and French sample.

	English Sample			
	Vocabulary	Lexical Access	Age 3	Age 4
Vocabulary				
Lexical Access	-.28*			
Conceptual Knowledge				
Age Three	.35	.06		
Age Four	.24	-.31	.28	
Age Five	.26	-.08	.32	.43*
PPVT Vocabulary				
Age Three	.63*	.04		
Age Four	.46*	.05	.63*	
Age Five	.50*	-.10	.55*	.60*
Sentence Repetition				
Age Three (logit)	.57*	-.11		
Age Four (logit)	.34	-.07	.59*	
Age Five (logit)	.35	.04	.52*	.75*
Nonword Repetition				
Age Three (logit)	.34	-.19		
Age Four (logit)	.28	-.03	.16	
Age Five (logit)	.15	-.34	.32	.45*
	French Sample			
	Vocabulary	Lexical Access	Age 3	Age 4
Vocabulary				
Lexical Access				
Conceptual Knowledge				
Age Three	.34	-.21		
Age Four	.18	-.06	.45*	
Age Five	-.05	-.12	.33*	.37*
PPVT Vocabulary				
Age Three	.50*	-.25		
Age Four	.52*	-.20	.53*	
Age Five	.47*	-.36*	.64*	.66*
Sentence Repetition				
Age Three (logit)	.49*	-.12		
Age Four (logit)	.40*	.006	.70*	
Age Five (logit)	.22	-.05	.55*	.66*
Nonword Repetition				
Age Three (logit)	.27	-.05		
Age Four (logit)	.15	-.01	.19	
Age Five (logit)	.10	.03	.24	.51*

Note. *significant after family-wise FDR correction.

Table 4. Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable conceptual knowledge.

	Conceptual Knowledge
Vocabulary Model	
+ Age	$\chi^2 = 55.96$, $p < .001$
+ Language Group	$\chi^2 = 2.04$, $p = .15$
+ Maternal Education	$\chi^2 = 10.06$, $p = .002$
+ Vocabulary	$\chi^2 = 7.04$, $p = .008$
+ Age X Vocabulary	$\chi^2 = 1.67$, $p = .20$
+ Group X Vocabulary	$\chi^2 = .03$, $p = .87$
+ Age X Group X Vocabulary	$\chi^2 = 1.65$, $p = .20$
AIC	-520.99
BIC	-480.28
LL	271.49
Lexical Access Model	
AIC	-513.94
BIC	-473.24
LL	267.97
Full Model	
AIC	-515.60
BIC	-460.09
LL	272.80

Table 5. Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable PPVT Vocabulary.

	PPVT Vocabulary
Vocabulary Model	
+ Age	$\chi^2 = 349.71$, $p < .001$
+ Language Group	$\chi^2 = 46.76$, $p < .001$
+ Vocabulary	$\chi^2 = 51.30$, $p < .001$
+ Age X Vocabulary	$\chi^2 = 1.81$, $p = .18$
+ Group X Vocabulary	$\chi^2 = .03$, $p = .86$
+ Age X Group X Vocabulary	$\chi^2 = 3.61$, $p = .06$
AIC	2346.31
BIC	2383.42
LL	-1163.16
Lexical Access Model	
AIC	2395.99
BIC	2433.10
LL	-1188.00
Full Model	
AIC	2345.82
BIC	2397.62
LL	-1158.91

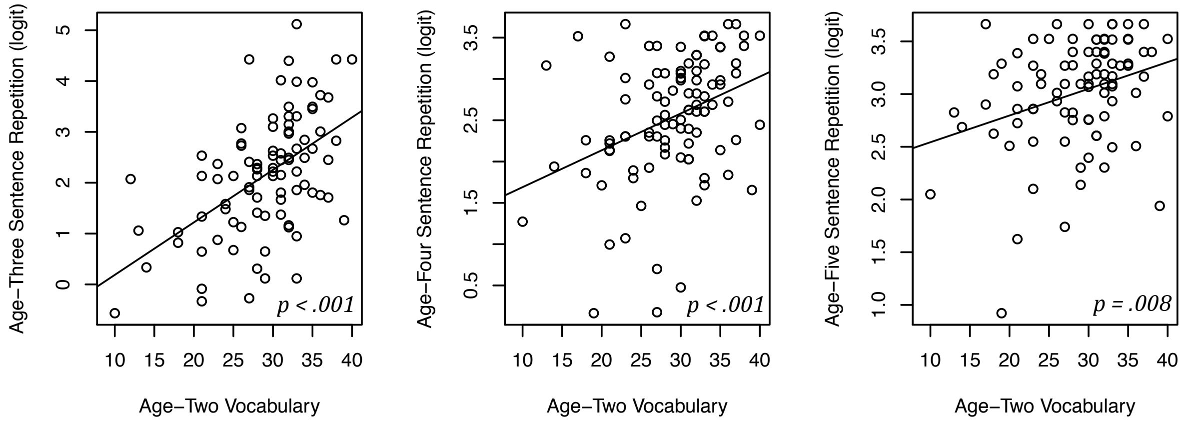
Table 6. Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable Sentence Repetition.

	Sentence Repetition
Vocabulary Model	
+ Age	χ^2 = 69.06, p < .001
+ Language Group	χ^2 = .72, p = .40
+ Child Sex	χ^2 = 4.92, p = .03
+ Vocabulary	χ^2 = 21.53, p < .001
+ Age X Vocabulary	χ^2 = 28.42 p < .001
+ Group X Vocabulary	χ^2 = .02, p = .89
+ Age X Group X Vocabulary	χ^2 = .70, p = .40
AIC	603.93
BIC	643.76
LL	-290.97
Lexical Access Model	
AIC	652.05
BIC	691.87
LL	-315.02
Full Model	
+ Age	χ^2 = 69.06, p < .001
+ Language Group	χ^2 = .72, p = .40
+ Child Sex	χ^2 = 4.92, p = .03
Vocabulary Entered First	
+ Vocabulary	χ^2 = 21.53, p < .001
+ Age X Vocabulary	χ^2 = 28.42 p < .001
+ Group X Vocabulary	χ^2 = .02, p = .89
+ Age X Group X Vocabulary	χ^2 = .70, p = .40
+ Lexical Access	χ^2 = 2.53, p = .11
+ Age X Lexical Access	χ^2 = .09, p = .77
+ Group X Lexical Access	χ^2 = 1.05, p = .31
+ Age X Group X Lexical Access	χ^2 = 1.94, p = .16
Lexical Access Entered First	
+ Lexical Access	χ^2 = .002, p = .96
+ Age X Lexical Access	χ^2 = .66, p = .42
+ Group X Lexical Access	χ^2 = .02, p = .90
+ Age X Group X Lexical Access	χ^2 = .20, p = .66
+ Vocabulary	χ^2 = 24.56, p < .001
+ Age X Vocabulary	χ^2 = 26.89, p < .001
+ Group X Vocabulary	χ^2 = .69, p = .40
+ Age X Group X Vocabulary	χ^2 = 1.57, p = .21
AIC	606.32
BIC	660.63
LL	-288.16

Table 7. Estimates of model fit for each candidate model and results from likelihood ratio tests of the addition of each parameter to the repeated measures model with the dependent variable Nonword Repetition.

	Nonword Repetition
Vocabulary Model	
+ Age	$\chi^2 = 65.69$, $p < .001$
+ Language Group	$\chi^2 = 39.08$, $p < .001$
+ Child Sex	$\chi^2 = 7.56$, $p = .006$
+ Vocabulary	$\chi^2 = 6.39$, $p = .01$
+ Age X Vocabulary	$\chi^2 = 2.23$, $p = .13$
+ Group X Vocabulary	$\chi^2 = .11$, $p = .74$
+ Age X Group X Vocabulary	$\chi^2 = .34$, $p = .56$
AIC	589.10
BIC	629.39
LL	-283.55
Lexical Access Model	
AIC	597.24
BIC	637.53
LL	-287.62
Full Model	
+ Age	$\chi^2 = 65.69$, $p < .001$
+ Language Group	$\chi^2 = 39.08$, $p < .001$
+ Child Sex	$\chi^2 = 7.56$, $p = .006$
Vocabulary Entered First	
+ Vocabulary	$\chi^2 = 6.39$, $p = .01$
+ Age X Vocabulary	$\chi^2 = 2.23$, $p = .13$
+ Group X Vocabulary	$\chi^2 = .11$, $p = .74$
+ Age X Group X Vocabulary	$\chi^2 = .34$, $p = .56$
+ Lexical Access	$\chi^2 = .46$, $p = .50$
+ Age X Lexical Access	$\chi^2 = .40$, $p = .53$
+ Group X Lexical Access	$\chi^2 = 1.10$, $p = .29$
+ Age X Group X Lexical Access	$\chi^2 = .23$, $p = .63$
Lexical Access Entered First	
+ Lexical Access	$\chi^2 = .09$, $p = .77$
+ Age X Lexical Access	$\chi^2 = .03$, $p = .86$
+ Group X Lexical Access	$\chi^2 = .26$, $p = .61$
+ Age X Group X Lexical Access	$\chi^2 = .64$, $p = .42$
+ Vocabulary	$\chi^2 = 7.00$, $p = .008$
+ Age X Vocabulary	$\chi^2 = 2.44$, $p = .12$
+ Group X Vocabulary	$\chi^2 = .60$, $p = .44$
+ Age X Group X Vocabulary	$\chi^2 = .29$, $p = .59$
AIC	594.91
BIC	649.85
LL	-282.45

Figure 1. Illustration of the interaction between age and vocabulary on sentence repetition.



Chapter 4, in part, is currently being prepared for submission for publication of the material. Smolak, E., Helm, J., Patrucco-Nanchen, T., Poulin-Dubois, D., Zesiger, P., & Friend, M. (in preparation). Convergence and divergence in prediction from vocabulary and speed of lexical access. The dissertation author was the primary investigator and author of this material.

CHAPTER 5:

General Discussion

The set of studies contained herein sought to investigate the stability of early language abilities in children from toddlerhood through the preschool years. Detailed examination of the stability of early language trajectories is important to the field of child language development for several reasons (Bornstein, Hahn, Putnick, & Suwalsky, 2014). First, if there is stability in language and other cognitive abilities across time this informs theories of the learning mechanisms underlying language acquisition, suggesting similar (and potentially domain-general) mechanisms are involved in basic (e.g., vocabulary) and more complex (e.g., morphosyntactic) components of the language system (Gómez & Gerken, 2000; Marchman & Bates, 1994). Second, consistency in children's relative performance across time has practical applications for early identification of children at risk for language delay or impairment. Indeed, early weakness in vocabulary and grammar within the first 2-3 years is associated with continued language weakness in adolescence (Rescorla, 2009) although these children do not all present with a later diagnosed language disorder (Wallace et al., 2015). Since children benefit from early, targeted intervention consisting of, for example, promoting parental engagement or shared book reading, early identification of risk and early intervention strategies have the potential to mitigate negative outcomes (Cates, Weisleder, & Mendelsohn, 2016; Hassinger-Das et al., 2016; Reese, Sparks, & Leyva, 2010). The three specific aims of this dissertation were to 1) Assess the relation between decontextualized vocabulary and downstream language at the group and individual levels, 2) Investigate differential prediction from vocabulary and lexical access and 3) assess generalization of findings in French- and English-speaking children.

The chapters in this disssertation provided evidence in support of the notion of the stability of early language and literacy trajectories. First, Chapter 2 followed up on recent results from large-scale studies that early vocabulary as assessed in the first two years of life predicts later literacy and academic achievement (Bleses et al., 2016; Duff et al., 2015; Lee, 2011; Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015). However, despite the evidence for stability between early language and later literacy, some studies revealed that the portion of variance in literacy achievement predicted by early vocabulary quite low (e.g., around 3 percent; Bleses et al., 2016). If vocabulary as early as age two predicts a small amount of variance in outcomes, it is unlikely to be of practical use in early identification of risk. However, there is a theoretical reason to expect greater stability between vocabulary and literacy outcomes that what was observed in some results. For example, the development of vocabulary is dependent at least in part on domain general learning mechanisms and statistical learning (Gómez & Gerken, 2000; Yu & Smith, 2012; Vlach & DeBrock, 2017). To the extent that academic achievement and literacy are also dependent on these learning mechanisms, we should expect more robust prediction from early vocabulary to later language and literacy. Indeed, a study utilizing factor analysis assessment of core language skill has revealed that there is strong stability in language over time (Bornstein et al., 2006). One potential reason for the modest prediction observed in the studies above is the use of a vocabulary measure that we hypothesize includes assessment of children's vocabulary knowledge that exists along a continuum from weak to strong (i.e., parent report of vocabulary). The contribution of Chapter 2 was to assess prediction from early vocabulary to kindergarten readiness at age four using a measure of vocabulary size that we hypothesize restricts assessment to robust, decontextualized vocabulary knowledge (Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015; Hendrickson, Poulin-Dubois, Zesiger, & Friend, 2017).

Results from this study revealed that decontextualized vocabulary size at age two was a robust predictor of later vocabulary and kindergarten readiness at age four. The measure of decontextualized vocabulary was revealed to be a relatively more important predictor than was parent report. Moreover, this effect was consistent across three samples: English speaking monolingual children, French speaking monolingual children, and French-English bilingual children. Finally, we expected that this effect of early vocabulary on kindergarten readiness could be part of a mediation model such that vocabulary at age two predicts vocabulary at age four which in turn is associated with kindergarten readiness. We tested a mediation model across samples, which did not yield evidence to support a full mediation effect however we acknowledge this could be due to limited power or to a conservative test of mediation. We return to this point below.

Strengths of this study were the inclusion of three samples of children tested across labs, differing geographically and linguistically. Results were remarkably similar across groups, suggesting generalizability of the findings. In the bilingual sample, similar results were observed but only for the dominant language. That is, vocabulary in the dominant, but not the non-dominant language, significantly predicted kindergarten readiness (an assessment that was also given in the dominant language). Thus this study presented preliminary evidence that, across different groups of children, strong early vocabulary development is critical for school success. An additional strength was the application of statistical methods for assessing the relative importance of different early assessments of vocabulary knowledge. Results from these analyses provide support for the relative importance of early, decontextualized vocabulary size for predicting vocabulary and kindergarten readiness at age four.

However, there were also limitations in this study to address. First, from the mediation results we concluded that there was not evidence for a mediation model and that there was a direct effect of early vocabulary on kindergarten readiness. However, many of the causal steps necessary for mediation were met. Indeed, in the English sample all steps for mediation were met before the application of a family-wise error correction. However, the Sobel tests for mediation were not significant. It is possible that we did not have the power to observe a full mediation effect. Thus, we only tentatively conclude there was not mediation but do suggest that future evidence with larger sample sizes investigate mediation models further. A second limitation was that we did not have a large enough sample size of bilingual children to conduct the same relative importance and regression mediation results as we did for the monolingual samples. Instead, we applied partial correlation approaches for the smaller sample. Our results from the bilingual sample (in the dominant language) parallel those from the monolingual samples however we note that future research should examine these questions in larger bilingual samples.

Chapter 3 follows from the results from Chapter 2 on the utility of early, decontextualized vocabulary for prediction to later vocabulary and kindergarten readiness at the group level. This study sought to extend these results by examining longitudinal prediction to general language ability at both the group level and the individual level in two groups of children: French monolingual and English monolingual children. We evaluated the strength of prediction from parent-reported compared to decontextualized vocabulary at ages 16 months and 23 months to general language ability at age three at the group level. Then, we divided children into two groups: those with average-to-high language abilities compared to those with low language abilities. We defined children with low language abilities as those who scored at least one standard deviation below their peers on a language factor composed of mean length of utterance, vocabulary, and sentence repetition. Our aim was to use a strong early predictor of this dichotomous outcome using receiver operating characteristic (ROC) curves.

First, evaluations of model fit revealed that the strongest prediction to general language skills at age three was obtained by a model with age-two decontextualized vocabulary as the sole predictor. Parent-reported vocabulary did not add an adequate amount of information to model fit above decontextualized vocabulary for us to conclude that full model was a better fit. Relatedly, we found that prediction from vocabulary at 16 months of age was not robust. Indeed, vocabulary at 23 months was a much better fit to the data. This was true both for English monolingual participants and French monolingual participants. Results from these analyses prompted us to narrow in on decontextualized vocabulary at 23 months as having the most potential as a predictor of the dichotomous language outcome although parent-reported vocabulary was a significant predictor of the age-three general language factor.

At the individual level, an optimal cutoff on the decontextualized vocabulary measure at age two (determined by the ROC curve) correctly identified 85 percent of children with low language at age three. Further, this cutoff correctly rejected 76 percent of children with average-to-high language at age three. These estimates, referred to as sensitivity and specificity, respectively, were higher than any other extant study with longitudinal identification. They were also higher than the estimates using parent reported vocabulary in the same study (74 percent and 77 percent). Thus, similar to the results we obtained for Chapter 2, results from Chapter 3 revealed that decontextualized vocabulary is a robust predictor of later language at both the group and individual levels and may have potential as an early screening measure for children at risk for low language in the preschool years. We note that we did not conclude that the MCDI could not also be useful for screening purposes but rather that an assessment of decontextualized vocabulary size may be particularly helpful in determining risk for lower language performance.

There were several strengths of the study contained in Chapter 2. First, we assessed children's general language ability at three years of age by constructing a language factor composed of mean length of utterance, vocabulary size, and sentence repetition performance. The benefit of using a factor analysis approach is that it removes unshared variance of the indicator variables (mean length of utterance, vocabulary, sentence repetition) to arrive at a measure reflecting the common variance shared across these variables. Thus, we can consider the general language factor to cover children's skills across a range of early language abilities. Second, we included examination of effects across different samples of monolingual children to assess generalization. In this way we can evaluate differences as well as similarities that emerged across samples. For example, of the 16-month assessments, parent reported vocabulary *production* was the strongest predictor in the English group whereas parent reported vocabulary *comprehension* was the strongest predictor in the French group, revealing variability across groups in the effects of parent reported vocabulary. However, in both groups, decontextualized vocabulary at 23 months was the strongest predictor of language skills at age three. Similarities in prediction from decontextualized vocabulary at 23 months in addition to similarities in the relation between mean length of utterance, vocabulary, and sentence repetition to the general language factor across groups allowed us to combine these two groups to take advantage of the increased sample size for the purposes of the ROC curves. This ROC curve approach was also a strength as it allowed us to present preliminary evidence on the utility of early decontextualized vocabulary for early identification of children at risk for lower language abilities.

However, this study also contained its weaknesses. First and foremost, we recruited a typically developing sample of children at 16 months of age with no known deficits. A small subsample (N = 8 out of 110) of children were diagnosed with speech-language difficulties or referred to speech-language services for a language/speech delay. We correctly identified 5 out of 8 of these children using the optimum cutoff on the decontextualized vocabulary measure. However, this number is too low to calculate accurate sensitivity and specificity estimates for the entire sample. Thus, we chose to also evaluate sensitivity and specificity of prediction to children with "low language" abilities compared to children with "average-to-high" language abilities. The sensitivity and specificity estimates for this binary outcome were good however we note that we do not conclude that the children identified with "low language" in this study had a clinical language delay/impairment. Rather, these children were in the low end of the distribution relative to their peers who were also enrolled in the study. The results from the present study provide preliminary evidence for the utility of decontextualized vocabulary as a screening measure however future research should be conducted with samples of children with higher risk for language disorders (e.g. a family history of speech/language disorders). Secondly, the outcome variable of interest used in this study was the general language factor, which included assessment of syntactic development and language processing in addition to vocabulary size. Given the correlations between these three indicator variables as well as the strong factor loadings for each, we conclude that we developed a language factor that represents children's language skill more generally. Nevertheless, there is research to suggest that although children with language impairment may have smaller vocabularies than their peers into school age and adolescence, vocabulary size has not been acknowledged as an effective tool, on its own, for identification of developmental language disorder in the preschool years and the school ages (Gray, Plante, Vance, & Henrichsen, 1999; Rice & Hoffman, 2015; Shahmahmood, Jalaie, Soleymani, Haresabadi & Nemati, 2016). However, in our research, vocabulary loaded with syntactic and language complexity assessments onto a general factor. Factor construction, in turn, focused on shared variance and reduced the influence of method variance, which may contribute to the ineffectiveness of vocabulary alone as a predictor. Thus, our identification of "low language" children was based on a general factor reflecting the shared variance across three measures of language skill.

Finally, the goal of Chapter 4 was to assess the hypothesis that vocabulary and speed of lexical access have both shared and differential prediction to language development. A body of work has investigated the importance of early lexical access/word processing (alone and in combination with parent report of vocabulary size) to language development (Fernald & Marchman, 2012; Fernald, Perfors, & Marchman, 2006; Marchman & Fernald, 2008). The first two chapters of this dissertation provided evidence for the robust prediction from early, decontextualized vocabulary to later language and kindergarten readiness above and beyond a parent report of vocabulary size. Still lacking however is the assessment of potential differential prediction from decontextualized vocabulary size and speed of lexical access. These components of language ability are related, but are also somewhat dissociable. In Chapter 4 we hypothesized that these components might yield differential prediction to the development of different types of language abilities throughout the preschool period. Specifically, the development of decontextualized vocabulary size is dependent in part on long-term retention of semantic knowledge built up from multiple experiences with word-referent pairings and is thus dependent upon the long-term memory system (Bion, Borovsky & Fernald, 2013; Horst & Samuelson, 2008; Yu & Smith, 2012; Vlach & DeBrock, 2017). On the other hand, lexical access is related to long-term knowledge but is highly dependent on processing efficiency, which is also linked to the working memory system (Fry & Hale, 1996; Kail & Salthouse, 1994; Marchman & Fernald, 2008). Thus, one way in which differential prediction may be observed is in considering prediction to linguistic outcomes that vary along a similar continuum from long-term knowledge and concepts to processing and working memory. The goal of this project was to gain more detailed information on the developmental trajectories of the language system throughout the early years of life. Results can help to clarify the different ways in which early language abilities build the foundation for the health of the language system.

However, on the whole, results from Chapter 4 did not support the notion that vocabulary and lexical access were dissociable in their prediction to outcomes ranging from long-term conceptual knowledge to working memory. As expected based on our hypothesis of a continuum from long-term knowledge to processing and working memory, vocabulary at age two predicted later vocabulary and conceptual development more strongly than did lexical access. Additionally, we found that age-two vocabulary more strongly predicted semantic knowledge than phonological working memory. However, the prediction from age-two vocabulary to conceptual knowledge was somewhat weak and the prediction to sentence repetition was strong which was in contrast to our expectations. Indeed, although we expected that our outcome measure of sentence repetition would be equally predicted by vocabulary and lexical access due to the nature of its task demands, only vocabulary predicted sentence repetition. Finally, our outcome measure of nonword repetition is considered to be an assessment of phonological short term/working memory. However, vocabulary predicted nonword repetition (accounting for a small amount of variance) and lexical access did not predict nonword repetition. An explanation for this finding lies, we think, in our assessment of lexical access. Due to the fact that this measure depends on two-year-old children's haptic response to two alternative image referents for a given word, we hypothesize that speed to touch reflects not only processing of the word's meaning but also deliberation mechanisms and motor planning. Thus, this speed measure may be an evaluation of children's executive control and hypothesis testing about word meanings in addition to lexical access. If this is the case it is not surprising that this measure would not load onto working memory components. However, directions for future research involve investigating whether this haptic measure does depend to some extent on executive control and deliberation mechanisms in addition to exploring differential prediction with a purer assessment of speed of word processing.

This study utilized a large sample of children composed of two groups differing linguistically and geographically. This is a strength of the study, since we were able to evaluate paths of prediction for a large sample and assess generalization of effects across language groups. There were observed differences across groups in performance with age: for example, French-speaking children responded more slowly on the haptic speed of lexical access measure than did English-speaking children. Moreover, their scores on a French receptive vocabulary measure were lower than English-speaking children (however, because we assessed raw scores on this instrument, it is possible this difference represents scale differences across the two adaptations). Finally, French-speaking children outperformed their English-speaking peers on a task of nonword repetition. Despite these performance differences, we generally found that paths of prediction were similar across groups, suggesting generalization of predictive relations despite linguistic differences. This is an important contribution because it suggests that similar learning and growth mechanisms underlie language acquisition across these groups. Additionally, this points to a possibility for similar screening and intervention strategies across children differing in linguistic background whether it is English or French. Future cross-linguistic work should investigate developmental trajectories across groups of children to further examine how these findings generalize to other languages.

In sum, the studies contained in this dissertation provided evidence that early language abilities including decontextualized vocabulary size and speed of lexical access are important building blocks for downstream language development and readiness for school. We provided evidence that early, decontextualized vocabulary is a robust predictor of later vocabulary, kindergarten readiness, and general language abilities, accounting for greater variance than parent report of early vocabulary. This finding supports idea that there is continuity in language abilities over time (Bornstein et al., 2006) and suggests that findings of modest variance in outcomes accounted for by vocabulary in prior work may reflect methodological limitations in early assessment. We also provided preliminary evidence for the utility of an early assessment of decontextualized vocabulary for predictively differentiating children with low language abilities from children with average-to-high language abilities. The results from this study revealed good values of sensitivity and specificity of the decontextualized vocabulary assessment, though we note the necessity of future research with samples of children with higher risk for language disorder. Finally, we extended recent work on longitudinal prediction and the relation between vocabulary and speed of lexical access by investigating differential prediction. Evidence from this study suggests that vocabulary knowledge was the primary predictor of the range of linguistic outcomes, though the strength of its prediction varied. Speed of lexical access as assessed in this study did not predict downstream linguistic development though we note the necessity for additional studies evaluating what skills are involved in touch latency and whether another assessment of lexical access/speed of word processing could yield the differential effects that we expected.

The detailed description of longitudinal trajectories and long-term prediction is necessary for researchers to develop a full picture of early language development, strategies for early identification, and the different patterns of deficits that can result in language impairment. Overall the results from the studies presented herein provide evidence that the early development of robust, decontextualized vocabulary knowledge provides a foundation for the later development of a diverse range of skills including kindergarten readiness, vocabulary knowledge, language complexity, and sentence processing/production. We built on extant work that suggests continuity in language development over time but which has found that vocabulary predicts a rather modest amount of variance in outcomes. Instead, we show that an easy-to-administer task of decontextualized vocabulary predicts a moderate amount of variance in outcomes and has the potential for early screening of children at risk for language delay. Finally, we provide an initial examination into more detailed trajectories of development by investigating exactly how different components of language in toddlerhood pave the way for downstream development. Future directions include application of these methods to samples of children with risk indicators for language disorder and expanding investigation into differential paths of prediction.

References

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*, 39-53.

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, *37*, 1461-1476.

Bornstein, M. H., Hahn, C. S., Bell, C., Haynes, O. M., Slater, A., Golding, J., ... & ALSPAC Study Team. (2006). Stability in cognition across early childhood: A developmental cascade. *Psychological Science*, *17*, 151-158.

Cates, C. B., Weisleder, A., & Mendelsohn, A. L. (2016). Mitigating the effects of family poverty on early child development through parenting Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in cognitive sciences*, *4*, 178-186.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*, 848-856.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*, 203-222.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*, 98-116.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237-241.

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*, 196-206.

Hassinger-Das, B., Ridge, K., Parker, A., Golinkoff, R. M., Hirsh-Pasek, K., & Dickinson, D. K. (2016). Building vocabulary knowledge in preschoolers through shared book reading and gameplay. *Mind, Brain, and Education*, *10*, 71-80.

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*, 723-735.

Hendrickson, K., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2017). Assessing a continuum of lexical–semantic knowledge in the second year of life: A multimodal approach. *Journal of Experimental Child Psychology*, *158*, 95-111.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*, 128-157.

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199-225.

Lee, J. (2011). Size matters: Early vocabulary as a predictor Marchman, V.A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339-366.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, F9-16.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, *86*, 1351-1370.

Reese, E., Sparks, A., & Leyva, D. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, *10*, 97-117.

Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. *Journal of Speech, Language, and Hearing Research*.

Rice, M. L., & Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2; 6 to 21 years of age. *Journal of Speech, Language, and Hearing Research*, *58*, 345-359.

Shahmahmood, T. M., Jalaie, S., Soleymani, Z., Haresabadi, F., & Nemati, P. (2016). A systematic review on diagnostic procedures for specific language impairment: The sensitivity and specificity issues. *Journal of Research in Medical Sciences: the Official Journal of Isfahan University of Medical Sciences*, *21*.

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217-230.

Wallace, I. F., Berkman, N. D., Watson, L. R., Coyne-Beasley, T., Wood, C. T., Cullen, K., & Lohr, K. N. (2015). Screening for speech and language delay in children 5 years old and younger: A systematic review. *Pediatrics*, 1-15.

Yu, C., & Smith, L. B. (2012). Modeling cross-situational word–referent learning: Prior questions. *Psychological Review*, *119*, 21.