Title
Neural Connectivity in Infants at Familial Risk for Autism Spectrum Disorder

Author
Tran, Xuan Amelia

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Neural Connectivity in Infants at Familial Risk for Autism Spectrum Disorder

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Neuroscience

by

Xuan A. Tran

2019

ABSTRACT OF THE DISSERTATION

Neural Connectivity in Infants at Familial Risk for Autism

by

Xuan A. Tran

Doctor of Philosophy in Neuroscience

University of California, Los Angeles, 2019

Professor Shafali Spurling Jeste, Chair

While autism spectrum disorder (ASD) is diagnosed based on behavioral symptoms at 3 years of age, the infant sibling study design has enabled the detection and characterization of atypical neural development during the first year of life, prior to the emergence of behavioral symptoms. Infants who have older siblings with ASD are at increased risk for ASD, language delay, and other neurodevelopmental delays. As such, it is important to identify as early as possible if an infant is on a trajectory towards atypical development in order to help guide close monitoring and implement targeted behavioral interventions. The body of work in this dissertation contributes to the field of infant sibling research by showing that with robust methods, electroencephalography (EEG) can be used to detect altered functional connectivity during the first year of life, starting as early as 3 months of age. Chapter 1 introduces known deficits in behaviors and neural connectivity in infants at risk for ASD, highlights methodological gaps in the field of EEG infant research, and outlines the goals of this dissertation. Chapter 2

addresses methodological considerations in the development of an EEG pre-processing pipeline, designed to maximize data quality and data retention for infant EEG. Chapters 3 through 5 present different aspects of a comprehensive study of functional connectivity during language processing in infants at risk for ASD, with focus on theta (4-6 Hz) and alpha (6-12 Hz) spectral power and phase coherence within putative language networks. Chapter 3 describes differences in coherence at 3-months of age in infants who show ASD symptoms at 18-months of age. Chapter 4 highlights altered trajectories in coherence development over the first year of life in infants who later have ASD symptoms at 18-months. At the same left fronto-central network that differentiated risk groups at 3-months of age, reduced average coherence over the first year of life is maintained in infants who showed ASD symptoms at 18 months. Chapter 5 characterizes connectivity as an endophenotype of ASD in familial risk infants using both the 3-month cross-sectional study design and the 3-12-month longitudinal study design. Connectivity measures that differentiate risk groups in Chapters 3-5 also relate to language ability and ASD symptoms at 18-months of age. Taken together, the body of work in this dissertation support the hypothesis that early differences in neural connectivity lay a foundation for and precede behavioral signs of neurodevelopmental disabilities in infants at risk for ASD.

The dissertation of Xuan A. Tran is approved.

Leslie Carver

Mirella Dapretto

Scott Makeig

Damla Senturk

Shafali Spurling Jeste, Committee Chair

University of California, Los Angeles

2019

DEDICATION

This dissertation is dedicated to the memory of my father, Quang Tran-Tan, who inspired me to pursue both science and medicine.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

ACKNOWLEDGEMENTS

Chapter 3 is a version of: Tran, X., Dickinson, A., McDonald, N., Scheffler, A., Frohlich, J., Marin, A., Kure Liu, C., Nosco, E., Senturk, D., Dapretto, M., Jeste, S.S. Functional connectivity during language processing in 3-month-olds infants at familial risk for ASD. Manuscript in preparation.

Chapter 4 is a version of: Tran, X., Scheffler, A., Dickinson, A., McDonald, N., Marin, A., Nosco, E., Senturk, D., Dapretto, M., Jeste, S.S. Longitudinal development of functional connectivity during language processing over the first year of life in infants at familial risk for ASD. Manuscript in preparation.

This work was supported by Autism Speaks Weatherstone Predoctoral Fellowship, the Achievement Rewards for College Scientists Foundation Award, the UCLA David Geffen School of Medicine Dean's Office, the National Institute of Child Health and Human Development grant P50 HD055784-08, the National Institute of General Medical Sciences grant R01 GM111378-01A1, and the National Institute of General Medical Sciences grant T32 GM008042.

I am thankful for all the children and families who participated in this research. This work would not have been possible without the unwavering support of my primary mentor, Dr. Shafali Jeste, and colleagues from the Jeste lab, the Dapretto lab, the Senturk lab, and the Swartz Center for Computational Neuroscience. I would like to thank members of my dissertation committee, Dr. Leslie Carver, Dr. Mirella Dapretto, Dr. Scott Makeig, and Dr. Damla Senturk, for their time and guidance. I thank Dr. Carlos Portera-Cailliau for his mentoring me throughout the UCLA-Caltech Medical Scientist Training Program. I thank Dr. Joel Frohlich, Dr. Makoto Miyakoshi, Dr. Abigail

Dickinson, Dr. Iman Mohammad-Rezazadeh, Dr. Sandy Loo, and Dr. Agatha Lenartowicz for mentoring me in EEG signal processing. I thank Dr. Aaron Scheffler for his support in statistical analyses. I would also like to thank Andrew Marin, Elizabeth Baker, Carolyn Ponting, Lisa Jackson, Erin Nosco, Dr. Nicole McDonald, and Dr. Amaya Miquelajauregui for their assistance in EEG and behavioral data collection. I thank my current and former labmates Emily Pompan, Meredith Ware, and Manjari Daniel for their help with literature review. I thank my friends in graduate school who have provided me with boundless social and professional support, namely Katherine Lawrence, Vidya Saravanapandian, Christine Olson, Dr. Cynthia He and Dr. Hua Chai. Most importantly I thank my family for their unconditional support of my personal and professional growth.

VITA

Academic Degrees

2012 B.S. in Neurobiology
University of California, Irvine
Irvine, CA

2012 B.S. in Chemistry
University of California, Irvine
Irvine, CA

Academic Positions

2017 Teaching Assistant
University of California, Los Angeles
Los Angeles, CA

Course: Behavioral and Cognitive Neuroscience
Professors: Kate Wassum, David Glanzman, Hakwan Lau

Course: Introduction to Functional Anatomy of Central Nervous
Professors: Mayumi Prins, Janel Le Belle

Fellowship and Award

2016 – 2019 Achievement Rewards for College Scientist Foundation Award
2017 – 2019 Autism Speaks Dennis Weatherstone Predoctoral Fellowship

Publications

Shahrestani P, **Tran X**, and Mueller L. (2012). Patterns of male fitness conform to predictions of evolutionary models of late-life. *Journal of Evolutionary Biology* 25(6):1060-1065. PMID: 22487207.

Shahrestani P, **Tran X**, and Mueller L. (2012). Physiology declines prior to death in Drosophila melanogaster. *Biogerontology* 13(5):537-545. PMID: 22960750.

Stewart JC, **Tran X**, and Cramer SC. (2014). Variability in performance of motor action selection task among older adults is explained by differences in brain function and structure. *Neuroimage* 86: 326-34. PMID: 24125791 - PMCID: PMC3896569.

Dodakian L, McKenzie AL, Le V, See J, Pearson-Fuhrhop K, Burke Quinlan E, Zhou RJ, Augsberger R, **Tran XA**, Friedman N, Reinkensmeyer DJ, and Cramer SC. (2017). A home-based telerehabilitation program for patients with stroke. *Neurorehabilitation and Neural Repair* 31(10-11):923-933. PMID: 29072556.

Abstracts

Stewart JC, **Tran X**, and Cramer SC. (2012). Neural correlates of motor action selection in older adults. Abstract for poster presentation; presented at the 42nd annual meeting of the Society for Neuroscience, New Orleans, LA.

Stewart JC, **Tran X**, and Cramer SC. (2013). Two distinct patterns of brain activation during motor action selection in older adults. Abstract for poster presentation; presented at the 23rd annual meeting of the Society for the Neural Control of Movement, San Juan, Puerto Rico.

Tran XA, Baker E, Frohlich J, Huberty S, Jeste SS. (2016). Baseline EEG recordings in young children with ASD: stimulus type matters. Abstract for poster presentation; presented at the 16th annual meeting of the International Society for Autism Research, Baltimore, MD.

Tran XA, Jeste SS, DiStefano C. (2016). Resting state EEG in young children with ASD pre- and post-intervention. Abstract for poster presentation; presented at the 46th annual meeting for the Society for Neuroscience, San Diego, CA.

Tran XA, Miquelajauregui A, Frohlich J, Jeste SS. (2017). Resting gamma power predicts language ability in infants at risk for ASD. Abstract for poster presentation; presented at the 17th annual meeting of the International Society for Autism Research, San Francisco, CA.

Tran XA, Kure Liu C, Dickinson AH, Dapretto M, Jeste SS. (2017). Functional connectivity predicts language ability in infants at risk for ASD. Abstract for poster presentation; presented at the 47th annual meeting for the Society for Neuroscience, Washington DC.

Tran XA, McDonald NM, Dickinson AH, Frohlich J, Dapretto M, Jeste SS. (2018). Functional connectivity during language processing in infants at familial risk for ASD. Abstract for oral presentation; presented at the 18th annual meeting of the International Society for Autism Research, Rotterdam, the Netherlands.

Tran XA, McDonald NM, Dickinson AH, Dapretto M, Jeste SS. (2018). Functional connectivity during language processing in infants at familial risk for ASD. Abstract for poster presentation; presented at the 48th annual meeting of the Society for Neuroscience, San Diego, CA.

Tran XA, Dickinson AH, McDonald NM, Dapretto M, Jeste SS. (2019). Functional connectivity during language processing in infants at familial risk for ASD. Abstract for panel presentation; presented at the 19th annual meeting of the International Society for Autism Research, Montreal, Canada.

Chapter 1: Introduction

1.1 Genetic basis of ASD

Autism spectrum disorder (ASD) is a prevalent neurodevelopmental disorder defined behaviorally based on impairments in social communication skills and the presence of restricted and repetitive patterns of behavior and interests (American Psychiatric Association, 2013). ASD has also been coined as “developmental disconnection syndromes” (Geschwind 2007). From genetics studies of ASD, we know that it is a condition characterized by disrupted connectivity. While defined behaviorally, a third of ASD cases have a genetic etiology. Upward to 10-20% of ASD cases are caused by single gene mutations associated with known genetic syndromes (Berg & Geschwind, 2012; Abrahams & Geschwind, 2008). During prenatal brain development, ASD genes converge to disrupt connectivity: ranging from disrupted synaptic connections at the molecular and cellular levels, to disrupted cortical and subcortical networks at the circuit level (de la Torre-Ubieta *et al.*, 2016; Chen *et al.*, 2015; Ebrahimi-Fakhari & Sahin, 2015; Berkel *et al.*, 2010; Durand *et al.*, 2007; Gilman *et al.*, 2011; Guilmatre *et al.*, 2009). At the molecular level, genetic mutations associated with ASD can affect protein translation, intracellular signaling, and synaptic signaling (synaptic adhesion, synaptic scaffolding) (Durand *et al.*, 2007; Berkel *et al.*, 2010; Gilman *et al.*, 2011; Ebrahimi-Fakhari & Sahin, 2015; Chen *et al.*, 2015). At the cellular level, genes implicated in ASD can alter neurogenesis and differentiation, neuronal migration, axogenesis and dendritic growth, synaptogenesis and pruning, excitatory/inhibitory balance, and gliogenesis (de la Torre-Ubieta *et al.*, 2016; Ebrahimi-Fakhari & Sahin, 2015; Chen *et al.*, 2015). At the circuit level, ASD associated genetic disruptions can cause microglia infiltration, decreased Purkinje cell size and number, decreased neuron size, increased neuronal density, and alterations to cortical minicolumns and cytoarchitecture (de la Torre-Ubieta *et al.*, 2016; Chen *et al.*, 2015). Taken together, these disruptions in synaptic pathways

can affect both small-scale synaptic connectivity and large-scale structural and functional connectivity of neural circuits. As such, electrophysiological (EEG) measures of functional neural connectivity are plausible biomarkers of risk for ASD. These connectivity measures infer which brain regions are functionally connected to form brain networks responsible for neural activities during both task performance and baseline resting state (Sporns, 2011).

Early disruptions in connectivity during infancy are likely to precede the atypical connectivity patterns generally observed in individuals with ASD (both with known and unknown genetic causes). As will be discussed below, given the genetic evidence of prenatal disruption in neural circuitry, we should be able to measure aberrant brain connectivity in the first year of life using methods such as EEG, prior to the onset of atypical behavioral signs.

1.2 Infant sibling model

Prospective longitudinal studies of infants with older siblings with ASD (Familial Risk infants: FR) have been the paradigm for the investigation of the earliest markers of ASD, as these infants are identified before birth (and therefore can be followed from the earliest time points) and are at elevated risk for developing ASD. Infants with 1 older sibling with ASD are at 20% risk of developing ASD, while infants with multiple older siblings with ASD are at 33-50% risk of ASD (Ozonoff *et al.*, 2011; Messinger *et al.*, 2015). The infant sibling model represents an interplay between shared polygenic risks and environmental factors (Risch *et al.*, 2014). Compared to simplex families with only 1 child with ASD, multiplex families with 2 or more children with ASD have lower rates of de novo genetic variants (Leppa *et al.*, 2016). Within multiplex families, children with ASD have greater amount of rare copy-number variants (CNVs) than their unaffected siblings (Leppa *et al.*, 2016). As a reflection of underlying polygenic risks, up to 70% of children with ASD within multiplex families do not inherit the same rare CNVs as their other siblings with ASD (Yuen *et al.*, 2015; Leppa *et al.*, 2016). Familial risk infants

constitute a heterogenous group with mixed outcomes, reflecting this wide background of polygenic risks. Among familial risk infants who do not develop ASD, upward to 28% have language delay and/or global developmental delay by 3 years of age (Messinger *et al.*, 2013; Landa *et al.*, 2012, Ozonoff *et al.*, 2014, Charman *et al.*, 2017). Some familial risk infants who do not meet criteria for ASD may still exhibit a range of subclinical ASD symptoms and neurodevelopmental delays, as part of the broader autism phenotype (Messinger *et al.*, 2013; Landa *et al.*, 2013; Ozonoff *et al.*, 2014; Charman *et al.*, 2017). Familial risk infants can be followed from birth - providing a valuable opportunity to study the earliest manifestation of atypical brain development. Due to their heightened risks, it is important to study familial risk infants earlier on during the first year of life to identify predictive markers of language impairments and social communication deficits that will help guide implementation of earlier behavioral interventions. Over the past two decades, the infant sibling model has been utilized by many researchers in the Baby Siblings Research Consortium (BSRC), the Infant Brain Imaging Study (IBIS) Network, and the British Autism Study of Infant Siblings (BASIS) Network to detect early behavioral and neural markers of ASD.

1.3 Robust behavioral evidence of ASD by the 2nd year of life

Behavioral studies of familial risk infants have identified the emergence of many atypical behaviors between the first and second year of life, particularly in the areas of social communication, restricted and repetitive behaviors, motor, and language (Jones *et al.*, 2014; Zwaigenbaum *et al.*, 2005; Landa & Garrett-Mayer, 2006, Yirmiya *et al.*, 2006; Mitchell *et al.*, 2006; Landa *et al.*, 2007; Yirmiya *et al.*, 2007; Iverson & Wozniak, 2007; Gamliel *et al.*, 2009; Paul *et al.*, 2011; LeBarton & Iverson, 2013; Hudry *et al.*, 2014; Lazenby *et al.*, 2016; West *et al.*, 2017). While language impairment is no longer part of the diagnostic criteria for ASD in the current fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (American

Psychiatric Association, 2013), delayed language onset or regression in language ability in young toddlers ASD are among the first behavioral warning signs that parents detect (Rapin & Dunn, 2003). Language is an affected domain among many children with ASD, with nearly a third of children over 5 years old remaining minimally verbal, and an additional quarter only able to produce words but not complete sentences (Anderson *et al.*, 2007; Tager-Flusberg & Kasari, 2013). As such, early markers of language delay are important to identify to help guide closer monitoring of at-risk children and deliver early, targeted language intervention.

Language deficits. At 5-months of age, familial risk infants do not discriminate between lexical stress patterns in speech; in contrast, LR infants have increased attention to trochaic stress patterns, which relates to increased word comprehension at 12 months (Ference & Curtin, 2013). A profile of low expressive language has been documented as early as 6 months of age in infants who show high amount of ASD symptoms at 24 months (Paul *et al.*, 2011). At 6 months of age, patterns of poorly phonated cries are detectable in infants who will develop ASD (Sheinkopf *et al.*, 2012). Atypical intonation patterns are observable by 12 months of age in infants who develop ASD (Macari *et al.*, 2012). Between 12-18 months, toddlers who develop ASD start showing impairments in both words production and expressive language (Mitchell *et al.*, 2006; Zwaigenbaum *et al.*, 2005; Landa & Garrett-Mayer, 2006; Iverson *et al.*, 2007; Estes *et al.*, 2015; Lazenby *et al.*, 2016). Deficits in receptive language are evident by 12-14 months of age, and include lower understanding of phrased speech (Mitchell *et al.*, 2006; Zwaigenbaum *et al.*, 2005; Landa & Garrett-Mayer, 2006; Estes *et al.*, 2015; Lazenby *et al.*, 2016). Between 12-18 months, infants who develop ASD also have lower words comprehension (Mitchell *et al.*, 2006; Lazenby *et al.*, 2016). Between 8-18 months of age, an atypical lack of language growth accompanying walk onset is observed in infants who develop ASD (West *et al.*, 2017). Toddlers with ASD often have more delays in receptive than expressive language (Luyster *et al.*, 2008; Ellis Weismer *et al.*, 2010; Hudry *et al.*, 2014; Kwok *et al.*, 2015), which results in a decreased

receptive advantage. This deficit in receptive language could be caused by toddlers with ASD having more trouble learning words from their environment. Of note, since language ability is heterogeneous among children with ASD, those with robust language skills at 36 months may not show any language deficits at 18 months (Talbott *et al.*, 2015). All in all, deficits in both receptive and expressive language can be clearly detected using standardized behavioral assessments by 12 months of age in infants with ASD, with earlier language delays being associated with more severe ASD profiles.

Social communication deficits. Between 12-14 months of age, deficits in initiation of social interaction are seen in infants who develop ASD, which include: decreased showing and pointing, fewer gestures, lower rate of joint attention initiation, and less gaze alternation (Barbaro & Dissanayake, 2013; Macari *et al.*, 2012; Talbott *et al.*, 2015; Landa *et al.*, 2007; Mitchell *et al.*, 2006; Zwaigenbaum *et al.*, 2005). Between 6 and 12 months of age, infants who develop ASD start to show decreased attention to faces as well as reduced response to their own names (Chawarska *et al.*, 2013; Feldman *et al.*, 2012; Nadig *et al.*, 2007). Deficits in social attention become more pronounced by 12 months of age, when infants who develop ASD have less social smiling, decreased gaze to faces, fewer directed vocalization, and less attention to their mothers and researchers in the same room (Ozonoff *et al.*, 2010; Luyster *et al.*, 2009; Hutman *et al.*, 2010; Wan *et al.*, 2012). As early as 12 months of age, infants who develop ASD also show deficits in imitation, which is critical for social learning (Feldman *et al.*, 2012; Zwaigenbaum *et al.*, 2005; Young *et al.*, 2011). Diminished imitation ability is also observed by 12 months of age in familial risk infants who have developmental delay but no ASD (Young *et al.*, 2011). In infants who develop ASD, impaired response to referential cues (i.e. pointing, gaze, verbal) are evident by 14-15 months of age (Landa *et al.*, 2007, Rozga *et al.*, 2011, Sullivan *et al.*, 2007, Yoder *et al.*, 2009). Between 14 and 17 months of age, infants who develop ASD show deficits in both shared positive affect and social referencing (Landa *et al.*,

2007; Cornew *et al.*, 2012). As seen, deficits in many aspects of social communication are evident by 12 months of age in infants with ASD, at the same time when language deficits start to emerge. Impairments in social communication can directly affect language learning, as infants who are less engaged in social interactions are less able to learn language from their caregivers.

Restricted and repetitive behaviors. As early as 12 months of age, atypical behaviors during object free-play (i.e. unusual visual exploration, rolling, rotating, spinning) are evident in infants who develop ASD (Ozonoff *et al.*, 2008). At 12 months of age, infants who develop ASD start showing repetitive behaviors during interactions with examiners (Elison *et al.*, 2014) and increased restricted and repetitive behaviors across all subtypes (stereotypical, self-injurious, compulsive, ritual/sameness, restricted) on parent report measures (Wolff *et al.*, 2014). Of note in infants with ASD, repetitive behaviors at 24 months of age correlate with both worse adaptive skills and lower socialization scores, concurrently (Wolff *et al.*, 2014). At 12-18 months of age, infants who develop ASD have increased amounts arm waving (Loh *et al.*, 2007). By 18 months of age, infants who develop ASD have decreased functional play and increased nonfunctional, repetitive play behaviors (Christensen *et al.*, 2010). Overall, profiles of restricted and repetitive behaviors are evident in infants with ASD by 12 months of age, and preoccupation with these behaviors may preclude infants from developing adaptive skills for social communication with their caregivers.

Motor deficits. At 6 months of age, infants who later develop ASD show evident of head lag when pulled to sit, limited motor control, and lower activity level (Flanagan *et al.*, 2012; Bryson *et al.*, 2007; Zwaigenbaum *et al.*, 2005). Motor skills at 6 months of age have been shown to predict behavioral outcomes, including expressive language and ASD diagnosis, in infants at familial risk for ASD (LeBarton & Landa, 2018). Specifically, infants who develop ASD have poorer motor skills at 6 months of age (LeBarton & Landa, 2018). Delayed walking onset

has been observed at 15-16 months in children who develop ASD (Iverson *et al.*, 2007). Between 12-18 months of age, children who develop ASD also show delays in fine and gross motor skills (Landa & Garrett-Mayer, 2006; Ozonoff *et al.*, 2010). In summary, motor deficits start to emerge by 6 months of age in infants with ASD, before the emergence of many other behavioral symptoms. Motor skills are specifically important to language development, since motor functions are necessary for speech production and the ability to walk also open greater opportunities to learn from the environment.

Taken together, these studies show that robust differences in behavior are evident by the second year of life in children with ASD, with less certainty in the first year despite the hypothesis that underlying neural networks should already be altered in infancy.

1.4 Motivations to study early brain development in ASD

The wide variety of behavioral impairments in ASD are likely caused by underlying differences in neural network structure and function, which affect how individuals process environmental inputs and generate appropriate responses. Studying early brain development in familial risk infants enable researchers to elucidate neural markers of atypical development. Using neural markers to identify which infants are likely to develop ASD and neurodevelopment delays will have great potentials in facilitating both earlier monitoring and earlier targeted intervention while the brain has the greatest plasticity. From genetic studies of ASD, we know that the formation of brain networks is affected very early on during fetal brain development, as such, atypical neural connectivity patterns should be detectable during infancy, before the emergence of ASD symptoms.

Thus far, studies of neural connectivity in familial risk infants have focused on structural and functional connectivity. Through diffusion tensor imaging, structural connectivity can be measured in terms of fractional anisotropy (FA) of white matter fiber tracts, which assesses the

diffusion of water molecules along axons. Brain regions are considered structurally connected if there are white matter fiber tracts connecting them. Greater structural connectivity can be indexed by higher FA, reflecting more robust white matter microstructure (axon diameter, myelination, and white matter fiber density) (Le Bihan *et al.*, 2001). Structural connectivity is known to vary dynamically throughout development, as white matter tracts develop at different rates throughout the brain during infancy and early childhood. Compared to structural connectivity, functional connectivity has higher temporal resolution in revealing integration and segregation of brain networks. Functional connectivity can be measured using both functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) and is calculated as the statistical interdependency between activities from different brain areas. There are many measures of functional connectivity, including phase coherence, phase locking, covariance, and correlation (Mohammad-Rezazadeh *et al.*, 2016). In typical development, functional connectivity has been shown to increase over the first year of life as neural networks mature and become more specialized, organized, and adept at integrating information (Bell & Fox, 1992; Cuevas & Bell, 2011; Xiao *et al.*, 2018).

1.5 Neuroimaging markers of ASD in familial risk infants

Neuroimaging studies of early brain development have identified atypical brain structures that relate to ASD traits. Starting at 6-months of age, structural imaging studies have identified atypical corpus callosum morphology in infants who go on to develop ASD. In infants who develop ASD, increased corpus callosum area and thickness at 6-12 months correlate with repetitive behaviors at 24 months (Wolff *et al.*, 2015). Structural neuroimaging of infants with later ASD diagnosis have also identified abnormalities in developmental trajectories of cortical surface area at 6-24 months that correlate with ASD severity at 2 years (Hazlett *et al.*, 2017). Hazlett et al identified an increase in cortical surface area at 6-12 months which precedes an

overgrowth in brain volume at 12-24 months; in which volume overgrowth correlates social affect impairment (Hazlett *et al.*, 2017). Subcortical volumes at 12 months of age relate to language skills at 24 months, with different relationships between brain and behavior discriminating infants with ASD from infants with language delay. In infants who develop ASD, larger subcortical volumes (thalamus, amygdala, caudate nucleus) predicted greater language skills; whereas smaller subcortical volumes predicted more normative language profiles in infants who develop language delay (Swanson *et al.*, 2017a). Hazlett et al and Swanson et al are complementary studies, showing how overgrowth in both total brain volumes and subcortical volumes are associated with ASD. In contrast, inhibition of brain overgrowth appears to be protective for ASD (Swanson *et al.*, 2017a). In addition to having atypical brain structures, infants who develop ASD also have increased extra-axial cerebrospinal fluid (CSF) throughout 6-24 months of age, with increased CSF being associated with worse motor ability at 6 months and greater ASD severity at 24-36 months (Shen *et al.*, 2013; Shen *et al.*, 2017). Altogether, these studies show that changes to brain structure are evident as early as 6 months of age and may be implicated in the development of core ASD symptoms.

Neuroimaging studies of familial risk infants have also identified aberrant structural and functional connectivity patterns that related to core ASD traits. As early as 6-weeks of age, infants show altered structural lateralization of language networks that relate to receptive advantage at 18-months and ASD severity at 36-months (Liu *et al.*, 2019). Compared to low risk controls, familial risk infants have lower FA in left superior longitudinal fasciculus (SLF) and higher FA in the right SLF. Across risk groups, increased left FA in dorsal language network at 6-weeks correlate with better receptive advantage at 18-months (Liu *et al.*, 2019). In low risk infants, strengthening of structural connectivity in the splenium between 6 and 24 months of life has also been shown to predict greater words production at 24-months (Swanson *et al.*, 2017b). Compared to low risk infants, familial risk infants who develop ASD show diverging trajectory in

structural connectivity development, with patterns of over-connectivity in many networks at 6-months of age that precede under-connectivity patterns observed by 24 months (Wolff *et al.*, 2012). Specifically, Wolff et al reported elevated FA in most white matter tracts at 6 months that precedes a slower developmental change in white matter microstructure and resulting in lower FA by 24 months amongst infants with ASD (Wolff *et al.*, 2012). Compared to low risk toddlers at 24 months, familial risk toddlers have inefficient local and global neural network based on structural connectivity (Lewis *et al.*, 2014). Functional neuroimaging of familial risk infants has characterized patterns of resting state connectivity at 6 months that relate to core ASD symptoms at 24 months (Emerson *et al.*, 2017). Taken together, these neuroimaging studies show that altered connectivity in many brain networks are present in the first year of life and may predispose infants towards developing language delay or ASD.

1.6 Strengths and limitations of EEG

Compared to MRI, EEG affords both practical and scientific advantages in the study of early brain development. EEG signal reflect the summation of post-synaptic transmission from millions of cortical pyramidal neurons. In contrast to MRI, EEG has high temporal resolution making it preferable for studying functional connectivity. Since MRI is more sensitive to motion artifacts, studies in infants must be done while the child is asleep since it is impractical to expect young children to remain still for an extended period while awake. EEG is more feasible for use with infants and young children because it can tolerate some movement and can be performed quickly in awake children. An additional advantage of EEG is that it is scalable, and portable EEG systems can be used in medical clinics and in the community. EEG is developmentally and temporally sensitive, and there exists a wide range of normative EEG data in typical development, starting from infancy through early childhood. And, most relevant to this work described in this dissertation, EEG is scientifically valuable because it captures neural

processes, such as neural synchrony and connectivity, that are thought to be disrupted in neurodevelopmental disorders. As a tool, EEG is not without limitations. EEG signal measured at the scalp reflects a summation of cortical electrical activity from multiple sources that have been volume conducted through the skull. As such, EEG has low spatial resolution. EEG signals are also vulnerable to noise from both physiological and non-physiological sources, and careful artifact removal is needed to preserve the underlying neural signals. Similar to MRI, EEG output measures of interest are dependent on both processing and analytic methods.

EEG studies of infants at risk for ASD have allowed us to examine disruptions in connectivity at the earliest possible developmental stage. To date, few studies have examined EEG connectivity in infants with elevated risk of developing ASD. While there are many measures of EEG functional connectivity, the measures of power and coherence have been most commonly examined in the ASD literature (Wang *et al.*, 2013; Luckhardt *et al.*, 2014; Schwartz *et al.*, 2017). Power is calculated in terms of the amplitude of a signal (the amount of EEG activity within a frequency band) and reflects baseline synchronization of underlying neural oscillations (Wang *et al.*, 2013). Coherence is the linear-correlation in relative amplitude and phase between signals from two electrodes, within a frequency band. While computationally simple, coherence provides insight into functional interactions between neural networks (Srinivasan *et al.*, 2007; Lukhardt 2014 Luckhardt *et al.*, 2014). Signals that are coherent in either amplitude or phase are assumed to have highly coordinated activity, or a strong functional connection, between underlying nodes of neurons that produced those signals (Fries, 2005; Srinivasan *et al.*, 2007; Duffy & Als, 2012). However, disadvantages to coherence include its nonstationarity limitations, its representation of only linear relationships, and its sensitivity to volume conduction (Mohammad-Rezazadeh *et al.*, 2016).

1.7 EEG markers of ASD in familial risk infants

A few studies have reported differences in spontaneous EEG spectral power between familial risk and low risk infants at 3-12 months of age (Tierney *et al.*, 2012; Gabard-Durnam *et al.*, 2015; Levin *et al.*, 2017). In a recent study, Levin and colleagues reported a pattern of reduced frontal power in familial risk infants across multiple frequency bands and described correlations between frontal alpha power at 3-months of age and expressive language at 12 months (Levin *et al.*, 2017). Lower frontal power across all frequency bands have been observed in 6-months old familial risk infants, in addition to altered trajectories in frontal spectral power across multiple frequency bands between 6-24 months (Tierney *et al.*, 2012). Compared to low risk controls, familial risk infants also show opposite trajectories in the development of frontal alpha power asymmetry between 6 and 18 months of age (Gabard-Durnam *et al.*, 2015). Overall, infants with shared familial risk for ASD have distinctive spectral power profiles across the first year of life, during spontaneous baseline activity, suggesting that spectral power may be an endophenotype of ASD.

Using task-based paradigms in older infants, three EEG studies have characterized differences in functional connectivity between familial risk and low risk infants (Orekhova *et al.*, 2014; Righi *et al.*, 2014; Keehn *et al.*, 2015). In two of these studies, familial risk infants showed atypical coherence patterns in the gamma frequency band at 12 months while listening to speech sounds (Righi *et al.*, 2014) and during face processing (Keehn *et al.*, 2015). Brain oscillations in the gamma frequency band (30-50 Hz) are involved in binding local neural circuitry and coordinating synchronous firing among neurons in different populations. Task related brain activities in the gamma band are involved in cognitive processes including memory, informational processing, object representation and language (Basar, 2013), and represents early states of sensory perception (Pantev 1991). Notably, Righi *et al* found that 12 months old infants with ASD have decreased coherence between frontal and temporal-parietal

regions during an auditory oddball paradigm with speech sounds, compared to infants without ASD; and regardless of ASD outcome, familial risk infants have lower coherence than low risk infants (Righi *et al.*, 2014). Keehn *et al* showed that 12 months old familial risk infants have atypical leftward lateralization of anterior-posterior gamma coherence during a face processing task, and this atypicality is greatest in infants who go on to receive ASD diagnosis (Keehn *et al.*, 2015). Orekhova *et al* examined a different measure of connectivity, using phase lag index, while 14-months old infants viewed dynamic social videos and found alpha band hyper-connectivity in frontal and central areas in familial risk infants who go on to develop ASD (Orekhova *et al.*, 2014). Of note, a recent study used the same methods by Orekhova *et al* in a new sample of infants and has failed to replicate group differences in connectivity between infants with and without ASD (Haartsen *et al.*, 2019). Importantly, both Orekhova *et al* and Haartsen *et al* describe the same relationship between alpha phase lag index at 14-months and repetitive behaviors at 36-months among familial risk infants who develop ASD (Haartsen *et al.*, 2019; Orekhova *et al.*, 2014). These prior works by Orekhova *et al*, Haartsen *et al*, and Levin *et al* have highlighted the importance of relating early neural markers to continuous measures of behavioral outcomes. Taken together, these studies have shown that atypical functional connectivity patterns are detectable by 12 months of age in infants at risk of ASD.

1.8 Methodological gaps

While EEG holds a lot of scientific promise, there are still methodological gaps in EEG-processing strategies regarding increases to both data retention and subject retention. Data loss from EEG data cleaning is common, which often leads to exclusion of subjects with noisy data from the final analyses (Table 1.1). Studies by Tierney *et al*, Gabard-Durnam *et al*, and Levin *et al* used the same minimum threshold of 10 seconds of good data from 120 seconds of spontaneous EEG recording as criteria to retain subjects for further analysis. With a low

threshold of 10 seconds of good data (good data from only 8% of recording length), 81-87% of subjects were retained for in the studies' final analyses (Tierney *et al.*, 2012; Gabard-Durnam *et al.*, 2015; Levin *et al.*, 2017). Averaged across all subjects, good data duration retention was 42-45% in 3-months old infants, and 32-51% in 6-24 months old children (Tierney *et al.*, 2012; Levin *et al.*, 2017). Studies that required subjects to have longer durations of good data tended to have lower subject retention after data cleaning. For example, Keehn *et al* required subjects to have at minimum 20% good trials and had a subject retention of 61% (Keehn *et al.*, 2015). Orekhova *et al* and Haartsen *et al* both retained subjects for further analyses if they had at least 120 seconds of good data out of 351 seconds recording (34% good data), this resulted in a subject retention rate ranging between 52% (Orekhova 2014) and 71% (Haartsen *et al.*, 2019). Since infant siblings represent a valuable and small subject pool, it is important to develop methodological techniques to maximize retention of subject's data.

Study	Minimum threshold	Data retained	Subject retained
Tierney 2012	10/120 s (8%)	39-61 s/120 s (32-51%)	122/140 (87%)
Gabard-Durnam 2015	10/120 s (8%)	Not reported	108/126 (86%)
Levin 2017	10/120 s (8%)	58-73 /120 s (42-45%)	39/48 (81%)
Orekhova 2014	120/351 s (34%)	Not reported	54/103 (52%)
Keehn 2015	10/50 trials (12 s, 20%)	Not reported	95/156 (61%)
Haartsen 2019	120/351 s (34%)	Not reported	101/143 (71%)

Table 1.1. *Data and subject retention from a select number of prior infant EEG studies.*

1.9 Scientific gaps

While structural and functional connectivity in infants at risk for ASD have been well characterized at baseline during sleep and spontaneous activity, little is known in this cohort in regard to task-specific connectivity during language processing. Structural connectivity in language networks during early infancy has been shown to relate later language ability (Liu *et al.*, 2019); and underlying differences in structural connectivity may contribute to altered functional connectivity in language networks. In the context of previously published studies, a fundamental question that remains is whether there are differences in early neural network connectivity during language processing that predict later ASD and language outcomes.

1.10 Goals of dissertation

This dissertation leverages data from a current longitudinal study on early biomarkers of ASD in familial-risk infants as part of the UCLA Autism Center for Excellence (ACE; NICHD 2P50HD055784-08). The ongoing study is in the seventh year of data collection and utilizes eye tracking to study social attention, EEG to examine cognitive processes of visual and auditory statistical learning, and MRI to evaluate passive language processing.

Since language is a shared affected domain among children with ASD and those with other delays, this dissertation is focused on studying functional connectivity during language processing in infancy, before the manifestation of behavioral symptoms. Research described in this dissertation is the first to relate EEG functional connectivity during an auditory language task to language outcome in familial risk infants. Identification of the earliest biomarkers of risk for core symptoms of ASD will lead the way for implementation of early behavioral interventions that may attenuate symptoms and even prevent the development of ASD (Zwaigenbaum *et al.*, 2015; Green *et al.*, 2015).

Chapter 2 addresses methodological considerations in the development of an EEG pre-processing pipeline with the goals of maximizing data quality and data retention for infant EEG. Chapters 3-5 share the same data collection and EEG processing methods, but each chapter characterizes different aspects of functional connectivity during language processing in infants at risk for ASD. Chapter 3 describes differences in connectivity at 3-months of age in infants who show ASD symptoms at 18-months of age. Chapter 4 highlights altered trajectories in connectivity development over the first year of life in infants who later have ASD symptoms at 18-months. Both chapters 3 and 4 are separate drafts of manuscripts that are in preparation for publication. Chapter 5 is focused on connectivity as an endophenotype, describing differences between familial risk and low risk infants, using both the 3-month cross-sectional study design and the 3-12-month longitudinal study design. As shared secondary analyses, chapters 3-5 also describe the relationship between early connectivity and behavioral outcomes.

Chapter 2: Methodological Considerations in Pre-Processing Infant EEG

2.1 Introduction to infant EEG

EEG is a valuable tool that is well suited to study infants. Unlike other imaging modalities like magnetic resonance imaging (MRI) or magnetoencephalography (MEG), EEG can tolerate infants' movements during the recording. While biological artifacts such as electromyogram (EMG) are recorded in the raw EEG, data pre-processing is designed to remove artifacts and preserve underlying neural signals. EEG is scalable and portable, allowing possible integration of EEG to the clinics, schools, and community settings. EEG is also developmentally sensitive, allowing researchers to measure subtle changes in neural activation as the infant's brain develop over the first year of life. In the research setting, EEG is recorded when the child is awake, allowing the study of many brain processes – from spontaneous activity to higher-order sensory and cognitive processing.

Infant EEG recordings are constrained by multiple factors that are inherent to the subject population. Fitting of the EEG net to the infant's head must be done quickly to minimize the child's discomfort which does not give the researcher ample time to adjust every single electrode for maximal conductivity. Research EEG recordings in infants must be kept short because infants usually can only tolerate a few minutes of a paradigm before becoming fussy or beginning to cry. Infants often move their heads, babble, or suck on a pacifier during the EEG recording, which can contribute high amount of EMG in the short recording. EEG recordings in infants are often done while the child is seated on their caregiver's lap. Young infants may lean their heads back against their caregivers, which can cause electrodes to shift and induce artifacts across many EEG channels. Older infants may tug on the net's straps or grab hold of the net's electrodes, which can induce high amount of artifact in the recording. Data processing

of infant EEG must be sensitive to the high-artifact nature of the data, as well as short paradigm duration.

In this chapter, methodological considerations in pre-processing of infant EEG data are discussed. In the development of the **Traditional Pipeline** and the **Revised Pipeline**, we have taken approaches to tackle biggest issues of infant EEG: electromyogram (EMG) and volume conduction. Traditional methods for removing EMG and other artifacts from infant EEG have often resulted in significant data loss, which can contribute to substantial subject loss from the research study. The **Revised Pipeline** was developed with the main goal of reducing data loss from infant EEG. Since infant EEG recordings already have short data length, it is important to try to and preserve as much clean neural signal as possible to increase the signal:noise ratio. Attributes of two processing pipelines, the **Traditional Pipeline** and the **Revised Pipeline** (Figure 2.1), are compared when both pipelines are used in parallel to clean data files from 3-6 months old infants. The **Revised Pipeline** has many advantages in terms of its reproducibility, processing speed, and maximal retention of clean file length – making it the ideal pipeline to be in the studies described in Chapters 3-5.

2.2 Methods

With the goal of addressing issues of EMG and volume conduction in EEG processing, two processing pipelines were developed and used in parallel to process continuous infant EEG data. The **Traditional Pipeline** was designed using prior published protocol for processing continuous EEG in young children with ASD (Dickinson *et al.*, 2018). The **Revised Pipeline** was designed to increase automation and reproducibility in the processing steps and improve output data quality. Both pipelines were designed with consultations from the Swartz Center for Computational Neuroscience and utilized the Matlab EEGLAB plugin (Delorme & Makeig, 2004). Overview of the processing steps for the **Traditional Pipeline** and the **Revised Pipeline**

are illustrated in Figure 2.1. To assess strengths and weaknesses of each processing pipeline, both the **Traditional Pipeline** and the **Revised Pipeline** were used in parallel to clean 166 continuous EEG files from 3-6 months old infants. 81 files were recorded from 3-months old infants and 85 files were recorded from 6-months old infants. The study design will be described in further details in Chapters 3 and 4.

EEG data acquisition

EEG was recorded for 2.5 minutes during an auditory language processing task using a 128-channel HydroCel Geodesic Sensor Net containing Ag/AgCl electrodes and sponges with saline electrolyte (Electrical Geodesics Inc., Eugene, OR). To improve each child's comfort, four of electrodes originally placed below and lateral to the eyes – channels 125-128 were removed from the net. Placement of electrodes conformed to the International 10-20 System (Jasper 1958). Net Amps 300 amplifier and Net Station 4.5.7 software on a Mac Pro desktop were used to record EEG (Electrical Geodesics Inc., Eugene, OR). Data was filtered online during recording using an analog bandpass elliptical filter between 0.1 and 100 Hz. EEG was sampled at 500 Hz. Data were referenced online during recording to a vertical reference in a location equivalent to Cz.

The significance of the language processing task will be later described in Chapters 3 and 4. The focus of this current chapter is on the EEG processing pipelines, and not on the domain of language processing. For each EEG processing step, I will describe key methodological considerations, main modifications between the **Traditional Pipeline** and the **Revised Pipeline**, along with details from each pipeline.

Filtering

Considerations*:* The choice of low pass filter must be appropriate for the EEG outcome measures of interest. EEGLAB Source Information Flow Toolbox (SIFT) is a useful tool for source-based EEG connectivity analyses (Delorme *et al.*, 2011). For future analysis with SIFT, a low-pass filter with wide transition band width needs to be chosen to help lower model order estimations. It is also important to remove 60 Hz line noise because the presence of 60 Hz noise in the data will cause high model order in SIFT. *Note:* EEGLAB CleanLine plugin (Mullen 2012), which used a multi-taper regression, was unable to remove all 60-Hz line noise from the data.

Modifications: The two pipelines differed in cut-off frequency and roll-off bandwidth for the low-pass filter. The **Traditional Pipeline** was low-pass filtered at 90 Hz with a narrow roll-off, while the **Revised Pipeline** was low-pass filtered at 50 Hz with a gentle, broad transition bandwidth.

Traditional Pipeline. Data was bandpass filtered at 1-90 Hz using a finite impulse response (FIR) filter with narrow roll-off (0.3 Hz) and strong attenuation (gain = -60 dB). 60 Hz electrical line noise was present and left in the data, with the intention of analyzing data below 50 Hz and above 70 Hz in the future.

Revised Pipeline. Data were high-pass filtered at 1.5 Hz using Blackman window FIR filter, with 1 Hz transition bandwidth and filter order of 2750. Due to the presence of 60 Hz line noise in the EEG recording, the data were further low-pass filtered at 50 Hz using Blackman window FIR filter, with 20 Hz transition bandwidth and filter order of 138. This pipeline was designed to pre-process data for both functional and effective connectivity analyses, including compatibility with SIFT.

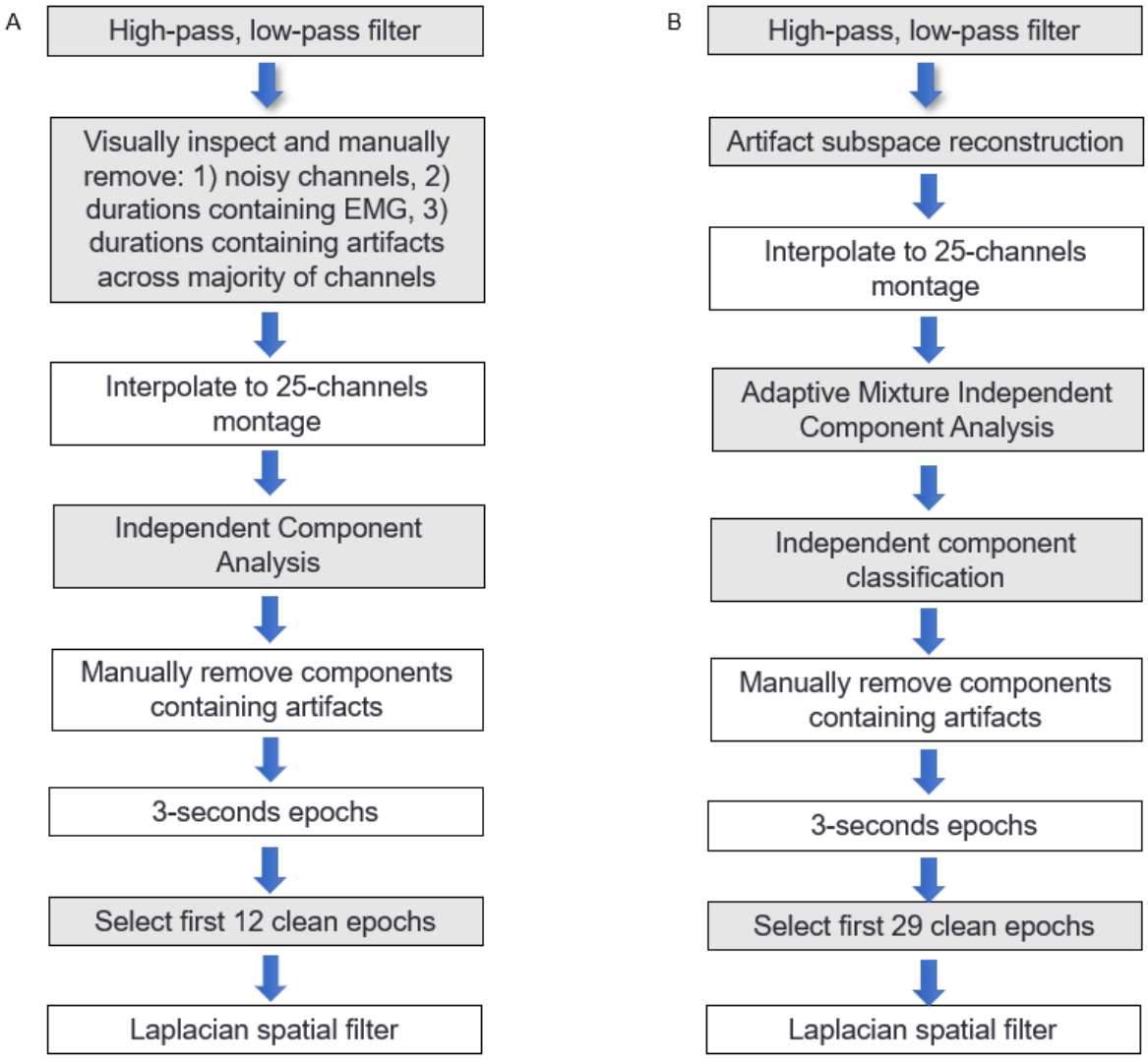

Figure 2.1. *Two pre-processing pipelines for continuous EEG data.* ***Figure 2.1a. Traditional Pipeline*** *involves manual data cleaning before running ICA.* ***Figure 2.1b. Revised Pipeline*** *utilizes ASR for automatic data cleaning before running AMICA. Highlighted steps are key differences between the pipelines.*

First-pass data cleaning

Considerations: Manual data cleaning has traditionally been used to process infant EEG; this method is time consuming but gives the user complete control over every decision to retain or reject portions of data. Manual data cleaning is subjective and user-dependent; thus, it is difficult to replicate precisely between multiple researchers, or even by the same researcher at another point in time. Automated data cleaning methods such as artifact subspace reconstruction (ASR) are reproducible, efficient, and have potential for preserving longer durations of underlying neural signals.

Modifications: The Traditional Pipeline used manual data cleaning, while in contrast the Revised Pipeline used ASR.

Traditional Pipeline. *Bad channel rejection*: data were visually inspected and bad channels with drift and artifacts were manually removed. *Bad duration rejection*: during visual inspection, durations of recording were manually removed if they contain EMG or artifacts affecting majority of channels (Figure 2.2). At this stage, eye blinks and eye movements were left in the data. After manual first-pass data cleaning, files needed to have durations of at least 37.5 seconds of good data to be retained for further processing.

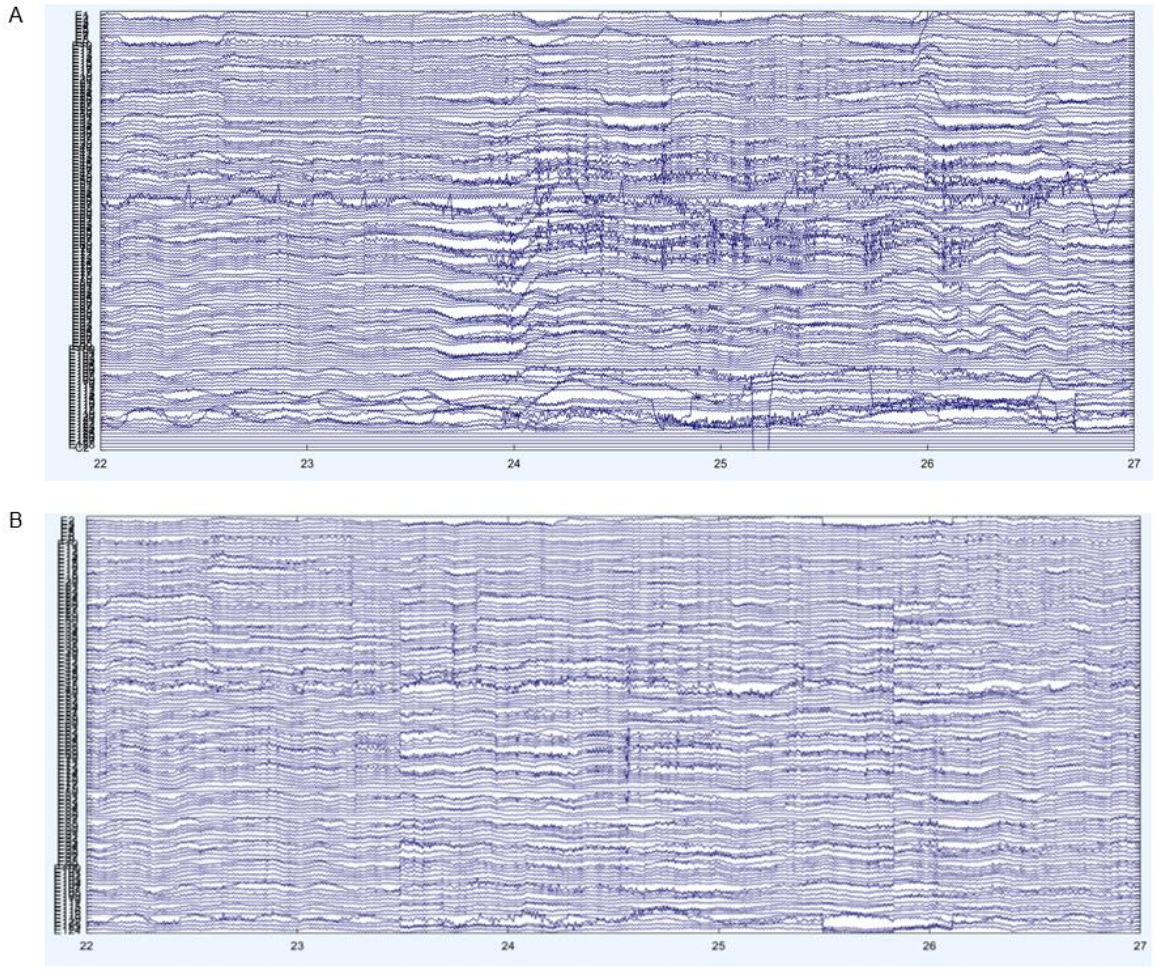

Figure 2.2. *Manual data cleaning example.* ***Figure 2.2a.*** *Filtered file before manual artifact removal.* ***Figure 2.2b.*** *Filtered file after manual removal of noisy channels and durations containing artifacts across many channels.*

Revised Pipeline. Artifact subspace reconstruction (ASR) was applied to clean continuous data using EEGLAB plugin *clean_rawdata* (Mullen *et al.*, 2015; Chang *et al.*, 2018). Mathematical equation for ASR is described in Figure 2.3. The algorithm for ASR is designed to use the available working memory on the computer to decide block-size for: 1) calculating calibrated reference data, and 2) calculating moving average during the data reconstruction step. Since a computer's working memory may vary each time ASR is run (depending on which

other processes the computer is running simultaneously), ASR default outputs will differ depending on the computer's available RAM. The algorithm for ASR was designed to maximize processing speed (ASR will run faster if it can use more working memory), and to prevent ASR from requiring more working memory than what is available on the computer. For ASR's outputs to be reproducible, the algorithm for ASR needs to be modified manually to run with a fixed random-access memory (RAM); the available RAM for the computer needs to be determined prior to this step. ASR computations for Chapters 2, 3, 4, 5 were performed on an MSI GS63VR Stealth Pro laptop, where ASR was run with fixed RAM allocation of 3200 megabytes.

Bad channel rejection. Channels were rejected if they were not well correlated with adjacent channels; specifically, channels with less than 75% correlation to their reconstruction based on other channels in the given time window were removed. Sliding time window of 1-second duration was used. Channels were rejected if they had flatline signal lasting longer than 3 seconds. A cut-off of 3 seconds was selected because phase coherence analyses require epoch length of at least 3 seconds. In Chapters 3-5, the **Revised Pipeline** was applied to pre-process data for phase coherence analyses.

Calibration of reference data. During the calibration step, ASR first selected the cleanest windows of channel data by applying infinite impulse response (IIR) spectral weighting (based on the inverse of a model for EEG power spectral density with 1/f trend and alpha peak at 8 Hz). The cleanest windows of channel data were concatenated to form the calibrated reference data. Principal component analysis (PCA) based on calibrated reference data was computed.

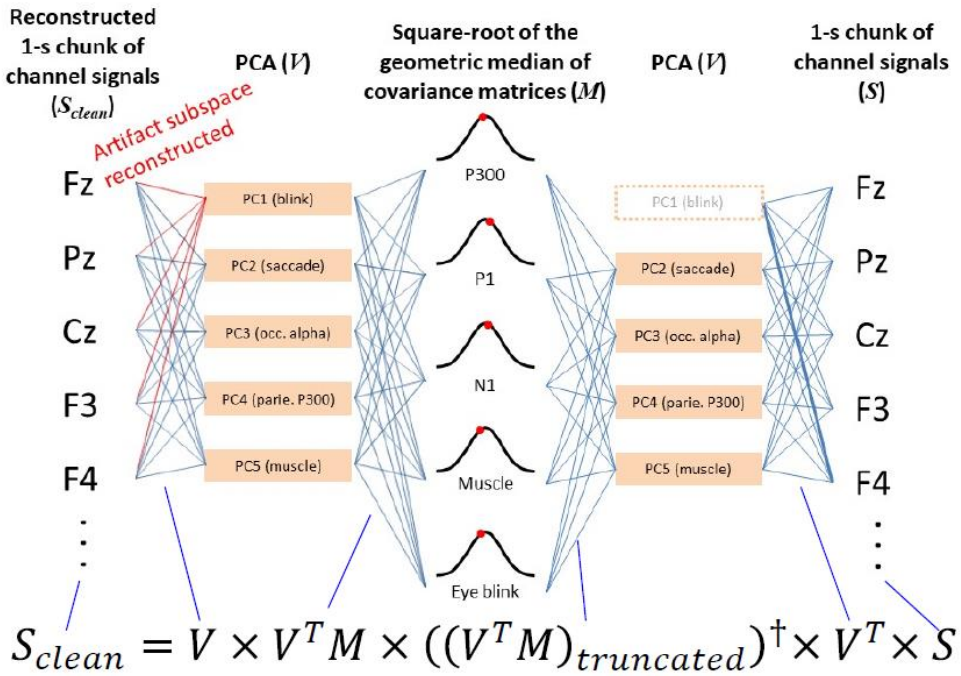

Figure 2.3*. Mathematical model for artifact subspace reconstruction (ASR). Figure by courtesy of Makoto Miyakoshi.*

Data rejection and interpolation. PCA was applied to each 1-second sliding window of raw data. Principal components (PCs) from each window of raw data are compared to PCs from the calibrated reference data. Data rejection and interpolation were done at the level of the PCs. Within each window, raw data's PCs with variance greater than 8 standard deviation from the calibration data's PCs were labeled as *artifact subspace*. Raw data's PCs labeled as artifact subspace were rejected and removed. During the reconstruction step, the artifact subspace was interpolated based on the calibration data's PCs. ASR performed data rejection and interpolation within 1-second sliding window throughout the entire EEG file. After interpolation was complete, data was back-projected from PC space into channel space.

Window rejection. Window rejection was done after ASR data rejection and interpolation have been completed. Within each 1-second sliding window, if more than 25% of channels had

z-scores root-mean-square amplitude outside the range of -3.5 to 5.5, then that window was rejected. Lowering the threshold for window rejection (i.e. 5-10%) would be much more conservative and would cause greater data loss from increased number of windows rejected. An example of a file cleaned with ASR is depicted in Figure 2.4.

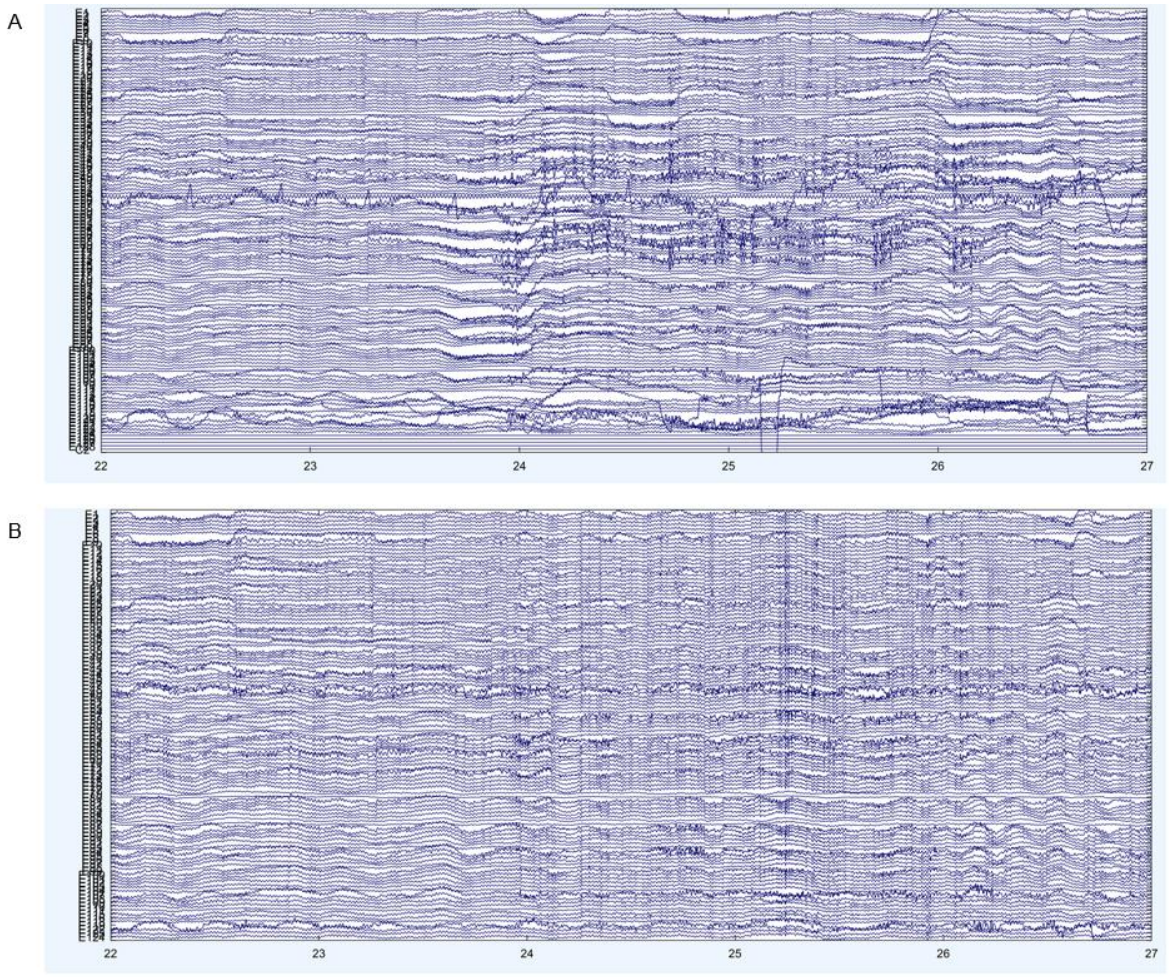

Figure 2.4. *ASR example.* ***Figure 2.4a.*** *Filtered file before ASR.* ***Figure 2.4b.*** *Filtered file after ASR*.

Both **Traditional Pipeline** and **Revised Pipeline** incorporated rejection of bad data durations. After data duration rejection, the remaining clean data segments were concatenated.

This type of concatenation caused discontinuities in the data and introduced high frequency noise (Figure 2.5). After pre-processing has been complete, further EEG analyses (spectral power, phase coherence …) should exclude any data epoch containing these discontinuities.

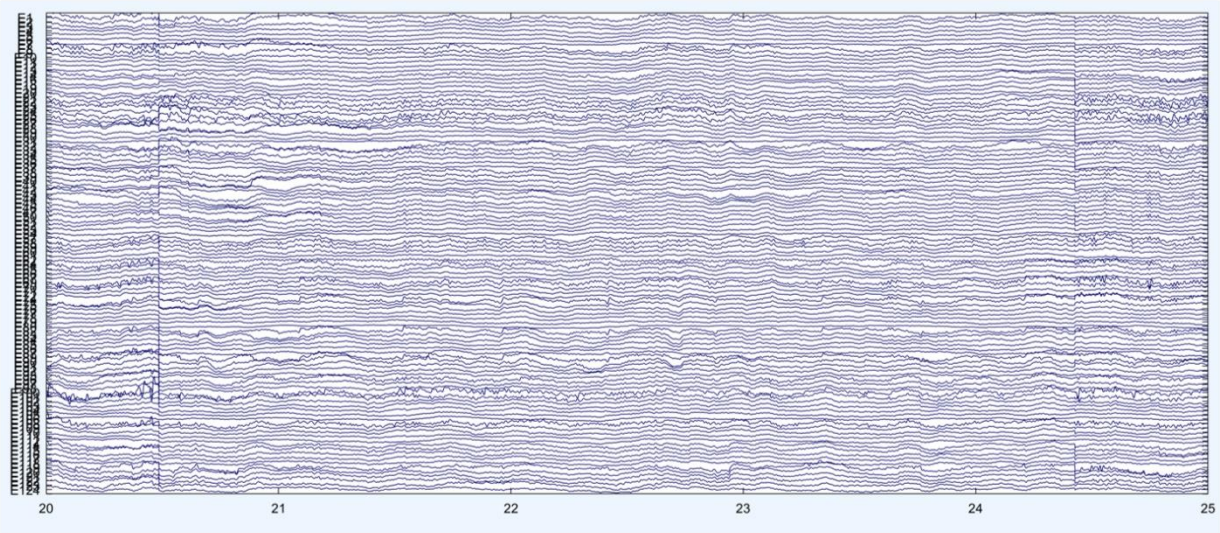

Figure 2.5. *Example of discontinuities in concatenated clean data, caused by removal of noisy durations of data.*

Data reduction

Considerations*:* In order for Independent Component Analysis (ICA) to run well and have good decomposition of the data, the following relationship between recording EEG channels and total data points in the file must be satisfied (Onton *et al.*, 2006):

$$(channel\ numbers)^2 \times 30 = \text{required data points.}$$

Since the number of data points in the file is dependent on the original data sampling rate and duration of the recording, the following parameters must be met prior to running ICA:

$$\frac{(channel\ numbers)^2 \times 30}{sampling\ rate\ (Hz)} = \text{minimum duration of good data (seconds).}$$

High-density 128-channels EGI nets are commonly used in to record EEG in awake infants in the research setting, due to ease in net application. If the data are obtained using a 128-channels net, sampled at 500 Hz, then a file length of at least 983 seconds (16.4 minutes) is required for to run ICA. At a fixed sampling rate of 500 Hz: 64-channels data requires 245.8 seconds of data (4 minutes), 32-channels data requires 61.4 seconds of data, and 25 channels data requires 37.5 seconds of data. These data lengths refer to the duration of good data after the first-pass in data cleaning has already been done, and not the duration of the raw data recording. ICA requires relatively clean data input for good decomposition quality. Inputting raw, messy data into ICA will yield a bad decomposition with few output components reflecting true brain activity.

If the researcher is interested in examining infants' neural activity specific to a task, one is constrained by both short data length and high number of recording channels. There are several common approaches to resolving the constraints of having short data length from high density EEG recordings: 1) performing dimension reduction with Principle Component Analysis (PCA) prior to running ICA; 2) selection of a smaller subset of channels from the original recording; 3) down-sampling the data by interpolating to a montage with fewer channels. Prior works in our lab examining spontaneous resting state activity in children with ASD and neurogenetic syndromes have used PCA to reduce the data to 24 dimensions prior to ICA (Frohlich *et al.*, 2016). A recent study by Artoni *et al* has highlighted several drawbacks to using dimension reduction with PCA prior to ICA, including reductions in both independent components (IC) stability and fewer number of dipolar "neural" IC (Artoni *et al.*, 2018). The selection of channel subsets was recommended as one of the first pre-processing steps in the Harvard Automated Processing Pipeline for Electroencephalography (HAPPE), to be done prior to bad channel rejection (Gabard-Durnam *et al.*, 2018). There are several drawbacks to selecting channel subsets prior to data-cleaning which includes: 1) selecting number of

channels used in the subset is arbitrary without knowing how the long the clean data duration will be; 2) noisy channels will be included in the channel subset selection, which will further reduce the number of channels containing true neural signal in the down-sampled dataset.

Considering these methodological constraints, both **Traditional Pipeline** and **Revised Pipeline** utilize data down-sampling after first-pass data cleaning has been completed. By knowing how much good data duration remains will allow the researcher to make an informed decision in choosing a channel montage to interpolate the data.

Modifications: None. Both **Traditional Pipeline** and **Revised Pipeline** shared the same data reduction parameters.

Traditional Pipeline: After manual removal of noisy channels and noisy durations of data, the remaining data from the original 128-channels montage was interpolated to 25-channels montage based on the International 10-20 System using interp_mont plugin (Desjardins, 2010; Jasper, 1958). A 25-channels montage was chosen because it requires only 37.5 seconds of good data for reliable ICA decomposition. Prior studies in our lab have used 30-seconds of clean data as the minimum cut-off for subject inclusion in further analyses (Frohlich *et al.*, 2016; McEvoy *et al.*, 2015).

Revised Pipeline: After cleaning with ASR, data from the original 128-channels montage was interpolated to 25-channels montage based on the International 10-20 System using interp_mont plugin (Desjardins, 2010; Jasper, 1958).

Independent component analysis

Considerations: ICA is useful for artifact removal and is also a required prerequisite for later analysis with SIFT. The data points requirement for good ICA decomposition was previously described in the *data reduction* step. ICA decomposition separates signals arising from neural sources as well as from artifact sources into individual ICs, allowing for inspection

and rejection of ICs containing artifacts. Among the different ICA algorithms, Adaptive Mixture Independent Component Analysis (AMICA) has been shown to outperform other ICA methods in scalp channels mutual information reduction and identifies the maximal number of dipolar brain IC (Delorme *et al.*, 2012).

Modifications: The two pipelines used different ICA algorithms. The **Traditional Pipeline** used extended infomax ICA, while the **Revised Pipeline** used AMICA.

Traditional Pipeline: Data decomposition using extended infomax ICA was performed and generated 25 independent components (ICs) per subject. Extended infomax ICA has been used in prior studies in the Jeste lab (Frohlich *et al.*, 2016; Dickinson *et al.*, 2018).

Revised Pipeline: 1-model Adaptive Mixture Independent Component Analysis (AMICA) (Palmer *et al.*, 2006; Palmer *et al.*, 2008) was performed and outputted 25 independent components (IC) per subject. AMICA learns the data structure without preset assumptions; prior studies have utilized AMICA in processing both infant Piazza *et al.*, 2016) and adult EEG data (Hsu *et al.*, 2018).

Second-pass data cleaning

Considerations: After ICA has been run, output ICs need to be inspected visually for remaining artifacts such as eye blinks, saccades, BCG (ballistocardiograms), and EMG. EEGLAB plugin ICLabel is a useful tool to aid in manual review and rejection of non-neural IC. IClabel is a machine-learning algorithm, trained on thousands of crowd-labeled adult EEG datasets (Pion-Tonachini *et al.*, 2017). IClabel accounts for each IC's activity power spectrum and scalp topography in its classification algorithm (Pion-Tonachini *et al.*, 2017; Pion-Tonachini *et al.*, 2019). IClabel classifies each IC with percent likelihood of "Brain", "Muscle", "Eye", "Heart", "Line noise", "Channel noise", or "Other".

Modifications: The **Revised Pipeline** utilized ICLabel, while the **Traditional Pipeline** did not use ICLabel.

Traditional Pipeline: Manual rejection of artifact ICs was done after careful inspection of each IC's activity power spectrum, scalp topography, and "component activity scroll". Neural ICs with 1/f trend in activity power spectrum, dipolar activity in the scalp topography, without BCG or EMG in "component activity scroll" were retained (Figure 2.6). ICs containing eye blinks (Figure 2.7a) or saccadic eye movement (Figure 2.7b) were identified and rejected based on localized frontal activity on the scalp topography and characteristic large deflections in the "component activity scroll". ICs containing BCG (Figure 2.7c) were identified and rejected based on cardiac 1-Hz activity on "component activity scroll", either posterior or diffuse spread of activity on scalp topography. ICs containing EMG (Figure 2.7d) were identified and rejected based on deviation from 1/f trend in the activity power spectrum, and presence of high-frequency artifact bursts in the "component activity scroll". ICs containing channel noise were also rejected; these were identified based on activity localized to one channel on the scalp topography (Figure 2.7d).

Revised Pipeline: Visual inspection of ICs was done as previously described in **Traditional Pipeline**, but with additional aid from ICLabel. Highlighted in yellow in Figure 2.6 and Figure 2.7 are IClabel's automatic classification of brain and artifact components. IClabel did not account for each IC's "component activity scroll". Manual visual inspection of each IC's "component activity scroll" was necessary to detect infrequent EMG bursts or subtle BCG. Careful review of ICs was still necessary, since some neural components from infant EEG were classified as "Other" by IClabel.

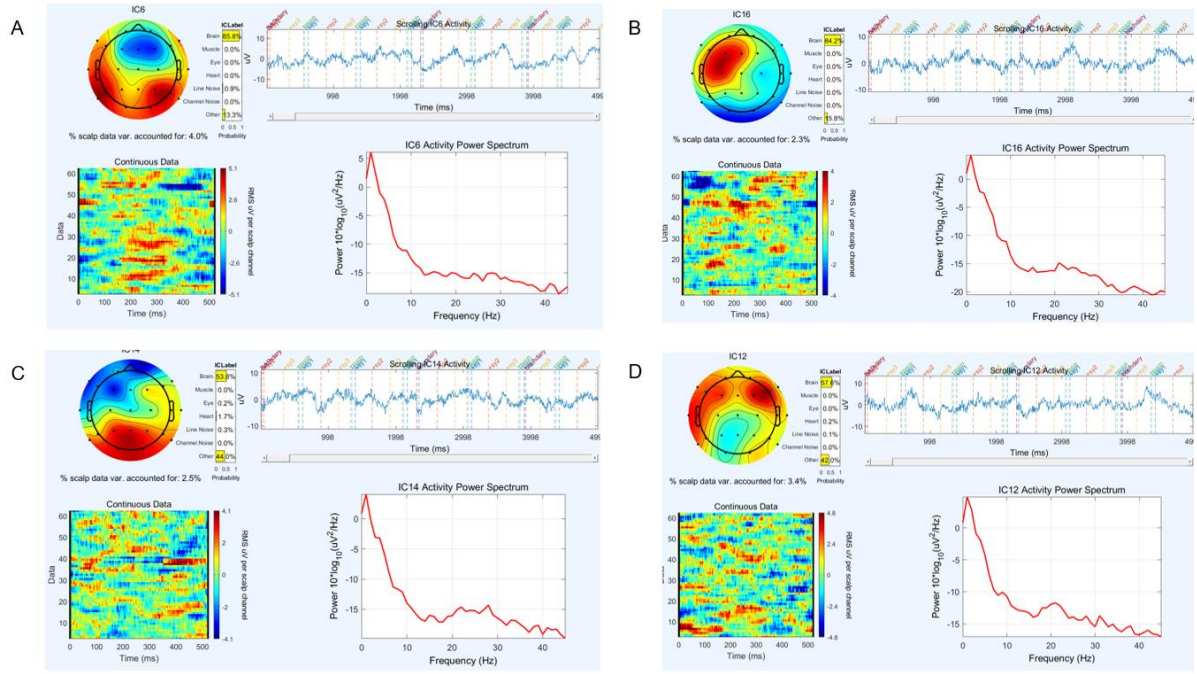

Figure 2.6. *Examples of independent components from neural sources.*

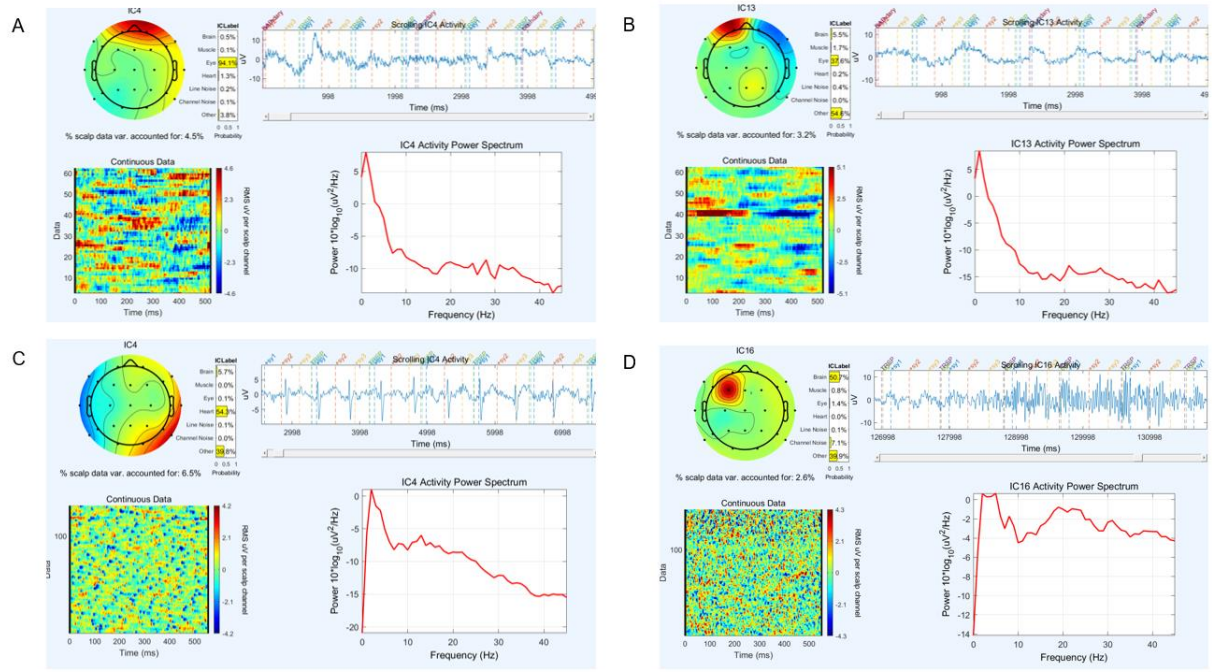

Figure 2.7. *Examples of independent components containing artifacts.* ***Figure 2.7a.*** *Eye blink.* ***Figure 2.7b.*** *Saccadic eye movement.* ***Figure 2.7c.*** *BCG from cardiac activity.* ***Figure 2.7d.*** *EMG burst.*

Epoching continuous data

Considerations: Epoch length must be appropriate for calculating the outcome measures of interest. In this case, phase coherence was one of our EEG measures of interest. Phase coherence analysis required epochs of at minimum 3-seconds in length. Since coherence analysis was sensitive to data length (i.e. different file length in the same subject would yield different coherence outputs), the same number of epochs was selected across all subjects. The number of epochs was dependent on the subjects' remaining file length after the *first-pass data cleaning step* of each pipeline.

Modifications: In the **Traditional Pipeline**, epochs contained discontinuities from concatenation of clean data (after the *first-pass data cleaning* step). In the **Revised Pipeline**, any epochs with discontinuities were discarded prior to final epoch selection. The two pipelines differed in the total number of epochs selected across all subjects.

Traditional Pipeline: Cleaned data were epoched into 3-second segments. The first 12 epochs of data were selected across all subjects.

Revised Pipeline: Cleaned data were epoched into 3-second segments. The first 29 epochs of data were selected across all subjects.

Laplacian spatial filter

Considerations: Volume conduction is a prevalent problem in high density EEG recordings and can cause neighboring channels to appear falsely to have the similar synchronous activity. In application to scalp-based analysis of high-density EEG obtained from young infants, Laplacian spatial filter is preferable to average referencing because it can mitigate effects of volume conduction.

Modifications: None. Both **Traditional Pipeline** and **Revised Pipeline** utilized the same Laplacian spatial filter.

Traditional Pipeline: A spherical spline Laplacian transform with head-circumference correction was applied to transform the cleaned data into current source density (Kayser 2006).

Revised Pipeline: Epoched data were passed through a Laplacian spatial filter to correct for volume conduction effects.

EEG spectral power analysis

Spectral power analysis was performed using 3-month EEG data that was pre-processed with either the **Traditional Pipeline** or the **Revised Pipeline**. Spectral power at 2-50 Hz was calculated using Welch's method in frontal and temporal-central regions of interest (ROIs) across the scalp, using custom scripts written in MATLAB as per prior protocol in the Jeste Lab (McEvoy et al., 2015). Left frontal ROI included channels F3, F7, and F9; left temporal-central ROI included channels T7, T9, and C3; right frontal ROI included channels F4, F8, and F10; right temporal-central ROI included channels T8, T10, and C4. For each 256-sample segment, the Fast Fourier Transform (FFT) was calculated on 128-point Hamming windows with 50% overlap. Absolute power was calculated by summing power estimates at every 0.5 Hz increment within each frequency band. All power values were log base 10 transformed.

2.3 Impact of Processing Pipelines on Data Quality

Bad channel rejection

Traditional Pipeline: All files underwent manual removal of bad channels. In the 3-month dataset, 12% of channels were manually removed (14 $\pm$ 5). In the 6-month dataset, 14%

of channels were manually removed (17 ± 7). Averaged across all 166 files, 86-88% channels were retained after manual bad channel rejection.

Revised Pipeline: ASR performed bad channel rejection on all files. With the *channel criterion* set to 0.75 and *flatline criterion* set to 3, ASR automatically rejected bad channels at a similar rate to manual channel rejection. In the 3-month dataset, ASR rejected 12% of channels (15 ± 8). In the 6-month dataset, ASR rejected 11% of channels (14 ± 10).

Bad duration removal

Traditional Pipeline: All 166 files from the 3-6-month datasets underwent manual removal of data durations containing noise and artifacts. In the 3-month dataset, on average 81 seconds were manually removed (range 3-124 seconds). In the 6-month dataset, on average 80 seconds were manually removed (range 16-208 seconds). Averaged across all 166 files, 47-48% of file length were retained after manual duration rejection (Table 2.1).

Revised Pipeline: Window rejection was enabled, and any 1-second window where deviations occurred in more than 25% of channels, after ASR interpolation, was removed. In the 3-month dataset, ASR performed window rejection on 7 out of 81 files. Among those 7 files: 6 files had 1-2.34 seconds rejected; 1 file had 50 seconds rejected. In the 6-month dataset, ASR only performed window rejection on 6 out of 85 files. In those 6 files, only 1-1.34 seconds were rejected from the entire file length. Averaged across all 166 files from the 3-6-month datasets, ASR retained 99.5-100% file length after data cleaning (Table 2.1).

IC retained

The two pipelines had similar number of retained ICs after ICA (**Traditional Pipeline**) and AMICA (**Revised Pipeline**), respectively (Table 2.1). More sophisticated techniques of

bootstrapping and clustering of ICs would be needed to compare the ICs' quality (such as stability and dipolarity of ICs) between the two pipelines (Artoni *et al.*, 2018; 2019).

Power Spectral Density

Power spectral density (PSD) plots were used to qualitatively compare data output from the **Traditional Pipeline** versus the **Revised Pipeline**. In the 3-month dataset, data processed with the **Revised Pipeline** generally had lower power across frontal and temporal-central ROIs compared to data processed with the **Traditional Pipeline** (Figure 2.8).

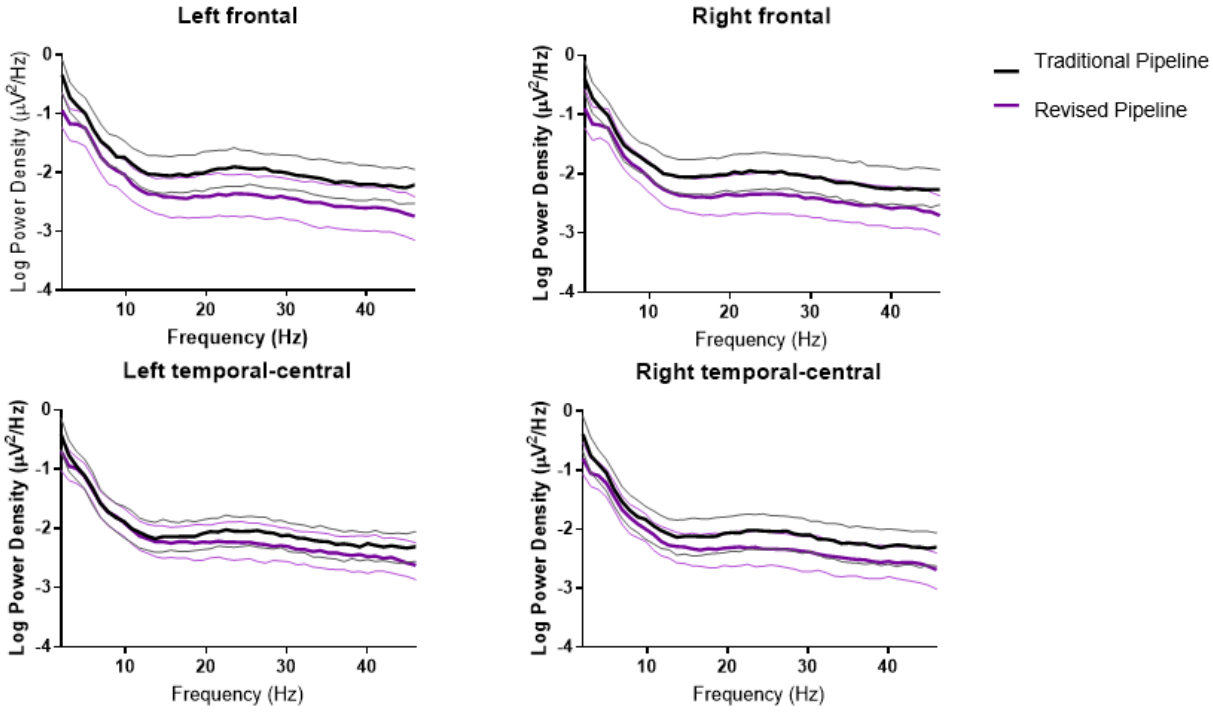

Figure 2.8. *PSD plots for 3-month data processed in parallel using the* ***Traditional Pipeline*** *and the* ***Revised Pipeline****. Log power density plots at frontal and temporal-central ROIs at 3-months (median and interquartile range). Black =* ***Traditional Pipeline****; purple =* ***Revised Pipeline****.*

2.4 Discussion

The goals of this chapter were two-fold: first, I wanted to highlight methodological considerations in choosing each EEG pre-processing step; and secondly, I wanted to introduce a processing pipeline that addressed issues of data loss, EMG, and volume conduction in high-density infant EEG data. In the prior infant EEG literature, traditional methods for removing EMG and other artifacts from initially short recordings have often resulted in both significant data loss and subject loss. Manual data cleaning is highly subjective and time consuming, making it difficult for data processing to be reproduced reliably between multiple researchers, even on the same dataset. The **Revised Pipeline** was developed with the goals of reducing data loss from infant EEG as well as establishing reproducible methods to enable possible future replication studies.

After manual data cleaning, **Traditional Pipeline** retained 47-48% of good data duration. Prior EEG studies in 3-6 months old infants have reported similar retention rate for good data duration (Tierney *et al.*, 2012; Levin *et al.*, 2017). In EEG data from 3-months old infants, Levin *et al* retained 42-45% of clean data duration (Levin *et al.*, 2017). In EEG data from 6-months old infants, Tierney *et al* retained 39-47% of clean data duration (Tierney *et al.*, 2012). In contrast, **Revised Pipeline** retained 99.5-100% of good data duration after using ASR in its first-pass data cleaning step. Because ASR removed bad data portions and interpolated them at the component level, very little bad data remains for window rejection. One of the greatest strengths of ASR is its ability remove artifact without compromising the file's length. Thus, while over 52-53% of data length was lost after manual cleaning using **Traditional Pipeline**, only 0-0.5% of data length was lost after ASR cleaning using **Revised Pipeline**. In preserving data length, ASR also increased subject retention (Table 2.1). After manual cleaning with the **Traditional Pipeline**, 11 files were rejected because they had fewer than 37.5 seconds of clean

data remaining. In contrast, the usage of ASR in **Revised Pipeline** allowed all files to be retained since the clean files' lengths were maximally preserved.

	Traditional Pipeline		**Revised Pipeline**	
	3-month	6-month	3-month	6-month
EEG files	81	84	81	84
Channels removed	14 ± 5 (12%)	17 ± 7 (14%)	15 ± 8 (12%)	14 ± 10 (11%)
Channel retained	110 ± 5 (88%)	107 ± 7 (86%)	109 ± 8 (88%)	110 ± 10 (89%)
Duration removed	81 ± 28 s (53%)	80 ± 33 s (52%)	0.8 ± 6 s (0.05%)	0.08 ± 0.3 s (0%)
Duration retained	71 ± 28 s (47%)	73 ± 28 (48%)	151 ± 7 s (**99.5%**)	151 ± 10 s (**100%**)
IC retained	10 ± 2 (range 4-16)	13 ± 3 (range 7-18)	12 ± 2 (range 7-19)	12 ± 2 (range 4-17)
Files rejected	7	4	0	0
File retained	74 (91%)	80 (95%)	81 (100%)	84 (100%)

Table 2.1. *Data quality and file retention after parallel pre-processing with* ***Traditional Pipeline*** *and* ***Revised Pipeline****.*

Overall, ASR proved to be effective in removing EMG and many artifacts from noisy infant EEG data while still preserving underlying neural signals and data length. Preservation of clean data length has many important implications: more data points allows for less drastic data reduction prior to ICA, which enables greater richness in the data to be retained. Data length retention also leads to increased subject retention, which is important in preserving statistical power in infant EEG studies that are constrained by small sample sizes. Compared to manually cleaned data, data cleaned with ASR had lower spectral power especially at higher frequency bands, suggesting a reduction in high frequency noise. The automation of ASR and ICLabel also support standardization and reproducibility in data processing between multiple researchers and across different research sites. The fast processing speed for ASR allows for efficient and quick data cleaning of large datasets with hundreds of files. One caveat is that ASR should only be used with continuous EEG data, since its algorithm for selecting reference calibration data and its 1-second sliding windows are both not compatible with stimulus-locked tasks.

Chapter 3: Functional connectivity during language processing in 3-month-olds infants at familial risk for ASD

3.1 Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder defined by impairments in social communication skills and the presence of restricted and repetitive patterns of behavior and interests (American Psychiatric Association, 2013). While ASD is often not diagnosed until 3-4 years of age, numerous prospective studies have characterized the emergence of behavioral symptoms between the first and second year of life (Jones *et al.*, 2014). The prospective infant-sibling study design enables infants with familial risk for ASD, defined by having at least one older sibling with ASD, to be studied from early infancy before they exhibit any overt behavioral symptoms. While 1-2% of children from the general population have ASD, as many as 20% of FR children meet criteria for ASD (Sumi *et al.*, 2006a; Elsabbagh & Johnson, 2010; Ozonoff *et al.*, 2011; Messinger *et al.*, 2015). Among FR infants who do not develop ASD, approximately 30% have atypical development by 3 years of age, including language delay and global developmental delay (Messinger *et al.*, 2013; Landa *et al.*, 2013; Ozonoff *et al.*, 2014; Charman *et al.*, 2017). Compared to low risk (LR) infants who do not have an older sibling with ASD, FR infants often exhibit deficits in motor, social communication, and language domains between the first and second years of life (Jones *et al.*, 2014). As early as 12 months of age, FR infants with ASD outcomes already exhibit lower receptive and expressive language compared to infants who do not have ASD (Mitchell *et al.*, 2006; Ozonoff *et al.*, 2014). Given that language is a shared affected domain among many FR children with ASD and other neurodevelopmental delays, studying the neural networks underlying language processing in early infancy in these children can provide valuable insight into the exact timing and mechanisms underlying atypical developmental trajectories. Since nearly a third of children with

ASD are minimally verbal (Tager-Flusberg & Kasari, 2013), identifying early neural markers of ASD and language delay is important in guiding delivery of early, targeted intervention while the brain still has the most plasticity for language learning.

Prior neuroimaging and electroencephalography (EEG) studies of FR infants have identified differences in early neural circuit development that precede behavioral ASD symptoms. Structural neuroimaging studies of FR infants with a later ASD diagnosis have identified abnormalities in developmental trajectories of cortical surface area and white matter pathways at 6-24 months that correlate with ASD severity at 24 months (Wolff *et al.*, 2012; Wolff *et al.*, 2015; Hazlett *et al.*, 2017). Functional neuroimaging of FR infants has revealed patterns of resting state connectivity at 6 months that relate to ASD symptoms at 24 months (Emerson *et al.*, 2017). Compared to MRI, EEG affords both practical and scientific advantages in the study of early brain development. EEG is scalable, portable, developmentally sensitive and has a relatively high tolerance for motion, making it an ideal modality to study infants. Unlike functional MRI, EEG is also a direct measure of electrophysiological brain activity, with sufficient temporal resolution to observe neural oscillations that change with cortical maturation (Uhlhaas *et al.*, 2010). Differences between FR and LR infants in spontaneous EEG spectral power have been observed in the first year of life (Tierney *et al.*, 2012; Gabard-Durnam *et al.*, 2015; Levin *et al.*, 2017). In a recent study, Levin and colleagues reported a pattern of reduced spontaneous frontal power in FR infants across multiple frequency bands, with frontal alpha power at 3 months of age predicting expressive language at 12 months (Levin *et al.*, 2017). Using task-based paradigms in older infants, several EEG studies have found differences in functional connectivity between FR infants and LR infants at 12-14 months that related to ASD symptoms and diagnosis at 3 years (Orekhova *et al.*, 2014; Righi *et al.*, 2014; Keehn *et al.*, 2015; Haartsen *et al.*, 2019). Notably, Righi and colleagues found that 12-month-old FR infants with ASD showed decreased EEG coherence between frontal and temporal-parietal regions during

exposure to speech sounds, compared to FR infants without ASD. Regardless of ASD outcome, FR infants had lower functional connectivity than LR infants at 12 months of age, but not at 6 months (Righi *et al.*, 2014). Although language is a core early developmental domain related to ASD, prior infant sibling studies have not examined language processing in infants younger than 6 months.

In the context of the previously published studies, a fundamental question that remains unanswered is whether there are differences in early neural network connectivity during language processing that predict later ASD and language outcomes. Through an ongoing prospective study of early brain development in FR infants, here we use EEG to study neural synchrony and connectivity during language processing at 3 months of age, well before behavioral indices of atypical language or social communication unfold. Language processing was measured as part of an auditory statistical learning (ASL) paradigm. In the application of ASL to word segmentation, infants use statistical associations between syllables to identify word boundaries implicitly. The ability to segment words from continuous speech using transitional probabilities (i.e., the greater co-occurrence of syllables within words than between words) has been documented in infants throughout the first year of life, from newborns (Flo *et al.*, 2019) to older infants (Saffran *et al.*, 1996; Aslin *et al.*, 1998; Saffran, 2001). ASL is essential to learning language (Jusczyk, 2002; Thiessen & Saffran, 2003; McNealy *et al.*, 2010). In typical development, speech segmentation ability at 7.5-12 months of age predicts later word production at two years and preschool language skills (Newman *et al.*, 2006).

For our primary question, we asked: Can EEG measures of neural synchrony and connectivity during language processing (based on spectral power and phase coherence) differentiate 3-month-old infants based on later ASD symptoms? As a follow-up, we explored whether EEG measures that differentiated groups also relate to 18-month language ability and ASD symptom severity. We hypothesized that the ASD-concern group would exhibit decreased

left-hemisphere connectivity during language processing and that atypical connectivity patterns found in the ASD-concern group at 3 months would relate to lower language ability and higher ASD symptoms at 18 months. As a scalable measure of aberrant brain development, functional connectivity has potential to guide early prediction and stratification of risk for language delay and ASD in this vulnerable population.

3.2 Materials and Methods

Participants

Infant participants were recruited to be part of a longitudinal study on early biomarkers of ASD, as part of the UCLA Autism Center of Excellence (ACE; NICHD 2P50HD055784-08). FR was defined by having at least one older sibling with a documented ASD diagnosis. Our LR group included infants who did not have any siblings with ASD. Exclusion criteria for the LR group included: 1) first or second-degree relative with ASD or other neurodevelopmental disorder (based on parental report), 2) history of any neurological syndromes or major birth trauma, and 3) gestational age below 37 weeks. Informed consent was obtained from parents of participants prior to assessment under protocols approved by the UCLA Institutional Review Board (IRB). To be included in this study, participants had to have usable EEG data at 3 months of age. A total of 74 participants (40 FR, 34 LR) completed the EEG task at 3 months; two additional children (1 FR, 1 LR) attended the 3-month visit but did not complete this task. There were no sex differences between the FR (16 female, 24 male) and LR (13 female, 21 male) groups. There were no differences in family income between the FR and LR groups, with the majority of families earning above $100,000 in annual income (45% of FR group, 62% of LR group). The majority of enrolled families had maternal education of at least college-level (75% of FR group, 91% of LR group), while the LR group had more mothers who completed graduate school (68%) compared to the FR group (25%).

Behavioral Measures

Several behavioral measures were collected at the 18-month visit.

Developmental abilities. Developmental ability was measured with the Mullen Scales of Early Learning (MSEL; Mullen, 1995). The MSEL is a standardized, norm-referenced developmental assessment that provides an overall index of ability, the Early Learning Composite (M=100, SD=15), and subscale scores for Receptive Language, Expressive Language, Visual Reception, Fine Motor, and Gross Motor skills (M=50, SD=10). We generated a nonverbal score by averaging the Visual Reception and Fine Motor t-scores, and a verbal score based upon the average of Receptive and Expressive Language t-scores.

Language skills. Language was assessed using the MSEL, as previously described, as well as using the MacArthur Communicative Development Inventory Words and Gestures checklist (CDI; Fenson *et al.*, 2007). The CDI is a standardized parent-report questionnaire used to track a child's emerging language and communication skills. On the CDI, parents are asked to select the number of words comprehended and words produce by their children from a 396-item vocabulary checklist.

ASD symptoms. ASD symptomatology was measured using the Autism Diagnostic Observation Scale Toddler Module (ADOS-T; Luyster *et al.*, 2009). The ADOS-T grouped children into 3 ranges of ASD concern: Little-to-No Concern, Mild-to-Moderate Concern, and Moderate-to-Severe Concern. A calibrated severity score (ADOS-T CSS), ranging from 1 to 10, was calculated based on ADOS-T overall score (Esler *et al.*, 2015). Based on ADOS-T CSS at 18-month, participants were divided into ASD-concern (CSS > 4) and No-ASD-concern (CSS < 4) groups. In this study, the ASD-concern group consisted of children whose scores on the ADOS-T originally placed them in Mild-to-Moderate and Moderate-to-Severe Concern groups. MSEL and ADOS-T data were available for 63 participants (36 FR, 27 LR; 14 ASD-concern, 49

No-ASD-concern); CDI data were available for 53 participants (30 FR, 23 LR; 10 ASD-concern; 43 No-ASD-concern).

EEG Stimuli

In this ASL paradigm, infants were passively exposed to a continuous stream of concatenated syllables that consisted of four different trisyllabic pseudo-words (Figure 3.1, from McNealy *et al.*, 2006; McNealy *et al.*, 2010; McNealy *et al.*, 2011). Pseudo-words were constructed from a set of 12 syllables and presented in random order, such that the transitional probability of hearing two adjacent syllables within words was 100%, while the transitional probability of hearing two adjacent syllables across word boundaries was 33% (McNealy *et al.*, 2006). These four pseudowords ("pabiku", "daropi", "tibudo, "golatu") were presented in a continuous speech stream without additional pauses between words. No pseudo-words were consecutively presented. Infants can implicitly learn the transitional probabilities between syllables and use these probabilities as cues to word boundaries.

Figure 3.1. *Pseudo-words presented during the auditory language processing task.*

EEG Data Acquisition

EEG was recorded using a 128-channel HydroCel Geodesic Sensor Net containing Ag/AgCl electrodes and sponges with saline electrolyte (Electrical Geodesics Inc., Eugene, OR). To improve infants' comfort, four of the electrodes originally placed below and lateral to the eyes (channels 125-128) were removed from the net. Placement of electrodes conformed to the International 10-20 System (Jasper, 1958). Net Amps 300 amplifier and Net Station 4.5.7

software were used to record EEG (Electrical Geodesics Inc., Eugene, OR). Data were filtered online during recording using an analog band-pass elliptical filter between 0.1 and 100 Hz. EEG was sampled at 500 Hz. Data were referenced online during recording to a vertex reference (channel Cz). The session was video-recorded to assist in subsequent data processing of the child's movement and behavior during the session. Testing procedures were conducted as per previously established protocols in infants and young children with developmental disabilities (Webb *et al.*, 2015). Approximately 2.5-minutes of continuous EEG was acquired for each participant during the presentation of the auditory stimuli, while infants sat on their parent's lap. Parents were instructed to not talk to the infants during the EEG recording, to help infants maintain upright sitting posture, and to prevent infants from touching the EEG net.

EEG Data Pre-Processing

EEG data were exported as MATLAB (Mathworks, Natick, MA) files and processed offline using the EEGLAB (v14.1.1b) signal processing environment (Delorme & Makeig, 2004), running under MATLAB R2017a (Figure 3.2). Continuous data were extracted for the auditory language processing task. Data were high-pass filtered at 1.5 Hz using Blackman window finite impulse response (FIR) filter, with 1 Hz transition bandwidth and filter order of 2750. Due to the presence of 60-Hz line noise in the EEG recording, the data were further low-pass filtered at 50 Hz using Blackman window FIR filter, with 20 Hz transition bandwidth and filter order of 138.

Artifact subspace reconstruction (ASR) was applied to continuous data using EEGLAB plugin *clean_rawdata* (Mullen *et al.*, 2015; Chang *et al.*, 2018). ASR was run using fixed random-access memory (RAM) allocation of 3200 megabytes. Within each 1-second sliding window, data portions with variance greater than 8 standard deviations from the calibration data were rejected and interpolated. Channels were rejected if they had flatline duration longer than 3 seconds. Channels with correlations less than 0.75 to its reconstruction based on other

channels in the given time window were also removed. Time windows where 25% of channels exceed z-scores of -3.5 to 5.5 on root-mean-square thresholding test were rejected. After cleaning with ASR, data from the original 128-channel montage was interpolated to a 25-channel montage based on the International 10-20 System using *interp_mont* plugin (Desjardins, 2010; Jasper, 1958).

Adaptive Mixture Independent Component Analysis (AMICA) (Palmer *et al.*, 2006; Palmer *et al.*, 2008) was performed and outputted 25 independent components (IC) per subject. The AMICA algorithm was used because it outperforms other ICA methods in scalp channels mutual information reduction and identifies the maximal number of dipolar brain IC (Delorme *et al.*, 2012). AMICA learns the data structure without preset assumptions; prior studies have utilized AMICA in processing both infant (Piazza *et al.*, 2016) and adult EEG data (Hsu *et al.*, 2018). ICs were visually inspected for remaining artifacts: eye blinks, saccades, BCG (ballistocardiograms), and electromyogram (EMG). Automatic IC classification was applied using EEGLAB plugin ICLabel to aid in manual review and rejection of non-neural IC. IClabel accounted for each IC's activity power spectrum and scalp topography in its classification algorithm (Pion-Tonachini *et al.*, 2017; Pion-Tonachini *et al.*, 2019).

Cleaned data were epoched into 3-second segments. Because phase coherence is sensitive to file length, the first 29 epochs of clean data were selected across all subjects for further analysis. Data were passed through a Laplacian spatial filter to account for possible volume conduction. Specifically, a spherical spline Laplacian transform with head-circumference correction was applied to transform the cleaned data into current source density (Kayser & Tenke, 2006).

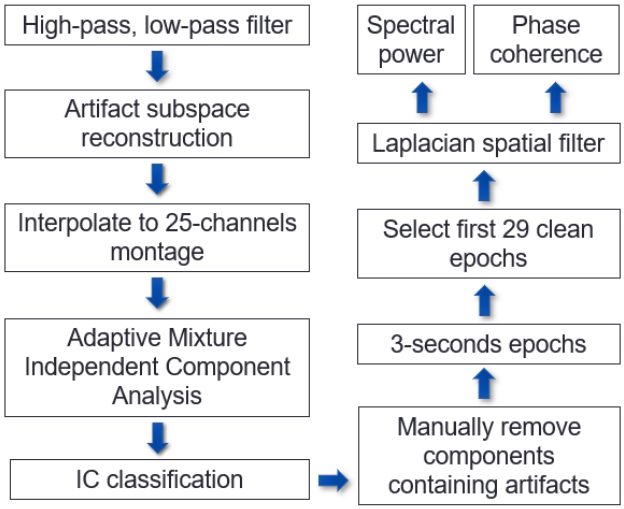

Figure 3.2. *Pre-processing pipeline for continuous EEG data. IC = independent component.*

EEG data quality

Averaged across 74 files from all subjects, ASR rejected 12% of channels (15 ± 8 channels) and 0.01% of file length (0.8 ± 5.8 seconds). The retained neural components were classified as "brain components" at 32-95% using IClabel. After manual removal of IC with artifacts, an average of 12 ± 2 (range: 7-19) brain components were retained per subject. There was no difference in data quality between ASD-concern and No-ASD-concern groups. All subjects who had EEG during the language processing task were used in the final analyses (100% subject retention).

EEG spectral power analysis

The current study focused on spectral power and phase coherence in the theta (4-6 Hz) and alpha (6-12 Hz) frequency bands. Oscillations in the theta band are linked with modulatory increase in attention and memory. Activity in the alpha band is developmentally sensitive and reflects underlying thalamocortical connectivity. Binned spectral power in the theta (4-6 Hz) and

alpha (6-12 Hz) frequency bands was calculated using Welch's method in frontal and temporal-central regions of interest (ROIs) across the scalp, using custom scripts written in MATLAB as per prior protocol in the Jeste Lab (McEvoy *et al.*, 2015). Left frontal ROI included channels F3, F7, and F9; left temporal-central ROI included channels T7, T9, and C3; right frontal ROI included channels F4, F8, and F10; right temporal-central ROI included channels T8, T10, and C4 (Figure 3.3a). For each 256-sample segment, the Fast Fourier Transform (FFT) was calculated on 128-point Hamming windows with 50% overlap. Absolute power was calculated by summing power estimates at every 0.5 Hz increment within each frequency band. Relative power was calculated by dividing absolute power at each frequency band by the total absolute power across 2-50 Hz. All power values were log base 10 transformed.

EEG coherence analysis

Connectivity was measured in the form of magnitude-square phase coherence across the 3-second period, in the theta (4-6 Hz) and alpha (6-12 Hz) frequency bands. Coherence is a measure of synchronization between two signals of the same frequency, and it quantifies the extent to which they share a constant oscillating frequency and phase difference. Neuronal sources share information by oscillating coherently (Fries, 2005). Within a frequency band, phase coherence is based on the correlation in phases between two electrodes' signals. Coherence for all channel pairs in each frequency band was calculated using EEGLAB *newcrossf* function, with a window size of 1024 samples (Delorme & Makeig, 2004). Each electrode-pair coherence value was calculated by first averaging coherence values across all time bins within each frequency bin, and then the mean coherence values was calculated by averaging within each frequency bin. Coherence values from all electrode pairs was compiled into a 25 x 25 matrix, such that each element (i,j) in each matrix represented the averaged

coherence between channel i and channel j, for a given subject. Phase coherence was calculated between 18 electrode pairs in language networks (fronto-temporal, fronto-central). Left frontal: F3, F7, F9; left temporal: T7, T9; left central: C3; right frontal: F4, F8, F10; right temporal: T8, T10; right central: C4 (Figure 3.3b). Only intrahemispheric fronto-temporal and fronto-central connections within putative language networks were chosen in order to examine functional networks supporting language processing.

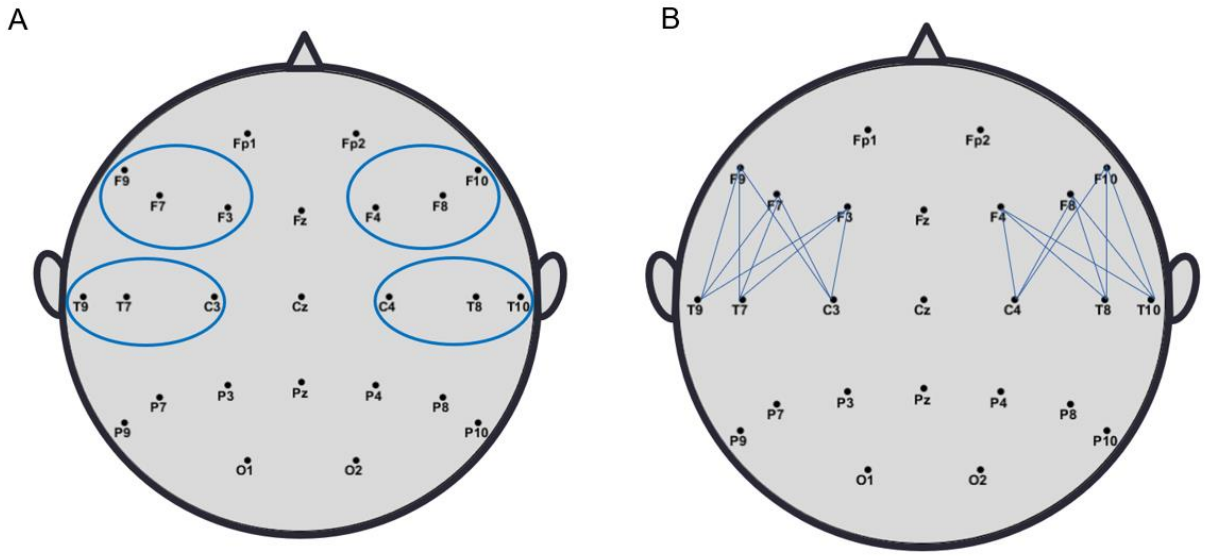

Figure 3.3. *Scalp map used in power and coherence analyses.* ***Figure 3.3a****. Frontal and temporal-central ROIs used in spectral power analysis.* ***Figure 3.3b****. Language network electrode pairs examined in phase coherence analysis.*

Statistical analysis

Descriptive analyses. Behavioral and EEG measures were plotted to check distributional assumptions. To follow-up on asymmetric distributions observed, normality of the data was more formally checked by Shapiro-Wilks tests. Prevalent deviations from normality was detected, which motivated our choice of nonparametric modeling methodology (via permutation

tests, Mann Whitney U-test and Spearman's rank-order correlations). Nonparametric tests were robust to deviations from distributional assumptions for the primary and secondary analyses. Nonparametric Mann-Whitney U-test was used to compare language scores and ADOS-T scores between ASD-concern and No-ASD-concern groups. Log power density (PSD) for ASD-concern and No-ASD-concern groups was graphed for qualitative characterization. At each point of the power spectrum, each group's median and 25-75th percentile was graphed.

Primary analyses. Nonparametric Mann-Whitney U-test was used to compare mean relative power at each ROI between ASD-concern and No-ASD-concern groups. Nonparametric permutation test was used to compare group mean in coherence values at each electrode pair between the ASD-concern group and the No-ASD-concern group. Within each frequency band, the permutation test shuffled diagnostic group membership of subjects to create new samples under the null hypothesis that the mean coherence between the two groups were the same at the 18 electrode pairs. False discovery rate (FDR) was used to correct for multiple comparisons for all 44 tests in the primary analysis (8 spectral power tests; 36 phase coherence tests) with alpha set at 0.05 (Benjamini & Hochberg, 1995). Effect size was calculated using Hedges' *g* since the ASD-concern and No-ASD-concern groups were dissimilar in size.

Secondary analyses. In a secondary follow-up analysis, nonparametric Spearman's rank-order correlations were used to relate 3-month measures that differentiated groups to 18-month MSEL expressive and receptive t-scores, CDI words comprehended and produced, CDI receptive advantage, and ADOS-T overall score. Spearman's correlations described the strength of association between two ranked variables. Shapiro-Wilk test, Levene's test, Mann-Whitney U-test, and Spearman's correlations were performed with the Statistical Package for Social Sciences (IBM SPSS Statistics, version 25).

3.3 Results

Outcome groupings

The ASD-concern group (n = 14) consisted of 12 FR and 2 LR infants. The No-ASD-concern group (n = 49) consisted of 24 FR and 25 LR infants (Table 3.1). Outcome groups were not matched in sex distribution; the ASD-concern group was predominantly male (2 female, 12 male) while the No-ASD-concern group had an even sex distribution (21 female, 28 male). There were no differences in family income between the ASD-concern and No-ASD-concern groups, with the majority of families earning above $100,000 in annual income (43% of ASD-concern group, 55% of No-ASD-concern group). The majority of enrolled families had maternal education of at least college-level (72% of ASD-concern group, 86% of No-ASD-concern group), while the No-ASD-concern group had more mothers who completed graduate school (49%) compared to the ASD-concern group (29%).

Developmental testing

At 18 months, FR infants had significantly lower MSEL standard scores, MSEL nonverbal t-scores, MSEL verbal t-scores, MSEL receptive language t-scores, and CDI words comprehended compared to LR infants. There was no significant difference between FR and LR in MSEL expressive language t-score, CDI words produced, or ADOS-T overall score and ADOS-T CSS. At 18-months, the ASD-concern group had significantly lower MSEL standard scores, MSEL nonverbal t-scores, MSEL verbal t-scores, MSEL receptive language t-scores, MSEL expressive language t-scores, CDI words comprehended, and CDI words produced compared to the No-ASD-concern group. Compared to the No-ASD-concern group, the ASD-concern group also had higher ADOS-T overall scores and higher ADOS-T CSS. Mean MSEL scores and ADOS-T scores are presented in Table 3.1; mean receptive language, expressive language, and receptive advantage scores are presented in Table 3.2.

	Familial-risk (FR)	Low-risk (LR)	ASD-concern	No-ASD-concern
n	36	27	14 (12 FR/ 2 LR)	49 (24 FR/ 25 LR)
Sex	13F/23M	10F/17M	2F/ 12 M	21F/ 28M
MSEL ELC standard score	89.6 ± 14.5* (52-116)	102.0 ± 14.9* (56-126)	74.6 ± 10.7* (52-85)	100.7 ± 11.7* (82-126)
MSEL nonverbal t-score	45.9 ± 7.5* (24.5-64.0)	52.0 ± 8.9* (25.0-66.0)	41.0 ± 9.7* (24.5-59.5)	50.7 ± 7.0* (34.0-66.0)
MSEL verbal t-score	43.1 ± 9.6* (21.0-63.5)	49.7 ± 10.3* (25.0-66.0)	31.9 ± 5.9* (21.0-40.5)	50.0 ± 7.4* (35.5-66.0)
ADOS-T overall score	7.4 ± 5.2 (1-18)	5.0 ± 4.3 (0-18)	14.3 ± 2.7* (10-18)	4.1 ± 2.5* (0-9)
ADOS-T CSS	3.2 ± 1.9 (1-7)	2.3 ± 1.4 (1-7)	5.6 ± 1.0* (4-7)	2.0 ± 0.8* (1-3)

Table 3.1. *Developmental outcomes and ASD symptoms at 18 months. Abbreviations: female (F); male (M); calibrated severity score (CSS). Mean, standard deviation, and range scores for each group are displayed. * Significant group differences respectively between FR versus LR, or ASD-concern versus No-ASD-concern.*

		Familial-risk (FR)	Low-risk (LR)	ASD-concern	No-ASD-concern
Receptive Language	MSEL *(t-score)*	40.0 ± 12.4* (20-63)	50.6 ± 14.7* (20-72)	27.6 ± 5.4* (20-37)	49.4 ± 12.3* (30-72)
	CDI *(words comprehended)*	143.3 ± 90.2* (21-370)	215.1 ± 116.5* (10-392)	57.0 ± 38.5* (10-150)	201.8 ± 99.8* (21-392)
Expressive Language	MSEL *(t-score)*	46.2 ± 9.8 (19-68)	48.9 ± 9.5 (27-65)	36.1 ± 9.4* (19-51)	50.6 ± 7.1* (33-68)
	CDI *(words produced)*	55.8 ± 75.8 (0-342)	63.0 ± 66.2 (3-249)	9.1 ± 11.8* (0-37)	70.8 ± 74.2* (2-342)

Table 3.2. *Language profiles at 18 months on the MSEL and CDI. Mean, standard deviation, and range scores for each group are displayed. * Significant group differences respectively between FR versus LR, or ASD-concern versus No-ASD-concern.*

A range of language ability was observed in both ASD-concern and No-ASD-concern groups at 18-months. The ASD-concern group included 12 infants whose scores fell below the average range on the MSEL (verbal t-score < 40; Charman *et al.*, 2017). While most infants in the No-ASD-concern group had typical language ability, 3 infants showed language delays (2 FR, 1 LR). On the CDI, the ASD-concern group's words production ranged from 0 to 37 words, whereas the No-ASD-concern group's words production ranged from 3 to 342 words.

Spectral power: ASD-concern versus No-ASD-concern

The ASD-concern group did not differ from the LR group in theta and alpha relative power at frontal and temporal-central ROIs (Mann-Whitney U-test $p > 0.05$) (Figure 3.4).

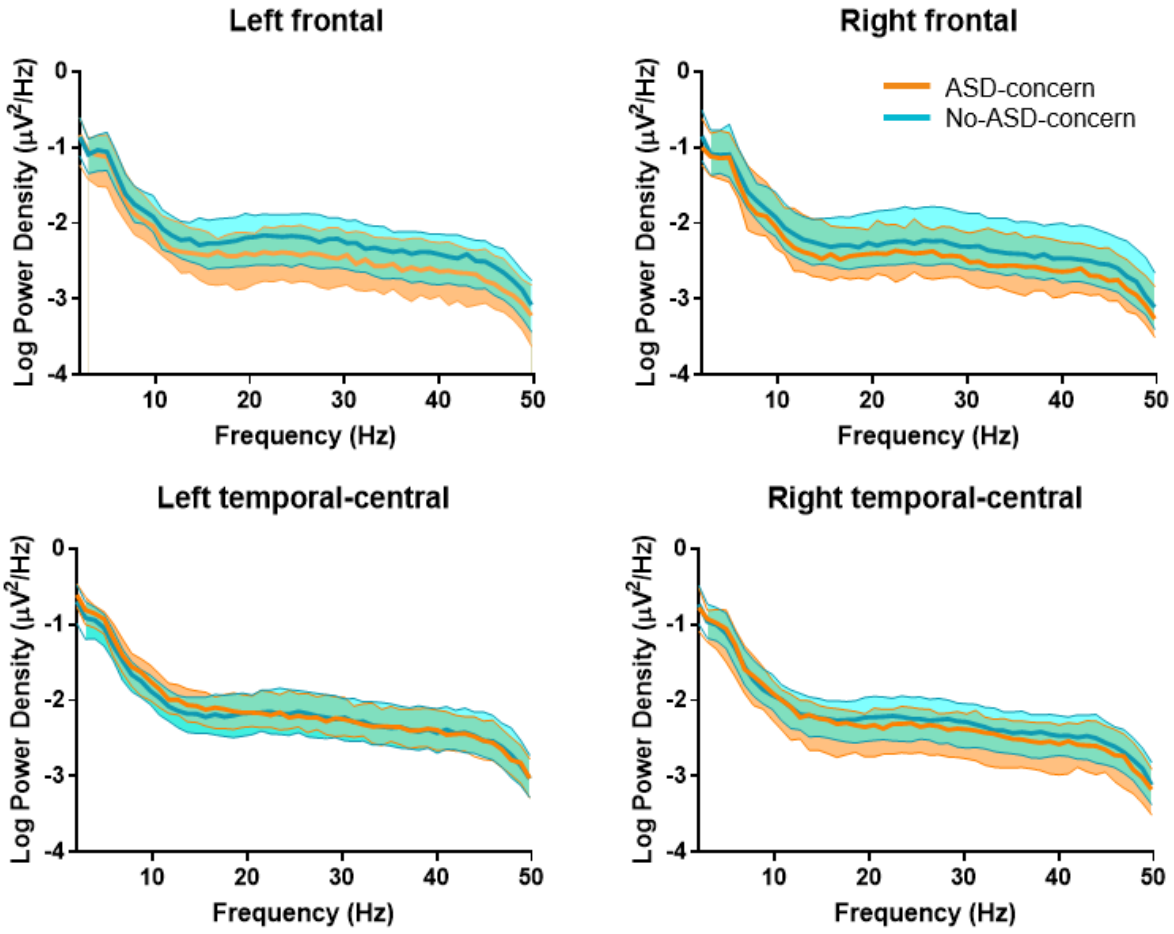

Figure 3.4. *PSD plots for ASD-concern and No-ASD-concern groups. Log power density plots at frontal and temporal-central ROIs at 3-months (median and interquartile range). Orange = ASD-concern; teal = No-ASD-concern.*

Phase coherence: ASD-concern versus No-ASD-concern

Theta phase coherence at left frontal-central electrode pair (F9-C3) differentiated the ASD-concern group from the No-ASD-concern group (Hedges' *g* = 0.6, p = 0.031), where the No-ASD-concern group had higher coherence (Figure 3.5a). At the same frontal-central electrode pair (F9-C3), alpha phase coherence also differentiated the ASD-concern group from

the No-ASD-concern group (Hedges' *g* = 0.7, p = 0.017), where the No-ASD-concern group similarly had higher coherence (Figure 3.5b). After FDR was applied and p-values were adjusted for multiple comparisons, the group differences in theta and alpha bands were not significant (adjusted p-values > 0.05). At the F9-C3 electrode pair for both theta and alpha coherence, the ASD-concern group was tightly clustered together while the No-ASD-concern group had a wide range in coherence values. Theta and alpha coherence at the other electrode pairs did not differentiate risk groups (p > 0.05).

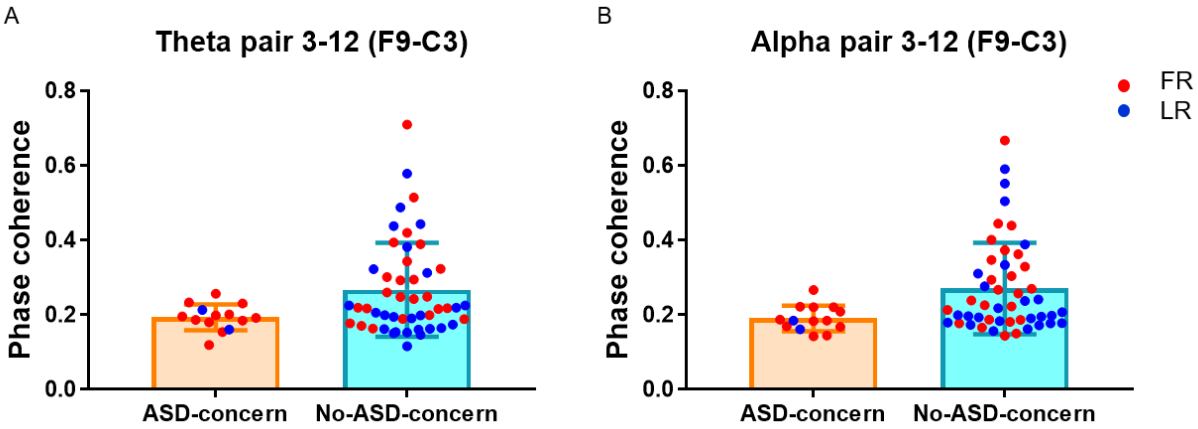

Figure 3.5. *Phase coherence for ASD-concern and No-ASD-concern groups.* ***Figure 3.5a.*** *Theta phase coherence at left frontal-central connection (F9-C3) differentiated the ASD-concern group from the No-ASD-concern group.* ***Figure 3.5b.*** *Alpha phase coherence at F9-C3 differentiated the ASD-concern group from the No-ASD-concern group.*

Phase coherence and behavioral outcomes:

Across all participants and within risk groups, theta coherence at the F9-C3 electrode pair did not correlate with 18-month language ability or ASD symptoms (p > 0.05). Alpha coherence at the F9-C3 electrode pair correlated with 18-month CDI words produced, across all participants (r = 0.315, p = 0.022) (Figure 3.6). Across all participants and within risk groups,

alpha coherence at the F9-C3 electrode pair did not correlate with 18-month ASD symptoms or other language measures ($p > 0.05$).

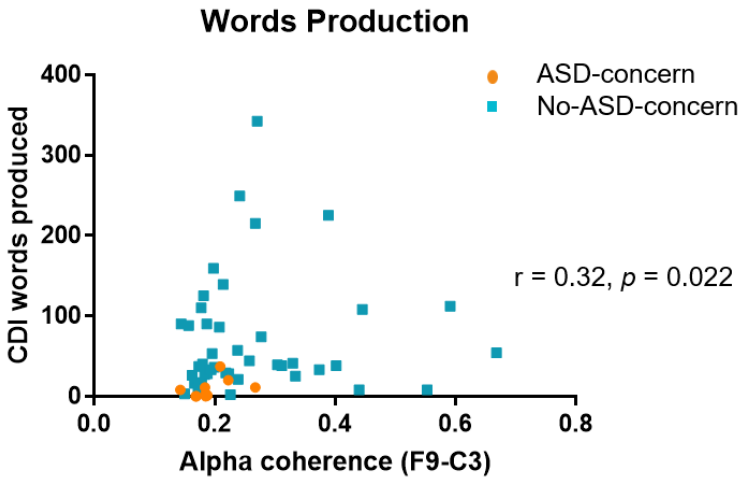

Figure 3.6. *Correlations between 3-month left frontal-central alpha coherence and 18-month words produced (CDI).*

3.4 Discussion

The present study is the first to examine spectral power and functional connectivity during language processing in 3-month-old infants with familial risk for ASD, with a focus on auditory statistical learning. While we found no significant differences based on risk group or outcome in EEG power, we did identify reduced connectivity in the left fronto-central area during language processing in infants who showed clinically significant ASD symptoms at 18 months. Specifically, left fronto-central phase coherence in both theta and alpha frequency bands was lower in the ASD-concern group compared to the No-ASD-concern group. Notably, this difference was driven by a subset of infants in the No-ASD-concern group who had very high coherence during language processing. Across risk groups, alpha coherence at 3 months correlated with word production at 18 months. Taken together, these findings support the

hypothesis that early differences in neural synchrony may lay a foundation for neurodevelopmental impairments in infants at risk for ASD, and these differences are able to be robustly quantified using EEG.

Language outcomes

Participants in our study exhibited language profiles at 18-months of age that were consistent with prior characterization of language deficits in infant-siblings and toddlers with ASD (Charman *et al.*, 2003; Luyster *et al.*, 2007; Luyster *et al.*, 2008; Ellis Weismer *et al.*, 2010; Mitchell *et al.*, 2006; Ozonoff *et al.*, 2014; Levin *et al.*, 2017). Most children who had ASD symptoms at 18 months also met criteria for language delay, with deficits in both receptive and expressive language. Of note, a small subset of children in the No-ASD-concern group also met criteria for language delay at 18 months; this may reflect broader autism phenotype in familial-risk children and specific language impairment in low-risk children (Messinger *et al.*, 2013; Landa *et al.*, 2013; Ozonoff *et al.*, 2014; Charman *et al.*, 2017; Tomblin *et al.*, 1997).

Auditory statistical learning

To learn word meanings, infants must first be able to segment individual words from continuous speech. Infants may utilize multiple cues to aid speech segmentation, including language-general cues like statistical word boundaries (Saffran *et al.*, 1996) and language-specific cues like stress patterns and phonotactic properties (Houston *et al.*, 2000; Houston *et al.*, 2004; Mattys & Jusczyk, 2001). Auditory statistical learning has been demonstrated behaviorally in infants as early as 5.5 months of age (Johnson & Tyler, 2010) as well as in newborns through functional neuroimaging (Flo *et al.*, 2019). In typically developing infants, speech segmentation ability relates to both word production at two years and preschool language skills (Newman *et al.*, 2006). Infants with lower speech segmentation ability may have

more trouble parsing words from continuous speech, which can contribute to deficits in word learning and result in language impairment. Auditory statistical learning is crucial for language learning and remains relevant beyond infancy. In high-functioning adolescents with ASD, impairments in auditory statistical learning measured with functional neuroimaging are related to impairments in communication (Scott-Van Zeeland *et al.*, 2010). Among adolescents with ASD who have a wide range of cognitive ability, auditory statistical learning assessed with event-related potentials relates to receptive language ability (Arnett *et al.*, 2018).

Reduced coherence during language processing in infants at risk for ASD

Our study identified unique connectivity profiles during language processing differentiated infants based upon emerging ASD symptoms. There may be a threshold over which connectivity during language processing promotes typical language development. Regardless of familial risk status, only infants in the No-ASD-concern group had high coherence values above a certain threshold. It is possible that infants who have stronger functional connectivity in language networks during exposure to speech streams are likely to perform better at word segmentation in novel contexts, and word segmentation ability in infancy lays the foundation for future word learning and production. Increased left-hemispheric connectivity during language processing may reflect different underlying neural mechanisms. Since coherence was calculated over the entire exposure phase of the paradigm, higher left frontal-central connectivity may reflect increased synchronization of neural networks in response to speech sounds. Infants' heightened attention to speech input may result in increased synchronization of left-hemispheric language networks. In typically developing infants, speech processing at 3 months of age is supported by activation of left temporal and left frontal cortex (Dehaene-Lambertz *et al.*, 2002; Dehaene-Lambertz *et al.*, 2006; Dehaene-Lambertz *et al.*, 2010; Shultz *et al.*, 2014). As language networks become more specialized between 1-4

months, infants' left temporal cortex gradually decreases response to non-speech sounds (Shultz *et al.*, 2014). Increased connectivity may also reflect increased synchronization of language networks during ASL. Prior ASL studies have shown that some infants are more successful than others in segmenting words from continuous speech (Saffran *et al.*, 1996; Aslin *et al.*, 1998; Saffran, 2001; Flo *et al.*, 2019). However, by measuring connectivity averaged across the exposure phase, this study did not directly test which infants were able to segment the speech stream during the exposure phase. Moreover, underconnectivity during language processing may both precede and predispose at-risk infants to later language impairment and ASD symptoms. Coherence values for the ASD-concern group clustered tightly together, suggesting that reduced coherence during language processing is a potential risk marker of atypical development.

While phase coherence differentiated risk groups, spectral power did not differ between infants with and without ASD symptoms at 18 months. To date, our study and Levin *et al* are the only two studies that have characterized EEG spectral power in 3-months old infants at familial risk for ASD. Results from our language processing study are complementary to the resting state findings by Levin et al, as we similarly did not find a difference in spectral power between infant groups stratified based on ASD outcomes; likewise, 3-month power in our study did not relate to 18-month language outcomes (Levin *et al.*, 2017).

Methodological strengths

Given the considerable data loss that hamper most infant EEG studies, we prioritized the development of a pre-processing pipeline designed to clean and retain data from EEG recordings that contained considerable movement artifact. Prior EEG studies of infant siblings retained 32-51% of data length after manual data cleaning (Tierney *et al.*, 2012; Levin *et al.*, 2017; Orekhova *et al.*, 2014; Haartsen *et al.*, 2019). By utilizing artifact subspace reconstruction

to remove and interpolate noisy windows of data, we were able to retain close to 100% of data length for all subjects after data cleaning. We also applied a Laplacian filter to our data, which removed widespread volume-conducted artifact from the scalp signal. Volume conduction presents a fundamental problem in high density EEG recordings, with neighboring channels appearing to have the same synchronous activity.

Limitations and future directions

Similar to many prior EEG studies of familial-risk infants, our small sample size limits statistical power, and it will be critical to replicate this study in a new cohort of infants. Most prior EEG studies in infants at high familial risk have not been replicated [only the British Autism Study of Infant Siblings has attempted replication in a new cohort (Orekhova *et al.*, 2014; Haartsen *et al.*, 2019)]. It remains unclear whether the discrepancies in reported findings across studies result from differences in data-analytic methods or rather reflect true trait differences in the sample. While our group difference findings in coherence did not survive correction for multiple comparisons, the crux of this work remain exploratory and hypothesis generating. Furthermore, since the present study only examined functional connectivity during the exposure phase of the word segmentation paradigm, during which infants were getting familiarized with nonsense words from a continuous speech stream, it is unclear whether increased connectivity reflected the infants' ASL or the infants' increased attention to speech stimuli. In future studies, functional connectivity should be examined during both the exposure phase and the test phase of an ASL paradigm in order to determine if increased connectivity during this task indeed provides a direct measure of word segmentation ability.

The present study revealed differences in functional connectivity during language processing can be detected as early as 3 months of age. Future studies should investigate the developmental trajectory of functional connectivity during language processing throughout the

first year of life, as the brain undergoes dramatic structural and functional maturation during this period. As the structural integrity of language networks are known to support language function, it is likely that infants who later develop ASD and language delay will have aberrant trajectories in both structural and functional connectivity of language networks during the first year of life.

Note: Chapters 3 and 4 are separate manuscripts in preparation for publications, thus there are some redundancy in the background and methods sections for these chapters.

Chapter 4: Longitudinal development of functional connectivity during language processing over the first year of life in infants at familial risk for ASD

4.1 Introduction

Autism spectrum disorder (ASD) is a common neurodevelopmental disorder affecting 1-2% of children in the general population (CDC 2014). ASD is often diagnosed by three years of age based on deficits in social communication and patterns of repetitive behaviors and restricted interests (American Psychiatric Association, 2013). Infants who have older siblings with ASD (familial risk or FR) are at increased risk for developing ASD, with rates approaching 20% (Sumi *et al.*, 2006b; Elsabbagh & Johnson, 2010; Ozonoff *et al.*, 2011; Messinger *et al.*, 2015). FR infants who develop ASD often show early behavioral symptoms of ASD between the first and second year of life (Jones *et al.*, 2014). As early as 12-months of age, deficits in both receptive and expressive language are evident among FR infants who develop ASD (Mitchell *et al.*, 2006; Ozonoff *et al.*, 2014). Using noninvasive techniques such as magnetic response imaging (MRI) and electroencephalography (EEG), aberrant brain development in FR infants can be studied during the first year of life, before the emergence of behavioral symptoms, which can, in turn, guide earlier detection of risk and earlier initiation of screening and interventions.

In recent years, many neuroimaging studies of FR infants have consistently characterized altered trajectories in the development of brain structures and neural network connections during the first year of life. In FR infants who later develop ASD, structural imaging studies have identified abnormalities in developmental trajectories of cortical surface area and

white matter pathways at 6-24 months that correlate with ASD severity at 24 months (Wolff *et al.*, 2012; Wolff *et al.*, 2015; Hazlett *et al.*, 2017). In low-risk infants (LR) who do not have older siblings with ASD, the rate of splenium structural development between 6 and 24 months has been shown to predict greater word production at 24-months (Swanson *et al.*, 2017b). Given that language is a shared affected domain among children with ASD and other neurodevelopmental delays, studying neural networks underlying language processing in early infancy can provide valuable insight into the emergence of atypical neurodevelopmental pathways. In a recent study, we identified a pattern of reduced functional connectivity during language processing, as measured by phase coherence, in 3-months old infants who later develop ASD symptoms by 18-months of age (Tran *et al*, in prep). Conversely, infants who had increased phase coherence in left hemispheric language networks at age 3-months showed greater expressive vocabulary at 18 months (Tran *et al*, in prep). In the context of known altered trajectories in structural connectivity development, it is still currently unknown how functional connectivity during language processing develops over the first year of life in infants at familial risk for ASD.

The goals of the current study were two-fold: 1) to examine whether developmental trajectories in phase coherence during language processing across the first year of life can differentiate infants based on later ASD symptoms, and 2) to investigate if change in coherence over the first year of life predicts language and ASD outcomes. Infants were divided into ASD-concern and No-ASD-concern groups based on their ASD symptom profiles at 18-months. The ASD-concern group was expected to maintain decreased left-hemispheric coherence over the first year of life. Atypical coherence trajectories found in the ASD-concern group between 3-12 months were expected to relate to worse language ability and increased ASD symptoms at 18-months.

4.2 Materials and Methods

Participants

Families were recruited to be part of a longitudinal study on early biomarkers of ASD in FR infants, as part of the UCLA Autism Center for Excellence (ACE; NICHD 2P50HD055784-08). FR was defined by having at least one older sibling with an ASD diagnosis. Our LR group included infants who did not have any siblings with ASD. Exclusion criteria for LR group included: (1) first or second-degree relative with ASD or other neurodevelopmental disorder (based on parental report), (2) history of any neurological syndromes or major birth trauma, and (3) prematurity. Informed consent was obtained from parents of participants prior to assessment under protocols approved by the UCLA Institutional Review Board (IRB). 87 subjects (51 FR, 36 LR) completed their visits at 3, 6, 9, and 12 months of age, out of which 85 subjects (50 FR, 35 LR) completed EEG during the language task for at least one timepoint. There were no sex differences between FR (21 female, 29 male) and LR (13 female, 22 male) groups.

Behavioral Measures

Several behavioral measures of ASD symptoms, overall developmental level, and language ability were assessed at 18-months.

Developmental abilities. Developmental ability was measured with the Mullen Scales of Early Learning (MSEL; Mullen, 1995). The MSEL is a standardized, norm-referenced developmental assessment that provides an overall index of ability, the Early Learning Composite (M=100, SD=15), and subscale scores for Receptive Language, Expressive Language, Visual Reception, Fine Motor, and Gross Motor skills (M=50, SD=10). We generated a nonverbal score by averaging the Visual Reception and Fine Motor t-scores, and a verbal score based upon the average of Receptive and Expressive Language t-scores.

Language skills. Language was assessed using the MSEL, as previously described, as well as using the MacArthur Communicative Development Inventory Words and Gestures

checklist (CDI; Fenson *et al.*, 2007). The CDI is a standardized parent-report questionnaire used to track a child's emerging language and communication skills. On the CDI, parents are asked to select the number of words comprehended and words produce by their children from a 396-item vocabulary checklist.

ASD symptoms. ASD symptomatology was measured using the Autism Diagnostic Observation Scale Toddler Module (ADOS-T; Luyster *et al.*, 2009). The ADOS-T grouped children into 3 ranges of ASD concern: Little-to-No Concern, Mild-to-Moderate Concern, and Moderate-to-Severe Concern. A calibrated severity score (ADOS-T CSS), ranging from 1 to 10, was calculated based on ADOS-T overall score (Esler *et al.*, 2015). Based on ADOS-T CSS at 18-month, participants were divided into ASD-concern (CSS > 4) and No-ASD-concern (CSS < 4) groups. In this study, the ASD-concern group consisted of children whose scores on the ADOS-T originally placed them in Mild-to-Moderate and Moderate-to-Severe Concern groups. MSEL and ADOS-T data were available for 71 subjects (44 FR, 27 LR; 17 ASD-concern, 54 No-ASD-concern); CDI data were available for 59 subjects (36 FR, 23 LR; 13 ASD-concern; 46 No-ASD-concern).

EEG Stimuli

In this study, language processing was measured as part of an auditory statistical learning (ASL) paradigm. In the application of ASL to word segmentation, infants use statistical associations between syllables to identify word boundaries implicitly. The ability to segment words from continuous speech using statistical word boundaries has been documented in infants throughout the first year of life, from newborns (Flo *et al.*, 2019) to older infants (Saffran *et al.*, 1996; Saffran *et al.*, 1999; Saffran, 2001). In typical development, speech segmentation ability at 7.5-12 months of age predicts later word production at two years and preschool language skills (Newman *et al.*, 2006).

In our ASL language processing task, infants were passively exposed to a continuous stream of concatenated syllables that consists of four different tri-syllabic pseudo-words (Figure 4.1, from McNealy *et al.*, 2006). Pseudo-words were constructed from a set of 12 syllables and presented in random order, such that the transitional probability of hearing two adjacent syllables within words was 100%, while the transitional probability of hearing two adjacent syllables across word boundaries was 33% (McNealy *et al.*, 2006). No pseudo-words were consecutively presented.

Figure 4.1. *Pseudo-words auditory stimuli presented during language processing task.*

EEG Data Acquisition

EEG was recorded using a 128-channel HydroCel Geodesic Sensor Net containing Ag/AgCl electrodes and sponges with saline electrolyte (Electrical Geodesics Inc., Eugene, OR). To improve each infant's comfort, four of the electrodes originally placed below and lateral to the eyes (channels 125-128) were removed from the net. Placement of electrodes conformed to the International 10-20 System (Jasper, 1958). Net Amps 300 amplifier and Net Station 4.5.7 software were used to record EEG (Electrical Geodesics Inc., Eugene, OR). Data were filtered online during recording using an analog band-pass elliptical filter between 0.1 and 100 Hz. EEG was sampled at 500 Hz. Data were referenced online during recording to a vertical reference in a location equivalent to Cz. The session was recorded with video to assist in subsequent data processing of the child's movement and behaviors during the session. Testing procedures were conducted per previously established protocols in infants and young children with developmental disabilities (Webb *et al.*, 2015). Approximately 2.5-minutes of continuous EEG was acquired for each subject at 3-months during an auditory language processing task. During

the task, infants sat on their parent's lap. Parents were instructed to not talk to the infants during the EEG recording, to help infants maintain upright sitting posture, and to prevent infants from touching the EEG net.

EEG Data Pre-Processing

EEG data were exported to a Matlab (Mathworks, Natick, MA) compatible format and processed offline using EEGlab (v14.1.1b) signal processing environment (Delorme & Makeig, 2004) running under Matlab R2017a (Figure 4.2). Continuous data were extracted for the auditory language processing task. Data were high-pass filtered at 1.5 Hz using Blackman window FIR filter, with 1 Hz transition bandwidth and filter order of 2750. Due to the presence of 60-Hz line noise in the EEG recording, the data were further low-pass filtered at 50 Hz using Blackman window FIR filter, with 20 Hz transition bandwidth and filter order of 138.

Artifact subspace reconstruction (ASR) was applied to continuous data using EEGLAB plugin clean_rawdata (Mullen *et al.*, 2015, Chang *et al.*, 2018). ASR was run using fixed random-access memory (RAM) allocation of 3200 megabytes. Within each 1-second sliding window, data portions with variance greater than 8 standard deviations from the calibration data were rejected and interpolated. Channels were rejected if they had flatline duration longer than 3 seconds. Channels with correlations less than 0.75 to its reconstruction based on other channels in the given time window were also removed. Time windows where 25% of channels exceed z-scores of -3.5 to 5.5 on root-mean-square thresholding test were rejected. After cleaning with ASR, data from the original 128-channels montage was interpolated to 25-channels montage based on the International 10-20 System using interp_mont plugin (Desjardins, 2010; Jasper, 1958).

Adaptive Mixture Independent Component Analysis (AMICA) (Palmer *et al.*, 2006, Palmer *et al.*, 2008) was performed and outputted 25 independent components (IC) per subject. AMICA algorithm was used because it outperforms other ICA methods in scalp channels mutual

information reduction and identifies the maximal number of dipolar brain IC (Delorme *et al.*, 2012). AMICA learns the data structure without preset assumptions; prior studies have utilized AMICA in processing both infant (Piazza *et al.*, 2016) and adult EEG data (Hsu *et al.*, 2018). IC were visually inspected for remaining artifacts: eye blinks, saccades, BCG (ballistocardiograms), and electromyogram (EMG). Automatic IC classification was applied using EEGLAB plugin ICLabel to aid in manual review and rejection of non-neural IC. IClabel accounted for each IC's activity power spectrum and scalp topography in its classification algorithm (Pion-Tonachini *et al.*, 2017, Pion-Tonachini *et al.*, 2019).

Cleaned data were epoched into 3-second segments. Because phase coherence is sensitive to file length, the first 29 epochs of clean data were selected across all subjects for further analysis. Data were passed through a Laplacian spatial filter to correct for possible volume conduction effects. Specifically, a spherical spline Laplacian transform with head-circumference correction was applied to transform the cleaned data into current source density (Kayser & Tenke, 2006).

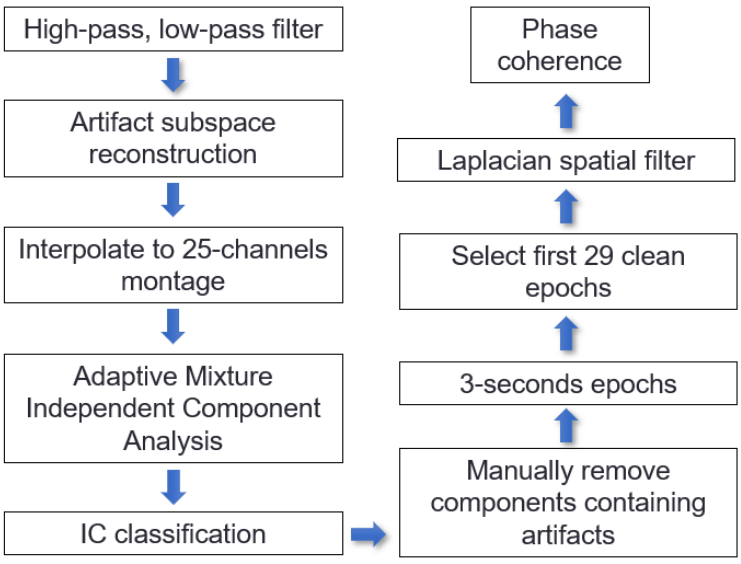

***Figure 4.2**. Continuous EEG pre-processing for functional connectivity analysis.*

EEG coherence analysis

Connectivity was measured in the form of magnitude-square phase coherence across the 3-second period, in the theta (4-6 Hz) and alpha (6-12 Hz) frequency bands. Coherence is a measure of synchronization between two signals of the same frequency, and it quantifies the extent to which they share a constant oscillating frequency and phase difference. Neuronal sources share information by oscillating coherently (Fries, 2005). Coherence for all channel pairs in each frequency band was calculated using EEGlab *newcrossf* function, with a window size of 1024 samples (Delorme & Makeig, 2004). Each electrode-pair coherence value was calculated by first averaging coherence values across all time bins within each frequency bin, and then the mean coherence values was calculated by averaging within each frequency bin. Coherence values from all electrode pairs was compiled into a 25 x 25 matrix, such that in each element (i,j) in each matrix represented the averaged coherence between channel i and channel j, for a given subject. Phase coherence was calculated between 18 electrode pairs in language networks (frontal-temporal, frontal-central). Left frontal: F3, F7, F9; left temporal: T7, T9; left central: C3; right frontal: F4, F8, F10; right temporal: T8, T10; right central: C4 (Figure 4.3). Only intrahemispheric fronto-temporal and fronto-central connections within putative language networks were chosen in order to examine functional networks supporting language processing.

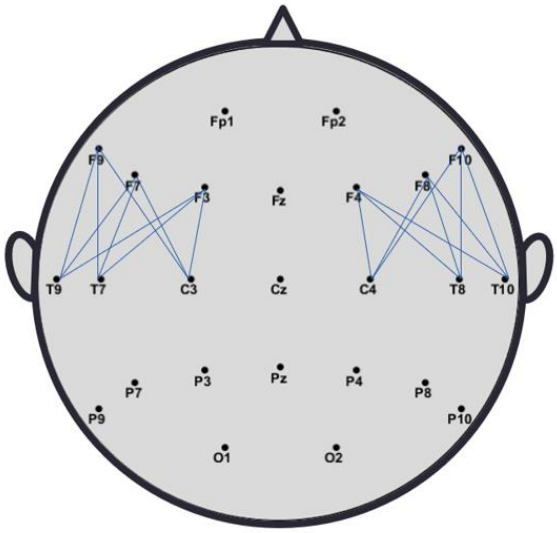

Figure 4.3. *Language network electrode pairs examined in phase coherence analysis.*

Statistical analysis

Descriptive analyses. Mann-Whitney U-test was used to compare 18-month language scores and ADOS-T scores between risk groups. Mann-Whitney U-test were performed with the Statistical Package for Social Sciences (IBM SPSS Statistics, version 25).

Primary analyses. At each electrode pair, a linear mixed effect model was used to model coherence (Y) as a function of time (t). The analysis was conducted in two stages. In the first stage, longitudinal trends for each group were modeled and contrasted to assess group-differences in coherence at each electrode. At each electrode pair and frequency band, a separate model was fit and a y-intercept and slope of the best fit line for each group was computed. The subject index (i) ranged from 1 through n, where n is the total number of subjects in each group. Subjects needed to have EEG data from at least one timepoint in order to be included in the model. Subject's age was centered at 7.5 months so that the intercept for

such group represented average coherence across the first year of life. For each electrode, the linear mixed effects model was specified as

$$Y_i(t) = \beta_1 Concern + \beta_2 NoConcern + \beta_3 Concern * t + \beta_4 NoConcern * t + \gamma_{0,i} + \gamma_{0,i} + \epsilon_i(t)$$

where Concern and NoConcern were dummy variables that identified group membership, β were fixed effects capturing group-specific slope and intercept effects, $\gamma_{0,i}$ and $\gamma_{0,i}$ were normally distributed subject-specific slope and intercept terms, and $\epsilon_i(t)$ was independent measurement error. Wald chi-square tests were used to compare differences in fixed effects between groups at each electrode pair and frequency band (Fox & Weisberg, 2019). The group contrasts were ASD-concern versus No-ASD-concern for the intercept and slope terms given by $\beta_1 - \beta_2$ and $\beta_3 - \beta_4$, respectively. From these contrasts, a subset of electrodes was identified in which slope or intercept terms varied between the two groups.

Secondary analyses. As part of follow-up secondary analyses, electrodes that were found to display overall group-differences or longitudinal group-differences were selectively considered to determine if subject-specific coherence trends at the identified electrodes were associated with cognitive outcomes. For stage two, at each electrode, subject-specific estimates were calculated using one of two models, agnostic to group membership. Group membership was omitted in this stage given that the primary concern was to associate subject-specific trends in coherence with cognitive outcomes. Including group-membership would introduce diagnostic information into the model and undercut the analytic goal of exploring connections between physiology and developmental outcomes. If the two groups displayed differences in longitudinal trends ($\beta_3 - \beta_4$), then the following model was fit,

$$Y_i(t) = \beta_0 + \beta_1 t + \gamma_{0,i} + \gamma_{0,i} + \epsilon_i(t)$$

where β were fixed effects capturing slope and intercept effects, $\gamma_{0,i}$ and $\gamma_{0,i}$ were normally distributed subject-specific slope and intercept terms, and $\epsilon_i(t)$ was independent measurement error. If the two groups did not display a difference in longitudinal trends but did display differences in overall group trends ($\beta_1 - \beta_2$) (and there was a non-significant main effect for time for each group), then the following reduced model was fit,

$$y_i(t) = \beta_0 + \gamma_{0,i} + \epsilon_i(t).$$

Subject-specific slope or intercept estimates were formed by summing the appropriate subject-specific terms with their corresponding fixed effects. Linear mixed effect modeling was performed using a custom script written in R (R Foundation for Statistical Computing, Vienna, Austria). Pearson's correlations were used to relate each subject's estimated slope and y-intercept to 18-month MSEL expressive and receptive t-scores, CDI words comprehended and produced, and ADOS-T overall score. Pearson's correlations were performed with SPSS.

4.3 Results

Outcome groupings

The ASD-concern group (n = 17) consisted of 15 FR and 2 LR infants. The No-ASD-concern group (n = 54) consisted of 29 FR and 25 LR infants.

Developmental testing

At 18-months, FR infants had significantly lower MSEL standard scores, MSEL nonverbal t-score, MSEL verbal t-score, MSEL receptive language t-scores, and CDI words comprehended compared to LR infants. Relative to LR infants, FR infants also had higher

ADOS-T overall scores, and higher ADOS-T CSS. There was no significant difference between FR and LR in MSEL expressive language t-score, CDI words produced. At 18-months, the ASD-concern group had significantly lower MSEL standard scores, MSEL nonverbal t-scores, MSEL verbal t-scores, MSEL receptive language t-scores, MSEL expressive language t-scores, CDI words comprehended, and CDI words produced compared to the No-ASD-concern group. The ASD-concern group also had higher ADOS-T overall scores and ADOS-T than the No-ASD-concern group. Mean MSEL standard scores and ADOS-T overall scores for each group are presented in Table 4.1; mean receptive language, expressive language, and overall verbal ability for the ASD-concern and No-ASD-concern groups are presented in Figure 4.4 and Table 4.2.

	Familial-risk (FR)	Low-risk (LR)	ASD-concern	No-ASD-concern
n	44	27	17 (15 FR/ 2 LR)	54 (29 FR/ 25 LR)
Sex	18F/26M	10F/17M	3F/14M	25F/ 29M
MSEL ELC standard score	89.1 ± 15.2* (52-116)	102.0 ± 14.9* (56-126)	73.8 ± 10.7* (52-85)	100.3 ± 12.0* (71-126)
MSEL nonverbal t-score	45.3 ± 8.0* (22.5-64.0)	52.0 ± 8.9* (25.0-66.0)	40.4 ± 10.0* (22.5-59.5)	50.2 ± 7.1* (34.0-66.0)
MSEL verbal t-score	43.1 ± 10.5* (21.0-63.5)	49.7 ± 10.3* (25.0-66.0)	31.6 ± 5.4* (21.0-40.5)	50.0 ± 8.0* (30.0-66.0)
ADOS-T overall score	7.7 ± 5.3* (1-20)	5.0 ± 4.3* (0-18)	14.4 ± 2.9* (10-20)	4.2 ± 2.5* (0-9)
ADOS-T CSS	3.3 ± 1.9* (1-8)	2.3 ± 1.4* (1-7)	5.7 ± 1.1* (4-8)	2.0 ± 0.8* (1-3)

Table 4.1. *Developmental outcomes and ASD symptoms at 18 months. Abbreviations: female (F); male (M); calibrated severity score (CSS). Mean, standard deviation, and range scores for each group are displayed. * Significant group differences respectively between FR versus LR, or ASD-concern versus No-ASD-concern.*

		Familial-risk (FR)	Low-risk (LR)	ASD-concern	No-ASD-concern
Receptive Language	MSEL *(t-score)*	40.8 ± 13.6* (20-72)	50.6 ± 14.7* (20-72)	27.8 ± 4.9* (20-37)	49.8 ± 12.8* (30-72)
	CDI *(words comprehended)*	147.1 ± 98.2* (21-370)	215.1 ± 116.5* (10-392)	78.2 ± 93.6* (10-368)	202.6 ± 99.3* (21-392)
Expressive Language	MSEL *(t-score)*	45.4 ± 10.1 (19-68)	48.9 ± 9.5 (27-65)	35.4 ± 8.7* (19-51)	50.3 ± 7.4* (30-68)
	CDI *(words produced)*	51.6 ± 69.9 (0-342)	63.0 ± 66.2 (3-249)	7.2 ± 10.9* (0-37)	69.9 ± 71.2* (2-342)

Table 4.2. *Language profiles at 18 months on the MSEL and CDI. Mean, standard deviation, and range scores for each group are displayed. * Significant group differences respectively between FR versus LR, or ASD-concern versus No-ASD-concern.*

A range of language ability was observed at 18-months in both ASD-concern and No-ASD-concern groups (Figure 4.4). The ASD-concern group included 15 infants whose scores fell below average range on the MSEL (verbal t-score < 40; Charman *et al.*, 2017) (Figure 4.4c).

While most infants in the No-ASD-concern group did have typical language ability, 4 infants showed language-delay (3 FR, 1 LR) (Figure 4.4c).

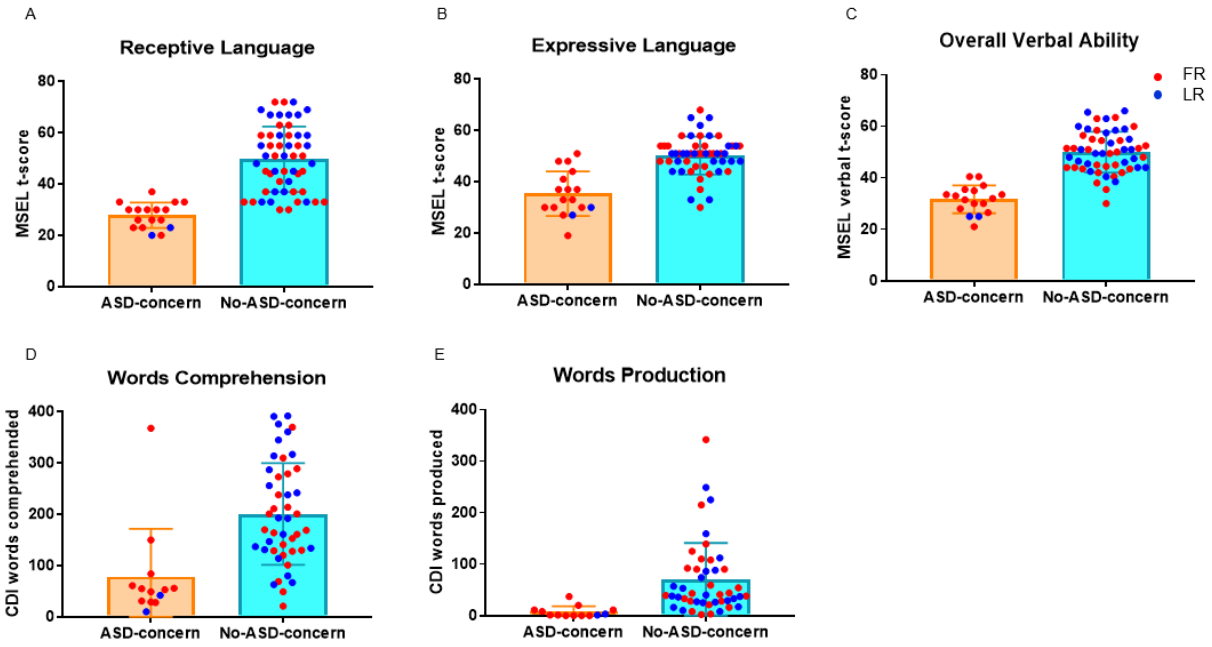

Figure 4.4. *Language profiles for ASD-concern and No-ASD-concern groups at 18 months.* ***Figures 4.4a-c****. MSEL receptive, expressive language, and overall verbal t-scores.* ***Figures 4.4d-e****. CDI words comprehension, words production.*

EEG data quality

Averaged across 297 files from all four timepoints, ASR rejected 9% of channels (12 ± 9 channels) and 0.002% of file length (0.34 ± 3.0 seconds). After manual removal of IC with artifacts, on average of 12 ± 2 (range: 4-19) brain components were retained per subject. There was no difference in data quality between ASD-concern and No-ASD-concern groups. Out of 297 EEG files recorded, 294 files passed data quality checks after pre-processing and were used in the final analyses (Table 4.3). 2 files were not processed due to errors during EEG

acquisition (missing paradigm markers) and 1 file failed data quality check because it lacked dipolar brain IC (after having 64 channels removed by ASR).

Timepoint	3-month	6-month	9-month	12-month
EEG files	74	78	72	73
Channels removed	15 ± 8 (12%)	14 ± 10 (11%)	10 ± 7 (8%)	8 ± 9 (6%)
Channel retained	109 ± 8 (range 84-122) (range 67-98%)	110 ± 10 (range 71-123) (range 57-99%)	114 ± 7 (range 70-122) (range 56-98%)	116 ± 9 (range 60-124) (range 48-100%)
Duration removed	0.8 ± 6 s (0.5%)	0.08 ± 0.3 s (0.1%)	0.2 ± 2 s (0.2%)	0.2 ± 0.7 (0.1%)
Duration retained	150 ± 7 s (range 110-161 s); (range 69-100%)	151 ± 10 s (range 98-162) (range 99-100%)	153 ± 8 s (range 104-163) (range 91-100%)	151 ± 11 s (range 91-162) (range 98-100%)
IC retained	12 ± 2 (range 7-19)	12 ± 2 (range 4-17)	11 ± 2 (range 5-16)	12 ± 3 (range 5-18)
Files rejected	0	1	0	2
File retained	74	77	72	71

Table 4.3. *EEG data quality and subject retention after pre-processing.*

Phase coherence: ASD-concern versus No-ASD-concern

Theta coherence slope differentiated the ASD-concern group from the No-ASD-concern group at left frontal-temporal F9-T9 electrode pair ($p = 0.049$). At the F9-T9, the ASD-concern group showed a decrease in coherence over time (negative slope) while the No-ASD-concern group showed an increase in coherence over time (positive slope) (Figure 4.5a). Theta coherence y-intercepts differentiated the ASD-concern group from the No-ASD-concern group at left frontal-central F9-C3 electrode pair ($p = 0.039$) and right frontal-temporal F10-T8 electrode pair ($p = 0.029$). The ASD-concern group had greater average coherence over the first year at F10-T8 (Figure 4.5b), and lower average theta coherence over the first year at F9-C3 (Figure 4.5c), compared to the No-ASD-concern group. Alpha coherence y-intercepts differentiated the ASD-concern group from the No-ASD-concern group at left frontal-central F9-C3 electrode pair ($p = 0.049$). The ASD-concern group had lower average alpha coherence over the first year at F9-C3 (Figure 4.5d). These four slope and y-intercept findings differentiated risk groups without correction for multiple comparisons. Slope and y-intercepts at the other electrode pairs within theta and alpha bands did not differentiate ASD-concern from No-ASD-concern ($p > 0.05$).

Phase coherence and behavioral outcomes

Across risk groups: alpha coherence y-intercept at F9-C3 correlated with MSEL expressive language t-score ($r = 0.264$, $p = 0.026$) (Figure 4.6a) and MSEL receptive language t-score ($r = 0.256$, $p = 0.012$) at 18-months (Figure 4.6b). Alpha coherence y-intercept at F9-C3 did not correlate with ASD symptoms or CDI language scores at 18-months ($p > 0.05$). Across all subjects, theta coherence slope at F9-T9, theta coherence y-intercepts at F10-T8 and F9-C3 did not correlate with 18-month ASD symptoms or expressive and receptive language scores the MSEL and CDI ($p > 0.05$).

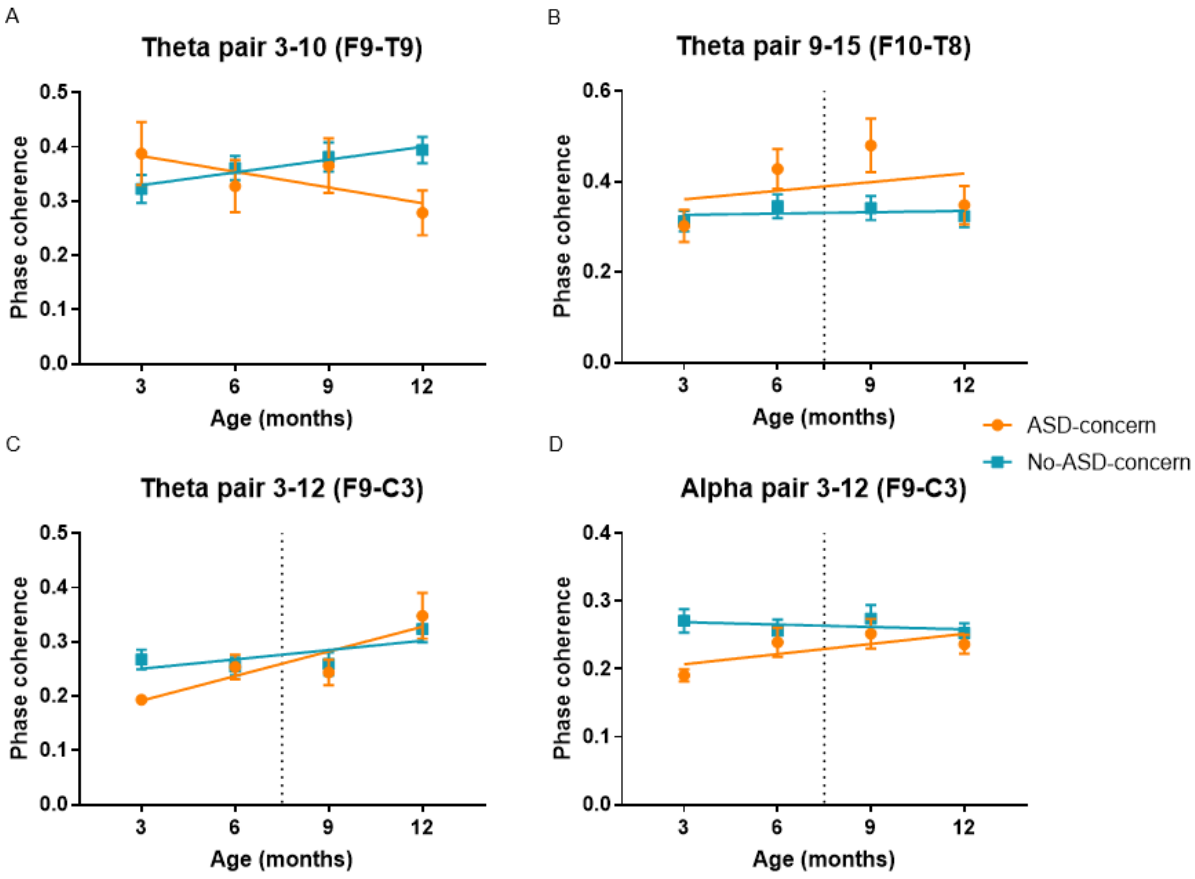

Figure 4.5. *Longitudinal trends in coherence between ASD-concern and No-ASD-concern groups.* ***Figure 4.5a****. Theta coherence slope at F9-T9 differentiated the ASD-concern group from the No-ASD-concern group.* ***Figures 4.5b-c****. Theta coherence y-intercepts at F9-C3 and F10-T8 differentiated the ASD-concern group from the No-ASD-concern group.* ***Figure 4.5d.*** *Alpha coherence y-intercept at F9-C3 differentiated the ASD-concern group from the No-ASD-concern group.*

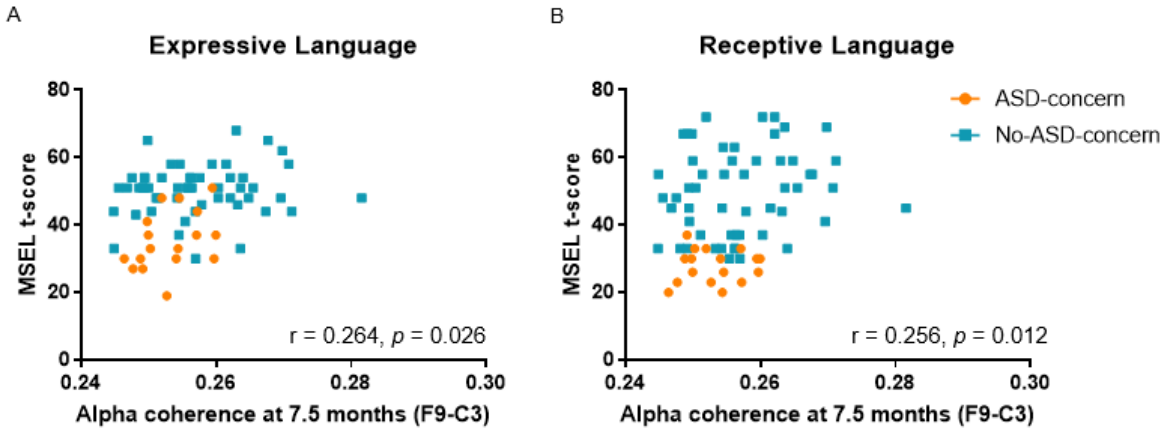

Figure 4.6*. Relationships between average coherence over the first year at F9-C3 and language outcome at 18-month.* ***Figure 4.6a****. Average alpha coherence over the first year at F9-C3 related to MSEL expressive language t-score****. Figures 4.6b****. Average alpha coherence over the first year at F9-C3 correlated with MSEL receptive language t-score.*

4.4 Discussion

This study is the first to examine developmental trajectories of functional connectivity during language processing over the first year of life in infants at familial risk for ASD. We found significant differences in coherence patterns over the first year of life that differentiated ASD-concern from No-ASD-concern groups. A pattern of decreasing left fronto-temporal theta coherence was observed over the first year of life in infants who later showed ASD symptoms at 18 months. Infants in the ASD-concern group also showed lower average left fronto-central coherence over the first year of life in both theta and alpha frequency bands. An opposite trend was observed in the right hemisphere, in which infants in the ASD-concern group had greater average right fronto-temporal coherence in the theta band over the first year of life. Across risk groups, average left fronto-central alpha coherence over the first year of life correlated with both expressive and receptive language ability at 18-months. These findings build upon our

previously reported work (Tran *et al*, in prep), showing that left fronto-central coherence is lower in the ASD-concern group at 3-months and continues to remain low throughout the first year of life.

Language outcomes

Most children in the ASD-concern group also met criteria for language delay at 18-months, with deficits in both receptive and expressive language. The language profiles of our ASD-concern group were consistent with prior reports of language deficits in infant-siblings and toddlers with ASD (Charman *et al.*, 2003; Luyster *et al.*, 2007; Luyster *et al.*, 2008; Ellis Weismer *et al.*, 2010; Mitchell *et al.*, 2006; Ozonoff *et al.*, 2014; Levin *et al.*, 2017). Of note, a small subset of children in the No-ASD-concern group also met criteria for language delay at 18 months. Language delay profiles without ASD symptomatology may reflect broader autism phenotype in FR children and specific language impairment in LR children (Messinger *et al.*, 2013; Landa *et al.*, 2013; Ozonoff *et al.*, 2014; Charman *et al.*, 2017; Tomblin *et al.*, 1997).

Auditory statistical learning

We chose to use auditory statistical learning as our language processing task because this domain is essential for language learning, both during infancy and beyond. Infants can rely on statistical word boundaries to help them segment a stream of continuous speech into individual words (Saffran *et al.*, 1996). The ability to perform auditory statistical learning has been documented in infants throughout the first year of life, starting in infants as young as 2-5 days old (Flo *et al.*, 2019). In typically developing infants, speech segmentation ability at 7.5-12 months of age relates to both word production at two years and preschool language skills (Newman *et al.*, 2006). Infants with worse speech segmentation ability may have more trouble

parsing words from continuous speech, which can result in reduced words-learning and cause eventual language impairment.

Reduced coherence during language processing in infants at risk for ASD

Our study identified unique connectivity profiles over the first year of life that stratified familial risk infants who are most likely to develop ASD from those who are least likely to develop ASD. Over the first year of life, the brain undergoes dramatic structural and functional maturation. Prior neuroimaging studies have shown that familial risk infants who develop ASD have diverging trajectories in the development of white matter tracts over the first and second years of life, which relate to future ASD symptoms such as repetitive behaviors (Wolff *et al.*, 2012; Wolff *et al.*, 2015). Among low risk infants, increasing structural connectivity in the splenium over the first and second years of life relate to better language ability, specifically with greater words production (Swanson *et al.*, 2017b). Since language function is dependent on the structural integrity of language network, it is expected that infants who later develop ASD and language delay should have aberrant trajectories in both structural and functional connectivity of language networks during the first year of life. Our longitudinal findings are consistent with this theory. Over the first year of life, infants who later develop ASD symptoms showed diverging trajectories in connectivity development as well as decreased average connectivity over the first year of life in left-hemispheric connections within putative language networks. From prior neuroimaging studies, we know that typically developing infants start out with bilateral temporal cortical activation in response to speech stimuli at birth that gradually shift to left-lateralized language networks during the first year of life (Dinstein *et al.*, 2011). In our study, infants without ASD concern show the neurotypical pattern of increasing functional connectivity in left frontal-temporal networks between 3 and 12 months in response to speech stimuli from our language processing task. In contrast, older infants and toddlers with ASD have been shown to have

atypically decreased left temporal activation in response to speech sounds (Eyler *et al.*, 2012). Similarly, infants with ASD concerns in our study show an altered trajectory of decreasing functional connectivity in left frontal-temporal networks between 3 and 12 months of age during language processing. Our coherence trajectory findings are consistent with prior neuroimaging works and shed light on the development of atypical language networks in ASD.

Our findings of decreased left hemispheric functional connectivity throughout the first year of life in infants with early concerns for ASD are consistent with prior literature documenting decreased left-hemisphere activation in response to language stimuli in infants and toddlers with ASD. Throughout the first year of life, left-hemispheric underconnectivity during language processing may both precede and predispose at-risk infants to later language impairment and ASD symptoms. In left fronto-central networks, average coherence over the first year of life correlated with overall verbal ability at 18-months, which highlighted the importance of left-lateralized language networks during language development. It is possible that infants who have stronger left-hemispheric functional connectivity in language networks during exposure to speech streams are likely to perform better at word segmentation in novel contexts, and word segmentation ability in infancy lay the foundation for future words learning and production.

Hemispheric effect

Opposite trends in theta coherence were observed between left-hemispheric and right-hemispheric connections for infants who later showed ASD symptoms at 18-months. Averaged over the first year of life, the ASD-concern group had both increased right fronto-temporal theta coherence and decreased left fronto-central theta coherence compared to the No-ASD-concern group. A recent neuroimaging study of a subset of infants from our cohort has documented altered structural lateralization of language networks in familial risk infants at 6-weeks of age (Liu *et al.*, 2019). Increased theta coherence in right fronto-temporal language networks may be

caused by underlying increased structural connectivity of right dorsal language tracts, as previously described by Liu *et al* in familial risk infants (Liu *et al.*, 2019). Increased functional connectivity in the right hemisphere during language processing during the first year of life, as seen in our study, may reflect the development towards altered lateralization of language networks in ASD. Prior neuroimaging studies have documented similar right-lateralized cortical activation during language processing in older infants and toddlers with ASD (Lombardo *et al.*, 2015; Redcay & Courchesne, 2008).

Limitations and future directions

One limitation of our current study is a small sample size. As such, it will be important to replicate this study in a new cohort of infants. Most prior EEG studies in familial-risk infants have not been replicated. To date, only one research group has attempted to replicate their infant EEG findings in a new cohort (Orekhova *et al.*, 2014; Haartsen *et al.*, 2019), and it remains unclear whether the heterogeneity in reported findings across studies are reflections of the subjects' true traits or are due to underlying differences in analysis methods. Furthermore, since the present study only examined coherence during the exposure phase of the language task, it is unclear whether increased coherence reflected the infants' ASL or the infants' increased attention to speech stimuli. Future studies may also integrate structural and functional connectivity measures across neuroimaging and EEG studies in the same cohort of infants, in order of identify cross-modality neural markers of language processing in infancy.

In conclusion, our study has shown that EEG can detect, and track altered functional connectivity patterns in language networks in familial risk infants throughout the first year of life, starting as young as 3 months of age. Early connectivity trajectories in language networks relate to both ASD symptoms profile and language ability at 18 months, suggesting that these early

neural markers may be useful in identifying infants who are most at risk of atypical development and who may benefit from early targeted behavioral intervention.

Chapter 5: Connectivity during language processing – an endophenotype of ASD risk

5.1 Introduction

In the previous two chapters, we have characterized the development of atypical functional connectivity patterns during language processing in infants who showed symptoms of autism spectrum disorder (ASD) at 18 months. Now in this chapter, we are shifting gears and taking a step back to examine functional connectivity profiles in infants with shared familial risk for ASD, irrespective of their future diagnostic outcomes. Familial risk infants (FR) are an interesting group to study because of their heterogenous outcomes, potentially reflecting a wide background of polygenic risks interacting with environmental factors. FR infants are at elevated risk of ASD, global developmental delay, language delay and other atypical developments; however, nearly half of FR infants are likely to have typical development (Ozonoff *et al.*, 2014; Messinger *et al.*, 2015; Charman *et al.*, 2017). While FR infants may inherit additive polygenic risks from their parents, siblings in the same family often do not inherit the same genetic variants. Within families with at least 2 children with ASD (multiplex families), children with ASD have greater amount of rare copy-number variants (CNVs) than their unaffected siblings (Leppa *et al.*, 2016). However, up to 70% of multiplex children with ASD do not inherit the same rare CNVs as their other siblings with ASD (Yuen *et al.*, 2015; Leppa *et al.*, 2016). While infants from simplex families (with only 1 child with ASD) are at 20% risk of ASD; infants from multiplex families are at even greater risk of ASD, with penetrance rates approaching 33-50% (Ozonoff et al., 2011; Messinger et al., 2015). Studying neural development in FR infants will allow us to elucidate of there are distinct neural markers of genetic, familial predisposition for ASD. Given the heterogenous developmental outcomes amongst FR infants, neural markers identified in the general FR group may be associated with either a pathological process towards atypical development or a protective compensatory process towards typical development.

Because of the heterogeneity in the FR group, prior studies have often focused on comparing either between FR infants with ASD to FR infants without ASD, or FR infants with ASD versus low risk infants without ASD. As early as 6 weeks of age, evidence of atypical lateralization of language networks is detectable in FR infants (Liu *et al.*, 2019). Between 6 and 24 months of age, FR infants have increased corpus callosum area and thickness compared to LR infants, in which FR infants with ASD show the greatest deviations from LR controls (Wolff *et al.*, 2015). FR infants without ASD appear to have intermediate trajectories in corpus callosum development, in between FR with ASD and LR (Wolff *et al.*, 2015). And while FR infants with ASD have increased extra-axial cerebral spinal fluid (CSF) between 6 and 24 months of age, there appear to be no difference in extra-axial CSF volume between LR controls and FR infants without ASD (Shen *et al.*, 2017). Using EEG, Righi and colleagues have described a pattern of reduced coherence in in FR infants that emerged between 6 and 12 months of age while infants listened to speech tones as part of auditory oddball paradigm (Righi *et al.*, 2014). Differences between FR and LR infants were detected at 12 months and not at 6 months, suggesting of different underlying trajectories in network development during late infancy (Righi *et al.*, 2014). To date, no prior study has investigated differences in functional connectivity between FR and LR infants during language processing between 3 and 12 months of age.

This chapter aims to answer two primary questions: 1) Can spectral power and phase coherence measured during language processing differentiate 3-month-old infants based on familial risk for ASD, and 2) do the patterns of coherence change over the first year of life distinguish FR from LR infants? We hypothesized that functional connectivity in language networks should be able to differentiate FR from LR infants both at 3-month of age and throughout the first year of life.

5.2 Methods

Participants

FR and LR infants were recruited to be part of a longitudinal study on early biomarkers of ASD, as part of the UCLA Autism Center of Excellence (ACE; NICHD 2P50HD055784-08). FR infants came from a mix of simplex and multiplex families. Simplex families had only 1 child with ASD, while multiplex families had at least 2 children with known ASD diagnosis. The study inclusion criteria were the same as previously described in Chapters 3 and 4. A total of 85 participants (50 FR, 35 LR) completed EEG during the language task for at least one timepoint, out of which 74 participants (40 FR, 34 LR) completed the EEG task at 3 months. Among the FR infants, 8 came from multiplex families. There were no sex differences between FR (21 female, 29 male) and LR (13 female, 22 male) groups.

Behavioral Measures

ASD symptoms, overall developmental level, and language ability were assessed at 18-months to characterize behavioral profiles of the participants. Behavioral measures using the Mullen Scales of Early Learning (MSEL), the MacArthur Communicative Development Inventory (CDI), and the Autism Diagnostic Observation Scale Toddler Module (ADOS-T) were the same as previously described in Chapters 3 and 4.

EEG Stimuli and Data Acquisition

Approximately 2.5 minutes of continuous EEG was recorded using 128-channel EGI net while each infant listened to a stream of concatenated syllables forming pseudowords. The auditory statistical learning paradigm and EEG acquisition parameters were the same as previously described in Chapters 3 and 4.

EEG Data Pre-Processing

Continuous EEG data was processed using EEGlab (v14.1.1b), following prior protocol described in Chapters 3 and 4. The processing pipeline utilized artifact subspace reconstruction (ASR), down-sampling to 25 channels, 1-model Adaptive Mixture Independent Component Analysis (AMICA), ICLabel, and Laplacian spatial filter.

EEG Data Quality

Averaged across all 294 files from all four timepoints, ASR rejected 9% of channels (12 $\pm$ 8 channels) and 0.002% of file length (0.34 $\pm$ 3.0 seconds). After manual removal of IC with artifacts, on average of 12 $\pm$ 2 (range: 4-19) brain components were retained per subject. There was no difference in data quality between FR and LR groups. All subjects who had EEG during the language processing task were used in the final analyses. Data quality and data retention at each timepoint were previously described in detail in Chapter 4.

EEG Power and Coherence Analyses

Relative power and phase coherence in the theta (4-6 Hz) and alpha (6-12 Hz) frequency bands were calculated using protocols previously described in Chapter 3. Spectral power analyses focused on frontal and central-temporal regions of interest (Figure 5.1a). Phase coherence analyses focused on 18 electrode pairs from putative language network (Figure 5.1b).

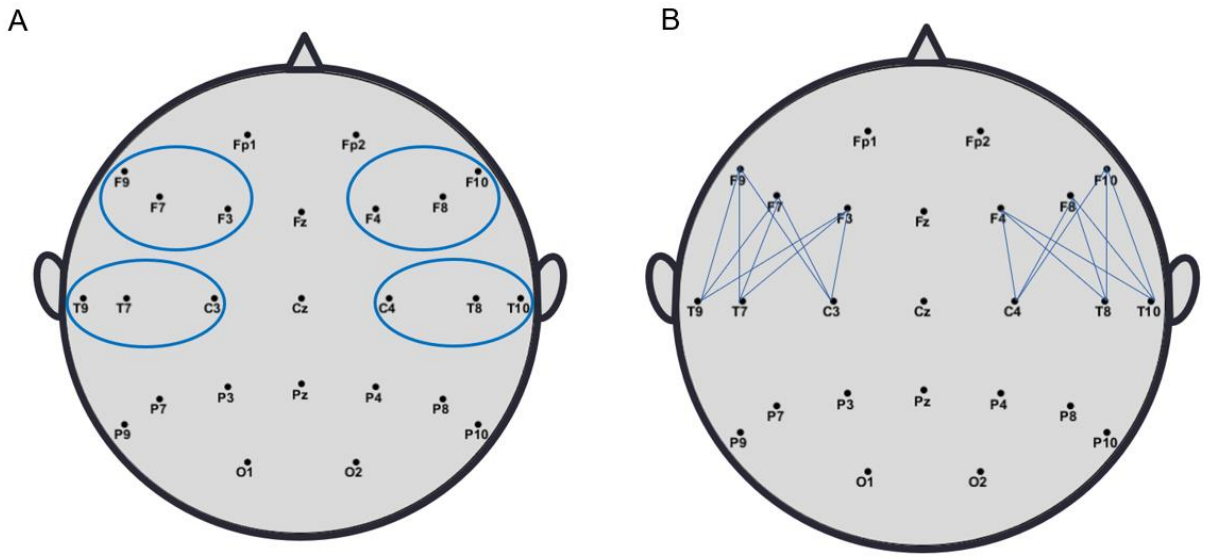

Figure 5.1. *Scalp map used in power and coherence analyses.* ***Figure 5.1a****. Frontal and temporal-central ROIs used in spectral power analysis.* ***Figure 5.1b****. Language network electrode pairs examined in phase coherence analysis.*

Statistical Analysis

These exploratory analyses were not corrected for multiple comparisons.

5.2.1 Cross-sectional study

Methods for risk-group comparisons were the same as previously described in Chapter 3. Comparison groups were FR versus LR, instead of ASD-concern versus No-ASD-concern. Analyses were focused on identifying group differences in spectral power and phase coherence at 3-months of age. Group differences in spectral power were assessed using the Mann-Whitney U-test. Group differences in phase coherence were assessed using permutation test.

5.2.2 Longitudinal study

Methods for risk-group comparisons were the same as previously described in Chapter 4. Comparison groups were FR versus LR, instead of ASD-concern versus No-ASD-concern. Analyses were focused on identifying group differences in coherence change over the first year of life, assessed in terms of fixed effects (slope and y-intercept) from linear mixed effect models.

5.3 Results

Developmental testing

At 18-months, FR infants had significantly lower MSEL standard scores, MSEL nonverbal t-score, MSEL verbal t-score, MSEL receptive language t-scores, CDI words comprehended, and receptive advantage compared to LR infants. Relative to LR infants, FR infants also had higher ADOS-T overall scores and higher ADOS-T CSS. There was no significant difference between FR and LR in MSEL expressive language t-score or CDI words produced. Mean receptive language, expressive language, overall verbal ability, and receptive advantage scores for FR and LR groups are presented in Figure 5.2.

A range of language ability was observed at 18-months among FR and LR groups (Figure 5.2). The FR group consisted of 16 children who met criteria for language-delay on the MSEL (verbal t-score less than 40; Charman *et al.*, 2017). Overall 36% of FR children met criteria for language-delay, among which 13 belonged in the ASD-concern group and 3 belonged in the No-ASD-concern group. The LR group consisted of 3 children who met criteria for language-delay on the MSEL. Overall 11% of LR children met criteria for language-delay, among which 2 belonged in the ASD-concern group and 1 belonged in the No-ASD-concern group.

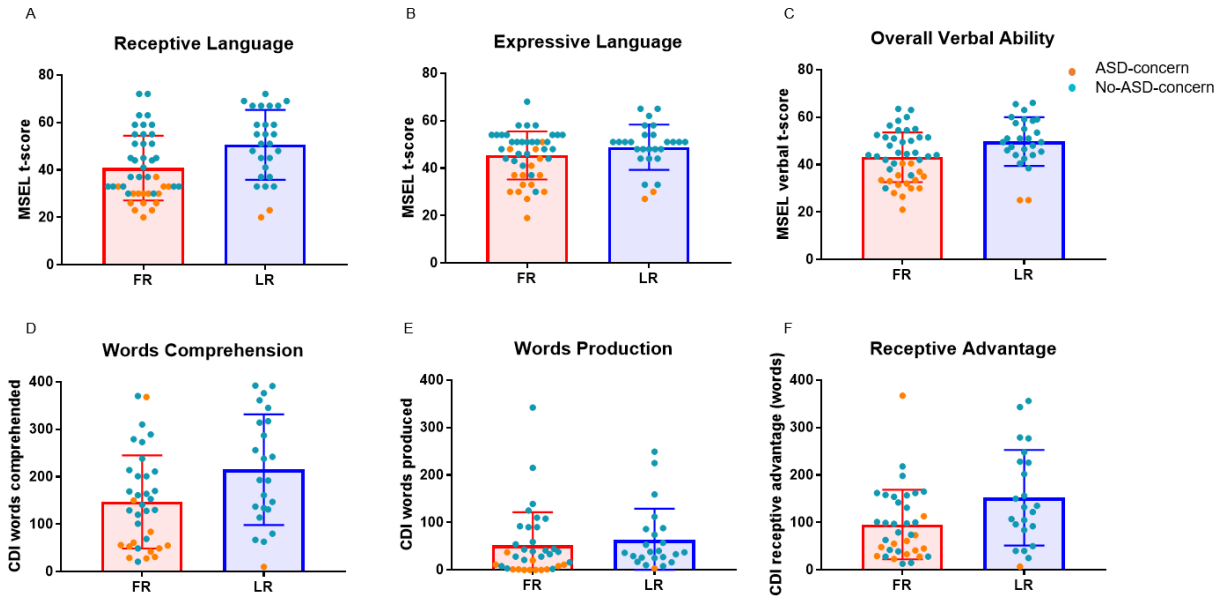

Figure 5.2. *Language profiles at 18 months on the MSEL and CDI. Receptive advantage was calculated as the difference between CDI words comprehended and words produced. Mean scores and standard deviation for FR and LR groups are displayed.*

5.3.1 Cross-sectional study

Spectral power at 3-month: Familial-risk versus Low-risk

The FR group did not differ from the LR group in theta and alpha relative power at frontal and temporal-central ROIs (Mann-Whitney U-test $p > 0.05$) (Figure 5.3).

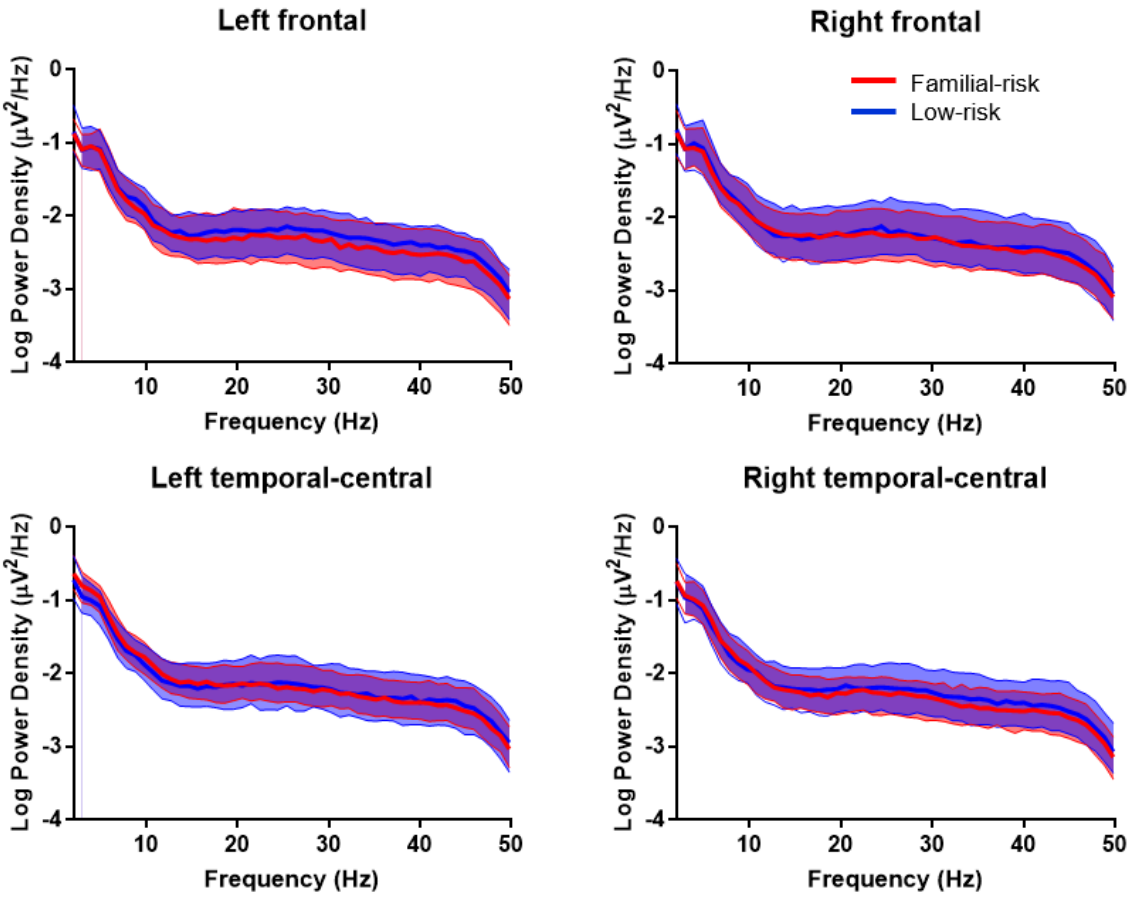

Figure 5.3. *PSD plots for FR and LR groups. Log power density plots at frontal and temporal-central ROIs at 3-months (median and interquartile range). Red = Familial-risk (FR); blue = Low-risk (LR).*

Phase coherence at 3-month: Familial-risk versus Low-risk

Theta phase coherence at right frontal-central electrode pair F10-C4 differentiated FR from LR (p=0.042), where LR had higher coherence (Figure 5.4a). Alpha phase coherence at left frontal-temporal electrode pair F7-T9 differentiated FR from LR (p=0.021) where LR had lower coherence (Figure 5.4b). After FDR was applied and p-values were adjusted for 36

comparisons (2 frequency bands x 18 electrode pairs), the group differences in theta and alpha bands were not significant (adjusted p-values > 0.05).

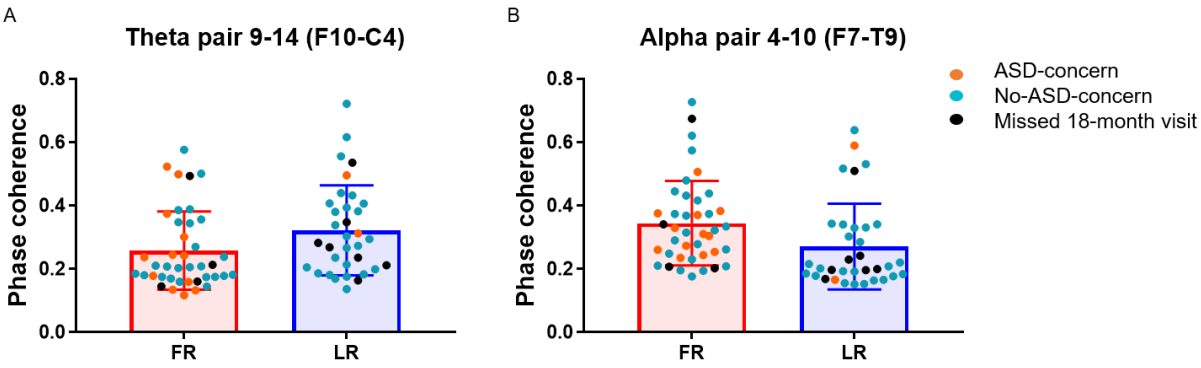

Figure 5.4. *Phase coherence for FR and LR groups.* ***Figure 5.4a.*** *Theta phase coherence at right frontal-central connection (F10-C4) differentiated FR from LR.* ***Figure 5.4b.*** *Alpha phase coherence at left frontal-temporal connection (F7-T9) differentiated FR from LR.*

5.3.2 Longitudinal study

Phase coherence at 3-12 months: Familial-risk versus Low-risk

Theta coherence slope differentiated FR from LR at right frontal-central F8-C4 electrode pair (p = 0.047). At F8-C4, FR had a slope of 0 while LR had a positive slope (Figure 5.5a). Theta coherence y-intercept differentiated FR from LR at right frontal-central F10-T10 electrode pair (p = 0.021). FR had greater average theta coherence over the first year at F10-T10, compared to LR (Figure 5.5b). Alpha coherence slope differentiated FR from LR at right frontal-central F10-C4 (p = 0.032). At F10-C4, FR had a positive slope while LR had a negative slope (Figure 5.5c).

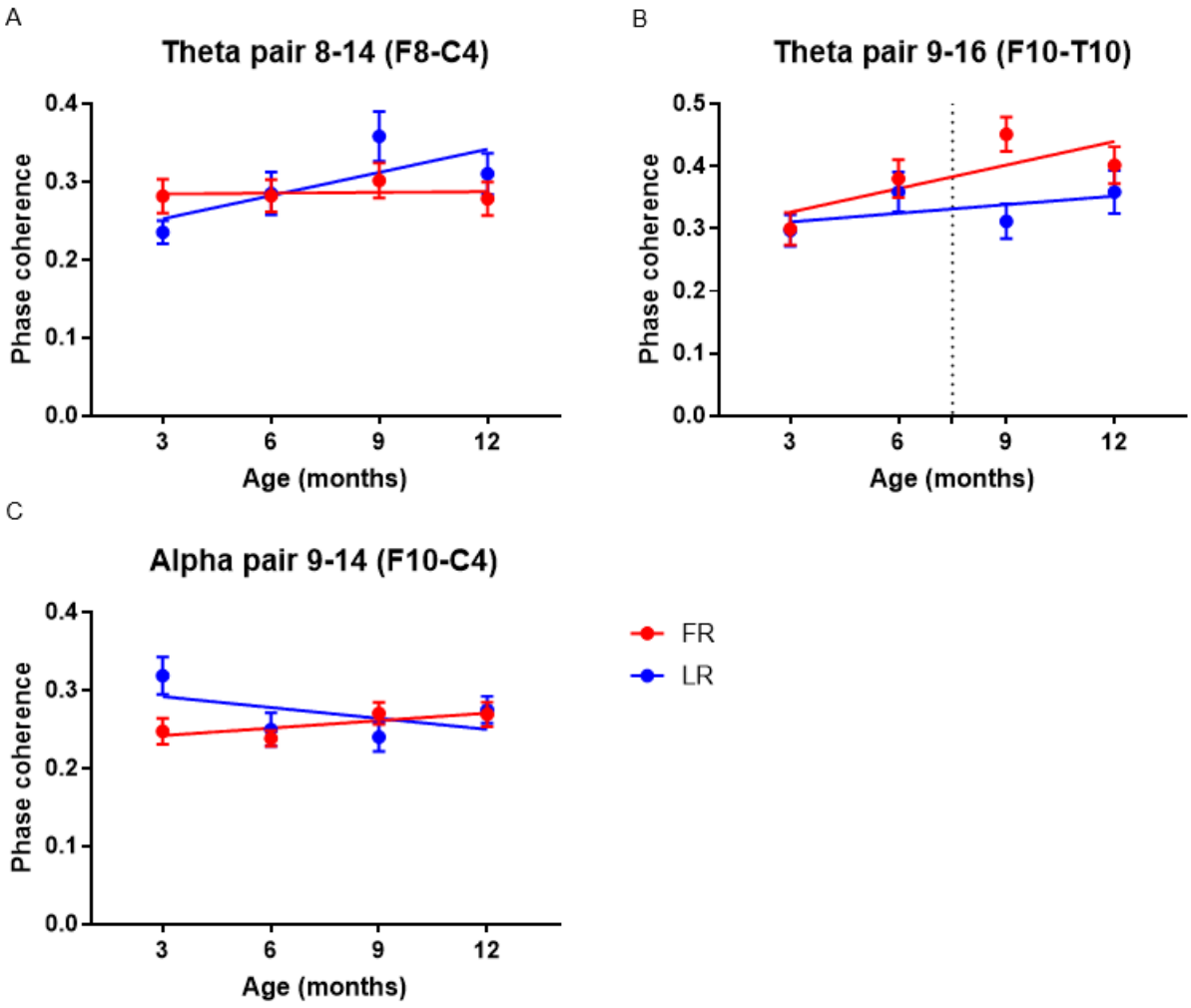

Figure 5.5. *Longitudinal trends in coherence between FR and LR.* ***Figure 5.5a.*** *Theta coherence slope at F8-C4 differentiated FR from LR.* ***Figures 5.5b.*** *Theta coherence y-intercepts at F10-T10 differentiated FR from LR.* ***Figure 5.5c.*** *Alpha coherence slope at F10-C4 differentiated FR from LR.*

5.4 Discussion

At 3 months of age, relative power in the theta and alpha bands did not differentiate infants at familial risk for ASD from those without familial risk. In contrast, phase coherence at 3 months of age did differentiate infants based on familial risk. Specifically, right fronto-central

theta coherence was lower in FR infants; while left fronto-temporal alpha coherence was higher in FR infants. Over the first year of life, FR and LR infants differed in right fronto-central coherence trajectories. In terms of right fronto-central theta coherence, LR infants had an increasing trajectory over the first year while coherence for FR infants remained stable and unchanged. Opposite trajectories were observed for right fronto-central alpha coherence: FR infants had increasing coherence while LR infants had decreasing coherence over the first year. Infants in the FR group also showed increased right fronto-temporal coherence over the first year of life in the theta band. Taken together, the functional connectivity networks that differentiated FR from LR were different from networks that differentiated ASD-concern from No-ASD-concern groups in Chapters 3 and 4.

Behavioral outcomes

Among FR children, over 36% met criteria for language-delay while 34% had elevated ASD symptoms at 18-months, with many children having both language-delay and high ASD symptoms. The rate of ASD-concern in our study was higher than the rate of ASD diagnosis in prior infant sibling studies, which could be due to our cohort being enriched in infants from multiplex families. In our study, 50% of multiplex infants developed ASD symptoms by 18-months of age – this rate is consistent with previously described ASD prevalence rates in multiplex families (Ozonoff *et al.*, 2011; Messinger *et al.*, 2015). Multiplex infants who have multiple siblings with ASD are known to have two-fold increased risk of developing ASD compared to simplex infants who only have one older sibling with ASD (Ozonoff *et al.*, 2011). Among LR children, 11% met criteria for language-delay and 7% had elevated ASD symptoms at 18-months, with similar overlap in children with both language-delay and ASD symptoms. The rate of ASD-concern in our LR cohort was higher than the prevalence of ASD in the general population, and this could be due to our study's small sample size and inclusion of first-born

infants in the LR group. As a group, FR infants had lower receptive language and reduced receptive advantage compared to LR infants. Among FR infants who did not have ASD symptoms at 18-months, approximately 11% had language-delay – this rate is consistent with prior characterizations of the broader autism phenotype (Messinger *et al.*, 2013; Landa *et al.*, 2013; Ozonoff *et al.*, 2014).

Functional connectivity during language processing as an endophenotype of ASD

Our study identified unique connectivity profiles during language processing that stratified infants based on familial risk and may serve as an endophenotype of ASD. During early infancy in infants as young as 3-months, FR infants had patterns of right hemisphere underconnectivity and left hemisphere overconnectivity that were suggestive of hemispheric differences during language processing. We have previously shown that reduced alpha coherence the left hemisphere at 3-months of age is associated with early ASD symptoms at 18-months (Chapter 3). It is possible that the increased left hemispheric alpha coherence at 3-months of age in the FR group maybe driven by a compensatory mechanism that is protective against ASD. The majority of FR infants with high alpha coherence at 3-months do not have ASD symptoms at 18-months; this supports our prior hypothesis that there may be a threshold over which left-hemispheric connectivity during language processing promotes typical development.

Over the first year of life, developmental trends in connectivity at different connections within right-hemisphere language networks also differentiated infants based on familial risk. FR and LR infants differed in coherence trajectories across multiple fronto-central connections in the right-hemisphere over the first year of life, showing that altered trajectories in functional connectivity during language processing are related to familial risk for ASD. Averaged over the first year of life, FR infants also had higher theta coherence over right fronto-temporal

connections which may be compensatory for the initial lower theta coherence observed at 3-months in right fronto-central connections. We have previously shown that increased fronto-temporal theta coherence in the right hemisphere over the first year of life is associated with early ASD symptoms at 18-months (Chapter 4). Our longitudinal results from the present study are complementary with our previously described findings from Chapter 4, showing that increased right hemispheric fronto-temporal theta coherence during the first year of life both characterizes familial ASD risk and serves as a marker of atypical neurodevelopment associated with ASD. Our findings are also consistent with previously described patterns of altered lateralization and increased structural connectivity in right-hemispheric dorsal language tracts among 6-weeks old FR infants (Liu *et al.*, 2019). It is possible that FR infants who have increased functional connectivity in the right hemisphere may have underlying structural overconnectivity of right-hemispheric language tracts. Increased functional connectivity in the right hemisphere during our language processing task may reflect altered lateralization of language networks in ASD. Prior neuroimaging studies have documented similar right-lateralized cortical activation during language processing in infants and toddlers with ASD (Lombardo *et al.*, 2015; Redcay & Courchesne, 2008). Since the brain undergoes dramatic structural and functional maturation during the first year of life, our findings highlight the importance of both early cross-sectional and longitudinal characterizations of infant brain development.

While phase coherence differentiated risk groups, spectral power at 3-months of age did not differentiate infants based on familial risk. Results from our language processing study are complementary to findings by Levin *et al* on Laplacian filtered resting state dataset. During resting state EEG, Levin *et al* describes patterns frontal power in FR infants from data processed using four parallel pipelines (average referenced, Laplacian referenced, binned versus unbinned power) and highlighted the sensitivity of EEG data to pre-processing and

analyses techniques (Levin *et al.*, 2017). Levin *et al* identifies differences in power between FR and LR infants using averaged referenced data but not with Laplacian filtered data. Similar to Levin *et al*'s Laplacian filtered data findings, we did not find a difference in spectral power between infants stratified based on familial risk (Levin *et al.*, 2017).

Limitations and future directions

One limitation of this study is the basis of all analyses on familial risk, while we know from prior studies that FR infants are a heterogenous group with many developmental outcomes. At 18-months, our FR group have a range of behavioral outcomes ranging from typical development to having language delay and high amount of ASD symptoms. Thus, the connectivity patterns we described for the FR group during the first year of life may be reflecting multiple divergent pathways towards both typical and atypical outcomes. In future studies, it will be important to further divide the FR group based on outcomes: FR infants with typical development, FR infants with ASD, and FR infants with other atypical development (including language delay). Comparing connectivity between FR subgroups will allow us to further characterize divergent developmental pathways among infants with similar genetic background.

Chapter 6: Summary and Conclusions

6.1 Summary

In Chapter 2, methodological considerations in the development of an EEG pre-processing pipeline were addressed in detail, with the goals of maximizing data quality and data retention for infant EEG. Attributes of two processing pipelines were compared when both pipelines were used in parallel to clean data files from 3-6 months old infants. The two pipelines differed in filtering parameters, mode for first-pass data cleaning (manual versus automatic artifact subspace reconstruction), type of independent component analysis used (extended infomax ICA versus adaptive mixture ICA), and mode for second-pass data cleaning (manual review of ICs versus use of automatic IC classification). The incorporation of automatic ASR and IC classification drastically reduced pre-processing time and established objective standards for reproducibility by multiple researchers. Manual data cleaning was labor intensive and inherently subjective, making it impossible to be reproducible between different researchers. Manual data cleaning resulted in over 50% data length loss, while cleaning with ASR allowed close to 100% of data length to be preserved. ASR proved to be effective in removing artifacts from noisy infant EEG data while still preserving underlying neural signals and data length. Preservation of clean data length has many important implications: more data points allows for less drastic data reduction prior to ICA, which enables greater richness in the data to be retained. Data length retention also leads to increased subject retention, which is important in preserving statistical power in infant EEG studies that are constrained by small sample sizes. Overall, the processing pipeline using ASR, AMICA, and automatic IC classification was shown to be superior due to its reproducibility, fast processing speed, and maximal retention of clean file length – making it the ideal pre-processing pipeline for infant EEG data with high number of artifacts and short recording lengths.

In Chapter 3, spectral power and phase coherence during language processing was examined in 3-months old infants at familial risk for ASD. Fronto-central and fronto-temporal areas of putative language networks were selected for power and coherence calculations. Power in theta and alpha bands did not differentiate risk groups at 3-months and did not relate to language and ASD symptoms at 18-months. Reduced coherence in left fronto-central networks in both theta and alpha bands was found in infants who later showed ASD symptoms at 18-months. Notably, this difference was driven by a subset of infants without ASD concerns, who had very high coherence during language processing. Left fronto-central alpha coherence at 3-months correlated with greater words production at 18-months. Coherence at 3-months was able to differentiate risk groups categorically and related to continuous measure of language outcome. There may be a threshold over which coherence during language processing promotes typical language development. As such, coherence measured during language processing can be used to identify infants who are most likely to have typical outcome (preserved language ability and little ASD symptoms).

Cross-sectional findings from Chapter 3 were examined further using a longitudinal study design in Chapter 4. In Chapter 4, developmental trajectories in phase coherence during language processing was examined throughout the first year of life in infants at familial risk for ASD. At the same left fronto-central network that differentiated risk groups at 3-months of age, reduced average coherence over the first year of life was maintained in infants who showed ASD symptoms at 18 months. At this left fronto-central network, average theta and alpha coherence over the first year correlated with greater receptive language, expressive language, and overall verbal ability at 18 months. There was also a hemispheric effect for theta coherence. Within fronto-temporal networks, infants with ASD symptoms at 18 months had decreased theta coherence trajectory in the left hemisphere, and greater average theta coherence over the first year of life in the right hemisphere. Our longitudinal findings capture the development of both

typical and atypical lateralization of language network during the first year of life. Infants without ASD concern show evident of neurotypical leftward lateralization of language networks, with increased functional connectivity in left frontal-temporal networks between 3 and 12 months in response to speech stimuli. In contrast, infants with ASD concerns show atypical lateralization of language networks, including both decreased functional connectivity in left frontal-temporal/central networks between 3 and 12 months of age, and increased connectivity in right frontal-temporal networks.

In Chapter 5, connectivity differences between familial-risk and low-risk infants were examined using both cross-sectional and longitudinal study designs in order to elucidate if connectivity may serve as an endophenotype of ASD risk. Fronto-central and fronto-temporal areas of putative language networks were selected for power and coherence calculations at 3 months; longitudinal trajectories in coherence development over the first year of life were also examined within these language networks. At 3-months of age, power in theta and alpha bands did not differentiate groups based on familial risk. Familial-risk infants had greater left fronto-temporal alpha coherence at 3-months, which may be a compensatory mechanism driven by familial-risk infants without ASD concerns. Over the first year of life, familial-risk infants and low-risk infants differed in right fronto-central coherence trajectories in both theta and alpha bands. Familial-risk infants also have increased right fronto-temporal coherence, averaged across the first year of life, which may be reflective of the atypical rightward lateralized language network associated with ASD. Our study identified unique connectivity profiles during language processing that stratified infants based on familial risk and may serve as an endophenotype of ASD risk. By 18 months, over a third of familial-risk infants in our study have either language delay or ASD symptoms, while the remaining familial-risk children appear to be typically developing. The familial-risk group is inherently heterogenous, and whether an infant in this group develops typically or atypically reflects complex interactions between the child's polygenic

risk and external environmental factors. Heterogeneity between cross-sectional and longitudinal coherence findings may reflect different sets of neural processes underlying typical and atypical development in our cohort.

6.2 Limitations and Future Directions

The studies described in this dissertation have some shared limitations. Analyses described were focused on scalp-based measures of spectral power and phase coherence, measures that have been traditionally examined in the field of EEG ASD research. Phase coherence is sensitive to volume conduction, which may falsely inflate the measured phase correlations between neighboring electrodes. To mitigate effects from volume conduction, we have utilized a Laplacian spatial filter. However, Laplacian filter may also remove some underlying neural activity with distributive patterns and current flow from tangential neural sources. Future studies should incorporate source-based measures of EEG connectivity, aligned to each subject's own structural MRI, in order to improve spatial resolution and bypass the volume conduction limitation of scalp-based measures.

Across the studies described in chapters 3-5, functional connectivity was examined during the exposure phase of the language processing task, when infants were getting familiarized with nonsense words from a continuous speech stream. It is unclear whether increased connectivity reflected the infants' increased attention to speech stimuli (sensory processing) or the infants' auditory statistical learning (cognitive processing). Distinguishing sensory processing from cognitive processing is important in targeting neural networks important to language processing and learning. In future studies, functional connectivity should be examined during both the exposure phase and the test phase of the ASL paradigm in order to determine if increased connectivity during this task is cognitive measure of word segmentation ability.

Despite 6 years of EEG data collection at UCLA, the final cohort of infants included in chapters 3-5 was still relatively small. A small sample size constrains analyses based on ASD outcomes, since only 20% of familial-risk infants are expected to develop ASD. While my studies were anchored in 18-month measures of ASD symptoms and not clinical ASD diagnosis, only 17 out of 85 infants had ASD concerns by 18 months. It will be critical to replicate findings from chapters 3-5 with a new cohort of infants. In order to have a large sample size, EEG data would need to be collected across multiple infant sibling research sites. Currently the British Autism Study of Infant Siblings (BASIS) is the only multi-site network that have shared EEG data. Robust structural and functional neuroimaging findings have come from the Infant Brain Imaging Study (IBIS); however, the IBIS network currently lacks EEG data collection. Most prior EEG studies in familial-risk infants have not been replicated (only BASIS has attempted replication in a new cohort) and it remains unclear whether the heterogeneity in reported findings across studies result from differences in analysis methods or due to true trait differences in the subject sample. In future studies, the pre-processing pipeline described in chapters 2-5 will be ideal for use in processing large amount of infant EEG data collected across multiple research sites.

6.3 Concluding Remarks

Taken together, the body of work in this dissertation support the hypothesis that early differences in neural synchrony may lay a foundation for neurodevelopmental impairments in infants at risk for ASD, and these differences are able to be robustly quantified using EEG. During language processing, infants at risk of ASD showed altered trajectories in language network connectivity over the first year of life – with detectable differences emerging as early as 3 months of age. By studying infants longitudinally, we were able to trace the development of typical leftward lateralization of language networks in infants without ASD concerns, and

atypical rightward lateralization of language networks in infants who develop ASD symptoms. It is possible that infants who have stronger left-hemispheric functional connectivity in language networks during exposure to speech streams are likely to perform better at word segmentation in novel contexts. As an early developmental domain, word segmentation is critical for words learning in infancy and lays the foundation for language processing and social communication.

As early as 3 months of age, we were able to identify EEG neural markers predictive of both future language outcomes and ASD symptom profiles. This work has huge potential for translational application to the clinic settings. From EEG measured during early infancy, we are able to identify infants who are likely to have atypical outcomes and would need closer monitoring throughout the first year of life. Having early neural predictors of atypical development during the prodromal period, before infants show any behavioral symptoms, can help guide the delivery of earlier, targeted behavioral intervention while the infants' brains still have the greatest plasticity.

References

Abrahams, B.S. & Geschwind, D.H. (2008) Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet*, **9**, 341-355.

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders*, Washington, DC.

Anderson, D.K., Lord, C., Risi, S., DiLavore, P.S., Shulman, C., Thurm, A., Welch, K. & Pickles, A. (2007) Patterns of growth in verbal abilities among children with autism spectrum disorder. *J Consult Clin Psychol*, **75**, 594-604.

Arnett, A.B., Hudac, C.M., DesChamps, T.D., Cairney, B.E., Gerdts, J., Wallace, A.S., Bernier, R.A. & Webb, S.J. (2018) Auditory perception is associated with implicit language learning and receptive language ability in autism spectrum disorder. *Brain and language*, **187**, 1-8.

Artoni, F., Delorme, A. & Makeig, S. (2018) Applying dimension reduction to EEG data by Principal Component Analysis reduces the quality of its subsequent Independent Component decomposition. *NeuroImage*, **175**, 176-187.

Artoni, F., Delorme, A. & Makeig, S. (2019) A visual working memory dataset collection with bootstrap Independent Component Analysis for comparison of electroencephalographic preprocessing pipelines. *Data in brief*, **22**, 787-793.

Aslin, R.N., Saffran, J.R. & Newport, E.L. (1998) Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological science*, **9**, 321-324.

Barbaro, J. & Dissanayake, C. (2013) Early markers of autism spectrum disorders in infants and toddlers prospectively identified in the Social Attention and Communication Study. *Autism : the international journal of research and practice*, **17**, 64-86.

Basar, E. (2013) A review of gamma oscillations in healthy subjects and in cognitive impairment. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, **90**, 99-117.

Bell, M.A. & Fox, N.A. (1992) The relations between frontal brain electrical activity and cognitive development during infancy. *Child development*, **63**, 1142-1163.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57**, 289-300.

Berg, J.M. & Geschwind, D.H. (2012) Autism genetics: searching for specificity and convergence. *Genome Biol*, **13**, 247.

Berkel, S., Marshall, C.R., Weiss, B., Howe, J., Roeth, R., Moog, U., Endris, V., Roberts, W., Szatmari, P., Pinto, D., Bonin, M., Riess, A., Engels, H., Sprengel, R., Scherer, S.W. & Rappold, G.A. (2010) Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation. *Nature genetics*, **42**, 489-491.

Bryson, S.E., Zwaigenbaum, L., Brian, J., Roberts, W., Szatmari, P., Rombough, V. & McDermott, C. (2007) A prospective case series of high-risk infants who developed autism. *Journal of autism and developmental disorders*, **37**, 12-24.

Chang, C.Y., Hsu, S.H., Pion-Tonachini, L. & Jung, T.P. (2018) Evaluation of Artifact Subspace Reconstruction for Automatic EEG Artifact Removal. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, **2018**, 1242-1245.

Charman, T., Drew, A., Baird, C. & Baird, G. (2003) Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of child language*, **30**, 213-236.

Charman, T., Young, G.S., Brian, J., Carter, A., Carver, L.J., Chawarska, K., Curtin, S., Dobkins, K., Elsabbagh, M., Georgiades, S., Hertz-Picciotto, I., Hutman, T., Iverson, J.M., Jones, E.J., Landa, R., Macari, S., Messinger, D.S., Nelson, C.A., Ozonoff, S., Saulnier, C., Stone, W.L., Tager-Flusberg, H., Webb, S.J., Yirmiya, N. & Zwaigenbaum, L. (2017) Non-ASD outcomes at 36 months in siblings at familial risk for autism spectrum disorder (ASD): A baby siblings research consortium (BSRC) study. *Autism research : official journal of the International Society for Autism Research*, **10**, 169-178.

Chawarska, K., Macari, S. & Shic, F. (2013) Decreased Spontaneous Attention to Social Scenes in 6-Month-Old Infants Later Diagnosed with Autism Spectrum Disorders. *Biol. Psychiatry*, **74**, 195-203.

Chen, J.A., Penagarikano, O., Belgard, T.G., Swarup, V. & Geschwind, D.H. (2015) The emerging picture of autism spectrum disorder: genetics and pathology. *Annual review of pathology*, **10**, 111-144.

Christensen, L., Hutman, T., Rozga, A., Young, G.S., Ozonoff, S., Rogers, S.J., Baker, B. & Sigman, M. (2010) Play and developmental outcomes in infant siblings of children with autism. *Journal of autism and developmental disorders*, **40**, 946-957.

Cornew, L., Dobkins, K.R., Akshoomoff, N., McCleery, J.P. & Carver, L.J. (2012) Atypical social referencing in infant siblings of children with autism spectrum disorders. *Journal of autism and developmental disorders*, **42**, 2611-2621.

Cuevas, K. & Bell, M.A. (2011) EEG and ECG from 5 to 10 months of age: developmental changes in baseline activation and cognitive processing during a working memory task. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, **80**, 119-128.

de la Torre-Ubieta, L., Won, H., Stein, J.L. & Geschwind, D.H. (2016) Advancing the understanding of autism disease mechanisms through genetics. *Nature medicine*, **22**, 345-361.

Dehaene-Lambertz, G., Dehaene, S. & Hertz-Pannier, L. (2002) Functional neuroimaging of speech perception in infants. *Science (New York, N.Y.)*, **298**, 2013-2015.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Meriaux, S., Roche, A., Sigman, M. & Dehaene, S. (2006) Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc Natl Acad Sci U S A*, **103**, 14240-14245.

Dehaene-Lambertz, G., Montavont, A., Jobert, A., Allirol, L., Dubois, J., Hertz-Pannier, L. & Dehaene, S. (2010) Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and language*, **114**, 53-65.

Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, **134**, 9-21.

Delorme, A., Mullen, T., Kothe, C., Akalin Acar, Z., Bigdely-Shamlo, N., Vankov, A. & Makeig, S. (2011) EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing. *Computational intelligence and neuroscience*, **2011**, 130714.

Delorme, A., Palmer, J., Onton, J., Oostenveld, R. & Makeig, S. (2012) Independent EEG sources are dipolar. *PloS one*, **7**, e30135.

Desjardins, J. (2010) Interpmont extension for EEGLAB. Free Software Foundation.

Dickinson, A., DiStefano, C., Lin, Y.-Y., Scheffler, A.W., Senturk, D. & Jeste, S.S. (2018) Interhemispheric alpha-band hypoconnectivity in children with autism spectrum disorder. *Behav. Brain Res.*, **348**, 227-234.

Dinstein, I., Pierce, K., Eyler, L., Solso, S., Malach, R., Behrmann, M. & Courchesne, E. (2011) Disrupted neural synchronization in toddlers with autism. *Neuron*, **70**, 1218-1225.

Duffy, F.H. & Als, H. (2012) A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls - a large case control study. *BMC medicine*, **10**, 64.

Durand, C.M., Betancur, C., Boeckers, T.M., Bockmann, J., Chaste, P., Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I.C., Anckarsater, H., Sponheim, E., Goubran-Botros, H., Delorme, R., Chabane, N., Mouren-Simeoni, M.C., de Mas, P., Bieth, E., Roge, B., Heron, D., Burglen, L., Gillberg, C., Leboyer, M. & Bourgeron, T. (2007) Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nature genetics*, **39**, 25-27.

Ebrahimi-Fakhari, D. & Sahin, M. (2015) Autism and the synapse: emerging mechanisms and mechanism-based therapies. *Current opinion in neurology*, **28**, 91-102.

Elison, J.T., Wolff, J.J., Reznick, J.S., Botteron, K.N., Estes, A.M., Gu, H., Hazlett, H.C., Meadows, A.J., Paterson, S.J., Zwaigenbaum, L. & Piven, J. (2014) Repetitive behavior in 12-month-olds later classified with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, **53**, 1216-1224.

Ellis Weismer, S., Lord, C. & Esler, A. (2010) Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of autism and developmental disorders*, **40**, 1259-1273.

Elsabbagh, M. & Johnson, M.H. (2010) Getting answers from babies about autism. *Trends Cogn Sci*, **14**, 81-87.

Emerson, R.W., Adams, C., Nishino, T., Hazlett, H.C., Wolff, J.J., Zwaigenbaum, L., Constantino, J.N., Shen, M.D., Swanson, M.R., Elison, J.T., Kandala, S., Estes, A.M., Botteron, K.N., Collins, L., Dager, S.R., Evans, A.C., Gerig, G., Gu, H., McKinstry, R.C., Paterson, S., Schultz, R.T., Styner, M., Network, I., Schlaggar, B.L., Pruett, J.R. & Piven, J. (2017) Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. *Science translational medicine*, **9**.

Esler, A.N., Bal, V.H., Guthrie, W., Wetherby, A., Ellis Weismer, S. & Lord, C. (2015) The Autism Diagnostic Observation Schedule, Toddler Module: Standardized Severity Scores. *Journal of autism and developmental disorders*, **45**, 2704-2720.

Estes, A., Zwaigenbaum, L., Gu, H., St John, T., Paterson, S., Elison, J.T., Hazlett, H., Botteron, K., Dager, S.R., Schultz, R.T., Kostopoulos, P., Evans, A., Dawson, G., Eliason, J., Alvarez, S. & Piven, J. (2015) Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *Journal of neurodevelopmental disorders*, **7**, 24.

Eyler, L.T., Pierce, K. & Courchesne, E. (2012) A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain : a journal of neurology*, **135**, 949-960.

Feldman, M.A., Ward, R.A., Savona, D., Regehr, K., Parker, K., Hudson, M., Penning, H. & Holden, J.J. (2012) Development and initial validation of a parent report measure of the behavioral development of infants at risk for autism spectrum disorders. *Journal of autism and developmental disorders*, **42**, 13-22.

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. Paul H. Brookes Publishing, Baltimore, MD.

Ference, J. & Curtin, S. (2013) Attention to lexical stress and early vocabulary growth in 5-month-olds at risk for autism spectrum disorder. *Journal of experimental child psychology*, **116**, 891-903.

Flanagan, J.E., Landa, R., Bhat, A. & Bauman, M. (2012) Head Lag in Infants at Risk for Autism: A Preliminary Study. *Am J Occup Ther*, **66**, 577-585.

Flo, A., Brusini, P., Macagno, F., Nespor, M., Mehler, J. & Ferry, A.L. (2019) Newborns are sensitive to multiple cues for word segmentation in continuous speech. *Developmental science*, e12802.

Fox, J. & Weisberg, S. (2019) *An R Companion to Applied Regression*. Sage, Thousand Oaks, California.

Fries, P. (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, **9**, 474-480.

Frohlich, J., Senturk, D., Saravanapandian, V., Golshani, P., Reiter, L.T., Sankar, R., Thibert, R.L., DiStefano, C., Huberty, S., Cook, E.H. & Jeste, S.S. (2016) A Quantitative Electrophysiological Biomarker of Duplication 15q11.2-q13.1 Syndrome. *PloS one*, **11**, e0167179.

Gabard-Durnam, L., Tierney, A.L., Vogel-Farley, V., Tager-Flusberg, H. & Nelson, C.A. (2015) Alpha asymmetry in infants at risk for autism spectrum disorders. *Journal of autism and developmental disorders*, **45**, 473-480.

Gabard-Durnam, L.J., O'Muircheartaigh, J., Dirks, H., Dean, D.C., 3rd, Tottenham, N. & Deoni, S. (2018) Human amygdala functional network development: A cross-sectional study from 3 months to 5 years of age. *Developmental cognitive neuroscience*, **34**, 63-74.

Gamliel, I., Yirmiya, N., Jaffe, D.H., Manor, O. & Sigman, M. (2009) Developmental trajectories in siblings of children with autism: cognition and language from 4 months to 7 years. *Journal of autism and developmental disorders*, **39**, 1131-1144.

Gilman, S.R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M. & Vitkup, D. (2011) Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron*, **70**, 898-907.

Green, J., Charman, T., Pickles, A., Wan, M.W., Elsabbagh, M., Slonims, V., Taylor, C., McNally, J., Booth, R., Gliga, T., Jones, E.J., Harrop, C., Bedford, R. & Johnson, M.H. (2015) Parent-mediated intervention versus no intervention for infants at high risk of autism: a parallel, single-blind, randomised trial. *The lancet. Psychiatry*, **2**, 133-140.

Guilmatre, A., Dubourg, C., Mosca, A.L., Legallic, S., Goldenberg, A., Drouin-Garraud, V., Layet, V., Rosier, A., Briault, S., Bonnet-Brilhault, F., Laumonnier, F., Odent, S., Le Vacon, G., Joly-Helas, G., David, V., Bendavid, C., Pinoit, J.M., Henry, C., Impallomeni, C., Germano, E., Tortorella, G., Di Rosa, G., Barthelemy, C., Andres, C., Faivre, L., Frebourg, T., Saugier Veber, P. & Campion, D. (2009) Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic

pathways in schizophrenia, autism, and mental retardation. *Archives of general psychiatry*, **66**, 947-956.

Haartsen, R., Jones, E.J.H., Orekhova, E.V., Charman, T. & Johnson, M.H. (2019) Functional EEG connectivity in infants associates with later restricted and repetitive behaviours in autism; a replication study. *Translational psychiatry*, **9**, 66.

Hazlett, H.C., Gu, H., Munsell, B.C., Kim, S.H., Styner, M., Wolff, J.J., Elison, J.T., Swanson, M.R., Zhu, H., Botteron, K.N., Collins, D.L., Constantino, J.N., Dager, S.R., Estes, A.M., Evans, A.C., Fonov, V.S., Gerig, G., Kostopoulos, P., McKinstry, R.C., Pandey, J., Paterson, S., Pruett, J.R., Schultz, R.T., Shaw, D.W., Zwaigenbaum, L. & Piven, J. (2017) Early brain development in infants at high risk for autism spectrum disorder. *Nature*, **542**, 348-351.

Houston, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R. & Cutler, A. (2000) Cross-language word segmentation by 9-month-olds. *Psychonomic bulletin & review*, **7**, 504-509.

Houston, D.M., Santelmann, L.M. & Jusczyk, P.W. (2004) English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, **19**, 97-136.

Hsu, S.H., Pion-Tonachini, L., Palmer, J., Miyakoshi, M., Makeig, S. & Jung, T.P. (2018) Modeling brain dynamic state changes with adaptive mixture independent component analysis. *NeuroImage*, **183**, 47-61.

Hudry, K., Chandler, S., Bedford, R., Pasco, G., Gliga, T., Elsabbagh, M., Johnson, M.H. & Charman, T. (2014) Early language profiles in infants at high-risk for autism spectrum disorders. *Journal of autism and developmental disorders*, **44**, 154-167.

Hutman, T., Rozga, A., DeLaurentis, A.D., Barnwell, J.M., Sugar, C.A. & Sigman, M. (2010) Response to distress in infants at risk for autism: a prospective longitudinal study. *Journal of child psychology and psychiatry, and allied disciplines*, **51**, 1010-1020.

Iverson, J.M., Hall, A.J., Nickel, L. & Wozniak, R.H. (2007) The relationship between reduplicated babble onset and laterality biases in infant rhythmic arm movements. *Brain and language*, **101**, 198-207.

Iverson, J.M. & Wozniak, R.H. (2007) Variation in vocal-motor development in infant siblings of children with autism. *Journal of autism and developmental disorders*, **37**, 158-170.

Jasper, H.H. (1958) The ten-twenty electrode system of the International Federation. *Electroencephalogr Clin Neurophysiol*, **10**, 371–375.

Johnson, E.K. & Tyler, M.D. (2010) Testing the limits of statistical learning for word segmentation. *Developmental science*, **13**, 339-345.

Jones, E.J.H., Gliga, T., Bedford, R., Charman, T. & Johnson, M.H. (2014) Developmental pathways to autism: A review of prospective studies of infants at risk. *Neuroscience and Biobehavioral Reviews*, **39**, 1-33.

Jusczyk, P.W. (2002) Some critical developments in acquiring native language sound organization during the first year. *Ann Otol Rhinol Laryngol Suppl*, **189**, 11-15.

Kayser, J. & Tenke, C.E. (2006) Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clin Neurophysiol*, **117**, 348-368.

Keehn, B., Vogel-Farley, V., Tager-Flusberg, H. & Nelson, C.A. (2015) Atypical hemispheric specialization for faces in infants at risk for autism spectrum disorder. *Autism research : official journal of the International Society for Autism Research*, **8**, 187-198.

Kwok, E.Y.L., Brown, H.M., Smyth, R.E. & Oram Cardy, J. (2015) Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, **9**, 202-222.

Landa, R. & Garrett-Mayer, E. (2006) Development in infants with autism spectrum disorders: a prospective study. *The Journal of Child Psychology and Psychiatry*, **47**, 629-638.

Landa, R.J., Gross, A.L., Stuart, E.A. & Bauman, M. (2012) Latent class analysis of early developmental trajectory in baby siblings of children with autism. *Journal of child psychology and psychiatry, and allied disciplines*, **53**, 986-996.

Landa, R.J., Gross, A.L., Stuart, E.A. & Faherty, A. (2013) Developmental trajectories in children with and without autism spectrum disorders: the first 3 years. *Child development*, **84**, 429-442.

Landa, R.J., Holman, K.C. & Garrett-Mayer, E. (2007) Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of general psychiatry*, **64**, 853-864.

Lazenby, D.C., Sideridis, G.D., Huntington, N., Prante, M., Dale, P.S., Curtin, S., Henkel, L., Iverson, J.M., Carver, L., Dobkins, K., Akshoomoff, N., Tagavi, D., Nelson, C.A., 3rd & Tager-Flusberg, H. (2016) Language Differences at 12 Months in Infants Who Develop Autism Spectrum Disorder. *Journal of autism and developmental disorders*, **46**, 899-909.

Le Bihan, D., Mangin, J.F., Poupon, C., Clark, C.A., Pappata, S., Molko, N. & Chabriat, H. (2001) Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging*, **13**, 534-546.

LeBarton, E.S. & Iverson, J.M. (2013) Fine motor skill predicts expressive language in infant siblings of children with autism. *Developmental science*, **16**, 815-827.

LeBarton, E.S. & Landa, R.J. (2018) Infant motor skill predicts later expressive language and autism spectrum disorder diagnosis. *Infant behavior & development*, **54**, 37-47.

Leppa, V.M., Kravitz, S.N., Martin, C.L., Andrieux, J., Le Caignec, C., Martin-Coignard, D., DyBuncio, C., Sanders, S.J., Lowe, J.K., Cantor, R.M. & Geschwind, D.H. (2016) Rare Inherited and De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. *Am J Hum Genet*, **99**, 540-554.

Levin, A.R., Varcin, K.J., O'Leary, H.M., Tager-Flusberg, H. & Nelson, C.A. (2017) EEG power at 3 months in infants at high familial risk for autism. *Journal of neurodevelopmental disorders*, **9**.

Lewis, J.D., Evans, A.C., Pruett, J.R., Botteron, K., Zwaigenbaum, L., Estes, A., Gerig, G., Collins, L., Kostopoulos, P., McKinstry, R., Dager, S., Paterson, S., Schultz, R.T., Styner, M., Hazlett, H. & Piven, J. (2014) Network inefficiencies in autism spectrum disorder at 24 months. *Translational psychiatry*, **4**, e388.

Liu, J., Tsang, T., Jackson, L., Ponting, C., Jeste, S.S., Bookheimer, S.Y. & Dapretto, M. (2019) Altered lateralization of dorsal language tracts in 6-week-old infants at risk for autism. *Developmental science*, **22**, e12768.

Loh, A., Soman, T., Brian, J., Bryson, S.E., Roberts, W., Szatmari, P., Smith, I.M. & Zwaigenbaum, L. (2007) Stereotyped motor behaviors associated with autism in high-risk infants: a pilot videotape analysis of a sibling sample. *Journal of autism and developmental disorders*, **37**, 25-36.

Lombardo, M.V., Pierce, K., Eyler, L.T., Carter Barnes, C., Ahrens-Barbeau, C., Solso, S., Campbell, K. & Courchesne, E. (2015) Different functional neural substrates for good and poor language outcome in autism. *Neuron*, **86**, 567-577.

Luckhardt, C., Jarczok, T.A. & Bender, S. (2014) Elucidating the neurophysiological underpinnings of autism spectrum disorder: new developments. *J Neural Transm (Vienna)*, **121**, 1129-1144.

Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., Bishop, S., Esler, A., Hus, V., Oti, R., Richler, J., Risi, S. & Lord, C. (2009) The Autism Diagnostic Observation Schedule-toddler module: a new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of autism and developmental disorders*, **39**, 1305-1320.

Luyster, R., Lopez, K. & Lord, C. (2007) Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of child language*, **34**, 623-654.

Luyster, R.J., Kadlec, M.B., Carter, A. & Tager-Flusberg, H. (2008) Language assessment and development in toddlers with autism spectrum disorders. *Journal of autism and developmental disorders*, **38**, 1426-1438.

Macari, S.L., Campbell, D., Gengoux, G.W., Saulnier, C.A., Klin, A.J. & Chawarska, K. (2012) Predicting developmental status from 12 to 24 months in infants at risk for Autism Spectrum Disorder: a preliminary report. *Journal of autism and developmental disorders*, **42**, 2636-2647.

Mattys, S.L. & Jusczyk, P.W. (2001) Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, **78**, 91-121.

McEvoy, K., Hasenstab, K., Senturk, D., Sanders, A. & Jeste, S.S. (2015) Physiologic artifacts in resting state oscillations in young children: methodological considerations for noisy data. *Brain imaging and behavior*, **9**, 104-114.

McNealy, K., Mazziotta, J.C. & Dapretto, M. (2006) Cracking the language code: neural mechanisms underlying speech parsing. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **26**, 7629-7639.

McNealy, K., Mazziotta, J.C. & Dapretto, M. (2010) The neural basis of speech parsing in children and adults. *Developmental science*, **13**, 385-406.

McNealy, K., Mazziotta, J.C. & Dapretto, M. (2011) Age and experience shape developmental changes in the neural basis of language-related learning. *Developmental science*, **14**, 1261-1282.

Messinger, D., Young, G.S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., Landa, R.J., Charman, T., Stone, W.L., Constantino, J.N., Hutman, T., Carver, L.J., Bryson, S., Iverson, J.M., Strauss, M.S., Rogers, S.J. & Sigman, M. (2013) Beyond autism: a baby siblings research consortium study of high-risk children at three years of age. *Journal of the American Academy of Child and Adolescent Psychiatry*, **52**, 300-308.e301.

Messinger, D.S., Young, G.S., Webb, S.J., Ozonoff, S., Bryson, S.E., Carter, A., Carver, L., Charman, T., Chawarska, K., Curtin, S., Dobkins, K., Hertz-Picciotto, I., Hutman, T., Iverson, J.M., Landa, R., Nelson, C.A., Stone, W.L., Tager-Flusberg, H. & Zwaigenbaum, L. (2015) Early sex differences are not autism-specific: A Baby Siblings Research Consortium (BSRC) study. *Molecular autism*, **6**, 32.

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I. & Bryson, S. (2006) Early language and communication development of infants later diagnosed with autism spectrum disorder. *Journal of developmental and behavioral pediatrics : JDBP*, **27**, S69-78.

Mohammad-Rezazadeh, I., Frohlich, J., Loo, S.K. & Jeste, S.S. (2016) Brain Connectivity in Autism Spectrum Disorder. *Current opinion in neurology*, **29**, 137-147.

Mullen, E. (1995) *Mullen Scales of Early Learning*. American Guidance Services, Inc., Circle Pines, MN.

Mullen, T.R., Kothe, C.A., Chi, Y.M., Ojeda, A., Kerth, T., Makeig, S., Jung, T.P. & Cauwenberghs, G. (2015) Real-Time Neuroimaging and Cognitive Monitoring Using Wearable Dry EEG. *IEEE transactions on bio-medical engineering*, **62**, 2553-2567.

Nadig, A.S., Ozonoff, S., Young, G.S., Rozga, A., Sigman, M. & Rogers, S.J. (2007) A prospective study of response to name in infants at risk for autism. *Arch Pediatr Adolesc Med*, **161**, 378-383.

Newman, R., Ratner, N.B., Jusczyk, A.M., Jusczyk, P.W. & Dow, K.A. (2006) Infants' early ability to segment the conversational speech signal predicts later language development: a retrospective analysis. *Dev Psychol*, **42**, 643-655.

Onton, J., Westerfield, M., Townsend, J. & Makeig, S. (2006) Imaging human EEG dynamics using independent component analysis. *Neurosci Biobehav Rev*, **30**, 808-822.

Orekhova, E.V., Elsabbagh, M., Jones, E.J., Dawson, G., Charman, T. & Johnson, M.H. (2014) EEG hyper-connectivity in high-risk infants is associated with later autism. *Journal of neurodevelopmental disorders*, **6**, 40.

Ozonoff, S., Iosif, A.-M., Baguio, F., Cook, I.C., Hill, M.M., Hutman, T., Rogers, S.J., Rozga, A., Sangha, S., Sigman, M., Steinfeld, M.B. & Young, G.S. (2010) A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *Journal of American Academy of Child and Adolescent Psychiatry*, **49**, 256-266.

Ozonoff, S., Young, G.S., Belding, A., Hill, M., Hill, A., Hutman, T., Johnson, S., Miller, M., Rogers, S.J., Schwichtenberg, A.J., Steinfeld, M. & Iosif, A.M. (2014) The broader autism phenotype in infancy: when does it emerge? *Journal of the American Academy of Child and Adolescent Psychiatry*, **53**, 398-407.e392.

Ozonoff, S., Young, G.S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Bryson, S., Carver, L.J., Constantino, J.N., Dobkins, K., Hutman, T., Iverson, J.M., Landa, R., Rogers, S.J., Sigman, M. & Stone, W.L. (2011) Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. *Pediatrics*, **128**, e488-495.

Ozonoff, S., Young, G.S., Goldring, S., Greiss-Hess, L., Herrera, A.M., Steele, J., Macari, S., Hepburn, S. & Rogers, S.J. (2008) Gross motor development, movement abnormalities, and early identification of autism. *Journal of autism and developmental disorders*, **38**, 644-656.

Palmer, J.A., Kreutz-Delgado, K. & Makeig, S. (Year) Super-Gaussian Mixture Source Model for ICA. *Proceedings of the Independent Component Analysis and Blind Signal Separation*. Springer Berlin Heidelberg, City. p. 854-861.

Palmer, J.A., Makeig, S., Kreutz-Delgado, K. & Rao, B.D. (Year) Newton method for the ICA mixture model. 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. City. p. 1805-1808.

Paul, R., Fuerst, Y., Ramsay, G., Chawarska, K. & Klin, A. (2011) Out of the mouths of babes: vocal production in infant siblings of children with ASD. *Journal of child psychology and psychiatry, and allied disciplines*, **52**, 588-598.

Piazza, C., Cantiani, C., Akalin-Acar, Z., Miyakoshi, M., Benasich, A.A., Reni, G., Bianchi, A.M. & Makeig, S. (2016) ICA-derived cortical responses indexing rapid multi-feature auditory processing in six-month-old infants. *NeuroImage*, **133**, 75-87.

Pion-Tonachini, L., Kreutz-Delgado, K. & Makeig, S. (2019) ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*.

Pion-Tonachini, L., Makeig, S. & Kreutz-Delgado, K. (2017) Crowd labeling latent Dirichlet allocation. *Knowledge and information systems*, **53**, 749-765.

Rapin, I. & Dunn, M. (2003) Update on the language disorders of individuals on the autistic spectrum. *Brain & development*, **25**, 166-172.

Redcay, E. & Courchesne, E. (2008) Deviant functional magnetic resonance imaging patterns of brain activity to speech in 2-3-year-old children with autism spectrum disorder. *Biol Psychiatry*, **64**, 589-598.

Righi, G., Tierney, A.L., Tager-Flusberg, H. & Nelson, C.A. (2014) Functional connectivity in the first year of life in infants at risk for autism spectrum disorder: an EEG study. *PloS one*, **9**, e105176.

Risch, N., Hoffmann, T.J., Anderson, M., Croen, L.A., Grether, J.K. & Windham, G.C. (2014) Familial recurrence of autism spectrum disorder: evaluating genetic and environmental contributions. *The American journal of psychiatry*, **171**, 1206-1213.

Rozga, A., Hutman, T., Young, G.S., Rogers, S.J., Ozonoff, S., Dapretto, M. & Sigman, M. (2011) Behavioral Profiles of Affected and Unaffected Siblings of Children with Autism: Contribution of Measures of Mother–Infant Interaction and Nonverbal Communication. *Journal of autism and developmental disorders*, **41**, 287-301.

Saffran, J.R. (2001) Words in a sea of sounds: the output of infant statistical learning. *Cognition*, **81**, 149-169.

Saffran, J.R., Aslin, R.N. & Newport, E.L. (1996) Statistical learning by 8-month-old infants. *Science (New York, N.Y.)*, **274**, 1926-1928.

Saffran, J.R., Johnson, E.K., Aslin, R.N. & Newport, E.L. (1999) Statistical learning of tone sequences by human infants and adults. *Cognition*, **70**, 27-52.

Schwartz, S., Kessler, R., Gaughan, T. & Buckley, A.W. (2017) Electroencephalogram Coherence Patterns in Autism: An Updated Review. *Pediatr Neurol*, **67**, 7-22.

Scott-Van Zeeland, A.A., McNealy, K., Wang, A.T., Sigman, M., Bookheimer, S.Y. & Dapretto, M. (2010) No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biol Psychiatry*, **68**, 345-351.

Sheinkopf, S.J., Iverson, J.M., Rinaldi, M.L. & Lester, B.M. (2012) Atypical cry acoustics in 6-month-old infants at risk for autism spectrum disorder. *Autism research : official journal of the International Society for Autism Research*, **5**, 331-339.

Shen, M.D., Kim, S.H., McKinstry, R.C., Gu, H., Hazlett, H.C., Nordahl, C.W., Emerson, R.W., Shaw, D., Elison, J.T., Swanson, M.R., Fonov, V.S., Gerig, G., Dager, S.R., Botteron, K.N., Paterson, S., Schultz, R.T., Evans, A.C., Estes, A.M., Zwaigenbaum, L., Styner, M.A., Amaral, D.G., Piven, J., Hazlett, H.C., Chappell, C., Dager, S., Estes, A., Shaw, D., Botteron, K., McKinstry, R., Constantino, J., Pruett, J., Schultz, R., Zwaigenbaum, L., Elison, J., Evans, A.C., Collins, D.L., Pike, G.B., Fonov, V., Kostopoulos, P., Das, S., Gerig, G., Styner, M., Gu, H., Piven, J. & Infant Brain Imaging Study, N. (2017) Increased Extra-axial Cerebrospinal Fluid in High-Risk Infants Who Later Develop Autism. *Biol. Psychiatry*, **82**, 186-193.

Shen, M.D., Nordahl, C.W., Young, G.S., Wootton-Gorges, S.L., Lee, A., Liston, S.E., Harrington, K.R., Ozonoff, S. & Amaral, D.G. (2013) Early brain enlargement and elevated extra-axial fluid in infants who develop autism spectrum disorder. *Brain : a journal of neurology*, **136**, 2825-2835.

Shultz, S., Vouloumanos, A., Bennett, R.H. & Pelphrey, K. (2014) Neural specialization for speech in the first months of life. *Developmental science*, **17**, 766-774.

Sporns, O. (2011) *Networks of the Brain*. MIT Press.

Srinivasan, R., Winter, W.R., Ding, J. & Nunez, P.L. (2007) EEG and MEG coherence: measures of functional connectivity at distinct spatial scales of neocortical dynamics. *Journal of neuroscience methods*, **166**, 41-52.

Sullivan, M., Finelli, J., Marvin, A., Garrett-Mayer, E., Bauman, M. & Landa, R. (2007) Response to joint attention in toddlers at risk for autism spectrum disorder: a prospective study. *Journal of autism and developmental disorders*, **37**, 37-48.

Sumi, S., Taniai, H., Miyachi, T. & Tanemura, M. (2006a) Sibling risk of pervasive developmental disorder estimated by means of an epidemiologic survey in Nagoya, Japan. *Journal of human genetics*, **51**, 518-522.

Sumi, S., Taniai, H., Miyachi, T. & Tanemura, M. (2006b) Sibling risk of pervasive developmental disorder estimated by means of an epidemiologic survey in Nagoya, Japan. *Journal of human genetics*, **51**, 518-522.

Swanson, M.R., Shen, M.D., Wolff, J.J., Elison, J.T., Emerson, R.W., Styner, M.A., Hazlett, H.C., Truong, K., Watson, L.R., Paterson, S., Marrus, N., Botteron, K.N., Pandey, J., Schultz, R.T., Dager, S.R., Zwaigenbaum, L., Estes, A.M. & Piven, J. (2017a) Subcortical Brain and Behavior Phenotypes

Differentiate Infants With Autism Versus Language Delay. *Biological psychiatry. Cognitive neuroscience and neuroimaging*, **2**, 664-672.

Swanson, M.R., Wolff, J.J., Elison, J.T., Gu, H., Hazlett, H.C., Botteron, K., Styner, M., Paterson, S., Gerig, G., Constantino, J., Dager, S., Estes, A., Vachet, C. & Piven, J. (2017b) Splenium development and early spoken language in human infants. *Developmental science*, **20**.

Tager-Flusberg, H. & Kasari, C. (2013) Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism research : official journal of the International Society for Autism Research*, **6**, 468-478.

Talbott, M.R., Nelson, C.A. & Tager-Flusberg, H. (2015) Maternal gesture use and language development in infant siblings of children with autism spectrum disorder. *Journal of autism and developmental disorders*, **45**, 4-14.

Thiessen, E.D. & Saffran, J.R. (2003) When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Dev Psychol*, **39**, 706-716.

Tierney, A.L., Gabard-Durnam, L., Vogel-Farley, V., Tager-Flusberg, H. & Nelson, C.A. (2012) Developmental Trajectories of Resting EEG Power: An Endophenotype of Autism Spectrum Disorder. *PloS one*, **7**, e39127.

Tomblin, J.B., Records, N.L., Buckwalter, P., Zhang, X., Smith, E. & O'Brien, M. (1997) Prevalence of specific language impairment in kindergarten children. *Journal of speech, language, and hearing research : JSLHR*, **40**, 1245-1260.

Uhlhaas, P.J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A. & Singer, W. (2010) Neural synchrony and the development of cortical networks. *Trends Cogn Sci*, **14**, 72-80.

Wan, M.W., Green, J., Elsabbagh, M., Johnson, M., Charman, T. & Plummer, F. (2012) Parent-infant interaction in infant siblings at risk of autism. *Res Dev Disabil*, **33**, 924-932.

Wang, J., Barstein, J., Ethridge, L.E., Mosconi, M.W., Takarae, Y. & Sweeney, J.A. (2013) Resting state EEG abnormalities in autism spectrum disorders. *Journal of neurodevelopmental disorders*, **5**, 24.

Webb, S.J., Bernier, R., Henderson, H.A., Johnson, M.H., Jones, E.J.H., Lerner, M.D., McPartland, J.C., Nelson, C.A., Rojas, D.C., Townsend, J. & Westerfield, M. (2015) Guidelines and best practices for electrophysiological data collection, analysis and reporting in autism. *Journal of autism and developmental disorders*, **45**, 425-443.

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*.

Wolff, J.J., Botteron, K.N., Dager, S.R., Elison, J.T., Estes, A.M., Gu, H., Hazlett, H.C., Pandey, J., Paterson, S.J., Schultz, R.T., Zwaigenbaum, L. & Piven, J. (2014) Longitudinal patterns of repetitive behavior in toddlers with autism. *Journal of child psychology and psychiatry, and allied disciplines*, **55**, 945-953.

Wolff, J.J., Gerig, G., Lewis, J.D., Soda, T., Styner, M.A., Vachet, C., Botteron, K.N., Elison, J.T., Dager, S.R., Estes, A.M., Hazlett, H.C., Schultz, R.T., Zwaigenbaum, L. & Piven, J. (2015) Altered corpus callosum morphology associated with autism over the first 2 years of life. *Brain : a journal of neurology*, **138**, 2046-2058.

Wolff, J.J., Gu, H., Gerig, G., Elison, J.T., Styner, M., Gouttard, S., Botteron, K.N., Dager, S.R., Dawson, G., Estes, A.M., Evans, A.C., Hazlett, H.C., Kostopoulos, P., McKinstry, R.C., Paterson, S.J., Schultz, R.T., Zwaigenbaum, L. & Piven, J. (2012) Differences in white matter fiber tract development

present from 6 to 24 months in infants with autism. *The American journal of psychiatry*, **169**, 589-600.

Xiao, R., Shida-Tokeshi, J., Vanderbilt, D.L. & Smith, B.A. (2018) Electroencephalography power and coherence changes with age and motor skill development across the first half year of life. *PloS one*, **13**, e0190276.

Yirmiya, N., Gamliel, I., Pilowsky, T., Feldman, R., Baron-Cohen, S. & Sigman, M. (2006) The development of siblings of children with autism at 4 and 14 months: social engagement, communication, and cognition. *Journal of child psychology and psychiatry, and allied disciplines*, **47**, 511-523.

Yirmiya, N., Gamliel, I., Shaked, M. & Sigman, M. (2007) Cognitive and verbal abilities of 24- to 36-month-old siblings of children with autism. *Journal of autism and developmental disorders*, **37**, 218-229.

Yoder, P., Stone, W.L., Walden, T. & Malesa, E. (2009) Predicting social impairment and ASD diagnosis in younger siblings of children with autism spectrum disorder. *Journal of autism and developmental disorders*, **39**, 1381-1391.

Young, G.S., Rogers, S.J., Hutman, T., Rozga, A., Sigman, M. & Ozonoff, S. (2011) Imitation from 12 to 24 months in autism and typical development: a longitudinal Rasch analysis. *Dev Psychol*, **47**, 1565-1578.

Yuen, R.K., Thiruvahindrapuram, B., Merico, D., Walker, S., Tammimies, K., Hoang, N., Chrysler, C., Nalpathamkalam, T., Pellecchia, G., Liu, Y., Gazzellone, M.J., D'Abate, L., Deneault, E., Howe, J.L., Liu, R.S., Thompson, A., Zarrei, M., Uddin, M., Marshall, C.R., Ring, R.H., Zwaigenbaum, L., Ray, P.N., Weksberg, R., Carter, M.T., Fernandez, B.A., Roberts, W., Szatmari, P. & Scherer, S.W.

(2015) Whole-genome sequencing of quartet families with autism spectrum disorder. *Nature medicine*, **21**, 185-191.

Zwaigenbaum, L., Bauman, M.L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., Mailloux, Z., Smith Roley, S., Wagner, S., Fein, D., Pierce, K., Buie, T., Davis, P.A., Newschaffer, C., Robins, D., Wetherby, A., Stone, W.L., Yirmiya, N., Estes, A., Hansen, R.L., McPartland, J.C. & Natowicz, M.R. (2015) Early Intervention for Children With Autism Spectrum Disorder Under 3 Years of Age: Recommendations for Practice and Research. *Pediatrics*, **136 Suppl 1**, S60-81.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J. & Szatmari, P. (2005) Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*, **23**, 143-152.