Title
Multi-Domain Synchrony Within Vocal Development

Author
Pretzer, Gina Marie

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

Multi-Domain Synchrony Within Vocal Development

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

in

Cognitive and Information Sciences

by

Gina Marie Pretzer

Committee in Charge:
Professor Teenie Matlock, chair
Professor Eric A. Walle
Professor Anne S. Warlaumont

2019

The Dissertation of Gina Marie Pretzer is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Eric A. Walle, Ph.D., Committee Member
Associate Professor of Psychological Sciences
University of California, Merced

Anne S. Warlaumont, Ph.D., Committee Member
Associate Professor, Department of Communication
University of California, Los Angeles

Teenie Matlock, Ph.D., Committee Chair
Professor of Cognitive and Information Sciences
McClatch Chair in Communication
University of California, Merced

University of California, Merced
2019

This dissertation is dedicated to my family for their unfailing love and support throughout this great journey.

I also dedicate this work to the many friends and colleagues who have stood by with words of encouragement, deep conversations, and funny cat videos to keep my spirits high.

This would not be possible without all of you.

Contents

List of Tables

List of Figures

Acknowledgements

First, I would like to thank my committee members – Teenie Matlock, Anne Warlaumont, and Eric Walle – for all their support and mentorship since I started my graduate studies at UC Merced.

Thank you, Teenie, for supporting me through all the comings and goings of the last couple years. I am so grateful that you were able and willing to bring me into your lab. I'm so glad we've had the opportunity to work together. Thank you, Anne, for launching me into the world of research and guiding me through the early years of my program. I am even more grateful for your continued mentorship and support throughout the later years. I've done so much and met so many amazing people in our field because of you. Thank you, Eric, for your suggestions and insights for projects, papers, and presentations over the last five years. Although I was a student in another department, I always felt that I could stop by and ask for ideas or advice.

I want to thank Jeff Yoshimi and Stephanie Shih for serving as members of previous committees throughout my studies. Thank you to Dan Hudock and Heather Ramsdell-Hudock for encouraging me to go into a research program. I never would have had the confidence to apply without you.

A very special thanks to all the Research Assistants who have spent countless hours coding vocalizations, processing audio files, organizing data, filling out spreadsheets, and going on pick up/drop off visits. Extra credit to Sara Mendoza and Jimel Mutrie for all your work to make things go smoothly for all of us. A million thanks to my fellow grad students, Sara Schneider, who has kept the lab running in tip top shape this past year, and Lukas Lopez, who has done the bulk of the training of the new research assistants.

I also thank the friends and colleagues at school and at work for challenging me to think critically and reflectively. I'm a better scientist and SLP because of you. So much love my to fellow speechies who inspire me on a daily basis. Thanks for talking and sharing ideas with me throughout this writing process.

My heartfelt gratitude goes out to my dear friends (old and new) and family who been by my side through all the ups and downs of this journey. You know who you are, and you know how much I value your support and encouragement. There are no words sufficient to express how deeply grateful I am to each and every one of you.

This work was funded in part by the National Science Foundation through grant numbers BCS-1529127 and SMA-1539129/1827744, by a James S. McDonnell Foundation Fellowship in Understanding Human Cognition, and by a Faculty Research Grant from the University of California, Merced Graduate and Research Council.

Curriculum Vitae

Gina M. Pretzer, MS, CCC-SLP
gpretzer@ucmerced.edu
ginapretzer.com

Education

Doctor of Philosophy, Cognitive and Information Sciences, University of California, Merced; School of Social Sciences, Humanities (2014-2019)

Advisor, Dr. Anne Warlaumont (2014-2017)
Dr. Teenie Matlock (2017-present)

Master of Science, Speech-Language Pathology, Idaho State University, School of Rehabilitation & Communication Sciences, Department of Communication Sciences and Disorders, (2012-2014)

Bachelor of Science, Communication Sciences and Disorder with an emphasis in Speech-Language Pathology, School of Rehabilitation & Communication Sciences, Department of Communication Sciences and Disorders, (2007-2012)

Manuscripts

Pretzer, G. M., Lopez, L. D., Walle, E. A., & Warlaumont, A. S. (in press). Infant-Adult Vocal Interaction Dynamics Depend on Infant Vocal Type, Child-Directedness of Adult Speech, and Timeframe.

Lopez, L. D., Walle, E. A., **Pretzer, G. M.,** & Warlaumont, A. S. (under review). Parent Social Engagement with Crawling and Walking Infants: Relations with Joint Attention and Language.

VPS, R., **Pretzer, G. M.,** Mendoza, S., Shedd, C., Kello, C. T., Gopinathan, A., & Warlaumont, A. S. (in preparation). A foraging approach to analysing infant and caregiver vocal behaviour.

Pretzer, G. M., Lopez, L.D., Walle, E. A., & Warlaumont, A. S. (in preparation). Locomotion, Language, and Interactions: How adult responsivity to infant vocalizations promote vocabulary development in walking and crawling infants.

Presentations (Presenter in bold face)

Pretzer, G. M. (2019, April). Infant Vocal Development for Early Interventionists at HopeSparks Children's Developmental Services Professional Development Seminar, Tacoma, WA.

Lopez, L. D., **Walle, E. A.,** Pretzer, G. M., & Warlaumont, A. S. (2018, July). Adult Responses to Infant Prelinguistic Vocalizations are Associated with Infant Vocabulary: A Home Observation Study. Presented in Walle & Lopez "Crawling and Walking Infants' Home Language Environments: Relations with Concurrent Vocabulary Size" at the 2018 International Congress of Infant Studies, Philadelphia, PA.

Walle, E. A., Lopez, L. D., Pretzer, G. M., & Warlaumont, A. S. (2018, July). Crawling and Walking Infants' Home Language Environments: Relations with Concurrent Vocabulary Size. Presented in Karasik & Walle "Walk, talk, play: Changes across the transition from crawling to walking in infancy" at the 2018 International Congress of Infant Studies, Philadelphia, PA.

Pretzer, G.M & Wood, P. (2017, October). Language Development in Bilingual Infants. Presented at CCCDS Training, Merced, CA.

VPS, R., Shea, T. M., Pretzer, G. M., Mendoza, S., Shedd, C., Kello, C., Gopinathan, A., & Warlaumont, A. S. (2017, August). Infant Vocalization as Foraging in an Acoustic Space. Paper presented at the Canadian-American-Mexican Graduate Student Physics Conference, Washington D.C.

Walle, E. A., Lopez, L. D., Johal, J., Pretzer, G., & Warlaumont, A. S. (2017, April). Parent Social Engagement with Crawling and Walking Infants: Relations with Joint Attention and Language. Paper presented at the Society of Research in Child Development biennial conference, Austin, TX.

Lopez, L. D., Pretzer, G., Walle, E. A., & Warlaumont, A. S. (2017, April). Parent Responding to Infant Canonical and Noncanonical Vocalizations: Relations with Infant Receptive and Productive Vocabulary. Poster presented at the Society of Research in Child Development biennial conference, Austin, TX.

Pretzer, G. M., Warlaumont, A. S., VanDam, M., MacWhinney, B., & Reding, C. (2016, November). HomeBank: A resource for sharing daylong audio recordings from children and families. Poster presented at the 2016 American Speech-Language and Hearing Association Convention, Philadelphia, PA.

Pretzer, G. M., Abney, D. H., Warlaumont, A. S., & Walle, E. A. (2016, November). Parental responses & language ability in walking & crawling infants. Technical Session presented at the 2016 American Speech-Language and Hearing Association Convention, Philadelphia, PA.

Pretzer, G. M., Warlaumont, A. S., & Walle, E. A. (2016, November). Human vs. machine coding of contingencies yields completely opposite results. Technical Session presented at the 2016 American Speech-Language and Hearing Association Convention, Philadelphia, PA.

Shea, T. M., Warlaumont, A. S., Kello, C. T., Noelle, D. C., Pretzer, G. M., & Walle, E. A. (2016, August). Deep autoencoding of naturalistic infant and parent vocalizations. Poster presented at the Workshop on Contemporary Neural Network Models, 15th Neural Computation and Psychology Workshop, Philadelphia, PA.

Pretzer, G. M., Warlaumont, A. S., & Walle, E. A. (2016, May). Bidirectional contingencies of infant-adult vocal interactions in daylong home recordings. Presented in J. Gros-Louis et al. "Caregiver-infant contingencies: uncovering social mechanisms of development" at the 2016 International Congress of Infant Studies, New Orleans, LA.

Pretzer, G. M., Warlaumont, A. S., Walle, E. A., Alvarez, E., & Flores, P. (2015, November). Examining infant-caregiver interactions in LENA home recordings of one-year-old infants. Technical Session presented at the 2015 American Speech-Language and Hearing Association Convention, Denver, CO.

Warlaumont, A. S, & Pretzer, G. M. (2015, November). How social interactions help build infants' speech production skills: New insights from daylong, naturalistic audio recordings. Presentation at UC Early Childhood Education Center Directors Meeting, Merced, CA.

Pretzer, G. M., **Warlaumont, A. S.,** & Walle, E. A. (2015, August). Infant vocalizations in the wild. Poster presented in M. Goldstein & C. Yu "Mechanisms of learning in social contexts," at the 2015 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), Providence, RI.

Pretzer, G. M., Warlaumont, A. S., & Walle, E. A. (2015, July). Phonetic abilities of walking and crawling infants. Poster presented at the 37th Annual Meeting of the Cognitive Science Society, Pasadena, CA.

Erickson, E., **Hudock, D.,** Ramsdell-Hudock, H., Altieri, N., O'Donnell, J., Vereen, L., & **Pretzer, G.** (2014, November). Evaluating listeners' reactions to disfluent speech before & after an intensive stuttering clinic. Poster presented at the American Speech Language and Hearing Association (ASHA) Annual Convention, Orlando, FL.

Bull, C., Hudock, D., Altieri, N., Seikel, J., & **Pretzer, G.** (2014, November). Male & female reactions to typical & atypical speech with the use of disclosure statements. Poster presented at the American Speech Language and Hearing Association (ASHA) Annual Convention, Orlando, FL.

Hudock, D., Pretzer, G., Altieri, N., & Seikel, J. A. (2014, March). Listeners' self-reported and physiological reactions to typical and atypical speech. Poster presented at Idaho State University's Division of Health Sciences Research Day.

Hudock, D., Altieri, N., Seikel, J., & Pretzer, G., (2013, November) Listeners' self-reported & physiological reactions to typical and atypical speech. Poster presented at the American Speech Language and Hearing Association (ASHA) Annual Convention, Chicago, IL.

Hudock, D., Altieri, N., Seikel, J., & **Pretzer, G.**, (2013, October) Listeners' self-reported & physiological reactions to typical and atypical speech. Poster presented at the Inter-Mountain Area Speech and Hearing (IMASH) Convention, Denver, CO.

Teaching

COGS 005 Introduction to Linguistics, Discussion sections, University of California, Merced (Spring 2015)

COGS 001 Introduction to Cognitive Science, Discussion sections, University of California, Merced (Fall 2014)

Credentials, Certifications & Training

Washington Speech-Language Pathology License, (July 2018-present)
American Speech-Language and Hearing Association Certificate of Clinical Competency in Speech-Language Pathology (CCC-SLP), (July 2017-present)
California Speech-Language Pathology Services Credential, in Language, Speech and Hearing (August 2017-March 2019)
California Teaching Credential, in Speech and Hearing Services (August 2016-December 2018)
California Provisional Speech-Language Pathology Services Credential, in Language, Speech and Hearing (August 2016-July 2017)
Thesis and Article Writing, Dissertation Boot Camp, Graduate Enrichment & Advancement Resources & Services (GEARS), Graduate Division, University of California, Merced (June 2015; January 2016)
Surviving the Classroom with 1st Generation Students, Teaching Matters, Center for Research on Teaching Excellence, University of California, Merced (August 2014)

Membership

California Teachers Association (CTA) (2017-2018)
California Speech-Language Hearing Association (CSHA) (2017-2018)
American Speech-Language and Hearing Association (ASHA) (2016-present)
International Congress of Infant Studies (ICIS) (2016-present)
New Investigators DARCLE (2016-2018)
  Co-Coordinator (2018)
Women in Cognitive Science (2015)
Cognitive Science Society (2015)
Daylong Audio Recording of Child Language Environments (DARCLE) (2015-2018)
  Board Member, New Investigators Liaison (2018)
National Student Speech Language and Hearing Association (NSSLHA), Region 10, Idaho State University Chapter, Member (2012-2014), National Member (2013-2017)
    Idaho State University Chapter President *(2013-2014)*

Awards

Graduate Research Fellowship Award, University of California, Merced (Spring 2018)
Graduate Research Fellowship Award, University of California, Merced (Fall 2017)
Summer Research Fellowship Award, University of California, Merced (Summer 2017)
Student Research Travel Award, ASHA Convention, Philadelphia, PA (November 2017)
Summer Research Fellowship Award, University of California, Merced (Summer 2016)
Graduate Research Fellowship Award, University of California, Merced (Spring 2016)
Graduate Research Fellowship Award, University of California, Merced (Fall 2015)

Summer Research Fellowship Award, University of California, Merced (Summer 2015)

Teaching Assistantship, University of California, Merced, School of Social Sciences, Humanities, and Arts, Department of Cognitive and Information Sciences (Fall 2014-Spring 2015)

IMASH Registration Scholarship, Denver, CO (October 2013)

Graduate Teaching Assistantship, Idaho State University, Department of Communication Sciences and Disorders (2013-2014)

Dean's List, Idaho State University (2012)

Service

Co-Coordinator for New Investigators DARCLE Group, Liaison between DARCLE and NI DARCLE groups; Submit meeting notes and readings for inclusion on DARCLE website (Fall 2017-Spring 2018)

Reviewer for Many Paths to Language (MPaL) Conference, Reviewed submissions for the 2017 MPaL Conference (July 2017)

Reviewer for Language in Infants and Toddlers Topic Area, Reviewed submissions for this content area for the 2016 ASHA Convention (April 2016)

SSHA Graduate Student Judge, Undergraduate Research Poster Competition; Judged undergraduate researchers' posters (February 2016)

Cognitive and Information Sciences Graduate and Faculty Weekly Meeting, Co-organizer for weekly departmental meetings; Coordinating student speaker schedules and distributing abstracts and materials prior to talks (Spring 2016)

NSSLHA, Planning and preparation for 2014 NSSLHA Workshop (Fall 2013 – Spring 2014)

Northwest Center for Fluency Disorders, Assisted with finding/securing grants for Intensive Stuttering Clinic; Assisting with creation of clinic manual for Intensive Stuttering Clinic (Fall 2013 – Spring 2014)

IMASH Convention, Session Moderator for ***"Forming Inter-Professional Collaborations to Treat Students Who Stutter"***; Volunteer at registration table, Denver, CO (October 2013)

FRIENDS of People Who Stutter Symposium, Assisted leading children's session, Pocatello, ID (April 2013)

Capitol Hill Day, Presented information to members of Idaho state legislature at NSSLHA information table, Boise, ID (February 2013)

Abstract

Language is inherently social; thus, language learning must be facilitated by social means. Infant-adult interactions are crucial for language development. However, one cannot view communication as an independent entity; rather, communication is simply one developmental domain within the dynamical scope of infant development. Specifically, language co-exists and develops in conjunction with the motor, cognitive, social, and emotional developmental domains. This dissertation will investigate the interdependent roles of all these developmental domains as they pertain to speech and language development.

The first investigation explores how infant-adult social interactions relate to the speech-relatedness of infant vocalizations. This study, which was recently accepted for publication in *Infant Behavior and Development*, examined two-event vocal sequences within the context of ultra-naturalistic home interactions captured via the LENA system. We found that the dynamics of infant-adult interactions were dependent upon the type of infant vocalization, the directedness of adult utterances, and the time lags between vocal events.

The second investigation builds upon the first by adding a locomotion dimension: How do these infant-adult interactions look in walking vs crawling infants? This question was previously explored with a smaller, unbalanced sample, but will be re-assessed with a larger, more equal sample of walking and crawling infants. We found differential responses to types of infant vocalizations between walking and crawling infants. We also found a marginally significant association between infant-directed speech and productive vocabulary.

The final component to this dissertation is an exploration of how milestone acquisition across a range of developmental domains (e.g., gross motor, fine motor, problem solving, social/emotional) relates to observable language behaviors within naturalistic daylong home audio recordings (e.g., infant vocalization count, turn count, and adult words heard). We found significant relations between infant vocalization types and adult utterance directions dependent upon age. Additionally, we found that Fine Motor, Social-Emotional, and Communication scores were predictive of vocabulary size. These three studies demonstrate evidence that supports the idea that language development is an active participant, along with the other developmental domains, within the dynamical system that is infant development.

This dissertation, Multi-Domain Synchrony Within Vocal Development, is submitted by Gina Marie Pretzer in 2019 in partial fulfillment of the degree Doctor of Philosophy in Cognitive and Information Sciences at the University of California, Merced, under the guidance of dissertation committee chair, Teenie Matlock.

Chapter 1

Introduction

Language is an integral part of being human. Language drives interactions, facilitates cooperation, and provides the framework for exchanging ideas and knowledge. From more overt methods, such as greeting with "hello," to subtly raising an eyebrow in response to an absurd idea, we are constantly exchanging messages with one another.

However instinctual or reflexive some facets of language may be, so much of how we communicate with others must be learned, whether through direct pedagogical methods (think about learning how to conjugate verbs in school) or through shared experiences and interactions with others. For many decades, researchers have studied how infants learn the complexities of language. We know that while generally predictable, language learning does vary between children. This highly complex pathway can be nonlinear and, at times, convoluted. Further, "breakdowns" can occur along this developmental sequence, resulting in language delays and disorders. When this occurs, specialists work with children to boost language development and mitigate any difficulties caused by these disorders.

As one of those specialists, I am strongly aware of the role that high-quality research plays in shaping one's clinical practice. Speech-Language Pathologists (SLPs) are charged with engaging in Evidence-Based Practice, which is "an approach in which current, high-quality research evidence is integrated with practitioner expertise and client preferences and values into the process of making clinical decisions" (American Speech-Language-Hearing Association, 2005). My interest in research started here, for what better way is there to incorporate evidence into your practice than to help uncover and contribute evidence yourself? This effect is bidirectional, as my research about prelinguistic vocal development has shaped my clinical career into one spent working in early intervention with infants and toddlers with speech delays and disorders.

In this dissertation, I will introduce three studies that were in part shaped by the idea that vocal development does not occur in a vacuum, but rather as part of a cooperative system of development between domains. Given that language learning is based within interactions and that speech is a very precise motor act, one should explore vocal development with an interdisciplinary mindset where social-emotional and motor development are also considered.

1.1 Infant Prelinguistic Vocal Development

Prelinguistic development is marked by a series of predictable progression starting from primitive cries and vocalizations at birth and culminating in true words (e.g., Buder, Warlaumont, & Oller, 2013; Koopmans-van Beinum & van der Stelt, 1968; Oller, 1980, 2000; Roug, Landberg, & Lundberg, 1989; Stark, 1980; Stoel-Gammon, 1989; Vihman, Macken, Miller, Simmons, & Miller, 1985). By 2-3 months of age, infants demonstrate adequate control of their articulators (i.e., the oral structures involved with vocal production, including the lips, tongue, and jaw) to produce more diverse protophone vocalizations, such as full vowels, squeals, yells, growls, whispers, raspberries, and marginal babbling (Buder et al., 2013; Oller, 1980; Roug et al., 1989; Stark, 1980). At around six months of age, infants produce canonical vocalizations (i.e., when a syllable has speech-like timing between full consonants and vowels; see Buder et al., 2013) more consistently. From then until 18 months of age, infants steadily increase the number of speech-related vocalizations, including first true words that appear around their first birthday (Oller, Eilers, Urbano, & Cobo-Lewis, 1997; Warlaumont & Ramsdell-Hudock, 2016). These prelinguistic milestones are foundational to later lexical development (Oller, 2000). For example, the phonetic structure of canonical babbling is generally similar to the phonetic structure of first true words (Stoel-Gammon, 1989; Vihman et al., 1985). The refinement and increased coordination of infants' vocalizations provide a foundation for communication.

1.2 Infant-Directed Speech

Prelinguistic evolution is linked to the rise of "caregiverese" (i.e., the use of highly salient speech; see Falk 2004). "Caregiverese" (often referred to as "motherese" or "parentese") features hyper-articulated vowels, exaggerated facial expressions and movements, less complex grammar and syntax, well-defined segmentation of words and phrases, and sing-song intonation and inflection. This combination of features scaffolds language acquisition and supports phonological and semantic bootstrapping during infant-caregiver interactions. The exaggerated intonation and stress patterns of "caregiverese" provide infants examples of how prosodic features affect the rhythm of words and phrases, helping infants discover the rhythms of their native language by 2 months of age. Fernald (1995) adds that "caregiverese" has been shaped by natural selection. Mothers use "context-specific intonation patterns" during routine caregiving activities, such as soothing, gaining attention, praise, prohibition, and modulating arousal. These acoustic features exploit an infant's propensity to respond differentially to various acoustic features of sounds. This is especially helpful for infants' immature auditory systems, as it allows them to better track a single speaker. Further, these modified acoustic properties compensate for the primitive perceptual, auditory, and cognitive skills early in development as the voice is more powerful than facial expressions in early infancy. "Caregiverese", also called Infant-Directed Speech (IDS) throughout this dissertation, serves many functions by promoting vocal, lexical, and social communicative development in infants.

1.3 Learning Through Infant-Adult Interactions

One key factor of prelinguistic vocal development is interactions between infants and caregivers. Prior research has found that infants alter the phonological structure of their vocalizations to match adult vocalizations. For example, Goldstein and Schwade (2008) found that when maternal responses to 9-month-old infants were manipulated such that half were fully-resonant vowels and half were consonant-vowel combinations, the infants in turn modified their responses to match the phonological structure of their mothers' vocalizations. Similarly, Bloom, Russell, and Wassenburg (1987) found that when adults maintained a conversational flow in interactions, 3-month-old infants produced a higher ratio of speech-like (i.e., syllabic) vocalizations to non-speech-like (i.e., non-syllabic) vocalizations.

There are numerous studies linking increased rates of infant speech-related vocalizations with contingent adult responses (e.g., Gros-Louis, West, & King, 2014; Gros-Louis & Miller, 2017; Goldstein, Schwade, and Bornstein, 2009; Warlaumont et al., 2014). For example, maternal responsivity within infant-adult interactions are highly implicated in the rates of canonical syllable production (Gros-Louis et al., 2014). Warlaumont and colleagues (2014) proposed a social feedback loop where infant and adult vocalizations are contingent upon each other. They found that infant speech-related vocalizations were more likely to elicit contingent adult responses, and that infant vocalizations were more likely to be speech-related if their previous speech-related vocalization received a contingent adult response.

Infant-adult interactions also promote lexical development (e.g., Gros-Louis et al., 2014; Tamis-LeMonda, Bornstein, & Baumwell, 2001; Ma, Golinkoff, Houston, and Hirsh-Pasek 2011; Weisleder & Fernald, 2013). Maternal responsivity, specifically to children's play and vocalizations, is connected to language milestones including first spontaneous words, acquiring 50 words, and combining words together (Tamis-LeMonda et al., 2001). We know that infant-directed speech (IDS) is a better facilitator of lexical development than other-directed speech (ODS; i.e., speech to other children adults, or even pets), even though there is evidence of vocabulary growth associated with overheard speech (Ma et al. 2011; Akhtar, Jipson, & Callanan, 2001; Floor & Akhtar 2006). IDS in one-to-one infant-adult interactions is a better facilitator of productive vocabulary development than IDS within groups (Ramirez-Esparza, Garcia-Sierra, & Kuhl, 2014; 2017a, 2017b). This is true for English and Spanish speaking children (see also Weisleder & Fernald, 2013). Additionally, infants who hear more IDS demonstrate better language processing skills.

Infants learn how to use their vocalizations to effect change in their environments while interacting with their caregivers. At a basic level, infant vocalizations elicit subsequent IDS (e.g., Warlaumont et al., 2014; Gros-Louis et al., 2014, 2017). Infants vocalize to gain attention and make requests. In this same manner, infants learn pragmatic (i.e., social usage) language skills.

When adults interact directly with infants, infants in turn learn how to navigate the pragmatic aspects of interpersonal interaction (e.g., Tamis-LeMonda, Kuchirko, & Song, 2014; Golinkoff, Can, Soderstrom, and Hirch-Pasek, 2014). Golinkoff and colleagues pose the question of quantity vs. quality of IDS. Rather than the sheer amount of IDS heard by an infant, does the type, or quality of IDS better influence infants' social language learning? Perceptually, IDS increases the salience of language input. IDS also highlights the speaker's underlying emotional state. Linguistically, IDS supports the development of phoneme discrimination (see Cristia, 2011). Further, the exaggerated prosodic features of IDS highlight the structure of words and phrases, such as grammatical boundaries between utterances or segmentation of words within phrases. IDS enables word learning in children and helps infants to better remember words they have heard. Infants glean and use the social information from IDS differently across their development. As their language-learning trajectory changes, aging infants are progressively able to harvest more information from the IDS they hear.

Adult responsivity to infant vocalizations can also support a child's later language abilities in children with developmental disabilities (Yoder & Warren, 1999; Harbison et al., 2018). Yoder and Warren (2002) found that interventions that supported increased parent responsivity were associated with increased language growth for children with intellectual disabilities, including Down syndrome and William's syndrome. Further, infants with higher rates of canonical syllable production before treatment benefited more from intervention strategies that encouraged high parent responsivity. Evidently, sensitive adult responses (i.e., responses that are attuned to the child's social and emotional regulation needs) to infant vocalizations can scaffold communication development (see also Gros-Louis et al., 2014; Field, 1980; Patten, Labban, Casenhiser, & Cotton, 2016). Note, however, that non-sensitive parent responses can actually be overstimulating to some children, including premature infants, children with Autism Spectrum Disorders, or significant cognitive delay. These are key facts that can inform a clinician's practice when working with families of young children with developmental disorders, various medical diagnoses, or even just speech and language delay.

1.4 The Present Work

There have been several common themes within my research and clinical practice over the past few years. Namely that (1) children learn and work best in their natural environments (i.e., in their own homes); (2) language development is inherently social, and, as such, one must consider the role of social and emotional development in conjunction with language; and (3) speech production is an ultra-precise motor task, thus, it is important to think about motor development in relation to speech production. The clinical side of this manifests as collaborating with Developmental Specialists and Infant Mental Health Therapists, as well as co-treating with Physical Therapists and Occupational Therapists. The research aspect of these ideas culminates in the present work.

The first section introduces a novel method of viewing naturalistic day long audio recordings of infant-adult interactions captured using the LENA system. Many studies exploring infant language development are conducted in laboratory settings or with in-home evaluators. This first study analyzes segments of daylong audio recordings, categorizing infant vocalization by speech-likeness and adult vocalizations by direction of addressee. Previous research shows that IDS is more likely to occur following infant vocalizations of any type within 1-2 seconds (Bornstein et al., 2015; Van Egeren et al., 2001). Further, adults are more likely to respond with non-overlapping sounds to speech-related vocalizations (i.e., canonical or marginal babbles) than non-speech-related vocalizations (e.g., reflexive and vegetative vocalizations, see Warlaumont et al., 2014), and that more mature consonant-vowel combinations (i.e., canonical vocalizations) are more likely to receive adult responses than less mature vowel only vocalizations (i.e. other protophone vocalizations; see Gros-Louis et al., 2014; Gros-Louis & Miller, 2018). This study extends upon those works.

The next chapter looks to add upon the first study by adding another dimension - locomotion. In this study, we expand the sample and differentiate between participants based on their locomotor status (i.e., are they walking or crawling as their primary mode of locomotion?). We also added receptive and productive vocabulary scores from the MacArthur-Bates Communicative Developmental Inventory (MCDI) as an outcome variable. Walle and Campos (2012) found that infants with walking experience (i.e., the ability to walk 10 feet without assistance) had significantly higher receptive and productive vocabularies, measured by MCDI, than age-matched infants who were crawling (i.e., the ability to travel a distance twice his or her body length on hands and knees). This current work provides support for those findings while adding information about how the relation between IDS and infant vocal maturity interacts with locomotor status and age.

The final section explores the relations between prelinguistic vocalizations, motor development, and infant-adult interactions within a longitudinal framework. Further, a clinical screening tool called the Ages and Stages Questionnaire was utilized to measure multi-domain milestone achievements. This study follows the daylong audio recordings of 41 infants across four dates within their first two years of life. Many of these infants were also enrolled in the research study that was completed by the infants in the two previous sections. Using the same coding methodologies of the first two sections on a subset of the recordings, this study examines how the maturity, or speech-likeness, of infant vocalizations and directionality of adult speech over time relate to acquisition of major communicative, motoric and social-emotional milestones from 3 to 18 months of age.

Chapter 2

Infant-Adult Vocal Interaction Dynamics Depend on Infant Vocal Type, Child-Directedness of Adult Speech, and Timeframe

2.1 Preface

In this chapter, I will present a study accepted for publication in *Infant Behavior and Development* and currently In Press that was co-authored by Lukas D. Lopez, Eric A. Walle, and Anne S. Warlaumont. This study explored the temporal contingencies between infant and adult vocalizations as a function of the type of infant vocalization, whether adult caregivers' vocalizations were infant-directed or other-directed, and the timescale of analysis. Excerpts of day-long audio recordings of 1-year-old infants were hand coded and then analyzed to test for temporal contingencies within infant-adult interactions at various time lags. At shorter time lags of 1-2 s, infant vocalizations predicted the occurrence of subsequent infant-directed adult vocalizations, and vice versa. The strength of the predictions were dependent upon the infant vocalization type. Additionally, at longer time lags, other-directed adult vocalizations were negatively associated with infant vocalizations, suggesting that the absence of other-directed speech may matter to infants. While there is some debate about the role of ODS in word learning, this finding might suggest that ODS also supports other areas of linguistic development, including learning the pragmatics of language. These results provide support for previous findings that adult responsivity and the speech-likeness of infant vocalizations are indeed related, and the time scale between vocal events influences these relations.

2.2. Introduction

A wealth of evidence indicates the importance of dynamic and contingent vocal interactions in fostering infant vocal communication development. However, much of this research has been conducted in laboratory settings (e.g., Akhtar, Dunham, & Dunham, 1991; Goldstein & Schwade, 2008; Gros-Louis, West, & King, 2014; Pelaéz, Virues-Ortega, & Gewirtz, 2011; Rollins, 2003), and studies occurring in the home often rely on a researcher being present (e.g., Bornstein, Putnick, Cote, Haynes, & Suwalsky, 2015; Shneidman, Arroyo, Levine, & Goldin-Meadow, 2013) or the use of automated coding (e.g., Warlaumont, Richards, Gilkerson, & Oller, 2014). The present study helps to address these shortcomings by using hand-coding of parent-infant interactions from day-long home audio recordings to examine the bidirectional relationship between infant and adult vocalizations at 12-13 months, a point in development when infants understand and begin to produce meaningful speech.

2.2.1 Infant Vocal Development

The emergence of infant prelinguistic communication is marked by a predictable progression (e.g., Buder, Warlaumont, & Oller, 2013; Koopmans-van Beinum & van der Stelt, 1968; Oller, 1980, 2000; Stoel-Gammon, 1989; Vihman, Macken, Miller, Simmons, & Miller, 1985). Infants begin vocalizing from birth, with primitive vocalizations and cries (Buder et al., 2013; Oller, 1980; Roug, Landberg, & Lundberg, 1989; Stark, 1980). Around 2-3 months of age, infants have gained sufficient control of their vocal tracts to start producing a wider variety of protophones such as full vowels, raspberries, squeals, growls, yells, whispers, and marginal babbling (Buder et al., 2013; Oller, 1980; Roug et al., 1989; Stark, 1980). By about age 6 months, infants begin to demonstrate canonical babbling, which is when a syllable has speech-like timing between consonant and vowel (Buder et al., 2013). From then until at least 18 months, the relative frequency of infant speech-related vocalizations increases steadily (Oller, Eilers, Urbano, & Cobo-Lewis, 1997; Warlaumont & Ramsdell-Hudock, 2016). These vocal milestones provide a foundation for lexical development (Oller, 2000). For example, the phonetic features of canonical babbling are typically similar to the phonetic features of first words (Stoel-Gammon, 1989; Vihman et al., 1985).

2.2.2 Contingent Infant-Adult Vocal Interactions

Importantly, infants acquire these vocal communication abilities in part through interactions with adults. Infant-adult vocal interactions allow parents to demonstrate appropriate conventions for communication, including pragmatic use of vocalizations and rhythms of dialogue (Baldwin & Meyer, 2007; Goldin-Meadow, Goodrich, Sauer, & Iverson, 2007; Golinkoff, 1986; Gros-Louis et al., 2014; Jaffe et al., 2001; Masur, 1982; Olson & Masur, 2011). Contingent adult responses, defined as adult vocalizations that immediately follow infant vocalizations with a likelihood that is greater than what would be expected by chance from the base rates of the infant and adult behaviors, appear to be especially helpful. For example, previous research has shown that infants alter the phonological structure of their vocalizations to match the phonological structure of previous adult vocalizations if those adult vocalizations were contingent responses to infant vocalizations, but not if they were non-contingent on the child’s behavior (Goldstein & Schwade, 2008; see also Bloom, Russell, & Wassenberg, 1987, and Bloom, 1988). Further support comes from correlations between sensitive caregiver responses to infant vocalizations (e.g., contingent responses, imitation of the infant’s vocalization, or commenting on an object within the infant’s visual field) and subsequent rates of canonical syllable production (Gros-Louis et al., 2014). There are also numerous studies linking contingent adult responding with increased frequency of infant speech-related vocalizations (Franklin et al., 2014; Goldstein, Schwade, & Bornstein, 2009; Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015; Gros-Louis et al., 2014; Nathani & Stark, 1996; Todd & Palmer, 1968; Warlaumont et al., 2014).

Additionally, prior research has shown that bidirectional infant-caregiver exchanges are important for other aspects of speech, language, and social development (e.g., Golinkoff et al., 2015; Hart & Risley, 1995; Tamis-LeMonda, Bornstein, & Baumwell, 2001). Parents provide models of speech sounds, word labels, and appropriate conversational behaviors, and embedding these models in contingent responses is particularly impactful (although the degree to which children are sensitive to various features of adult language input changes with age; Adamson & Bakeman, 2006). Further, adult responses that label objects promote word learning (both receptively and expressively) in infants even younger than 9 months of age (see Goldstein, Schwade, Briesch, & Syal, 2010; Gogate, Maganti, and Laing, 2013; Matatyaho & Gogate, 2008; Wu & Gros-Louis, 2014).

Adult responsivity to infant vocalizations can also help mediate the relationship between infant prelinguistic vocalizations and infants' later language abilities in children with developmental disabilities (Yoder & Warren, 1999; Harbison et al., 2018). Yoder and Warren (2002) found that interventions targeting parent responsivity were associated with growth in language abilities for children with intellectual disabilities, including Down syndrome and William's syndrome. Moreover, infants with higher rates of pretreatment canonical syllable production benefitted most from intervention supporting high parent responsivity. Clearly, sensitive adult responses (i.e., responses that are attuned to the child's social and emotional regulation needs; see Gros-Louis et al., 2014; Field, 1980; Patten, Labban, Casenhiser, & Cotton, 2016) to infant vocalizations can provide important support for communication development.

Infant-adult vocal interactions involve not only adult responses to infant vocalizations but also infant responses to adult vocalizations. Previous studies have demonstrated, across cultures, that infants as young as 4-months-old actively engage in turn taking with adults in their environment, with increased likelihood of infant vocalization following adult utterances (e.g., Bornstein et al., 2015; Van Egeren, Barratt, & Roach, 2001) and that infant and maternal contingencies are significantly correlated (Bornstein et al., 2015). At a larger timescale, Gros-Louis and colleagues (2014) found that contingent maternal responses predicted maternal-directed infant vocalizations in following months. Considering that adult responses often take the form of speech, and that adult vocalizations (at least those that are non-overlapping with infant vocalizations) often occur in response to and encouragement of further infant speech-like vocalizations, we would predict that infant vocalizations are more likely to be speech-like when they occur in response to adult input. Consistent with this, Warlaumont and colleagues (2014) found evidence for a social feedback loop wherein infant vocalization is contingent upon adult vocalization and vice versa. Following up on this finding, Gros-Louis and Miller (2018) distinguished between vowel only and consonant-vowel speech-related infant vocalizations. They found that parents were more likely to respond contingently to consonant-vowel than vowel only infant vocalizations. They also found that 10-month-old infants were significantly more likely to produce vowel-like vocalizations following an adult response to the infant's prior vowel-like vocalization and that 12-month-old infants were significantly more likely to produce consonant-vowel vocalizations following an adult response to the infant's prior consonant-vowel vocalization.

2.2.3 Infant-Directed Versus Other-Directed Speech

In cultures where infants are regularly provided with infant-directed speech (IDS) from their caregivers, IDS is particularly beneficial for speech and language development compared to adult-directed speech (ADS). IDS exaggerates the phonological features of speech, thereby highlighting important aspects of the speech signal crucial for recognition of vowels and consonants and increasing infant attention to adult vocalizations (Baldwin & Meyer, 2007; Cristia, 2011; Fernald, 1985, 1989; Fernald & Kuhl, 1987; Kuhl et al., 1997; Papousek & Papousek, 1989). This slower rate of speech and exaggerated pronunciation permits more effective modeling of words for infants (Song, Demuth, & Morgan, 2010) and is associated with infant word learning (Golinkoff & Alioto, 1995; Ma, Golinkoff, Houston, & Hirsh-Pasek, 2005; Ramírez-Esparza, Garcia-Sierra, & Kuhl, 2014; Shneidman et al., 2013; Weisleder & Fernald, 2013) and appreciation of pragmatics (Golinkoff et al., 2015). Moreover, depressed mothers, who typically have less pitch alteration, were found to have infants with larger productive vocabularies when the mother demonstrated more pitch modification (Porritt, Zinser, Bachorowski, & Kaplan, 2013)

Although research has shown that IDS promotes language development in various cultures, the role of other-directed speech (ODS), including ADS, is less clear. Many studies, including those sampling from day-long home recordings featuring abundant quantities of ODS, exclude ODS from all analyses (e.g., Ramírez-Esparza et al., 2014). Those studies that do analyze ODS have shown that the quantity of ODS to which a child is exposed seems to have little effect on language learning (Kuhl, Tsao, & Liu, 2003; Oller, 2010; Weisleder & Fernald, 2013). While those studies have not reported beneficial effects of ODS on language learning, other work has demonstrated that toddlers can learn new words from overheard speech (Akhtar, Jipson, & Callanan, 2001). Additionally, Floor and Akhtar (2006) reported evidence of word learning after overhearing ODS by 18-month-old infants.

Furthermore, although modification of speech directed to infants appears to share many properties across cultures, there are cultures in which substantial differences in pragmatic, linguistic and acoustic properties of IDS have been reported. For example, within the Kwara'ae culture, young infants are primarily exposed to a "modified register" of the language that share features of IDS seen in other cultures but has different pragmatic features and emphases (Watson-Gegeo & Gegeo, 1986). Similarly, Kaluli adults have been reported to produce little IDS but have been reported to frequently use a high-pitched register to "translate" for their preverbal infants (Ochs & Schieffelin, 1984). Moreover, pitch modification in IDS has been found to be higher for more highly-educated American mothers (Broesch & Bryant, 2015). Even significant differences in exposure to IDS and in the form that IDS takes do not totally preclude language learning, raising the possibility that ODS might play a substantial role in language learning and making it worth analyzing alongside IDS.

It is also possible that in IDS-heavy cultures, reduced exposure to ODS could be beneficial to language learning under certain circumstances. Specifically, reduction in ODS could reduce background noise, allowing infants to more effectively learn from IDS and their own vocalizations, and as well as signal to infants that adults are paying attention to them. Thus, while infants may or may not gain immediate phonological or lexical information from overhearing ODS, there may be other functions of ODS within these exchanges and *absence* of ODS may itself be valuable at least for some cultures.

Prior research indicates that IDS shows patterns of sequential association with child vocalizations — in fact, the majority of the studies of infant-adult interaction cited above focused exclusively on adult vocalizations that were infant-directed and excluded vocalizations directed toward another adult. It remains to be examined whether ODS also shows patterns of sequential association wherein an infant's vocalization predicts the subsequent occurrence of overheard adult speech or whether overheard adult speech predicts the subsequent occurrence of an infant vocalization. Better understanding of the temporal contingencies between infant vocalizations, IDS, and ODS may help clarify both the pragmatic properties of IDS and ODS within a given culture, and thereby inform our understanding of the roles each can play in communication development. Study of day-long home audio recording have the advantage of offering ample opportunities to sample both IDS and ODS.

2.2.4 Studying Early Vocal Interactions in Day-long Home Recordings

Much of the prior research on prelinguistic development has been conducted in laboratories or short sessions in the infant's home with a researcher present. However, wearable audio recording systems, such as the LENA system (LENA; LENA Research Foundation, Boulder, Colorado, United States), now allow researchers to collect full-day observations of infant-adult interactions in the home. The LENA system is comprised of pocket-sized audio recorders called digital language processors (DLP), specially designed clothing to hold DLPs, and software to process and analyze collected audio data. This methodology allows researchers to study how infants and their caregivers interact across a broad and ecologically valid setting with minimal observer effects. A recent study using first-person and third-person point of view recordings revealed that when using first-person viewpoint cameras in home recordings, there is a higher frequency of "less socially desirable maternal behaviors" (e.g., being distracted or critical) observed than in third-person viewpoint recordings (Lee, Skinner, Bornstein, Radford, Campbell, Graham, & Pearson, 2017). Another study comparing day-long audio-only recordings to hour-long video recordings made by parents during the same day found higher rates of language production during the video samples than during the rest of the portions of the day when the infant was awake (Bergelson, Amatuni, Dailey, Koorathota, & Tor, 2018). Thus, the absence of a researcher may provide a more ecologically valid sample than one collected in a laboratory context. Day-long recordings provide a naturalistic means to study vocalization types and contexts beyond those that commonly occur in the lab or when a researcher is present. For example, infants' distress vocalizations may be more prevalent in day-long home recordings and adults are faced with numerous situations in which they must decide whether to vocalize to the infant or engage with other adult conversational partners or other children.

Previous research by Warlaumont et al. (2014) examined infant and adult vocalization contingencies utilizing the LENA system's automatic labeling of parent and infant vocalizations, which compared to labeling by trained human listeners is less accurate but also less time intensive and less subject to inter and intra-rater variability. The researchers found that infant speech-related vocalizations were more likely to be followed by adult vocalizations than children's non-speech-related vocalizations. Further, children's subsequent rates of production of speech-related vs. not-speech-related sounds were related to these adult responses at both immediate and longer-term timescales. However, analyzing such recordings in detail beyond speaker diarization requires either custom vocalization analysis algorithms that are not yet widely available (e.g., Oller et al., 2010) or human coding/transcription. Even so, hand coding by human listeners remains the gold standard for characterizing many aspects of vocalizations, including whether an utterance contains canonical syllables (Warlaumont & Ramsdell-Hudock, 2016) and to whom adult speech is directed (Schuller et al., 2017). Furthermore, hand-coding is necessary to ensure accurate labeling of the onset and offset of speaker vocalizations, especially during periods of overlap. Thus, utilizing hand coding can allow researchers to examine day-long recordings of infants and their families in more detail and with greater accuracy than that provided by existing automated methods.

2.2.5 Overview of the Present Study

This study analyzed hand-coded segments from day-long LENA recordings of 12- to 13-month-old infants focusing on periods of high child vocalization rate (i.e., volubility). We examined the temporal relationships between infant vocalizations and IDS. Additionally, we examined the relationship between infant vocalizations and ODS (i.e., adult speech to other adults or to children other than the target infant), which is often excluded in language development studies. We also distinguished between a range of infant vocalization types, including canonical vocalizations (i.e. babbling or first true words containing at least one adult-like consonant-vowel or vowel-consonant transition; see Oller, Eilers, & Cobo-Lewis, 1997 and Buder et al. 2013), non-canonical babble (including marginal babbling and other non-canonical, non-reflexive, non-vegetative sounds), and reflexive sounds (laughs and cries). We assessed the degree to which each infant vocalization type predicted subsequent adult vocalizations, and vice versa, using three time windows of analysis, 1 s, 2 s, and 5 s. This design provided a comprehensive analysis of bidirectional contingencies between child vocalizations and adult vocalizations of different types in naturally occurring and unsupervised contexts at a point in development when infants typically begin to produce utterances with recognizable meanings. Finally, we tested for mean pitch, pitch variability, and mean intensity differences between IDS and ODS, to establish a broad sense of the acoustic differences between IDS and ODS in this sample; this is the first analysis of pitch and intensity for IDS and ODS from day-long audio recordings.

2.3 Methods

2.3.1. Participants

Nineteen infants (12 female) aged 12- to 13-months were selected from a larger, ongoing study of 62 recruited infants. All families were recruited from the San Joaquin Valley in central California. Families from the larger study were excluded because their recording was shorter than the 10-hour requirement or was not returned (n = 34), the primary language spoken in the home was Spanish (i.e., > 50% of the time, according to the parents' estimates; n = 15), the infant was outside the targeted age range (n = 8), or the recording was split over multiple days (n = 7). The requirement to record 10-hours within a single day was designed to ensure that the samples were drawn from a large range of different contexts that an infant typically experienced over the course of a day.

Eight infants were from families with both parents having a college degree (including 2 graduate degrees), 5 infants from families with one parent having a college degree (including 1 graduate degree), 5 infants from families with neither parent having a college degree, and one infant's family did not report parental educational experience. Fourteen infants were reported to be of Caucasian descent, 10 infants of Hispanic descent, and 2 infants of Asian descent. No diagnoses of developmental disorders were reported for any of the infants.

2.3.2. Recording Procedure

Each family was mailed a LENA recorder and vest specially designed with a pocket on the chest to hold the recorder. The recorder captured the infant's vocalizations, as well as linguistic and other auditory input from the infant's home environment. Parents were instructed to turn on the audio recorder when the child awoke in the morning and slide it into the pocket of the provided vest. The infant then wore the vest for the entire day, with the exception of baths, naps, or car rides when the vest could be placed nearby while the recording continued. Parents were also allowed to pause the recording for privacy reasons. The recorder automatically shut off after 16 hours of audio collection, though some families did not or were unable to record for the entire 16 hours.

2.3.3. Selection of High Infant Volubility Samples

The recordings were first processed using the LENA Pro system's software. The LENA software's automatic labeling system applies a set of mutually exclusive sound source labels to the entirety of each recording. The automated labels are: infant wearing the recorder, male adult, female adult, other child, electronic sounds, noise, and silence, with all labels except silence being further divided into "near" (i.e. relatively loud) and "far" sounds. LENA's Automatic Data Extractor (ADEX) was used to identify the three most voluble (i.e., containing the highest number of "near" infant vocalizations) 5-minute samples for each infant's day-long recording. However, an identified sample was replaced by the next highest-volubility sample if: the highest volubility 5-minute sample was within 30 minutes of another of the highest-volubility samples; the sample did not have at least 10 infant speech-related vocalizations (i.e., canonical babbling, marginal babbling, or other protophone vocalizations; see Buder et al., 2013 for definitions) as judged by a human listener; or it sounded as if the infant had an object (e.g., a pacifier) obstructing her mouth for a significant portion of the sample. These selection criteria ensured that distinct observational times from the day were included, the automatically assessed high volubility of the sample was valid, and the infant's vocal expressivity was not impeded.

2.3.4. Utterance Identification and Classification by Human Listeners

For each segment, infant and adult vocalizations were marked using the EUDICO Linguistic Annotator (ELAN, 2018; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Researchers were trained to identify and code infant and adult vocalizations using training procedures created by the authors. Researchers marked vocalizations by listening to the recording and pausing, replaying, and jumping backward and forward as needed to mark the onsets and offsets of infant vocalizations according to the following criteria:

1. *Mark any sound of nontrivial loudness that you think was made by the vocal tract (i.e. larynx, throat, mouth, lips, teeth, nasal passage, etc.).*
2. *Be as accurate as possible when setting utterance boundaries - try to be within tenths of a second.*
3. *Annotations should run from the onset of sound to the offset of sound. [For adult vocalizations, it was also noted that ends of phrases/sentences are often natural ends of vocalizations.]*

Next, vocalizations were coded for the speaker (i.e., infant or adult) and type of utterance. Researchers were trained to classify infant vocalizations within four categories: Reflexive (laugh and cry), Vegetative (e.g., burp, hiccup, cough, yawn, heavy breathing), Canonical (i.e., a syllable that has speech-like timing between consonant and vowel; see Buder et al., 2013), and Non-Canonical/Non-Reflexive (including marginal babbling and other non-canonical protophones; here forward referred to simply as “Non-Canonical”). Vegetative infant vocalizations were subsequently excluded. The adult vocalizations were coded as Infant-Directed, Other-Directed (i.e., speaking to another person who is not the target infant), or unknown, based on the intended addressee. We did not distinguish to whom ODS was directed. At times, ODS was directed to other adults (e.g., one adult discussing dinner plans with another). At other times, ODS was directed to animals (e.g., telling the dog to go outside) or to other children. In the event that adult vocalizations were judged to be directed to other children as well as the target infant, these vocalizations were classified as IDS. Transcribers listened to each vocalization a maximum of three times during the coding process to balance having ample opportunity to hear the vocalization against over-analysis and second-guessing.

Inter-rater reliability for infant vocalization type and adult utterance direction was calculated using percent agreement and Cohen’s kappa. Reliability checks were completed on six randomly selected 5-minute samples (one sample each from approximately 30% of the participants). Reliability coders re-coded the originally marked infant vocalizations and the direction of the adult vocalizations, using the same criteria as the original coders. Interrater reliability was adequate for infant vocalization type (percent agreement = 83.40%; Cohen’s $k = .58$) and substantial for Adult Utterance Direction (percent agreement = 89.90%; Cohen’s $k = .76$) (see Landis & Koch, 1977). The disagreements in infant vocalizations codes were evenly spread across all vocalization types. Specifically, disagreements between canonical and non-canonical syllables represented 55% of the overall discrepancies and disagreements between non-canonical and reflexive vocalizations represented the remaining 45% of disagreements. There were no disagreements between canonical and reflexive infant vocalizations.

2.3.5. Temporal Contingency Analyses

Data were converted into a single series of events to test for temporal contingencies between child and adult utterances of various types. The event series allowed for the identification of vocalizations immediately preceding or immediately following any given utterance (Bakeman, 1997; Yoder & Symons, 2010).

Pauses between vocalizations and overlapping vocalizations were first defined in order to convert the original set of utterance boundaries into a single event series (Lloyd, Yoder, Tapp, & Staubitz, 2015). A pause (*P*) was defined as any duration greater than *P* between the end time of the current vocal event and the start time of the next vocal event. *P* was set to either 1 s, 2 s, or 5 s to comprehensively explore different possible timescales at which vocal contingencies might operate (Van Egeren et al., 2001). Figure 2.1 illustrates how pauses were inserted for the case where *P* = 1 s.

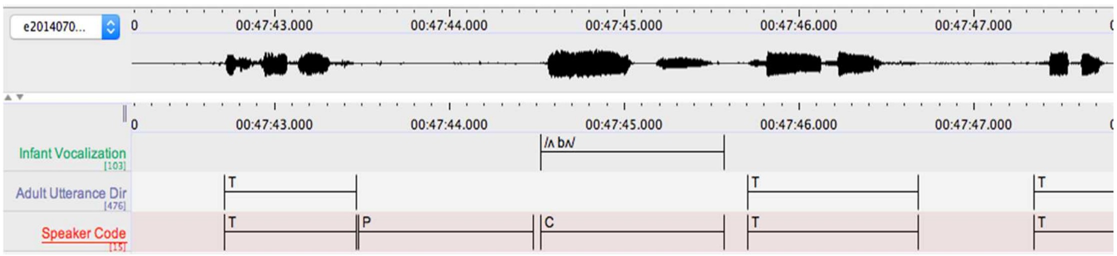

Figure 2.1. A coded segment from a participant's recording, with lags between vocalizations of greater than or equal to 2 s coded as a pause (*P*). Lags between vocalizations that are less than 2 s are not coded. /ʌ bʌ/: phonetic transcription of an infant canonical (C) vocalization. T: an infant-directed adult vocalization. Speaker Code: a tier indicating how event series were constructed. The event series for this example was T P C T T.

Previous work has found that adults frequently respond to infants with less than 1 s lag (corresponding to *P* = 1 s in the present analytic approach) and that infants are sensitive to contingent responses operating within this timeframe (e.g., Keller, Lohaus, Völker, Cappenberg, & Chasiotis, 1999; Van Egeren et al., 2001; Warlaumont et al., 2014). However, it is possible that in some cases 1 s might not be long enough to capture temporal contingencies between speakers. Thus, we also conducted all temporal contingency analyses with 2 s maximum lag durations (which, using our analysis approach, corresponded to *P* = 2 s). Research on infant-adult vocalizations has found that 2 s lags capture infant-adult bidirectional vocal interactions and contingent responses across cultures (Van Egeren et al., 2001; Gros-Louis, West, Goldstein, & King, 2006; Hilbrink, Gattis, & Levinson, 2015; Bornstein et al., 2015). A 2 s pause definition has also been used in research on conversational turn-taking with slightly older children (Yoder, Davies, & Bishop, 1994). Henning and Striano (2011) found that younger infants were sensitive to slightly longer lag times of 3 s. These 3- and 6- month old infants altered their affect (i.e., smiling) based on the lag time before their mothers' responses. Moreover, in addition to including time lags of 1 s and 2 s, we also chose to include a longer time lag of 5 s in our analyses to reduce the likelihood of missing potentially relevant temporal contingencies, especially given research showing that response lags tend to increase around the onset of first words (Hilbrink et al., 2015) and given that the

LENA software, which is now being widely used by researchers and interventionists, uses a 5 s lag in defining conversational turns.

For the scope of this study, we operationalize "follow" as meaning an adult utterance occurring within a specified time lag after an infant vocalization with no intervening infant or adult vocalizations and vice versa. For example, when assuming P = 5 s, if the child vocalizes and the parent vocalizes 2 s and 4 s after the child, only the first adult vocalization would be considered to follow the child vocalization and thus end the 5 s time window until the next infant vocalization occurs, creating a new 5 s time window in which a subsequent adult vocalization may or may not follow. Likewise, if an infant produces a vocalization and then vocalizes again 1 s later and the adult vocalizes 1 s after the second infant vocalization, the adult vocalization would only be considered to follow the second infant vocalization, not the first, even though it did also occur within 5 s of the first of the two infant vocalizations.

In contrast to the automatic labeling algorithm used by LENA software that excludes overlapping utterances, overlapping utterances were included in our analyses, with the first speaker's vocalization coded as the first event and the later occurring vocalization coded as the second event (see Figures 2.2-2.4 on the next page). For instances when the earlier onset vocalization extended greater than or equal to P (where P is the minimum pause duration in seconds) past the end time of the later onset event, the first event was split into two vocalizations (event 1 and 3). When the later onset overlapping vocalization extended past the end of the initial vocalization for less than P, the initial vocalization was coded as the first event and the overlapping vocalization was coded as the second event.

Next, we created pairs of 2-event sequences to analyze the temporal contingencies between vocalization types. The 2-event sequences were created by combining the type of each vocal event (except for the final vocal event within the 5-minute segment) with the type of the following vocal event (e.g., an Infant Reflexive utterance, R, followed by a child-directed, T, would be given a sequence code of RT). These sequence codes were compiled for all three 5-minute segments for all infants.

Finally, we ran generalized linear mixed effects regression models to assess the relations between the first event type and the second event type for each 2-event sequence. All models included participant ID as a random effect and assumed a Binomial distribution, since the dependent variables were all binary. Analyses were programmed in R (R Core Team, 2018) using the lme4 (Bates, Maechler, Bolker, & Walker, 2018), lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017), and car (Fox et al., 2018) packages.

We first analyzed how the first event in each event sequence pair being a Canonical, Non-Canonical Non-Reflexive, or Reflexive infant vocalization uniquely predicted whether the subsequent event was an Infant-Directed adult utterance, coded as 1 if so and 0 if not. The Canonical, Non-Canonical, and Reflexive variables were each dummy coded as 1 if the first event was the relevant infant vocalization type and 0 otherwise. If the first event in the sequence was any other type of event (i.e., if it was a pause or an adult utterance), then all three infant vocal type fixed effect variables were coded as 0. We also ran the analyses with the same predictors but predicting whether the second event in each pair was an Other-Directed adult utterance. The analyses were run

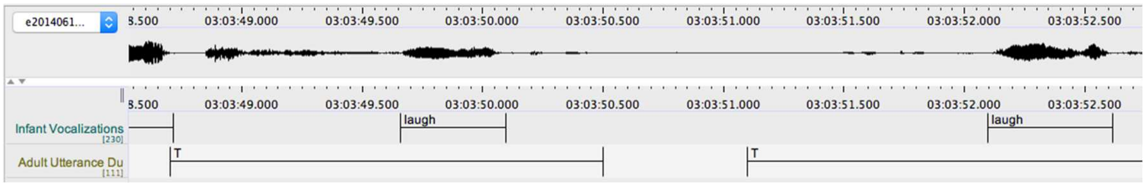

Figure 2.2. A coded segment from a participant's recording, depicting examples of infant laughs overlapping infant-directed vocalizations (labeled "T". If the overlapping vocalization was coded as the first event, the later onset overlapping vocalization was coded as the second event. With P = 1, 2, or 5 s, the event series would be T R T R.

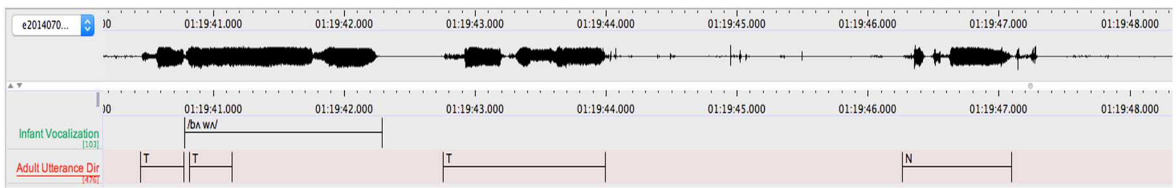

Figure 2.3. A coded segment from a participant's recording, depicting an infant-directed vocalization (labeled "T") overlapping a (phonetically transcribed) canonical infant vocalization where a first event begins before and extends beyond the other. If the overlapping vocalization was coded as the first event, the later onset overlapping vocalization was coded as the second event. When P = 1 s, the event series would be: T C T C T P P N. With P = 2 s, the event series would be: T C T T P N. With P = 5 s, the event series would be: T C T T N.

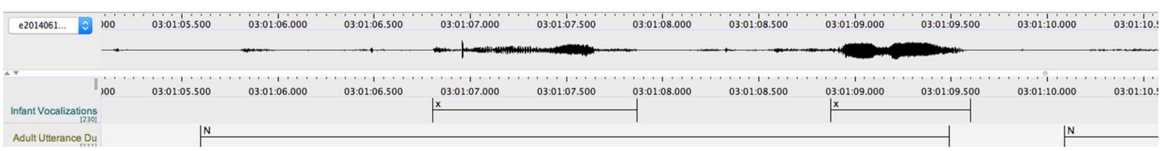

Figure 2.4. A coded segment from a participant's recording, depicting non-canonical infant vocalizations (labeled "x") overlapping other-directed speech (labeled "N"). If the overlapping vocalization was coded as the first event, the later onset overlapping vocalization was coded as the second event. With P = 1 s, the event series would be: N X N X N. With P = 2 and 5 s, the event series would be: N X X N.

separately for each of the minimum pause durations (using a 1-, 2-, or 5-second definition of "pause"). For each significant effect, we ran additional regression models to perform post hoc comparisons of the strength of the observed effects; a model was built which directly compared first events of one type to the first events of the contrasting type (ignoring other event types) in their likelihood of being followed by the second event type in question. The same approach was used to test temporal contingencies in the reverse direction; i.e., how well Infant-Directed and Other-Directed adult vocalizations as first events in the sequenced event pairs were uniquely associated with subsequent infant vocalizations of the three types. Although other researchers have used odds ratio analyses (e.g., Van Egeren et al., 2001; Bornstein et al., 2015), we chose to run mixed-effects regression analyses for their ability to easily include multiple predictor variables within the same analysis. Pilot analyses using odds ratio methods were generally consistent with the mixed effects regression results reported below. Separate analyses were run for each adult vocalization type and for each pause duration.

2.3.6. Acoustic Analyses

We used Praat (Boersma & Weenink, 2018) to automatically estimate each adult vocalization's pitch contour. We used the autocorrelation method with a time step of 0.01 s, a pitch floor of 75 Hz, and a pitch ceiling of 1000 Hz. All other parameters were set to default values. The mean and standard deviation of the pitch contour were then calculated. We used mel units rather than Hz for these calculations to account for the nonlinearity pitch perception (and fundamental frequency production). We also used Praat to compute the mean intensity in dB SPL of each adult vocalization. We then ran three mixed effects linear regression models to test for differences between ID and OD vocalizations' mean pitch, pitch variability, and intensity, controlling for infant ID as a random effect. Intensity was included because it is very easy to measure and could conceivably be a very readily accessible cue to infants. Note, however, that intensity is from the perspective of the infant-worn microphone, not from the perspective of the adult speaker. Utterances less than 0.05 s were excluded, as were utterances for which Praat's output was undefined, which can happen in pitch analyses if no voicing is detected within the utterance.

2.3.7. Code and Data Availability

All analysis scripts and the 2-event sequence data are available at https://github.com/gpretzer/WWScripts. Additionally, 5 minute audio segments that did not contain last names or especially private episodes and their annotations from infants whose caregivers agreed to share the recordings publicly are available on FigShare (Pretzer, Warlaumont, Lopez, & Walle, 2018). Full LENA recordings from all participants who gave consent are available on HomeBank (VanDam et al., 2016) within the Warlaumont Corpus (Warlaumont, Pretzer, Mendoza, & Walle, 2016).

2.4. Results

This section presents the findings of our current study. In the first subsection, we provide descriptive statistics including the relative frequencies of infant and adult vocal events from the entire data set. The next two subsections detail the relation between infant vocalizations and subsequent adult vocalizations as well as the relations between adult vocalizations and subsequent infant vocalizations. The final subsection presents acoustic analyses that consider both frequency and intensity of IDS and ODS vocalizations.

2.4.1. Frequencies of Infant and Adult Vocalization Events

Table 2.1 shows the total count for each infant and adult vocalization type for each minimum pause duration. Adult IDS was about 40% more frequent than adult ODS, but both were amply represented in the dataset. Non-canonical non-reflexive vocalizations were the most frequent type of infant vocalization, being about 3.5 times more common than canonical infant vocalizations and about 5.5 times more common than reflexive infant vocalizations. The number of pause events was, unsurprisingly, greatly dependent on pause duration. The numbers of vocalization event types changed slightly as a function of pause duration, due to the way pause duration effected how overlapping events were treated.

Table 2.1
Event counts per vocalization type across all infants, given 1 s, 2 s, or 5 s pause duration

	Adult IDS	Adult ODS	Infant Canonical	Infant Non-Canonical	Infant Reflexive	Pause
1 s	2127	1541	630	2197	418	7302
2 s	2114	1526	613	2183	399	2919
5 s	2107	1519	604	2179	375	770

2.4.2. Infant Vocalizations Predict Subsequent Adult Vocalizations With Various Pause Durations

Table 2.2 shows the results for analyses of how strongly each of the three infant vocalization types uniquely predicted that the next event would be an adult Infant-Directed utterance when pause duration (P) was 1 s. Given the many planned tests, we used a conservative α cutoff (p = .001) to indicate statistical significance. Canonical (b = 0.99, p < .001), Non-Canonical (b = 0.70, p < .001), and Reflexive (b = 1.70, p < .001) infant vocalizations each predicted a subsequent Infant-Directed adult vocalization. Further post-hoc comparisons with a conservative α cutoff (p = .001) revealed that the Reflexive vocalizations more strongly predicted that an Infant-Directed adult utterance would follow than either Canonical or Non-Canonical vocalizations did. Although not included in Table 2, because the comparison between Canonical and Non-Canonical vocalizations here is of particular interest given results reported in prior literature (Gros-Louis & Miller, 2018) it is worth noting that Canonical infant vocalizations were marginally more likely than Non-Canonical to be followed by an Infant-Directed adult vocalization (b = 0.32, p = .005) There were no statistically significant relations between Canonical, Non-Canonical, or Reflexive infant vocalizations and subsequent Other-Directed adult vocalizations.

Similarly, analyses of P = 2 s (see Table 2.2) revealed that Canonical (b = 0.57, p < .001), Non-Canonical (b = 0.28, p < .001), and Reflexive (b = 1.19, p < .001) infant vocalizations predicted a subsequent Infant-Directed adult vocalization. Post-hoc analyses found that Reflexive infant vocalizations predicted subsequent Infant-Directed adult vocalizations more strongly than did Non-Canonical infant vocalizations. Again, we make an exception to our p = .001 cutoff to note that Canonical infant vocalizations were marginally more likely than Non-Canonical to be followed by an Infant-Directed adult vocalization (b = 0.30, p = .007) However, unlike with the 1 s pause duration analyses,

Table 2.2
Infant Vocalizations Predicting Subsequent IDS or ODS

Pause	First Event	Second Event	*b*	95% CI Lower Bound	95% CI Upper Bound	*p*	Statistically Significant Post-Hoc Comparisons
1 s	Canonical	IDS	0.57	0.66	1.32	<.001	
1 s	Non-Canonical	IDS	0.28	0.48	0.92	<.001	
1 s	Reflexive	IDS	1.19	1.32	2.09	<.001	Reflexive > Canonical; Reflexive > Non-Canonical
1 s	Canonical	ODS	-0.25	-0.15	0.69	0.03	
1 s	Non-Canonical	ODS	-0.48	-0.32	0.25	0.73	
1 s	Reflexive	ODS	-0.38	-0.44	0.58	0.53	
2 s	Canonical	IDS	0.57	0.23	0.90	<.001	
2 s	Non-Canonical	IDS	0.28	0.06	0.50	<.001	
2 s	Reflexive	IDS	1.19	0.80	1.58	<.001	Reflexive > Non-Canonical
2 s	Canonical	ODS	-0.25	-0.70	0.16	0.05	
2 s	Non-Canonical	ODS	-0.48	-0.77	-0.21	<.001	
2 s	Reflexive	ODS	-0.38	-0.92	0.10	0.01	
5 s	Canonical	IDS	0.13	-0.21	0.47	0.19	
5 s	Non-Canonical	IDS	-0.09	-0.31	0.13	0.19	
5 s	Reflexive	IDS	0.70	0.29	1.11	<.001	Reflexive > Non-Canonical
5 s	Canonical	ODS	-0.67	-1.11	-0.26	<.001	
5 s	Non-Canonical	ODS	-0.88	-1.14	-0.60	<.001	
5 s	Reflexive	ODS	-0.83	-1.38	-0.34	<.001	

we also detected that Non-Canonical infant vocalizations were significantly associated with decreased likelihood of subsequent Other-Directed adult utterances ($b = -0.48$, $p < .001$).

Finally, when P = 5 s (Table 2.2), only Reflexive infant vocalizations significantly predicted subsequent Infant-Directed adult utterances ($b = 0.70$, $p < .001$). Post-hoc analyses revealed that Reflexive vocalizations were significantly stronger predictors of subsequent Infant-Directed speech than Non-Canonical vocalizations were ($p < .001$). Conversely, analyses of the 5 s minimum pause duration revealed that Canonical ($b = -0.67$, $p < .001$), Non-Canonical ($b = -0.88$, $p < .001$), and Reflexive ($b = -0.83$, $p < .001$) infant vocalizations all were associated with decreased likelihood of the following event being an Other-Directed adult utterance.

2.4.3. Adult Vocalizations Predict Subsequent Infant Vocalizations With Various Pause Durations

Table 2.3 shows the results for analyses of how each of the three infant vocalization types predicted that the previous event would have been an adult Infant-Directed utterance. When the maximum lag was $P = 1$ s, Infant-Directed adult utterances were significantly associated with subsequent Canonical ($b = 0.51$, $p < .001$), Non-Canonical ($b = 0.39$, p < .001), and Reflexive ($b = 1.52$, $p < .001$) infant vocalizations. Post-hoc comparisons determined that the Reflexive infant vocalizations were a stronger predictor of a preceding infant directed adult utterance than Canonical ($p < .001$) and

Table 2.3
IDS or ODS Preceding Infant Vocalizations

Pause	First Event	Second Event	*b*	95% CI Lower Bound	95% CI Upper Bound	*p*	Statistically Significant Post-Hoc Comparisons
1 s	IDS	Canonical	0.51	0.15	0.86	<.001	
1 s	IDS	Non-Canonical	0.39	0.16	0.61	<.001	
1 s	IDS	Reflexive	1.52	1.13	1.90	<.001	Reflexive > Canonical; Reflexive > Non-Canonical
1 s	ODS	Canonical	0.35	-0.07	0.75	0.01	
1 s	ODS	Non-Canonical	0.12	-0.15	0.39	0.12	
1 s	ODS	Reflexive	0.15	-0.39	0.63	0.33	
2 s	IDS	Canonical	0.21	-0.15	0.73	0.05	
2 s	IDS	Non-Canonical	0.11	-0.11	0.34	0.09	
2 s	IDS	Reflexive	0.98	0.58	1.38	<.001	Reflexive > Canonical; Reflexive > Non-Canonical
2 s	ODS	Canonical	-0.15	-0.59	0.25	0.23	
2 s	ODS	Non-Canonical	-0.39	-0.67	-0.13	<.001	
2 s	ODS	Reflexive	-0.27	-0.79	0.21	0.08	
5 s	IDS	Canonical	-0.13	-0.48	0.21	0.23	
5 s	IDS	Non-Canonical	-0.16	-0.38	0.06	0.02	
5 s	IDS	Reflexive	0.53	0.11	0.94	<.001	Reflexive > Non-Canonical
5 s	ODS	Canonical	-0.50	-0.92	-0.10	<.001	
5 s	ODS	Non-Canonical	-0.76	-1.03	-0.49	<.001	
5 s	ODS	Reflexive	-0.66	-1.18	-0.17	<.001	

Non-Canonical ($p < .001$) infant vocalizations. There were no significant relations between Other-Directed adult vocalizations and subsequent Canonical, Non-Canonical, and Reflexive infant vocalizations.

Analyses with P = 2 s (see Table 2.3) revealed that preceding Infant-Directed adult utterances were significantly associated with subsequent Reflexive infant vocalizations, and that this was a stronger association ($b = 0.98$, $p <.001$) than with Canonical and Non-Canonical infant vocalizations ($p < .001$). Further, we found that preceding Other-Directed adult vocalizations were negatively associated with subsequent Non-Canonical infant vocalizations ($b = -0.39$, $p < .001$).

Finally, we ran analyses with P = 5 s (see Table 2.3). Reflexive infant vocalizations ($b = 0.53$, $p < .001$) succeeded Infant-Directed adult utterances with a stronger association than Non-Canonical infant vocalizations. Second event Canonical ($b = -0.50$, $p < .001$), Non-Canonical ($b = -0.76$, $p < .001$), and Reflexive ($b = -0.66$, $p < .001$) infant vocalizations were negatively associated with first event Other-Directed adult utterances.

2.4.4 Acoustic Analyses of IDS and ODS

Table 2.4 shows results of mixed effects regressions predicting acoustic features as a function of whether an adult vocalization was Infant-Directed or Other-Directed. Positive b values indicate that the measure was greater for ID vocalizations. Mean pitch (in mel) was significantly higher for ID vocalizations ($b = 23.28$, $p < .001$), as was vocalization intensity (in dB; $b = 5.57$, $p < .001$). Pitch variability, as measured by standard deviation in mel of a vocalization's pitch contour, did not significantly differ for ID vocalizations compared to OD vocalizations ($b = -3.09$, $p = .06$).

Table 2.4

Acoustic Differences for Infant-Directed vs. Other-Directed Adult Vocalizations

Acoustic feature	IDS mean (SD)	ODS mean (SD)	*b*	*p*
Mean pitch (mel)	238.20 (73.53)	210.35 (81.27)	23.28	<.001
Pitch standard deviation (mel)	45.77 (36.65)	53.78 (48.16)	-3.09	.06
Mean intensity (dB)	70.23 (7.41)	64.46 (8.09)	5.57	<.001

2.5. Discussion

2.5.1. Infant Vocalizations Predict Subsequent Infant-Directed Adult Vocalizations, and Vice Versa

All three types of infant vocalizations (Reflexive, Canonical, and Non-Canonical Non-Reflexive) predicted that the following event would be an Infant-Directed adult vocalization, as long as the maximum lag between events was relatively short (1 or 2 s). When the allowed time lag was 5 s, a significant tendency for Infant-Directed adult vocalizations to follow infant vocalizations was detected only for reflexive infant vocalizations. This result is consistent with prior research (Van Egeren et al., 2001;

Bornstein et al., 2015) showing that adults' infant-directed vocalizations toward younger infants are more likely to occur following infant vocalizations (both distress and non-distress), and that such responses typically occur within 1 or 2 s.

When looking at first-order sequential contingencies, these sequential relations between infant vocalizations and infant-directed adult vocalizations appear to be bidirectional. All three types of infant vocalizations were more likely to occur following infant-directed adult vocalizations than at other times. This finding is also consistent with prior research, which has typically found fairly symmetrical patterns of temporal association (Jaffe et al., 2001; Van Egeren et al., 2001).

2.5.2 Differences Between Infant Vocalization Types.

Reflexive vocalizations were stronger predictors of subsequent IDS than non-canonical vocalizations. This is consistent with prior work utilizing hand-coding of younger infants' vocal interactions with their caregivers during shorter observation windows (Van Egeren et al., 2001), which has also found strong relationships between infant distress vocalizations of younger, 4-month-old infants, and infant-directed adult utterances. Although it is possible that reflexive vocalizations in 4-month-olds are different than in 12-month-olds (e.g., younger infants may produce a higher proportion of cries to laughs than older infants; younger infants may produce more reflexive vocalizations relative to speech-related vocalizations than older infants; adults may respond more sympathetically to cries from a 4-month-old than to cries from a 12-month-old etc.), it is still interesting to note that non-speech-related vocalizations are such strong predictors of adult responsivity. On the other hand, this finding seems to differ from Warlaumont et al.'s (2014) finding that adult responses were more likely to follow speech-related than reflexive infant vocalizations. Two methodological differences could account for this discrepancy. First, Warlaumont et al. (2014) relied on the automatic labels provided by the LENA system, which may have resulted in the speech-related vs. non-speech-related distinction not mapping as faithfully onto protophones vs. reflexive sounds as the present study's human listener based labels. Moreover, the use of automatic labeling necessitated that Warlaumont and colleagues ignore vocalizations overlapping with other sound sources, which might be especially likely when adults are responding to infant cries. Second, Warlaumont et al. (2014) analyzed the entire day-long audio recordings, whereas the present study analyzed only high-infant-volubility samples from within the day-long recordings. We would speculate that the difference in how overlapping vocalizations were treated across the studies is the main factor, but this and other suppositions should be subject to testing in future research.

Canonical vocalizations were also more likely than non-canonical vocalizations to be followed by IDS (p's > .001 but < .01), when considering 1 s and 2 s lags. This corresponds with previous work by Gros-Louis and colleagues (2006) finding that mothers in a laboratory setting were more likely to respond to infant consonant-vowel sounds (i.e., canonical sounds) than to vowel-like sounds (which we include in our definition of non-canonical vocalizations). These results also align with those of Gros-Louis and Miller (2018), who found that adults were more likely to provide vocal responses following consonant-vowel vocalizations (equivalent to our canonical vocalizations) than following vowel only vocalizations (included within our non-canonical vocalization).

Our results, particularly the bidirectionality of the findings, also fit with work by Goldstein, King, and West (2003) indicating that contingent maternal responses increased subsequent infant vocalizations that were more mature, namely canonical syllables. In light of such prior research, the bidirectionality observed in the present study may indeed reflect a true contingency of infant vocal type on preceding adult vocalization, though we cannot be certain because the same methodological issue with inferring causality discussed later also applies here.

2.5.3. Other-Directed Adult Vocalizations are Negatively Associated with Infant Vocalizations

Our examination of first-order sequential contingencies with 5 s maximum lag between vocal utterances (i.e. a 5 s minimum pause duration) found that all three infant vocalization types were associated with a decreased likelihood that an Other-Directed adult vocalization would follow. This relationship was again bidirectional, with Other-Directed adult vocalizations predicting decreased likelihood of any type of infant vocalization as the next event.

This is a novel and interesting finding, suggesting that other-directed adult utterances, such as adult-directed adult vocalizations, are not unrelated to infant behavior. Prior work (Kuhl, Tsao, & Liu, 2003; Oller, 2010; Shneidman et al., 2013; Weisleder & Fernald, 2013) has found that unlike infant-directed adult speech, other-directed adult speech is not predictive of children's language learning. However, the present results suggest that OD speech should not be totally discounted. One plausible explanation for our results is that a reduction in adults' other-directed vocalizations may cue the infant that adults are paying attention to them. Infants might also vocalize to prompt adults in the environment to focus attention on the infant, which would in turn reduce other-directed adult speech as adults. Alternatively, infants might be more likely to vocalize at times when they are better able to hear their own vocalizations above the background noise of other conversations. Yet another possible explanation for the negative association between ODS and infant vocalization is that it is a byproduct of the positive temporal association between infant vocalizations and IDS (and potentially between successive infant vocalizations, not analyzed here). It is also possible that the human coders used in the present study were better able to hear infant vocalizations (especially short and quiet sounds that would have been categorized as Non-Canonical Non-Reflexive) without other-directed adult conversation taking place in the background. Regardless, the present findings make a case for the importance of considering not only infant-directed adult vocalizations but also other-directed adult vocalizations when considering the full range of contexts that infants experience over the course of a day. This may be particularly relevant when studying cultures that vary in the forms and/or frequency of infant-centered speech (see next section). Incorporation of video or other data could also help provide a clearer picture of the physical locations of people and other aspects of infants' physical and social contexts (see Fausey, Jayaraman, & Smith, 2016).

It is also noteworthy that the detection of the relationship (albeit a negative one) between other-directed adult vocalizations and infant vocalizations necessitated the use of a longer, 5 s criteria for a pause. The 2 s maximum lag revealed only a negative association for one of the infant vocalization types (Non-Canonical Non-Reflexive), and no negative associations with infant vocalization types were found for the 1 s maximum lag. Thus, while focusing on shorter timescales appears appropriate when examining infant-directed adult behavior occurring during typical parent-infant dyadic play sessions, it is important to consider longer timescales of possible associations between infant and adult behavior when examining naturalistic observations in which infants may compete with other individuals for adults' attention (and vice versa).

2.5.4 Acoustic Differences Between Infant-Directed and Other-Directed Adult Speech

In addition to finding qualitative differences in how IDS and ODS related temporally to infant vocalizations, we also observed differences in the acoustic features of IDS and ODS. On average, the mean pitch of infant-directed adult vocalizations was higher than that of other-directed adult vocalizations, consistent with the large body of prior work analyzing pitch of infant- versus adult-directed speech in more constrained settings (Fernald, Taeschner, Dunn, & Papousek, 1989; Broesch & Bryant, 2015).

We also found that infant-directed adult vocalizations had greater intensity (a.k.a. amplitude and the acoustic correlate of perceived loudness) as recorded by the infant-worn microphone compared to other-directed adult vocalizations; this is not something that has been found in prior research on IDS acoustics. One possible explanation could be that observed IDS was produced in closer proximity to the infant than overheard ODS. Regardless of the cause, the intensity of adult vocalization could provide both an additional benefit for children's communication development and an additional cue to infants about which speech is intended for their benefit. The focus here on child-centered naturalistic recordings thus permitted the identification of a new acoustic feature distinguishing infant-directed speech in the real world, from the child's perspective.

Finally, unlike prior work (Fernald et al., 1989), we did not find greater pitch variability for infant-directed vocalizations. Our pitch measurements were based on automated analyses, which may be less accurate, and our data were from a wider range of contexts. Either of these factors could account for the lack of pitch variability effects seen here. Nevertheless, it suggests the possibility that caregivers' pitch variability may be less of a cue and aid to infant communication development than increased pitch and intensity overall.

The acoustic differences observed between infant-directed and other-directed adult utterances could help explain why infant-directed adult vocalizations show positive temporal association with infant vocalizations. The louder, infant-directed vocalizations may be more salient to infants. Moreover, their higher pitch, which more closely matches infants' own voices, may be beneficial for infants' learning to map their own vocalizations onto adult targets. Thus, the acoustic features that distinguish ID speech may serve to stimulate infant vocalization. In turn, infant vocalizations, especially those indicating distress and those that are more mature, may cue adults to modify their vocalization acoustics in order to distract, teach, or soothe the infant. It also seems plausible that different types of infant vocalization may elicit different types of acoustic modifications by adults, both in their infant-directed and in their other-directed speech, and vice versa. Other research automatically analyzing pitch of infant and adult vocalizations within day-long audio recordings (but not distinguishing child-directed from other-directed adult speech) has found evidence for pitch matching between 12–30 month old children and their caregivers (Ko, Seidl, Cristia, Reimchen, & Soderstrom, 2015).

2.5.5. The Difficulty of Inferring Causality from First-Order Temporal Contingencies

Our findings that patterns of temporal association were predominantly bidirectional, including the associations between reflexive vocalizations (where cries greatly outnumbered laughs) and infant-directed adult vocalizations, highlights issues of inferring causality from first-order temporal associations. Though possible, it seems unlikely that the high tendency for infant-directed adult vocalizations to be followed by infant cries is due to infant cries being *caused* by adult vocalizations. A more likely explanation is that infant cries tended to occur in bouts, with adults responding to infant cries within such bouts. More formally, where P(B | A) = the probability of vocal event B occurring given that vocal event A has just occurred, it is unclear if the pattern of infant vocalizations tending to follow adult vocalizations (i.e., P(infant | adult) > P(infant | not adult)), is due to P(infant | infant → adult) > P(infant | infant → not adult) or whether it is due to P(infant | infant → any event) > P(infant | not infant → any event) and P(adult | infant) > P(adult | not infant). This latter scenario would be the case if infants tend to produce their vocalizations over a protracted period of time or in bouts while adults have a tendency to respond to the earlier infant vocalizations within those series. A symmetrical set of indistinguishable possibilities also applies for explaining the P (adult | infant) > P (adult | not infant) pattern. These issues were made salient by the comprehensiveness of the present study's consideration of different infant vocal types (particularly cries, which were the most frequent type of reflexive vocalization) and of both infant preceding adult and adult preceding infant patterns.

2.5.6. Additional Future Directions

Although we have referred to the present study as a "comprehensive" study of first-order temporal contingencies between infant and adult vocalizations (because we included three main categories of infant vocalizations, both infant-directed and non-infant-directed input from adults and analyzed three different possible timescales at which such contingencies could operate), there exist many possible useful extensions. For instance, Golinkoff and colleagues (2015) have discussed the idea of quality vs. quantity in regard to IDS. The present study identified the *quantity* of vocalizations in temporal relation to infant behavior. However, future studies could examine certain *qualitative* aspects of the IDS utterances, such as the pragmatics (e.g., imperative, interrogative, prohibition, etc.) or function (i.e., is the utterance labeling an object, expanding upon an infant utterance, recasting an infant utterance, etc.) of the IDS utterances within day-long naturalistic recordings. Although previous work has examined some qualitative aspects of IDS, including part vs. whole within word labels (Gogate et al., 2013), adjusting levels of sensitivity (Zukow-Goldring, 1996), or directives (Reddy, Liebal, Hicks, Jonnalagadda, & Chintalapuri, 2013), such research has not used data from day-long home recordings. Similarly, infant vocalizations could also be broken down into a much more fine-grained set of types, such as distinguishing between different types of non-canonical babble.

We also encourage future work examining three-event sequences and/or use other analytic approaches or experimental designs to investigate the causal underpinnings of the observed bidirectional associations. Some prior studies have addressed three-event sequential patterns (Warlaumont et al., 2014; Gros-Louis & Miller, 2018; Harbison et al., 2018), though their methods would need to be adapted in order to distinguish the specific alternatives posed above. A related approach would be to apply Granger causality analysis to consider a larger range of timescales and possible patterns of temporal association across speakers and vocalization types (see Xu, Abney, & Yu, 2017, for an example of this approach applied to multi-modal parent-infant interactions).

Additionally, the methods used in the present study could also be applied to recordings of children at other ages (e.g., Jaffe et al., 2001; Tamis-LeMonda et al., 2001), cultures (Lee, Jhang, Relyea, Chen, & Oller, 2018), or populations, such as children with clinical or medical diagnoses, or who are at-risk for language disorders (see Yoder & Warren, 2002). For example, different cultures have different norms regarding the frequency and nature of child-directed speech and speech directed to other children and adults. As mentioned above, it has been reported that in Kaluli culture, infant-directed speech is rare and is not typically used to elicit an infant response, whereas these infants instead hear a good deal of high-pitched other-directed but infant-centered speech in which caregivers speak "for" younger (under 6-months-old) prelinguistic infants (Ochs & Schieffelin, 1984). It would be interesting to separately code infant-directed, infant-centered other-directed (both by the adults speaking-for and others' speech to the adults doing the speaking-for), and not-infant-centered other-directed adult Kaluli speech to assess the temporal contingencies between each adult vocalization type and infant vocal behavior, and the relation of these contingencies with specific acoustic features. It has been asserted that Kaluli caregivers don't usually speak for a prelinguistic infant as an interpretation of infant vocalizations, but rather in response to other contextual factors and non-vocal infant actions. Thus, one might predict that infant-centered other-directed speech would not have the same temporal associations with infant vocalization as was observed for infant-directed speech in the present study.

Even within industrialized, Western cultures there exists substantial variability in the frequency of adult input directed to the child and in how much adults facilitate infant-adult conversational and proto-conversational exchanges (Hart & Risley, 1995). The temporal contingency analyses controlled for base rates of adult and infant vocalizations and individual or subgroup differences were not examined. It is possible that base rates of adult vocalization may correlate with the strength of infant-adult temporal contingencies, but it is also possible that contingencies can be strong despite relatively low base rates. This would be an interesting topic for future work comparing both within and across cultures. Additionally, as discussed in the section above, this present study does not allow us to make any conclusions about causality. In order to make definitive claims about causality would require tightly controlled experiments manipulating IDS and ODS and testing for differences in infant vocalizations. Finally and relatedly, future work could relate patterns of contingency, base rates of behaviors, and acoustic modifications in infant-adult vocal interactions from within day-long audio recordings to measurable language, social, emotional and other outcomes.

2.6. Conclusion

Sampling day-long home audio recordings provides a set of vocalizations that are more representative of infants' typical experiences than those used in many prior studies. Manually annotating segments of day-long home audio recordings enabled us to assess the relationship between various types of infant vocalizations and *infant*-directed and *other*-directed adult speech. Our findings support prior findings of temporal relationships between infant-directed adult speech and infant vocalizations of all types, indicating that such patterns generalize to natural home environments. Moreover, we identified additional patterns of parent-infant vocal interactions, particularly that other-directed adult vocalization and infant vocalization are negatively associated with each other. Finally, that all the detected contingencies were found to be bidirectional highlights some of the issues with inferring causality from first-order temporal contingencies.

Chapter 3

Locomotion, Language, and Interactions: How adult responsivity to infant vocalizations relates to vocabulary development in walking and crawling infants.

3.1 Preface

Infants undergo a series of dynamic changes during the first year and a half of their life, starting at nearly complete dependence upon others for all functions, culminating with the ability to self-transport and rudimentarily converse around their first birthday. These changes are influenced by interactions with adults in their environments. Indeed, infant-adult vocal interactions have been found to promote prelinguistic development. There is also evidence of an interaction between walking and expressive and receptive language development (see Walle & Campos, 2014). This chapter presents a study that explored the relationships between prelinguistic development, adult responsivity, vocabulary development, and locomotor status (i.e., walking or crawling as primary mode of locomotion). The current study is in preparation for submission and co-authored by Lukas D. Lopez, Eric A. Walle, and Anne S. Warlaumont. In this study, we expanded the participant pool from the prior chapter to include 41 infants from the ongoing Walking and Language study led by the Walle Interpersonal Development Lab. In addition to our previous research questions, we asked how do motor development and language learning influence one another? We added an outcome measurement in the MacArthur-Bates Communicative Developmental Inventory (MCDI). We ran a series of generalized linear mixed effects models to test for associations between infant vocalization types and infant-directed adult speech (IDS) or other-directed adult speech (ODS). We also tested for interactions between locomotor status (i.e., walking or crawling) and type of infant vocalization. Finally, we explored how differential responses to walking and crawling infants are associated with receptive and productive vocabulary. We found similar patterns of significance within infant-adult interactions to those found in the study reported in Chapter 2. We also found that for walking infants, less mature infant vocalizations (i.e., Non-Canonical and Reflexive vocalizations) predicted more Infant Directed Speech (IDS) and that the more mature infant vocalizations (i.e., Canonical vocalizations) were associated with decreased Other-Directed Speech (ODS). For crawling infants, Reflexive vocalizations predicted IDS and all vocalization types predicted decreased ODS. Although we did not replicate Walle and Campos' (2014) finding of significantly higher vocabulary in walking infants than crawling infants, we did find (using a conservative α cutoff (p = .001) due to the number of tests) marginally significant (p = .002) differences in the relations between IDS and Productive vocabulary for walkers and crawlers.

3.2 Introduction

During the first year and a half of life, infants' vocalizations increase in maturity and variability, starting with reflexive vocalizations (e.g., cries) and very primitive speech-related vocalizations (e.g., quasivowels), progressing into more speech-like vocalizations (e.g., full vowels and gooing), and expanding into more complex protophones (e.g. squeals, growls, and marginal babbling; see Buder, Warlaumont, & Oller, 2013). Next, we see the onset of canonical babbling (i.e., syllables with an adult-speech-like transition between a consonant and full vowel), which increases in frequency and speech-likeness over the second half of the first year of life. Finally, first true words appear around the infant's first birthday. Overlapping this burst of new sounds and first words is a series of predictable motor milestones, including locomotion and self-propulsion. Typically, infants are able to sit unsupported at around 5-6 months of age and are transitioning to crawling between 8-10 months of age, with independent walking emerging around the first birthday (Adolph & Berger, 2005; Bayley, 1936; Bayley, 1969; Frankenburg et al, 1992).

As infants refine their motor abilities, their interactions with their environments and the adults therein are also changing (e.g., Bringen, Emde, Campos, & Applebaum, 1995; Clearfield, 2011; Karasik, Tamis-LeMonda, & Adolph, 2013; Kretch, Franchak, & Adolph, 2013; Walle & Campos, 2014). Once walking, infants are more likely to engage in social interactions with the adults than their crawling counterparts (e.g., Karasik, Tarmis-LaMonda, & Adolph, 2013). These changes in infant-adult interactions are noteworthy because infant-caregiver interactions influence the child's prelinguistic development (Bloom, 1998; Nathani & Stark, 1996; Brody et al., 2004; Goldstein & Schwade, 2008; Gros-Louis et al., 2014; McCune & Vihman, 2001; Papousek & Papousek, 1989; Stoel-Gammon, 1998; Wu & Gros-Louis, 2014). We do know that locomotor status (i.e., walking or crawling) is associated with vocabulary development (He, Walle, & Campos, 2015; Walle & Campos, 2014; Walle & Warlaumont, 2015; West, Leezenbaum, Northrup, & Iverson, 2017). The present work examines the link between motor and language development further. We explore how locomotion and parent responsivity to infant vocalizations are interrelated.

3.2.1 Locomotor Development

Motor development among infants is dynamic and variable; however, the process of learning locomotion is defined by predictable milestones. Initially, infants are dependent upon others to move from one place to another, but they learn to independently transport themselves by age 12 months old (Adolph & Berger, 2005; Bayley, 1936; Bayley, 1969; Frankenburg et al, 1992). Infants must learn to coordinate muscle movements of the arms, legs, and trunk, as well as gain strength and endurance for locomotion. Typically developing children progress through these stages (sitting, crawling, pulling to stand, and first steps) over a period of six to eight months starting around 4-6 months of age.

3.2.2 Prelinguistic Development

Similar to motor development, infants follow a stable path of vocal development, marked by predictable patterns of skill acquisition. Oller (2000) categorizes infant phonations as either speech-related (Protophones) or non-speech-related. Non-speech-related sounds, such as reflexive and vegetative signals are present regardless of an infant's speech capabilities. Protophones, however, develop along a predictable timeline. At birth, infant vocalization is limited to reflexive and vegetative signals and quasi-vowel protophones. Infants then start to expand the range of duration, loudness, pitch, and voice quality present in their vocal repertoires. They also begin to engage in marginal babbling (e.g., syllables with either no full vowel or an elongated transition from consonant to vowel/quasi-vowel; see Buder, Warlaumont, & Oller, 2013; Oller, 2000). Next in prelinguistic development is the canonical babbling stage, which typically emerges from 6-8 months of age. Canonical babbling is distinguishable from marginal babbling in that the syllable contains an adult-like timing in the transition between the margin (true consonant onset and/or offset) and nucleus (full vowel) of the syllable.

3.2.3 Language and Locomotor Development

Remarkably, one the most salient language development milestones, first true words, is achieved at or near the same time one of the most significant motor milestones - walking independently. Although there are documented links between motor development and cognitive development, relatively little was known until recently about the relationship between locomotion and language (Bringen, Emde, Campos, & Applebaum, 1995; Clearfield, 2011; Karasik, Tamis-LeMonda, & Adolph, 2013; Kretch, Franchak, & Adolph, 2013; Walle & Campos, 2014; He, Walle & Campos, 2015; Walle & Warlaumont, 2015). We do know that language deficits and motor deficits co-occur in children (Bishop, 2002; Diamond, 2000; Hill, 1998; Ojemann, 1984; Viholainen et al., 2002). Viholainen and colleagues (2002) report that between 60-90% of children with developmental disabilities have comorbid language and motor deficits. Some researchers suggest underlying mechanisms connecting the development of language and motor abilities.

Walle and Campos (2014) questioned whether the onset of walking is related to language development in a two-study investigation, comprised of longitudinal parent reports via a bi-weekly online survey regarding motor and language abilities and a naturalistic observational study of infant-parent dyads. Both studies utilized the MCDI, which is a checklist that parents complete to indicate words that their child can say (Productive Language) and understand (Receptive Language; see Fenson et al., 1994, 2000). Walle and Campos (2014) first found that both age and walking experience were significant predictors of both Receptive and Productive Language. Walking infants had significantly higher receptive and productive vocabularies than their age matched peers who were still crawling. The findings were replicated cross-sectionally across the United States and cross-linguistically in a study of typically developing Chinese infants exposed to Mandarin (He, Walle, & Campos, 2015).

Walle and Warlaumont (2015) used daylong LENA (LENA; LENA Research Foundation, Boulder, Colorado, United States) recordings to examine how 13 age-matched walking and crawling infants' vocal productions, language input, and turn-taking were related to their vocabulary development. Some of those infants were included in the present set of analyses. They did not detect a significant difference in vocabulary size between locomotor groups. However, they did find that greater number of adult words heard, as measured by LENA's automatic analyses, corresponded with increased Receptive and Productive vocabulary (reported by parents via the MCDI) for walking, but not crawling, infants. Further, infant vocalization counts and turn-taking counts also corresponded with increased Productive vocabulary, but only for walking infants.

West and colleagues (2017) conducted a longitudinal study in which 91 children at risk for Autism Spectrum Disorders (ASD; i.e., a full biological sibling had received a diagnosis of ASD) and 25 children with no family history of ASD were followed for to observe changes in receptive and productive language across the transition from crawling to walking. The children at risk for ASD were later categorized by whether they had a diagnosis of ASD or other language delay. The authors found that only the infants at risk for ASD that later received a diagnosis of ASD did not demonstrate an increase in vocabulary following the onset of walking.

3.2.4 Overview of the Present Study

Given the relationship between locomotor development and vocabulary development, further exploration is needed to identify more specifically how these are interrelated. While this exact relationship is unknown, some researchers credit changes in social interactions following the acquisition of walking. Another possibility an underlying motor function for both movement and speech, as evidenced by children with disorders impacting both speech and movement. Although there has been great progress towards understanding relationships between infant developmental domains with regards to language development, very little work has focused on uncovering potential interrelationships between locomotor development, language development, and infant-caregiver vocal interactions. The present study analyzed hand-coded segments from daylong audio recordings of 11-13-month-old infants during high infant volubility (i.e., high child vocalization rates) times during the day, using the same procedures as in Pretzer et al. (in press). Infant vocalizations were classified by type (i.e., canonical, non-canonical/non-reflexive, and reflexive; see Buder et al., 2013) and adult utterances were categorized as either infant-directed (IDS) or other-directed (ODS). We investigated whether (1) adults responded differently to diverse vocalization types (canonical, non-canonical/non-reflexive, or reflexive), (2) parents responded differently to crawling and walking infants, and (3) how different patterns of responses related with infants' language development, specifically receptive and expressive vocabulary acquisition, for walking and crawling infants.

3.3 Methods

3.3.1 Participants

Forty-one infants (23 female) aged 11-13 months (walkers mean = 12.7 months, SD = 0.40; crawlers mean = 12.6 months, SD= 0.51) were selected from an ongoing study of 132 recruited infants living in the San Joaquin Valley in central California. Of the 46 participants, 21 infants had at least two weeks of walking experience, while the other 20 were had less than two weeks of walking experience.

Families from the larger study were excluded from the present analyses because their recordings were too short (i.e., <10 hours) or took place over multiple days, Spanish was the primary language spoken at home (i.e., >50% of the time per parent report; n= 10), the infant was outside the targeted age range (n=5), or recordings were not completed (due to technical error or not being returned by family; n=21). We chose to include recordings of at least 10 hours in length to ensure that segments drawn from the recording were taken from a larger variety of contexts.

Eighteen infants were from families with both parents having a college degree (including 11 graduate degrees), 13 infants were from families with one parent having a college degree (including 1 graduate degree), and 15 infants were from families with neither parent having a college degree. Thirty-four infants were reported to be of Caucasian descent, 24 infants of Hispanic descent, and 3 infants of Asian descent. One participant disclosed a congenital heart defect at birth, but there were no reported diagnoses of developmental delays or serious ongoing medical conditions for any other participants.

3.3.2 Recording Procedure

Families were either mailed a LENA recorder and vest specially designed vest to hold the recorder or had study materials dropped off at their homes. The recorders were used to capture the infant's vocalizations, vocalizations of nearby speakers, and other environmental sounds (e.g., electronic noise). Parents were instructed to turn on the recorder and place it in the chest pocket of the provided vest before donning on the infant. The infant then wore the vest for the entire day, with the exception of baths, naps, or car rides, when the recorder could be placed nearby to continue recording vocalizations. Families could pause the recorder for privacy reasons. The recorder automatically shut off after 16 hours; however, some families chose to stop the recorder prior to the 16-hour mark.

Additionally, parents were also asked to complete and return a participant history questionnaire that asked about demographic information and locomotor milestones. They were also asked to complete the MCDI: Words and Gestures (Fenson et al., 2007) to measure receptive and productive vocabulary.

3.3.3 Selection of High Infant Volubility Samples

The audio recordings were first processed using LENA Pro system's software. The software's automatic labeling system applied a set of mutually exclusive labels to identify the speakers (e.g., target child) or sound sources (e.g., electronic sounds) continually throughout the entire recording. The automated labels are: infant wearing the recorder, male adult, female adult, other child, electronic sounds, noise, and silence, with all labels except silence being further divided into "near" (i.e. relatively loud) and "far" sounds. We used LENA's Automatic Data Extractor (ADEX) to segment each recording into 5-minute sections and then identified the three most voluble (i.e., containing the most infant vocalizations identified by the automated labeling system) 5-minute segments. However, high volubility segments identified for coding were replaced by the next most voluble segment if the segment started within 15 minutes after the end of the previous segment or it sounded as if an object were obstructing the infant's mouth (e.g., a pacifier) during the majority of the segment, per human listener. These criteria were established to allow for distinct observational times throughout the day, to validate that these segments truly were the most voluble times of day per LENA's automatic labeling system, and to ensure that the infant's utterances were not impeded.

3.3.4 Identification and Classification of Infant and Adult Utterances

Trained researchers used the EUDICO Linguistic Annotator (ELAN, 2018; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006) to identify and code infant and adult vocalizations using an annotation system devised by the authors (Pretzer et al., in press). First, at team of researchers, comprised of the first two authors and UC Merced undergraduate student research asssistants, marked vocalizations by listening to the recording and pausing, replaying, and jumping backward and forward as needed to mark the onsets and offsets of infant vocalizations according to the following criteria:

1. *Mark any sound of nontrivial loudness that you think was made by the vocal tract (i.e. larynx, throat, mouth, lips, teeth, nasal passage, etc.).*
2. *Be as accurate as possible when setting utterance boundaries - try to be within tenths of a second.*
3. *Annotations should run from the onset of sound to the offset of sound. [For adult vocalizations, it was also noted that ends of phrases/sentences are often natural ends of vocalizations.]*

Next, vocalizations were coded by speaker and by type. Infant vocalizations were coded *V* for vegetative (e.g., cough, hiccup, burp), *R* for reflexive (i.e., laugh or cry), *X* for non-canonical/non-reflexive (i.e., a protophone vocalization that does not contain a canonical syllable, including marginal babbling; see Buder et al., 2013), or *C* for canonical (i.e., containing at least one syllable with a speech-like transition between a full consonant and vowel). However, vegetative infant vocalizations were not included in these analyses. Adult vocalizations were coded by the perceived direction of their utterance (i.e., whether the utterance was directed toward the target infant or directed to any other person or animal present). IDS was coded as *T* (i.e., to the target child), and ODS was coded as *N* (i.e., not to the target child). Segments containing Spanish language were coded by native or fluent Spanish speakers.

3.3.5 Event Series Analyses

We created an event series of all infant and adult vocalizations using the same methods by Pretzer and colleagues (in press) with this current data set. Once coded, segments were excluded from analysis if: 1) they were deemed to be mostly in Spanish by human coders; 2) there were fewer than 10 infant speech-related utterances (i.e., canonical or marginal babbling or other protophones as judged by a human listener (see Buder et al., 2013 for definitions); or 3) the segment was more than 10 seconds less than 5 minutes (e.g., some of the highest volubility segments occurred at the end of a recording and were not of full 5 minute duration). Infant and adult vocalizations from each coded segment were converted into a single event series to identify temporal contingencies between infant and adult utterances. First, overlapping vocalizations and pauses between vocalizations were identified. A pause was defined as the duration greater than *P* between the end of the first vocal event and the beginning of the subsequent vocal event (i.e., the empty space between vocalizations). For these analyses, we set the lag time between vocalizations at $P = 1$ s, as previous work has identified 1 s as a timescale at which infants are sensitive to contingent vocalizations (e.g., Gros-Louis, West, Goldstein, & King, 2006; Keller, Lohaus, Völker, Cappenberg, & Chasiotis, 1999; Pretzer et al., in press; Van Egeren et al., 2001; Warlaumont et al., 2014). Thus, any space between vocalizations of at least 1 s in duration was designated as a "pause".

Unlike LENA's automatic labeling system, we chose to include overlapping utterances where the first speaker's vocalization was coded as event 1 and the second speaker's vocalization was coded event 2. For cases in which event 1 extended greater than or equal to *P* beyond the end of event 2, then event 1 was split into event 1 and event 3. In cases where event 2 extended less than *P* beyond the end of event 1, the vocalizations remained event 1 and event 2, respectively.

Next, we paired adjacent vocalizations from the single event series into 2-event sequences to test for temporal contingencies between vocal events. Starting with the first vocalization of the event series, the vocal event was combined with the next vocalization in the event series. For example, if a canonical vocalization (C) was followed by an infant-directed adult utterance (T), the 2-event sequence would be CT.

We ran generalized linear mixed effects models using a binomial distribution with Infant ID as a random effect to analyze the relationship between the first event type and second event type for each sequence pair. All analyses were programmed in R (R Core Team, 2018) using the lme4 (Bates, Maechler, Bolker, & Walker, 2018), lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017), and car (Fox et al., 2018) packages.

We then ran generalized linear mixed effects models to explore relations between infant and adult vocalizations for walking and crawling infants, separately. Finally, we ran more generalized linear mixed effects models to test for associations between locomotor status (i.e., walking or crawling), infant vocalization type, adult responsivity (i.e., IDS vs ODS), and reported receptive and productive vocabulary scores gathered from the MCDI.

3.3.6 Code and Data Availability

All analysis scripts and the 2-event sequence data are available at https://github.com/gpretzer/LRScripts. Additionally, example 5-minute audio segments that did not contain last names or especially private episodes and their annotations from infants whose caregivers agreed to share the recordings publicly will be available on FigShare. Full LENA recordings from participants who gave consent are available on HomeBank (VanDam et al., 2016) within the Warlaumont Corpus (Warlaumont, Pretzer, Mendoza, & Walle, 2016).

3.4 Results

This section presents the findings of our current study Subsection 3.4.1 reports the frequencies of specific vocalizations and sequences of infant vocalization followed by adult vocalization (i.e., CT, XT, RT, CN, XN, RN) for walking and crawling infants. We then test whether child vocalization categories predict IDS or ODS (Subsection 3.4.2). The next subsection (3.4.3) relates how infant vocalizations differ between locomotor groups. Subsection 3.4.4 reports relations and interactions between locomotion and child or adult vocalization type and MCDI-P or MCDI-R scores. The final subsection (3.4.5) expands upon the previous analyses by examining how adult responses to infant vocalizations (IDS or ODS and specific vocal event sequences) relate to MCDI-R or MCDI-P scores.

3.4.1 Descriptive Statistics

We hand-coded 240 5-minute segments from 46 infants' daylong audio recordings. Infant vocalizations were classified by type and adult vocalizations were coded based on direction of utterance (see Section 3.3.4 above). Table 3.1 reports counts for each type of infant and adult vocalization for walking and crawling infants. 53% of all vocalizations were produced by an infant, and 85% of those infant vocalizations were speech-related (i.e., canonical or non-canonical/non-reflexive). Table 3.2 reports counts for specific sequences of infant vocalizations followed by adult vocalizations. Within the two-event sequence pairs, 70% of adult vocalizations that followed an infant vocalization were IDS, while the remaining 30% were ODS.

Table 3.1
Event Counts per Vocalization Type for Crawling and Walking Infants

	Adult IDS	Adult ODS	Infant Canonical	Infant Non-Canonical	Infant Reflexive	Pause
Crawlers	2113	1700	1113	2665	947	15412
Walkers	2670	1868	1123	2727	490	15074
Total	4783	3568	2236	5392	1437	30486

Table 3.2
Frequency of Infant-Adult Vocal Sequences for Crawling and Walking Infants

	Canonical-IDS	Canoncial-ODS	Non-Canonical-IDS	Non-Canonical-ODS	Reflexive-IDS	Reflexive-ODS
Crawlers	269	76	398	209	252	89
Walkers	256	114	506	259	123	41
Total	525	190	904	468	375	130

3.4.2 Infant vocalizations predict subsequent IDS and ODS

Table 3.3 shows the results for how each of the three infant vocalizations types uniquely predicted that the subsequent vocal event would be IDS. Given the number of planned tests, we used a Bonferroni correction to set a conservative α cutoff ($p = .001$) to indicate statistical significance. We found that Canonical ($b = 1.36$, $p < .001$), Non-

Canonical ($b = 1.02$, $p < .001$), and Reflexive ($b = 1.59$, $p < .001$) infant vocalizations each predicted subsequent IDS. Further post-hoc comparisons (again using a conservative α cutoff, $p = .001$) revealed that both Canonical ($b = 0.39$, $p < .001$) and Reflexive ($b = 0.63$, $p < .001$) vocalizations were significantly stronger predictors of IDS than Non-Canonical vocalizations.

Table 3.3 also presents the results for how each of the three infant vocalizations types uniquely predicted subsequent ODS. Non-Canonical infant vocalizations ($b = 0.24$, $p < .001$) predicted subsequent ODS.

Table 3.3
Infant Vocalizations Predicting Subsequent IDS or ODS

First Event	Second Event	*b*	95% CI Lower Bound	95% CI Upper Bound	*p*	Statistically Significant Post-Hoc Comparisons
Canonical	IDS	1.36	1.17	1.55	<.001	Canonical > Non-Canonical
Non-Canonical	IDS	1.02	0.87	1.16	<.001	
Reflexive	IDS	1.59	1.36	1.80	<.001	Reflexive > Non-Canonical
Canonical	ODS	0.16	-0.11	0.42	0.04	
Non-Canonical	ODS	0.24	0.06	0.41	<.001	
Reflexive	ODS	0.24	-0.09	0.55	0.01	

3.4.3 IDS and ODS Predict Subsequent Infant Vocalizations

Table 3.4 reveals how first event IDS utterances were predicted by second event infant vocalizations of each type. IDS was significantly associated with subsequent Canonical ($b = 0.78$, $p < .001$), Non-Canonical ($b = 0.76$, p < .001), and Reflexive ($b = 1.37$, $p < .001$) infant vocalizations. Post-hoc comparisons determined that Reflexive vocalizations were a stronger predictor of a preceding IDS utterance than Canonical ($b = 0.71$, $p < .001$) and Non-Canonical ($b = 0.69$, $p < .001$) vocalizations.

Table 3.3 also shows how first event ODS utterances were uniquely predicted by subsequent infant vocalizations of each type. First event ODS were significantly predicted by second event Non-Canonical ($b = 0.34$, $p < .001$) infant vocalizations.

3.4.4 Infant and Adult Vocalization Frequencies for Walking and Crawling Infants

We ran a generalized linear model to examine relations between infant and adult vocal types based on locomotor status. We found a negative association between Reflexive infant vocalizations ($b = -0.66$, $p < .001$) and locomotor status.

Table 3.4
IDS or ODS Preceding Infant Vocalizations

First Event	Second Event	*b*	95% CI Lower Bound	95% CI Upper Bound	*p*	Statistically Significant Post-Hoc Comparisons
IDS	Canonical	0.77	0.57	0.98	<.001	
IDS	Non-Canonical	0.76	0.61	0.91	<.001	
IDS	Reflexive	1.37	1.14	1.59	<.001	Reflexive > Canonical; Reflexive > Non-Canonical
ODS	Canonical	0.21	-0.06	0.46	0.01	
ODS	Non-Canonical	0.33	0.16	0.50	<.001	
ODS	Reflexive	0.19	-0.15	0.50	0.05	

We then ran generalized linear mixed effects models with a Poisson distribution and InfantID as a random effect to explore relations between infant and adult vocalizations for walking and crawling infants, separately (Table 3.5). At the 5-minute level, we found that for walkers, Non-Canonical ($b = 0.22$, $p < .001$) and Reflexive ($b = 0.17$, $p < .001$) infant vocalizations predicted more IDS, while Canonical ($b = -0.19$, $p < .001$) infant vocalizations predicted fewer IDS utterances in the 5-minute segment. For crawling infants, only Reflexive infant vocalizations ($b = 0.28$, $p < .001$) significantly predicted IDS. Next, we found that for walking infants, Canonical infant vocalizations ($b = -0.15$, $p < .001$) were negatively associated with ODS in five-minute segments. For crawling infants, all infant vocalization types predicted less ODS (Canonical, $b = -0.41$, $p < .001$; Non-Canonical, $b = -0.13$, $p < .001$; Reflexive, $b = -0.21$, $p < .001$).

3.4.5 Associations Between Infant and Adult Vocalizations and MCDI Scores in Walking and Crawling Infants

First, we ran linear regression models with locomotor status predicting Recepttive ($p = 0.39$) and Productive (p =0.13) scores, but found no significant relations. Next, we ran linear regression models with the interaction between locomotor status and each type of infant vocalizations as predictors of MCDI score to examine how Receptive and Productive vocabulary were associated with locomotor status, but found no significant relations ($p = .74$ and $p = .56$; see Figure 3.1). We then ran similar models, but using the interaction between locomotor status and direction of adult utterances to predict Receptive and Productive vocabulary scores (see Table 3.6), but found no significant associations (all p's > .001). The interaction between total IDS vocalizations heard and locomotor status was a marginally significant predictor of Productive vocabulary scores ($b = 0.63$, $p = 0.005$). Next, we ran generalized linear mixed effects models with Poisson

Table 3.5

Relations between Infant Vocalizations and Adult Vocalizations for Walking and Crawling Infants

Locomotor Status	Infant Vocalization	Adult Vocalizations	*b*	*p*
Crawling	Canonical	IDS	0.06	0.08
Crawling	Non-Canonical	IDS	0.04	0.15
Crawling	Reflexive	IDS	0.29	<.001
Crawling	Canonical	ODS	-0.41	<.001
Crawling	Non-Canonical	ODS	-0.13	<.001
Crawling	Reflexive	ODS	-0.21	<.001
Walking	Canonical	IDS	-0.18	<.001
Walking	Non-Canonical	IDS	0.22	<.001
Walking	Reflexive	IDS	0.17	<.001
Walking	Canonical	ODS	-0.15	<.001
Walking	Non-Canonical	ODS	-0.07	0.02
Walking	Reflexive	ODS	-0.02	0.53

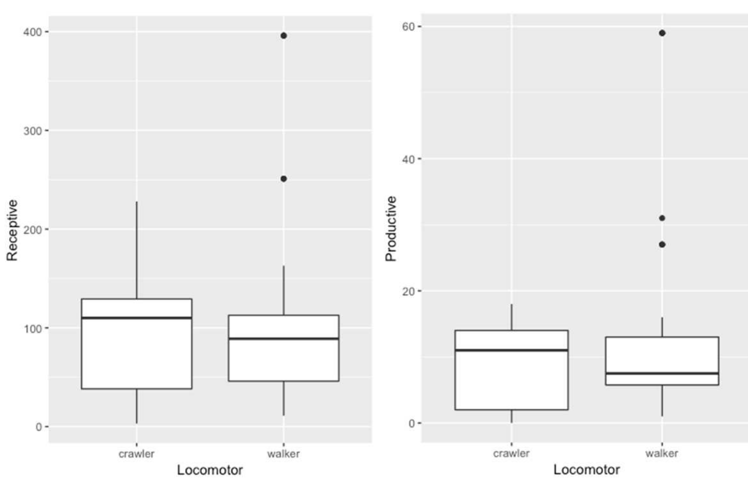

Figure 3.1. (Left) A boxplot showing the distribution of MCDI-Receptive scores for crawling and walking infants. (Right) A boxplot showing the distribution of MCDI-Productive scores for crawling and walking infants. Locomotor status did not significantly predict Receptive or Productive vocabulary size.

distributions and using InfantID as a random effect to look at these relations in walking and crawling infants, separately. Locomotor groups were analysed separately due to the size of the sample. Again, there were no significant results ($p < .001$), but we did find a marginally significant relation between total IDS utterances predicting Productive vocabulary in both walking ($b = 0.46$, $p = .002$) and crawling ($b = 0.85$, $p = 0.002$) groups. These marginal results are interesting to note, as we know that IDS is related to vocabulary acquisition.

Table 3.6

Infant and Adult Vocalization Types Predicting Receptive and Productive Vocabulary Size

Locomotor Status	MCDI Subtest	Vocalization Type	*b*	*p*
Crawling	Receptive	Canonical	0.06	0.85
Crawling	Receptive	Non-Canonical	0.19	0.54
Crawling	Receptive	Reflexive	0.24	0.40
Crawling	Receptive	IDS	0.56	0.03
Crawling	Receptive	ODS	0.32	0.22
Crawling	Productive	Canonical	0.11	0.72
Crawling	Productive	Non-Canonical	0.59	0.06
Crawling	Productive	Reflexive	0.71	0.01
Crawling	Productive	IDS	0.85	0.002
Crawling	Productive	ODS	0.36	0.17
Walking	Receptive	Canonical	0.25	0.18
Walking	Receptive	Non-Canonical	0.15	0.44
Walking	Receptive	Reflexive	-0.06	0.77
Walking	Receptive	IDS	0.43	0.01
Walking	Receptive	ODS	0.17	0.32
Walking	Productive	Canonical	0.26	0.15
Walking	Productive	Non-Canonical	-0.03	0.86
Walking	Productive	Reflexive	-0.001	0.99
Walking	Productive	IDS	0.46	0.002
Walking	Productive	ODS	0.20	0.18

3.4.6 Associations Between IDS and ODS and Vocabulary in Walking and Crawling Infants

Given that the total number of IDS utterances heard by an infant was a marginally significant predictor of their Productive vocabulary for both walking and crawling infants, we decided to examine the total number of IDS and ODS responses to specific infant vocalizations, i.e. sequences where infant vocalizations were followed by IDS or ODS within 1 s, to see whether they would predict vocabulary scores. We looked at proportions of Canonical-IDS, Canonical-ODS, Non-Canonical-IDS, Non-Canonical-ODS, Reflexive-IDS and Reflexive-ODS vocal sequences in 5-minute segments rather than counts due to the highly variable base rates of each sequence of vocal events. Table 3.7 reports the associations between infant-adult vocal sequences and vocabulary, as measured by the MCDI, for walking and crawling infants. We found that for walkers, Canonical-ODS (b = 0.001, p < .001), Non-Canonical-IDS (b = 0.002, p < .001), Non-Canonical-ODS (b = 0.001, p < .001), and Reflexive-IDS (b = 0.011, p < .001) vocal sequences predicted receptive vocabulary size, while Canonical-ODS (b = -0.001, p < .001) and Reflexive-ODS (b = -0.041, p < .001) vocal sequences were negatively associated with receptive vocabulary size. For crawling infants, we found that Canonical-IDS (b = 0.003, $p < .001$), Non-Canonical-IDS (b = 0.002, $p < .001$), and Non-Canonical-ODS vocal sequences predicted receptive vocabulary size, while Canonical-ODS (b = -0.003, $p < .001$) vocal sequences were negatively associated with vocabulary size.

Table 3.7

Infant-Adult Vocal Sequences Predicting Receptive and Productive Vocabulary Size

Locomotor Status	MCDI Subtest	Vocal Sequence	*b*	*p*
Crawling	Receptive	Canonical-IDS	0.003	<.001
Crawling	Receptive	Non-Canonical-IDS	0.002	<.001
Crawling	Receptive	Reflexive-IDS	0.001	0.001
Crawling	Receptive	Canonical-ODS	-0.003	<.001
Crawling	Receptive	Non-Canonical-ODS	0.006	<.001
Crawling	Receptive	Reflexive-ODS	0.001	0.041
Crawling	Productive	Canonical-IDS	0.007	<.001
Crawling	Productive	Non-Canonical-IDS	0.005	<.001
Crawling	Productive	Reflexive-IDS	0.004	<.001
Crawling	Productive	Canonical-ODS	-0.007	<.001
Crawling	Productive	Non-Canonical-ODS	0.006	<.001
Crawling	Productive	Reflexive-ODS	0.011	<.001
Walking	Receptive	Canonical-IDS	-0.001	<.001
Walking	Receptive	Non-Canonical-IDS	0.002	<.001
Walking	Receptive	Reflexive-IDS	0.011	<.001
Walking	Receptive	Canonical-ODS	0.001	<.001
Walking	Receptive	Non-Canonical-ODS	0.001	<.001
Walking	Receptive	Reflexive-ODS	-0.041	<.001
Walking	Productive	Canonical-IDS	0.001	0.001
Walking	Productive	Non-Canonical-IDS	<.001	0.92
Walking	Productive	Reflexive-IDS	-0.003	<.001
Walking	Productive	Canonical-ODS	<.001	0.75
Walking	Productive	Non-Canonical-ODS	<.001	0.29
Walking	Productive	Reflexive-ODS	-0.003	<.001

Additionally, we found that for walking infants, only Reflexive-IDS (b = 0.008, $p < .001$) vocal sequences predicted productive vocabulary size, and Reflexive-ODS sequences (b = -0.003, $p < .001$) were negatively associated with productive vocabulary size. However, for crawling infants, every vocal sequence type was significantly related to Productive vocabulary size. Canonical-IDS (b = 0.007, $p < .001$), Non-Canonical-IDS (b = 0.005, $p < .001$), Non-Canonical-ODS (b = 0.006, $p < .001$), Reflexive-IDS (b = 0.004, p < .001), and Reflexive-ODS (b = 0.011, $p < .001$) vocal sequences all predicted

Productive vocabulary scores with only Canonical-ODS vocal sequences (b = -0.007, $p <$ .001) negatively associated with Productive vocabulary.

3.5 Discussion

3.5.1 Bidirectional Associations Between Infant Vocalization Type and Adult Utterance Direction

We used identical analyses to Pretzer and colleagues (in press) to see if their findings about patterns of interaction in infant-adult vocal exchanges would remain true in a much larger sample size, but only using a 1 s time lag. We found that all three infant vocalization types (i.e., Canonical, Non-Canonical, and Reflexive) significantly predicted that the subsequent vocal event in the sequence would be IDS. Further, Non-Canonical infant vocalizations (b = 0.24, $p <$.001) predicted subsequent ODS. These results closely match patterns of significance found by Pretzer and colleagues (in press). They also support previous work (Van Egeren et al., 2001; Bornstein et al., 2015) that shows that IDS is more likely to occur following infant vocalizations (whether distress or non-distress signals). Canonical vocalizations were better predictors of subsequent IDS than Non-Canonical vocalizations, which is particularly interesting as it supports findings by Gros-Louis and Miller (2018), who found that adults were more likely to vocally respond to consonant-vowel vocalizations (analogous to our Canonical vocalizations) than to vowel only vocalizations (included in our Non-Canonical vocalizations). We also found that both Canonical and Reflexive vocalizations were significantly stronger predictors than Non-Canonical vocalizations, and that Reflexive vocalizations were marginally stronger predictors than Canonical (p =.003), similar to Pretzer and colleagues (in press). These results appear to coincide with Van Egeren and colleagues (2001) who found strong relations between infants' distress vocalizations (i.e., cries) and IDS. However, they also appear to differ with Warlaumont and colleagues (2014) who reported that adult responses were more likely to follow speech-related (i.e., Canonical or Non-Canonical) infant vocalizations. As discussed in Pretzer et al. (in press), these differences may be attributed to several factors, including that the Warlaumont at al. (2014) study used the automatically assigned LENA labels from entire daylong recordings, where we used hand coded labels from segments of daylong recordings. The automatically assigned labels may not as reliably identify speech-related vocalizations as human coders' assigned labels. Secondly, this hand coding method also allowed for analysis and inclusion of overlapping variables, which were excluded by Warlaumont and colleagues (2014) as they used the automatically assigned labels.

First event IDS vocalizations were predicted by subsequent Canonical, Non-Canonical, and Reflexive vocalizations, with Reflexive vocalizations being a stronger predictor than either Canonical or Non-Canonical vocalizations. Again, these results support findings by Pretzer and colleagues (in press). These patterns of contingencies appear to be bidirectional, consistent with prior research that shows symmetrical patterns of temporal association (Goldstein, West, & King, 2003; Jaffe et al., 2001; Van Egeren et al., 2001). However, we must take caution when inferring causality from first order contingencies (see Pretzer et al., in press).

When examining relations between infant vocalizations and ODS, we found that Non-Canonical predicted subsequent ODS. It is also interesting that Non-Canonical infant vocalizations were predicted to follow ODS. In day to day life, these event sequences might present as an infant vocalizing to herself while a nearby adult speaks to another child or adult present. This sequence might also manifest as an infant using less mature speech-related sounds as a bid for adult attention while that adult is speaking to another person present. Orthographic transcriptions of adult speech, more detail about the acoustic characteristics of infant vocalizations, and/or video of the episodes would help to make more definite claims as to the significance of these particular event sequences.

3.5.2 Walking and Crawling Infants Hear and Produce Different Vocalization Types at Varying Rates

We found that more Non-Canonical (walkers) and Reflexive (walkers and crawlers) infant vocalizations were associated with more IDS in a five-minute segment. Surprisingly, we found a negative association between Canonical vocalizations and IDS for walking infants. This difference in walkers might be explained by proximity. Walking infants might not be around adults when they are producing higher rates of Canonical vocalizations. We also found that Canonical (walkers and crawlers), Non-Canonical (crawlers), and Reflexive (crawlers) all predicted decreased ODS in a five-minute segment. This resonates with previous research that generally indicates that ODS does not strongly support children's language learning, extending prior research primarily focused on vocabulary development.

3.5.3 Parent Responsivity is Related to Infant Vocalization Type, Vocabulary, and Locomotion

We examined the relationship between locomotor status and MCDI scores and found that while crawlers had slightly higher vocabulary sizes than walking infants, this difference was not significant (Figure 3.1). This is particularly interesting because Walle and Warlaumont (2015) found the opposite, but still non-significant, result (walkers > crawlers) with a smaller, earlier portion of this dataset. Further, these two studies appear to go against Walle and Campos (2014) and He, Walle, and Campos (2015) who found significantly higher receptive and productive vocabularies for walking infants than crawling infants. Demographic differences could account for these discrepancies. The families in Walle and Campos' (2014) study were families from the San Francisco Bay Area with an average income of approximately $115,000 annually for Study 1 and $90,000 annually for Study 2 and average parent educational experience of a bachelor's degree. These families reported that English, Spanish, Korean, Tamil, German, French, Hindi, Japanese, Mandarin Chinese, Russian were spoken in the home. In the present study, families were from the San Joaquin Valley, which has a significantly higher Spanish-speaking population. Although the infants in this study were reported to hear Spanish less than 50% of the time, many infants still heard at least some Spanish on a regular basis. Other languages heard included French and Japanese. Further, these families had an average income of approximately $50,000 annually and reported that the average parent educational experience was either completion of high school or one college degree per household.

We next looked at how each infant and adult vocalization type was associated with vocabulary scores for walking and crawling infants. We found marginally significant relations (p =.002) between total IDS vocalizations heard and Productive vocabulary for both walking and crawling infants. Although marginal, this is a very interesting result. We know from prior research that IDS supports vocabulary development. We also know that object labeling is a highly effective method of teaching new words.

Then, we looked at the specific vocal event sequences in a 5 min segment and compared it to that segment's infant's vocabulary scores, where C = Canonical, X = Non-Canonical, and R = Reflexive infant vocalizations and T = IDS and N = ODS adult utterances. We looked at the proportions of sequenced events rather than counts to account for varying base rates of each sequence between infants and between segments. Receptive Vocabulary scores were higher for walkers with higher proportions of CN, XT, XN, RT vocal sequences and crawlers with higher proportions of CT, XT, XN vocal sequences. We also found lower Receptive Vocabulary for walkers with higher proportions of CT and RN vocal event sequences and for crawlers with higher proportions of CN sequences. At first glance, there are definite contrasts in patterns of significance between walking and crawling infants, but there are no discernable systems of predictable results, such as consistency within adult response type, infant vocalizations speech-likeness, or by locomotor status.

Similarly, we looked at how specific vocal event sequences predicted Productive vocabulary in walking and crawling infants. For walking infants, we found that RT and RN vocal sequences predicted Productive vocabulary. For crawling infants, every sequence event but CN predicted Productive vocabulary. Again, we see disparate patterns of significance between walking and crawling infants. We could have expected that higher Receptive or Productive vocabulary scores would be associated with higher proportions of CT, XT, and RT than with higher proportions of CN, XN, and RN vocalizations. It is interesting that there is a negative association between CT sequences and Receptive vocabulary in walking infants, as IDS is known to support language development. Specifically, parent labeling of objects is a highly effective way for infants to learn new words, so it would be reasonable to expect that infants who received more IDS responses would have higher vocabularies.

3.5.4 Limitations of the present study

Although this study presents new findings about the relations between locomotor ability, the speech-likeness of infant vocalizations, and reported vocabulary size, there are some limitations to this work. One such limitation is that we only use selected segments (for a total of 15-30 minutes depending on the number of included segments per infant) out of the entire daylong recording. While we observe infant-adult interactions in a highly naturalistic setting, these interactions are not representative of an entire day, but rather the times of day at which the infants produced the highest number of vocalizations. Another potential limitation is that our classifications are dependent upon human listeners and hand coding is highly subjective. Infant vocalizations are particularly difficult to classify given the great variation of vocalizations produced by any given infant. Human errors may be present in the coding. Further, high infant volubility segments do not guarantee high adult volubility. In some segments, adult vocalizations were sparse, resulting in a wide range of frequency for the vocal event sequence pairs we analyzed in this study. Examining segments with high adult words counts could potentially provide more examples of these vocal sequences, which might alter some of the patterns observed here.

Additionally, examining two event sequences can imply causality; however, the two vocal events may incidentally be paired together. For example, a TR (IDS followed by Reflexive vocalization) sequence would appear to imply that an IDS vocalization elicited a cry or laugh. While this certainly may be the case (e.g., the adult says something that makes the infant laugh or cry), this sequence of events could be a result of infant reflexive vocalizations occurring in bouts or being of long duration. Thus, we cannot claim that the first vocal event causes the second vocal event, but rather we state that the vocalizations are related and predictive.

Finally, the MCDI scores could be interpreted with some caution as this instrument gathers all information via parent report. While parent report can be reliable, the scores might not accurately reflect true receptive and productive vocabulary sizes. Frank and colleagues (2017) explain that while parent report could be subject to bias, the MCDI is structured such that it is a valid and reliable source for estimating infant vocabulary. By providing parent reporters with a list of highly salient items to an age and language, bias is reduced compared to a free recall method of reporting vocabulary. It should also be noted that since the MCDI is *only* a measure of vocabulary size, it is not a measure of overall language ability. A true and detailed measure of language ability would ideally be acquired from a standardized assessment administered by a language development expert, such as a Speech-Language Pathologist.

3.5.5 Future Directions

This work provides evidence that locomotor capabilities and vocal development are related. First steps and first words occur within the same time frame (approximately one year of age). However, future work could look for relations between other significant motor and vocal developmental milestones. For example, infants are crawling and producing canonical babbles more consistently around 6 months of age. As early as three months of age, we see another overlap between starting to roll over and starting to produce early protophone vocalizations more consistently. It would also be interesting to explore correlations between later motor milestones (such as running or jumping with both feet off the ground) and later developing features of expressive language (including combining words into phrases two or more words in length).

Another avenue for exploration could be a more detailed, qualitative look at the IDS directed toward walking versus crawling infants. Lopez and colleagues (under review) found that an interaction between adult's labeling utterances and infant's canonical productions had a positive association with receptive vocabulary in 1-year-old infants from a nearly overlapping sample as in the present study. Additionally, an interaction between adult responses that incorporated sounds from the infant's previous vocalization (e.g., imitations) and canonical infant vocalizations were positively associated with expressive vocabulary for those same year-old infants. Given that adults respond differentially to same-aged walking and crawling infants (see Karasik, Tamis-LeMonda, & Adolph, 2013), a deeper look into the types of vocal responses could potentially reveal some very interesting social interaction patterns with walking and crawling infants.

Finally, in regard to classifying infant vocalizations, future researchers could work towards a more standardized and objective method of categorizing types of infant vocalizations. Currently, the LENA system's algorithms can differentiate between speakers and categorize infant vocalizations as cries or non-cries. Hand coding protophone vocalizations is highly subjective, but an algorithm that could specifically classify infant protophone vocalizations would be a large step towards a more objective method of classification.

3.6 Conclusion

We found similar patterns of significance within infant-adult interactions to those found in previous work by Pretzer and colleagues (in press). We also found that for walking infants, less mature infant vocalizations (i.e., Non-Canonical and Reflexive vocalizations) predicted more Infant Directed Speech (IDS) and that the more mature infant vocalizations (i.e., Canonical vocalizations) were associated with decreased Other-Directed Speech (ODS). For crawling infants, Reflexive vocalizations predicted IDS and all vocalization types predicted decreased ODS. Although we did not replicate Walle and Campos' (2014) finding of significantly higher vocabulary in walking infants than crawling infants, we did find marginally significant (p = .002) relations between IDS and Productive vocabulary for both walkers and crawlers.

We found an emerging set of results supporting a relationship between locomotion and language development. It may be no coincidence that first steps and first words emerge around the same time -- a child's first birthday. Anecdotally, infants show a "pause" in language development while they approach true first steps. Afterwards, language "catches up" and we get first words or more consistent words. The developmental systems in play focus on one or the other when on the cusp of achieving big milestones.

It is interesting to imagine how locomotor development and language development are interrelated. One might consider the physiological changes brought about by the transition from walking to crawling. The infant's upright posture and increased core strength adds respiratory strength and control resulting in greater breath support. Also, upright posture puts the speech production system in a more advantageous posturing for vocalization. Visually, upright infants have a whole new world to look at and explore (e.g., Kretch et al., 2014). This can lead to more opportunities to see things and learn about them. Parents can label objects in the infant's expanded environment, leading to vocabulary growth. Further, if children are able to more efficiently self-locomote, they have more opportunity to explore their environments, which also leads to more learning. With their hands free from crawling or cruising, infants are able to bring objects to their parents or lead their parents to things (Karasik et al., 2013). This facilitates even more opportunity to label, narrate, and describe the infant's surroundings, which supports language learning. Finally, we should consider the social implications of walking, such as joint attention and engagement (e.g., Walle, 2016). Additionally, once walking, infants can travel distances to initiate interactions with others. They are not reliant upon adults to come to interact with them, but rather can take a more proactive role in the interactions that drive language development.

Chapter 4

Multi-Domain Synchrony in Speech and Language Development

4.1 Preface

This chapter describes two explorations of subsets of data from a larger ongoing NSF funded study entitled *Infant Vocalizations as Foraging for Caregiver Responses* (IVFCR). Anne S. Warlaumont, Christopher T. Kello, and Ajay Gopinathan are the principal investigators of the IVFCR Project. The IVFCR project explores how infants traverse their acoustic space to "forage" for adult responses by analyzing daylong audio recordings of infants and their families at 3, 6, 9, and 18 months of age. During my two years as a Graduate Student Researcher on this project, I assisted with developing data collection procedures and training Research Assistants in the study procedures. One of my main contributions to the IVFCR Project was to suggest the Ages and Stages Questionnaire (ASQ), which is a parent report measure that identifies whether infants have met specific age expected milestones across developmental domains. The ASQ is a commonly used screening tool used by early interventionists (including Speech-Language Pathologists, Physical Therapists, Occupational Therapists, and Developmental Specialists), pediatricians, and other professionals who work with young children. Not only is the ASQ easy to administer (via parent report), but it is a valid and comprehensive examination of milestones across all developmental domains (e.g. Gollenberg, Lynhc, Jackson, McGuinness, & Msall). The ASQ is well known among clinicians and healthcare professionals, which could make any resulting studies utilizing this data more accessible to practitioners.

First, we examined adult responsivity to different infant vocalization types, as done before, but for infants at four time points. Segments of a subset of the recordings were hand coded as the in the two previous chapters. This longitudinal exploration revealed significant relations between infant vocalization types and adult utterance directions dependent upon age. For example, there are significant associations between IDS and the most common vocalization types at younger ages (i.e., Non-Canonical and Reflexive) and between IDS and all vocalization types for older infants (i.e., 9- and 18-month-olds). These findings will be submitted for publication once we have a full set of coded segments for each participant. Next, we explored how parent report of achieved developmental milestones related to LENA measured variables. We ran a variety of tests to search for associations between motor scores and child vocalization counts and social/emotional scores with conversational turn counts for 38 infants who had completed all 4 recordings. Key findings include that Fine Motor, Social-Emotional, and Communication scores were predictive of vocabulary size. These findings will be written up and submitted for publication once all participants have completed all their recording

4.2 Introduction

Over the last two chapters we have seen that language does not develop in a vacuum - there are other factors at play. Specifically, social and motor development are highly implicated in language development. In Chapter 2, we explored infant-adult social interactions. We investigated how infant vocal type predicted adult responsivity, and vice versa. We found that at shorter timeframes (1 or 2 s), infant vocalizations predicted subsequent IDS, and vice versa. In Chapter 3, we found emerging results that suggest that locomotor development is related to vocabulary development. For walking infants, less mature infant vocalizations (i.e., Non-Canonical and Reflexive vocalizations) predicted more Infant Directed Speech (IDS), while more mature infant vocalizations (i.e., Canonical vocalizations) were associated with decreased Other-Directed Speech (ODS). For crawling infants, however, only Reflexive vocalizations predicted IDS, while all vocalization types predicted decreased ODS. Although we could not replicate previous findings of significantly higher vocabulary in walking infants than crawling infants (Walle & Campos, 2014), we did find marginally significant (p = .002) relations between IDS and Productive vocabulary for both walking and crawling infants.

This current work is a continuation and elaboration of the two previous studies (see Chapter 2 and Chapter 3). We utilized the same hand-coding scheme used in both prior studies to analyze 5 minute segments of high infant volubility (selected with the same criteria). In fact, some of the infants whose recordings were coded in this study were also enrolled in the same Walking and Language study conducted in the Walle Interpersonal Development lab. The IVFCR project, and thus the data for this study, differ in that recordings were collected at four different ages, rather than one single recording. Further, this work includes developmental milestone information in the motor and social/emotional domains collected from a screening tool used by many clinicians and healthcare professionals. We investigated whether 1) speech-related infant vocalizations predict adult responsivity and how this differs across development; 2) ASQ developmental domains are predictive of vocabulary size; 3) ASQ measures of motor development are related to infant volubility (infant vocalization counts determined by LENA's algorithm); 4) ASQ measures of social/emotional development are related to conversational turn exchanges (conversational turn counts measured by LENA's algorithm); and 5) infant volubility predicts vocabulary size.

4.3 Methods

4.3.1 Participants

56 infants (25 female) were recruited from the San Joaquin Valley in Central California via posted flyers, community outreach events, and presentations to expecting mothers at Mercy Medical Center Merced's "Stork Tour". To take part in the study, families had to meet the following criteria: 1) Participants must enroll in study prior to age 3 months; 2) Participants must live within approximately 30-40 miles of Merced; and 3) Participants must expect to be present and living in the area for all four recordings at 3-, 6-, 9-, and 18- months of age. Participants were excluded from the present analyses if they dropped out of the study (n=9), they did not follow recording instructions (i.e.,

collecting at least 10 hours in a single day; n=1), the infant was outside of the target age range (i.e., within two weeks of turning 3, 6, 9, or 18 months old) at the date of recording (n=1), or all four recordings were not completed (due to technical error or not being returned by the family; n=1). Four infants were not included in this sample as they are yet to have completed their 18 month recording or have not had all questionnaire data entered into our records (as of March 2019). Five infants were from primarily Spanish speaking families and completed study questionnaires (e.g., ASQ, MCDI) in Spanish.

Twenty-four infants were from families with both primary caregivers having a college degree (including 20 graduate degrees), 16 were from families with one primary caregiver having a college degree (including 3 graduate degrees), and 14 were from families with neither primary caregiver having a college degree. Fifty-seven infants hear English as their primary language, seven hear Spanish, and one infant hears primarily Thai. Families reported that their infants were also exposed to American Sign Language, French, German, Japanese, and Portuguese. Thirty-two infants were reported to be of Caucasian descent, 26 infants of Hispanic descent, four of Asian descent, two of Black/African American descent, three of Native American descent, two of other descent, and one declined to answer. Families reported diagnoses of hip dysplasia, torticollis, allergic colitis, GERD, and tracheomalacia, but otherwise, there were no reported diagnoses of developmental delay or ongoing health concerns.

4.3.2 Recording Procedures

Families completed daylong recordings using the LENA Pro system at 3-, 6-, 9-, and 18- months of age. Trained researchers brought study materials to the participant's home or to a location in the community. At each drop off visit, the researchers would review recording procedures and the accompanying paperwork and answer any questions asked by the family. They also gave monetary compensation to the family at the drop off visit.

Families were instructed to pick a "typical" day for the infant to make the recording (in many cases on the weekend when the infant was home with her family) where they could collect at least 10 hours of recording in a single day. Families were recommended to start the recording ***no later*** than 8am that morning and end the recording ***no earlier*** than 7pm that evening. Recorders could be paused for privacy reasons, but families were asked to keep paused time under an hour to ensure at least 10 hours of recording.

Parents placed the LENA recorder into the pocket of a specially designed vest which was worn for the duration of the recording. Families could remove the vest during car rides, naps, or bath time, but were asked to leave it running nearby in order to capture any vocalizations produced or heard by the infant. After the recording, the vest was removed and the recorder was turned off (if it had not run down the battery and turned off automatically after 16 hours of recording).

In addition to completing the audio recording, parents were asked fill out accompanying paperwork at each visit. After every recording, parents were asked to fill out an "End of Recording Questionnaire" detailing the context of the recording day (e.g., times of recording, who was present, any difficulties with recording, etc.). Parents were also asked to complete The Ages and Stages Questionnaire (ASQ; Bricker et al., 1999) at every age recording. The ASQ is a clinical screening tool where parents report which developmental milestones their infants have achieved in the developmental domains of: Communication, Gross Motor, Fine Motor, Social-Emotional, and Problem Solving. They also completed the STIM-Q Cognitive Home Environment questionniare at ages 6-, 9-, and 18- months. The STIM-Q is an instrument developed by Dreyer, Mendelsohn, and Tamis-LeMonda (2010) to measure a family's cognitive home environment. Finally, parents completed the MacArthur-Bates Communicative Developmental Inventory: Words and Sentences on the same day as the 18-month recording (MCDI; Fenson et al., 2007).

After the recording was completed, the researchers returned for a Pick Up Visit where they collected all study materials and debriefed the participants about the recording experience. Once collected, the recordings were uploaded with LENA Pro software and all questionnaire data were input into metadata sheets for later use. Following the 18-month recording, researchers compiled reports of the number of child vocalizations produced, adult words heard, and conversational turns taken at each recording. These reports were taken to the family as a Final Visit where the family was also given a final compensation if all recordings were made as instructed and all study materials were returned.

4.4 Data Analyses

This section details the two methods of data analysis used in this present work. The first subsection presents our findings from analyzing hand coded segments from every recording for six infants. The second subsection describes how we used the information gathered from the automatic analysis of each LENA recording, ASQ, and MCDI for each infant. Prior to running any analyses, each audio recording was processed using LENA Pro's software. The software program uses an algorithm to automatically assign labels to indicate speakers (e.g., target infant) or sound sources (e.g., electronic noise) throughout the duration of the recording. The LENA software also provides counts for total infant vocalizations produced, total adult words heard, and total conversational turns.

4.4.1 Part 1 - Coding Interactions

For the first part of our data analyses, we used the same coding scheme mentioned in the two previous chapters. Please see Section 3.3.3 for a description of how we selected high infant volubility samples, Section 3.3.4 for details about the identification and classification of infant and adult vocalizations, and Section 3.3.5 for information about how we took the hand-coded infant and adult vocalizations and constructed an event series and two-event sequences for analysis. Once high-volubility segments were identified for coding, the list was randomized and listed without infant ID in order to prevent coder bias based on infant age or identity. Because the entire collection of segments had not been coded, this process resulted in an uneven number of recordings per age and per participant. Just as in Chapters 2 and 3, we ran generalized linear mixed effects models to test for associations between infant vocal type and the direction of adult.

4.4.2 Part 2 - ASQ exploration

The second part of our data analyses focused on identifying relations between the automatic counts tallied by LENA Pro software and the reported milestone achievements across domain, per the ASQ's parent report. The ASQ provides a score for each developmental domain, with scores falling into one of three categories: 1) within normal limits (white), 2) at risk (light gray), and 3) demonstrating a delay (dark gray). Scores were tallied for every domain for all four ASQs for each infant. Since the scores are not continuous across ages or domains (i.e., a score of 30 at 3 months is not equivalent to a score of 30 at 6 months and a score of 30 in communication is not equivalent to a score of 30 in problem solving), we adjusted the score by subtracting it from the cutoff score (i.e., the score at which a child is considered to demonstrate a delay in that domain.

We ran a series of linear mixed effects models to test whether 1) ASQ developmental domains are predictive of vocabulary size; 2) ASQ measures of motor development are related to infant volubility (infant vocalization counts determined by LENA's algorithm); 3) ASQ measures of social/emotional development are related to conversational turn exchanges (conversational turn counts measured by LENA's algorithm); and 4) infant volubility predicts vocabulary size. Given convergence errors with several of these models, only a portion of these results were usable and reported below. Future explorations and analyses of these data should integrate other models to address these predictions.

4.5 Results

This section reports the results from both tracks of data analysis. In subsection 4.5.1, we present descriptive information including the frequencies of all infant (i.e., canonical [C], non-canonical/non-reflexive [X], and reflexive [R] vocalizations) and adult (i.e., IDS [T] and ODS [N]) vocal events. We also report the proportion of specific infant-adult vocal sequences at each age (i.e., CT, CN, XT, XN, RT, RN). The next subsection (4.5.2) presents the results from our analyses of the hand coded segments where we explored how infant vocalizations predict adult responsiveness. Subsection 4.5.3 reports how each domain in the ASQ uniquely predicted productive vocabulary at each age recording.

4.5.1 Descriptive Statistics

We hand-coded 99 5-minute segments from 27 infant's daylong audio recordings. Infant vocalizations were classified by type and adult vocalizations were coded based on direction of utterance (see Section 3.3.4). We included 31 3-month-old recordings, 35 6-month-old recordings, 22 9-month-old recordings, and 11 18-month-old recordings. Table 4.1 reports average counts for each type of infant and adult vocalization for infants at each age (3, 6, 9, and 18 months). Table 4.2 reports proportions of each type of infant and adult vocalizations at each age. Selected segments were placed in a randomized order for coding, resulting in an uneven number of recordings for each age.

4.5.2 Infant vocalizations predict adult responsivity at all ages

Table 4.3 shows the results for how each of the three infant vocalizations types uniquely predicted that the subsequent vocal event would be IDS at a given age. As in the

Table 4.1

Infant and Adult Vocalization Counts by Age

Age	Canonical	Non-Canonical	Reflexive	IDS	ODS	Unknown
3 months	72	1925	232	786	306	22
6 months	149	1743	229	511	403	38
9 months	201	894	138	451	410	24
18 months	310	1165	237	420	412	26

Table 4.2

Proportions of Infant and Adult Vocalizations by Age

Age	Canonical	Non-Canonical	Reflexive	IDS	ODS	Unknown
3 months	72	1925	232	786	306	22
6 months	149	1743	229	511	403	38
9 months	201	894	138	451	410	24
18 months	310	1165	237	420	412	26

previous chapter, the pause duration, *P*, was set to 1 s. Given the number of planned tests, we used a Bonferroni correction to set a conservative α cutoff ($p = .001$) to indicate statistical significance. At 3-months old, Non-Canonical ($b = 0.54$, $p < .001$) and Reflexive ($b = 1.19$, $p < .001$) infant vocalizations were significant predictors of IDS, but no infant vocalizations significantly predicted subsequent ODS. First event IDS utterances were significantly predicted by second event Canonical ($b = 1.13$, $p < .001$), Non-Canonical ($b = 0.53$, $p < .001$), and Reflexive ($b = 1.04$, $p < .001$) infant vocalizations, but there were no associations between first event ODS and second event infant vocalizations of any type. We do not report any post-hoc comparisons between infant vocalizations types as these models failed to converge.

Similarly, at age 6-months, Non-Canonical ($b = 0.57$, $p < .001$) and Reflexive ($b = 1.01$, $p < .001$) infant vocalizations were significant predictors of IDS, but no infant vocalizations significantly predicted subsequent ODS. There were no significant relations between first event IDS or ODS with any second event vocalization type. By 9-months old, Canonical ($b = 0.83$, $p < .001$), Non-Canonical ($b = 0.88$, $p < .001$) and Reflexive ($b = 1.10$, $p < .001$) infant vocalizations were significant predictors of IDS, while Non-Canonical ($b = 0.66$, $p < .001$) infant vocalizations significantly predicted subsequent ODS. First event IDS was predicted by subsequent Non-Canonical infant vocalizations ($b = 0.76$, $p < .001$), but there were no significant relations between first event ODS and second event infant vocalizations of any kind.

Again at 18-months old, Canonical ($b = 1.45$, $p < .001$), Non-Canonical ($b = 0.68$, $p < .001$) and Reflexive ($b = 1.64$, $p < .001$) infant vocalizations were significant predictors of IDS, but this time both Non-Canonical ($b = 0.66$, $p < .001$) and Reflexive ($b = 0.92$, $p = .001$) infant vocalizations significantly predicted subsequent ODS. Second

Table 4.3

Infant Vocalizations Predicting Subsequent IDS and ODS

Age	First Event	Second Event	*b*	*p*
3 months	Canonical	IDS	0.62	0.10
3 months	Non-Canonical	IDS	0.54	<.001
3 months	Reflexive	IDS	1.19	<.001
3 months	Canonical	ODS	-18.26	0.30
3 months	Non-Canonical	ODS	-0.04	0.78
3 months	Reflexive	ODS	-0.73	0.21
6 months	Canonical	IDS	0.90	0.003
6 months	Non-Canonical	IDS	0.57	<.001
6 months	Reflexive	IDS	1.01	<.001
6 months	Canonical	ODS	0.27	0.47
6 months	Non-Canonical	ODS	-0.01	0.94
6 months	Reflexive	ODS	-0.41	0.22
9 months	Canonical	IDS	0.83	<.001
9 months	Non-Canonical	IDS	0.88	<.001
9 months	Reflexive	IDS	1.10	<.001
9 months	Canonical	ODS	-0.19	0.55
9 months	Non-Canonical	ODS	0.76	<.001
9 months	Reflexive	ODS	0.76	0.02
18 months	Canonical	IDS	1.45	<.001
18 months	Non-Canonical	IDS	0.68	<.001
18 months	Reflexive	IDS	1.64	<.001
18 months	Canonical	ODS	-0.14	0.61
18 months	Non-Canonical	ODS	0.66	<.001
18 months	Reflexive	ODS	0.92	0.001

event Reflexive vocalizations predicted first event IDS ($b = 0.88$, $p < .001$) and ODS ($b = 0.96$, $p < .001$), but Non-Canonical infant vocalizations only predicted first event IDS ($b = 0.61$, $p < .001$).

Table 4.4

Ages and Stages Questionnaire Scores at 18 Months Predict Productive Vocabulary at 18 Months

Developmental Domain	*b*	*p*
Communication	0.007	<.001
Gross Motor	0.003	0.002
Fine Motor	0.007	<.001
Problem Solving	-0.001	0.47
Social-Emotional	0.014	<.001

4.5.3 Multi-Domain Development Predicts Vocabulary Size

We ran a generalized linear model with a Poisson distribution to examine how an infant's score in a given developmental domain would uniquely predict vocabulary size at 18 months. We found that scores in the Fine Motor ($b = 0.007$, $p < .001$), Communication ($b = 0.007$, $p < .001$), and Social-Emotional ($b = 0.014$, $p < .001$) domains predicted higher vocabulary scores on the MCDI at 18 months old (see Table 4.4). Additional models were run to further explore these relations but failed to converge and were not included in this work.

4.6 Discussion

4.6.1 Predicting Interactions

As in Chapter 3, we used identical analyses to Pretzer and colleagues (in press) to see if their findings about patterns of interaction in infant-adult vocal exchanges would remain true for infants at four different ages. However, we did not report post-hoc analyses of comparisons between vocal types as several models failed to converge. At 3 and 6 months, Non-Canonical and Reflexive were significant, but at 9 and 18 months, all vocal types were significant predictors of IDS. This might be partially due to the fact that at earlier ages, canonical vocalizations are relatively infrequent (see Table 4.1 for frequency of vocalization types across ages).

There were no significant predictors of ODS at 3 or 6 months. Non-Canonical infant vocalizations were significant at 9 and 18 months and Reflexive vocalizations were significant at 18 months. This might indicate less ODS at younger ages; perhaps because during high volubility times, infants are more demanding of their parents' care and attention, resulting is less ODS. The results at later ages are similar to those of Pretzer and colleagues (in press).

First event IDS was predicted by any second event infant vocalization for 3-month-olds. It might appear that at 3 months of age, infants are consistently vocalizing (regardless of type) in response to IDS or this may be a case of infants vocalizing naturally in bouts with adults interspersing IDS into the infant bouts. For older infants, first event IDS and ODS were predicted by second event Non-Canonical infant vocalizations at 9 and 18 months, and Reflexive infant vocalizations at 18 months. These are similar to the 12-month-old infants in Pretzer et al. (in press).

4.6.2 Vocabulary Size is Predicted by Various Developmental Domains

We found several correlations between developmental domains and MCDI score. First, we found a significant relation between a child's Communication score and Vocabulary size. It is reassuring that we find a direct association between a vocabulary measure (i.e., the MCDI) and a tool used to predict language abilities (i.e., the ASQ). Given the precedent in prior research to treat MCDI scores as a measure of language development, this is a valuable relation to find. Next, we found that Fine Motor scores predicted MCDI scores. This result is intuitive, as speech production is one of the most intricate fine motor tasks we do, and the MCDI version administered at 18 months combines productive and receptive vocabulary. The relation between Social-Emotional abilities and vocabulary is also a natural fit, since language is inherently social and language is learned within social contexts. We found a marginal relationship (p = .002) between vocabulary and the Gross Motor domain. This is not surprising, given all the large body motor systems implicated in verbal communication (i.e., posture and stable body positioning, core strength for diaphragmatic support, etc.). It was surprising to see no significant association between Problem Solving scores (which are meant to represent cognitive development) and Vocabulary score, given that cognition and language are interconnected. Overall, these findings speak to the recruitment of other developmental domains within language development.

4.6.3 Limitations of the Present Study

While we performed two separate, in depth analyses of this data set, this work is by no means an exhaustive exploration of all the recordings and questionnaire responses collected by the IVFCR study. We only coded and examined segments from 27 of the 47 infants who completed this study, and the randomized order of segments for coding prevented us from having the same number of segments for each age or having vocalizations from all four recordings for a single infant. Thus, we used an incomplete and unbalanced dataset.

Secondly, we attempted to look at certain relations in more depth (i.e., *Motor Milestones and Vocal Production, Social-Emotional Milestones and Conversational Interactions, and Predicting Vocabulary Based on Infant Volubility*), but failure to converge rendered those models unusable. We intended to explore the relation between ASQ motor scores and LENA infant volubility, ASQ social-emotional scores and LENA turn counts, and MCDI vocabulary scores with LENA infant volubility. Given the unbalanced composition of the dataset and relatively small number of coded segments used in this analysis, it is likely that a more complete dataset would eliminate these convergence errors, allowing us to investigate these very interesting questions.

Finally, and perhaps most importantly, we must use some caution when interpreting ASQ and MCDI scores. The ASQ is an un-standardized screening tool which by definition does not give a fully comprehensive view of an infant's development in a specific domain at any given age. The ASQ provides information about achieving the most salient milestones, which allows healthcare professionals to flag children who are at risk or demonstrating signs of developmental delays. For the purposes of our study, information about achieving the most common milestones in a given developmental domain gives us sufficient information to explore multi-domain interactions within language development. Additionally, the ASQ and MCDI are both completed via parent report. There is documented support for the use of the ASQ (e.g., Schonhaut, Armijo, Schönstedt, Alvarez, & Cordero, 2013) and the MCDI (Fenson et al., 2007; Frank et al., 2017), as valid and reliable tools. Both instruments are structured to account for parental biases caused by over or under reporting their children's abilities on these types of measures, which supports our use of the ASQ and MCDI was a valid means of collecting important developmental data when circumstances did not allow for full developmental evaluations by clinical specialists.

4.6.4 Future Directions for Research

This study scratches the surface of what we can learn from such a deep dataset. First, and most obviously, the collection of recordings were intended for various acoustic analyses and explorations of the vocal foraging hypothesis. This work has been started by VPS and colleagues (in preparation).

However, there are still many avenues for exploration using our methodologies. We continue to hand-code segments from these LENA recordings, which will allow us to better understand the differences in interactions at various ages. Second, we could expand our use of the parent report measures collected with each LENA recording. We could explore the other domains of the ASQ (i.e., Communication and Problem Solving) in relation to child vocalization counts, adult words heard, and conversational turns as reported by the LENA software. There is also another parent report tool, the STIM-Q, that was not included in the present work. It would be very interesting to see how a standardized measure of home environment compares what we find from daylong home audio recordings.

Third, 19 infants from the IVFCR project also participated in the Walking and Language study at 12 months old (see sections 3.3.1 and 3.3.2 for more information about the Walking and Language study). Given that those recordings were collected in a nearly identical manner as the IVFCR recordings, we could potentially use those as a fifth data point for those infants, giving us recordings at 3-, 6-, 9-, 12- and 18- months of age to hand code selections from to get an even more in-depth longitudinal picture of how infant-adult interactions change as the infant ages. Although the Walking and Language study did not use the ASQ, participants filled out the MCDI, which could allow us to look at vocabulary growth from 12- to 18- months old.

Fourth, the IVFCR project included a clinical screening tool that was selected with the intention of combining a primarily research-oriented methodology (i.e., using LENA recordings) with a primarily clinically-oriented measure that is familiar to practitioners. There is great potential for sharing these results, particularly from a more complete, later iteration of this project, with practitioners who work directly with infants and their families.

Fifth, a number of our models failed to converge and were thus not included in this present work. In the future, these analyses should be redesigned to better fit the data in order to test the hypothesis we were unable to address in this current work.

Finally, we intend to share all hand-coded data, recordings, and questionnaire information from participants who gave consent with HomeBank (VanDam et al., 2017). HomeBank is an online repository where researchers who use LENA can store and share their data with other researchers. Currently, the data for most participants who approved sharing of their data are available on HomeBank (Warlaumont et al., 2016). Further, the transcripts of these hand-coded segments and scripts used for analyses will be archived and available to share with other researchers. It is vitally important to embrace the idea of open science and data sharing whenever possible.

4.7 Conclusion

This study provides novel insight to the multi-dimensionality of language development over time. We found that communication develops in conjunction with motor and social-emotional skills. As we found above, these domains are intrinsically linked and should be considered together when thinking about language development. Further, we are the first to combine the ASQ with LENA recordings. As stated before, this tool was selected with the intention to make this area of research more accessible to clinicians and other healthcare professionals. Finally, these results can inform not only clinical practices, but the way we think about language development in our research.

Chapter 5
Discussion

Language acquisition is a dynamic process integrating cognition, social interactions, and the motor coordination and sequencing of articulatory structures. At the outset of this dissertation, I presented the idea that language development, specifically infant vocal development, is a process recruiting multiple developmental domains. While researchers have started looking at multi-domain interactions, we have only begun to address the synchrony between domains during language development. Further still, there are vital connections between the clinical world and research world that should not be ignored.

This dissertation explored these ideas using a combination of approaches used by researchers and clinicians. Chapter 2 presented a new human coding method to look at infant-adult interactions in their ultra-naturalistic home environments. Chapter 3 continued this hand-coding scheme while incorporating a new input variable -- locomotion -- and an outcome variable in the MCDI. Chapter 4 further advanced these ideas while adding in a social component. This chapter also incorporated a clinical screening tool, the Ages and Stages Questionnaire (ASQ), to look for more specific indicators of motor and social-emotional development. Here, I will summarize the most important observations and results from each chapter.

5.1 Introducing a Human Coding Methodology to Observe Infant-Adult Interactions in the Wild

In Chapter 2, we introduced a human coding scheme to look at how infants and adults interact at home by examining 5-minute segments with high infant volubility from daylong LENA audio recordings. Trained coders hand-labeled infant vocalizations according to type (i.e., Canonical, Non-Canonical/Non-Reflexive, and Reflexive) and adult vocalizations by direction of speech (i.e., IDS or ODS). Previously, researchers using LENA recordings have relied heavily upon automatically labeled vocal events. The LENA Pro system's algorithm automatically identifies and classifies each vocalization with a mutually exclusive code designating the speaker (e.g., target infant, nearby female, nearby male, etc.) or source of environmental sound (e.g., electronic noise, silence, etc.). However, I wanted more in depth information than the LENA labels could provide. Given my strong interest in prelinguistic development, it was important that the target child's vocalizations were labeled with more detail than simply "cry" vs. "non-cry", leading to the Canonical, Non-Canonical, and Reflexive designations used in this dissertation. Although LENA's labels can differentiate between male and female speakers, I was more interested in to whom adult vocalizations were addressed than the gender of the speaker. Rather than relying on acoustic analyses to identify whether an adult vocalization was directed to a child or not, coders relied upon context, semantic information, and other prosodic features of IDS (e.g., utterance length, inflection, etc.) to classify an adult vocalization as infant-directed or other-directed

We employed the same coding scheme in Chapter 3 and Chapter 4. However, hand-coding the segments of recordings from younger infant was quantitatively and qualitatively different (e.g., few canonical vocalizations in 3- and 6- month-old infants, true words produced by 18-month-old infants) and subjectively more difficult (i.e., many recordings had [sometimes multiple] other young children present during the recording). However, across all three studies, we found evidence that adult responsivity is directly related to infant vocalizations. In younger (i.e., 3- and 6-month-old infants), more mature vocalizations receive higher rates of IDS, while at older ages (i.e., 9, 12, and 18 months), all infant vocalizations predict IDS.

5.2 Locomotor Status Relates to Interactions and Vocabulary Development

In Chapter 3, we built upon the results of Chapter 2 and added a new dimension for analysis. When thinking about motor development, walking emerges around the same time as first words. Therefore, we looked at locomotion as a predictor for vocabulary development. Although we were unable to replicate previous findings by Walle and Campos (2014) or He and colleagues (2015), we did identify patterns of marginal (p = .002) significance between IDS and Productive vocabulary.

In early intervention, we often see co-morbid motor and communication disorders. In some cases, the impact on communication is motoric in nature (e.g., Cerebral Palsy), while in others cases, the impact appears to be cognitive-based but concomitant to a motor disorder. Childhood Apraxia of Speech, which is a deficit in motor planning and execution for verbal expression, is a well known example of a motor-speech disorder. However, even articulation and phonological disorders have some level of motor involvement, whether it is due to weakness, restricted range of motion, discoordination of articulators, or structural abnormalities that impair movement for speech production. Vocabulary is a measurable facet of language development which allows us to get some idea of how the social interaction and motor pieces work together to support language learning.

5.3 Multi-domain recruitment for language development

In Chapter 4, we continued to build upon the results of Chapters 2 and 3, by looking at motor development as a whole, rather than just locomotor status. We also included social-emotional development as another area of exploration. This study is especially unique and exciting because of the Ages and Stages Questionnaire (ASQ). This is a screening tool used by many pediatricians and early interventionists. The ASQ is something that many SLPs are familiar with - especially in my line of work. Using the ASQ gave us access to information about the most salient milestones in motor and social-development. This makes sense because communication, when broken down far enough, is a combination of a social activity and motor processes. Using the ASQ, MCDI, and LENA Pro system is a novel way to integrate clinical and research oriented tools to look at communicative. We found similar patterns of infant-adult interactions in the 9 and 18 month old infants as did Pretzer and colleagues (in press) with 12 month old infants (also see Chapter 2). Additionally, we found that Fine Motor, Gross Motor, and Communication scores from the ASQ were each predictive of vocabulary size. This finding may be of particular interest to clinicians and other professionals who work with young children who have motor or communication disorders.

5.4 Conclusion

This dissertation focused on the synchronous development between motor, social, and language domains in infants who are learning to communicate. I have presented three studies that explored the idea that since speech is a motor act learned within social interactions, we should look at how motor and social development line up with language development. This dissertation also introduced a human coding methodology used to observe infant-adult interactions both at a single time point and longitudinally. Finally, this dissertation combined a well-known clinical screening tool (ASQ) with research oriented tools (LENA, MCDI). Future work should continue to investigate multi-domain synchronous development as it relates to infant vocalizations. Additionally, future research should be structured with the intention to share findings not only with other researchers (e.g., via the Open Science Foundation, HomeBank, etc.) but with clinical professionals as well, in order to inform our knowledge in the most complete way possible.

References

Adamson, L. B., & Bakeman, R. (2006). Development of displaced speech in early mother–child conversations. *Child Development*, *77*(1), 186-200.

Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of child language*, *18*(1), 41-49.

Adolph, K. E., & Berger, S. E. (2005). Physical and motor development. *Developmental Science: An advanced textbook, 5,* 223-281.

Adolph, K. E., Berger, S. E., & Leo, A. J. (2011). Developmental continuity? Crawling, cruising, and walking. *Developmental Science, 14*(2), 306-318.

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development*, *72*(2), 416-430.

Bakeman, R. (1997). *Observing interaction: an introduction to sequential analysis* (2nd ed.). New York: Cambridge University Press.

Baldwin, D. & Meyer, M. (2007). How inherently social is language?. *Blackwell handbook of language development, 87-106.*

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2018). lme4: Linear mixed-effects models using 'Eigen' and S4 [R Package].

Bayley, N. (1936). The development of motor abilities during the first three years: A study of sixty-one infants tested repeatedly. *Monographs of the Society for Research in Child Development,1*(1), 1-26.

Bayley, N. (1969). Scales of infant development. *Manual. NY: The Psychological Corporation.*

Bishop, D. V. M. (2002). Motor immaturity and specific speech and language impairment: Evidence for a common genetic basis. *American Journal of medical Genetics (Neuropsychiatric Genetics), 114,* 56-63.

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2018). Day by day, hour by hour: Naturalistic language input to infants. *Developmental science*, e12715.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language, 14*(02), 211-227. doi:10.1017/S0305000900012897

Bloom, K. (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language, 15*(03), 469-480.

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 11 November 2018 from http://www.praat.org/

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. D. (2015). Mother-infant contingent vocalizations in 11 countries. *Psychological Science*, *26*(8) 1272-1284.

Bringen, Z., Emde, R. N., Campos, J. J., & Appelbaum, M. I. (1995). Affective reorganization in the infant, the mother, and the dyad: The role of upright locomotion and its timing. *Child development, 66*(2), 499-514.

Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across Western and traditional cultures. *Journal of Cognition and Development*, *1*, 31–43.

Buder, E. H., Warlaumont, A. S., & Oller, D. K. (2013). An acoustic phonetic catalog of prespeech vocalizations from a developmental perspective. *Comprehensive perspectives on child speech development and disorders: Pathways from linguistic theory to clinical practice,* 103-134.

Clearfield, M. W. (2011). Learning to walk changes infants' social interactions. *Infant Behavior and Development, 34,* 15-25.

Cristia, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *The Journal of the Acoustical Society of America, 129,* 3271-3280.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development, 71*(1), 44-65.

Dreyer, B. P., Mendelsohn, A. L., & Tamis-LeMonda, C. S. (2010). Stim-Q Cognitive Home Environment. *New York University Web site*.

ELAN [Computer software] (2018). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese?. *Behavioral and Brain Sciences, 27*(4), 491-503.

Fausey, C. M., Jayaraman, S., & Smith, L. B. (2016). From faces to hands: Changing visual input in the first two years. *Cognition*, *152*, 101-107.

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.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). MacArthur-Bates Communicative Development Inventories: User's guide and technical manual (2nd ed.). Baltimore, MD: Brookes.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, *21*(1), 95-116.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant behavior and development*, *8*(2), 181-195.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message?. *Child Development*, 1497-1510.

Fernald, A. (1995). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391-428). New York, NY: Oxford University Press.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant behavior and development*, *10*(3), 279-293.

Fernald, A., Taeschner, T., Dunn, J., & Papousek, M. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501.

Field, T. (1980). Supplemental stimulation of preterm neonates. *Early Human Development*, *4*(3), 301-314.

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations?. *Infancy*, *9*(3), 327-339.

Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., … R-Core (2018). car: Companion to Applied Regression [R Package].

Franklin, B., Warlaumont, A. S., Messinger, D., Bene, E., Nathani Iyer, S., Lee, C.-C., Lambert, B., & Oller, D. K. (2014). *Language Learning and Development*, *10*(3), 279-296.

Gogate, L. J., Maganti, M., & Laing, K. B. (2013). Maternal naming of object wholes versus parts to preverbal infants: A fine-grained analysis of scaffolding at 6–8 months. *Infant Behavior and Development*, *36*(3), 470-479.

Goldin-Meadow, S., Goodrich, W., Sauer, E., & Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Developmental science*, *10*(6), 778-785.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences, 100*(13), 8030-8035.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19,* 515-522.

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: Five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Development*, *80*(3), 636-644.

Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010). Learning while babbling: Prelinguistic object-directed vocalizations indicate a readiness to learn. *Infancy*, *15*(4), 362-391.

Golinkoff, R. M. (1986). 'I beg your pardon?': The preverbal negotiation of failed messages. *Journal of child language*, *13*(3), 455-476.

Golinkoff, R. M. & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language, 22*, 703-726.

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to me: The social contexts of IDS and its effects on early language acquisition. *Current Directions in Psychological Science, 24*(5), 339-344.

Gros-Louis, J., & Miller, J. L. (2018). From 'ah'to 'bah': social feedback loops for speech sounds at key points of developmental transition. *Journal of child language*, *45*(3), 807-825.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development, 30*(6), 509-516.

Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy, 19*(4), 385-408.

Harbison, A. L., Woynaroski, T. G., Wade, J. W., Warlaumont, A. S., & Yoder, P. J. (2018). A new measure of child vocal reciprocity in children with autism spectrum disorder. *Autism Research*. doi: 10.1002/aur.1942

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Brookes Publishing.

Henning, A., & Striano, T. (2011). Infant and maternal sensitivity to interpersonal timing. *Child Development*, *82*(3), 916-931.

Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: A longitudinal study of mother-infant interaction. *Frontiers in Psychology*, 1492. doi: 10.3389/fpsyg.2015.01492

Hill, E. L. (1998). A dyspraxic deficit in specific language impairment and developmental coordination disorder? *Developmental Medicine & Child Neurology.*

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development, 66*(2), 1-149.

Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2011). Transition from crawling to waking and infants' actions with objects and people. *Child Development, 82*(4), 1199-1209.

Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2013). Crawling and walking infants elicit different verbal responses from mothers. *Developmental Science,* 1-8.

Keller, H., Lohaus, A., Völker, S., Cappenberg, M., & Chasiotis, A. (1999). Temporal contingency as an independent component of parenting behavior. *Child Development*, *70*(2), 474-485.

Ko, E.S., Seidl, A., Cristia, A., Reimchen, M., & Soderstrom, M. (2016). Entrainment of prosody in the interaction of mothers with their young children. *Journal of Child Language*, *43*(2), 284–309.

Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In *Precursors of early speech* (pp. 37-50). Palgrave Macmillian UK.

Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2013). Crawling and walking infants see the world differently. *Child Development,* 1-16.

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., ... & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*(5326), 684-686.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9096-9101.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest: Tests in Linear Mixed Effects Models [R Package].

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. doi:10.2307/2529310

Lee, C.C., Jhang, Y., Relyea, G., Chen, L. M., & Oller, D. K. (2018). Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings. *Infant Behavior and Development, 50,* 140-153.

Lee, R., Skinner, A., Bornstein, M. H., Radford, A. N., Campbell, A., Graham, K., & Pearson, R. M. (2017). Through babies' eyes: Practical and theoretical considerations of using wearable technology to measure parent-infant behavior from the mothers' and infants' view points. *Infant Behavior and Development, 47*, 62-71.

Lloyd, B. P., Yoder, P. J., Tapp, J., & Staubitz, J. L. (2015). The relative accuracy and interpretability of five sequential analysis methods: A simulation study. *Behavior Research Methods*, *48*, 1482–1491.

Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult- directed speech. *Language Learning and Development, 7,* 209-225.

Masur, E. F. (1987). Imitative interchanges in a social context: Mother-infant matching behavior at the beginning of the second year. *Merrill-Palmer Quarterly (33)*4, 453-472.

Matatyaho, D. J., & Gogate, L. J. (2008). Type of maternal object motion during synchronous naming predicts preverbal infants' learning of word–object relations. *Infancy*, *13*(2), 172-184.

McCune L., & Vihman, M. M. (2001). Early phonetic and lexical development. *Journal of Speech, Language, and Hearing Research, 44,* 670-684.

Nathani, S., & Stark, R. E. (1996). Can conditioning procedures yield representative infant vocalizations in the laboratory? *First Language, 16*(48), 365-387.

Ochs, E., & Schieffelin, B. B. (1984). "Language acquisition and socialization: Three developmental stories and their implications." In R. A. Shweder & R. A. LeVine (Eds.), *Culture theory: Essays on mind, self, and emotion*, (pp. 276–320). Cambridge University Press.

Ojemann, G. A. (1984). Common cortical and thalamic mechanisms for language and motor functions. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 246*(6), R901-R903.

Oller, D. K. (1980). The emergence of speech sounds in infancy. *Child phonology, 1,* 93-112.

Oller, D. K. (2000). The emergence of the speech capacity. Mahwah, NJ: Erlbaum.

Oller, D. K. (2010). All-day recordings to investigate vocabulary development: A case study of a trilingual toddler. *Communication Disorders Quarterly*, *31*(4), 213-222.

Oller, D. K., Eilers, R. E., Urbano, R., & Cobo-Lewis, A. B. (1997). Development of precursors to speech in infants exposed to two languages. *Journal of child language*, *24*(02), 407-425.

Oller D. K., Niyogi, P., Gray, S., Richards, J. A., Gikerson, J., Xu, D., Yapanel, U., and Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America, 107*(30), 13354-13359.

Olson, J., & Masur, E. F. (2011). Infants' gestures influence mothers' provision of object, action and internal state labels. *Journal of Child Language, 38*(5), 1028-1054.

Papousek, M., & Papousek, H. (1989). Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language, 9*(6), 137-157.

Patten, E., Labban, J. D., Casenhiser, D. M., & Cotton, C. L. (2016). Synchrony Detection of Linguistic Stimuli in the Presence of Faces: Neuropsychological Implications for Language Development in ASD. *Developmental neuropsychology*, *41*(5-8), 362-374.

Pelaéz, M., Virues-Ortega, J., & Gewirtz, J. L. (2011). Reinforcement of vocalizations through contingent vocal imitation. *Journal of Applied Behavior Analysis*, *44*(1), 33-40.

Pretzer, G. M., Lopez, L. D., Walle, E. A., & Warlaumont, A. S. (in press). Infant-Adult Vocal Interaction Dynamics Depend on Infant Vocal Type, Child-Directedness of Adult Speech, and Timeframe. *Infant Behavior and Development.*

Pretzer, G., Warlaumont, A., Lopez, L., & Walle, E. (2018). Infant and adult vocalizations in home audio recordings. doi: 10.6084/m9.figshare.6108392.v1

Porritt, L. L., Zinser, M. C., Bachorowski, J. A., & Kaplan, P. S. (2014). Depression diagnoses and fundamental frequency-based acoustic cues in maternal infant-directed speech. *Language Learning and Development*, *10*(1), 51-67.

Ramírez-Esparza, N., Garcia-Sierra, A., & Kuhl, P. K. (2014). Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science, 17*(6), 880-891.

Reddy, V., Liebal, K., Hicks, K., Jonnalagadda, S., & Chintalapuri, B. (2013). The emergent practice of infant compliance: An exploration in two cultures. *Developmental Psychology*, *49*(9), 1754.

Rollins, P. R. (2003). Caregivers' contingent comments to 9-month-old infants: Relationships with later language. *Applied Psycholinguistics*, *24*(2), 221-234.

Roug, L., Landberg, I., & Lundberg, L.-J. (1989). Phonetic development in early infancy: a study of four Swedish children during the first eighteen months of life. *Journal of Child Language*, *16*(1), 19–40. doi: 10.1017/S0305000900013416

R Core Team (2018). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org.

Schonhaut, L., Armijo, I., Schönstedt, M., Alvarez, J., & Cordero, M. (2013). Validity of the ages and stages questionnaires in term and preterm infants. *Pediatrics*, *131*(5), e1468-e1474.

Schuller, B., Steidl, S., Batliner, A., Bergelson, E., Kajewski, J., Janott, C., Amatuni, A., Casillas, M., Seidl, A., Soderstrom, M., Warlaumont, A., Hidalgo, G., Schnieder, S., Heiser, C., Hohenhorst, W., Herzog, M., Schmitt, M., Zian, K., Zhang, Y., Trigeorgis, G., Tzirakis, P., & Zafeiriou, S. (2017). The INTERSPEECH 2017 computational paralinguistics challenge: Addressee, cold, & snoring. *Proceedings of INTERSPEECH 2017* (pp. 3442-3446). doi:10.21437/Interspeech.2017-43

Shneidman, L. A., Arroyo, M. E., Levine, S., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language, 40*, 672-686.

Song, J. Y., Demuth, K., & Morgan, J. L. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America, 128,* 389-400.

Stark, R. E. (1980). Stages of speech development in the first year of life. *Child phonology, 1,* 73-90.

Stoel-Gammon, C. (1989). Prespeech and early speech development of two late talkers. *First Language, 9,* 207-224.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72*(3), 748-767.

Todd, G. A., & Palmer, B. (1968). Social reinforcement of infant babbling. *Child Development*, 591-596.

VanDam, M., Warlaumont, A. S., Bergelson, E., Cristia, A., Soderstrom, M., De Palma, P., & MacWhinney, B. (2016). HomeBank: An Online Repository of Day-long Child-Centered Audio Recordings. *Seminars in Speech and Language*, *37*(02), 128-142.

Van Egeren, L. A., Barratt, M. S., & Roach, M. A. (2001). Mother-infant responsiveness: Timing, mutual regulation, and interactional context. *Developmental Psychology*, *37*(5), 684-697.

Vihman M. M., Macken, M. A., Miller, M., Simmons, H., & Miller, J. (1985). From babbling to speech: A re-assessment of the continuity issue. *Language, 61*(2), 297-445.

Viholainen, H., Ahonen, T., Cantell, M., Lyytinen, P., & Lyytinen, H. (2002). Development of early motor skills and language in children at risk for familial dyslexia. *Developmental Medicine & Child Neurology, 44*(11), 761-769.

Walle, E. A. (2016). Infant social development across the transition from crawling to walking. *Frontiers in psychology*, *7*, 960.

Walle, E. A. & Campos J. J. (2014). Infant language is related to the acquisition of walking. *Developmental Psychology, 50,* 336-348.

Walle, E., & Warlaumont, A. S. (2015). Infant Locomotion, the Language Environment, and Language Development: A Home Observation Study. In *CogSci*.

Warlaumont, A. S., Pretzer, G. M., Walle, E. A., Mendoza, S., & Lopez, L. D. (2016). Warlaumont HomeBank Corpus. doi:10.21415/T54S3C.

Warlaumont, A. S., & Ramsdell-Hudock, H. L. (2016). Detection of total syllables and canonical syllables in infant vocalizations. *Proceedings of INTERSPEECH 2016,* (pp. 2676-2680). doi: 10.21437/Interspeech.2016-1518

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science, 25*(7), 1314-1324.

Watson-Gegeo, K. A., & Gegeo, D. W. (1986). Calling-out and repeating routines in Kwara'ae children's language socialization. *Language socialization across cultures*, *3*, 17-50.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experiences strengthens processing and builds vocabulary. *Psychological Science, 24*(11), 2143-2152.

West, K. L., Leezenbaum, N. B., Northrup, J. B., & Iverson, J. M. (2017). The relation between walking and language in infant siblings of children with autism spectrum disorder. Child development.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006) (pp. 1556-1559).

Xu, T. L., Abney, D. H., & Yu, C. (2017). Discovering multicausality in the development of coordinated behavior. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 1369-1374).

Yoder, P. J, Davies, B., & Bishop, K. (1994). Reciprocal sequential relations in conversations between parents and children with developmental delays. *Journal of Early Intervention*, *18*(4), 362-379.

Yoder, P. J., & Symons, F. J. (2010). *Observational measurement of behavior.* New York: Springer Pub.

Yoder, P. J., & Warren, S. F. (1999). Maternal responsivity mediates the relationship between prelinguistic intentional communication and later language. *Journal of Early Intervention*, *22*(2), 126-136.

Yoder, P. J., & Warren, S. F. (2002). Effects of prelinguistic milieu teaching and parent responsivity education on dyads involving children with intellectual disabilities. *Journal of Speech, Language, and Hearing Research*, *45*(6), 1158-1174.

Zukow-Goldring, P. (1996). Sensitive caregiving fosters the comprehension of speech: When gestures speak louder than words. *Early Development and Parenting: An International Journal of Research and Practice*, *5*(4), 195-211.