



UNIVERSITY OF PADOVA

Department of Developmental Psychology and Socialization*

Master Degree in Developmental and Educational Psychology

Final dissertation

**The Effects of Technology Use on Infant Language
Development: Study of an Italian Cohort**

Supervisor

Professor Silvia Benavides-Varela

Co-supervisor

Dott.ssa Tamara Bastianello, PhD.

***Candidate: Adriana Grazia-Maria
Angelina Loschiavo***

Student ID number: 2108660

Academic Year 2025/2026

Abstract

The use of technology and screen media has become increasingly prevalent in early childhood. International guidelines discourage screen media for children under two years, aside from videochats with caregivers, so that sedentary behaviour does not become a standard part of the infant's routine (Ponti et al., 2023). Despite these recommendations, parents have reported that infants between the ages of 0 and 2 are exposed to more than an hour of screen media per day (Alroqi et al., 2022; Balbinot et al., 2016) and that infants under the age of two own their own cell phones and tablets (Alroqi et al., 2022). A growing body of research cautioned that early and prolonged exposure to screen media may negatively affect several domains of child development, including language development (Operto et al., 2020). This is particularly concerning because language is used for communication and self expression (Crystal, 2018), and its development can be impacted by the quantity and quality of the language the infant is exposed to (Hart et al., 1977), including the ways they interact with technology and screen media. However, the measurements for infant language development as well as technology use and screen media exposure have relied primarily on parent reports, which may be subject to artifacts such as social desirability bias and recency effects.

This thesis project has three main aims. The first is to validate the use of a novel, objective measure, the Language ENvironment Analysis (LENA) system for its ability to reliably measure television and electronic sounds (TV/ES) in an Italian sample. The second is to describe the acoustic environment that Italian infants are exposed to in a typical day, and to describe the pattern of the infant's language development. The third is to investigate the effect of TV/ES exposure on infant language development.

Twenty-eight infants were followed longitudinally across four key milestones in linguistic development, i.e., pre-babbling, babbling, first word production, and vocabulary spurt phase. During each recording session, families were given the LENA device to record the auditory environment of the infant and completed the Italian version of the MC-BDI (Caselli et al., 2015) to confirm the infant's linguistic phase of development. A reliability analysis based on 2,100 minutes of recordings revealed no significant differences between LENA-derived and manually coded TV/ES exposure, with a weak to moderate correlation between the human coder and the LENA device. Infant media exposure was shown to vary across the first two years of life, as well

as peak in specific times of the day. Mixed-effects models on child's language development indicated significant changes across the phases of linguistic development, and a significant interaction between child vocalizations and their exposure to TV/ES. Infants with a higher exposure to TV/ES had lower vocalization counts in the vocabulary spurt phase as compared to all other phases of development. These findings support using LENA to measure TV/ES exposure on Italian-speaking infants and support previous literature that describes how increased exposure to TV/ES may have a negative effect on infant language development. This highlights the importance in considering screen-time and related environmental factors when examining early language development.

Acknowledgements

I would like to express my gratitude to my supervisor, Prof. Silvia Benavides-Varela for her guidance, encouragement, and support for me throughout this project. I am especially thankful for the trust she placed in me to pursue an original research idea, and to explore my own scientific interests. I am also grateful to Dr.ssa Tamara Bastianello, for always making time to answer all of my questions, and providing thoughtful feedback at every step of this process. I am also thankful for the opportunities she gave me to contribute to manuscript preparation, experiences that strengthened my skills in literature review and academic writing. These opportunities have been invaluable in preparing me for the next stages of my academic career.

I would also like to express my thanks to my family and friends, who have supported me throughout this process. To Chiara Nascimben for her guidance in hands-on research experience in the lab and assistance with statistical analyses. To my parents, who supported me both emotionally and financially, and who've always encouraged me to chase my dreams. Most importantly, to my partner, who has been there every step of the way throughout this journey. I am thankful for his patience and for believing in me throughout the entirety of this process.

And to everyone who has supported, encouraged, guided, and believed in me throughout this journey, I thank you. This accomplishment would not have been possible without the kindness of so many people, and I will carry these experiences with me into the next chapter of my career.

Table of Contents

Chapter 1: Infant Language Development	8
Language and Communication	8
Theories of Language Development	10
Behaviourism	10
Nativism	13
Interactionism	16
Constructivism and Neuroconstructivism	19
Linguistic Milestones	21
Gestation & Birth: Perceptual Processes	21
Prelinguistic Production: The Cry, Coo, and Vocal Play Stages	22
Babbling	27
Nonverbal Communication to Promote Language Comprehension and Production	29
Protowords and First Words	31
The Vocabulary Spurt	32
Summary	34
Chapter 2: The Role of the Adult and Sources of Interruption in Infant Language Development	35
The Role of The Adult	35
Quantity and Quality of Linguistic Input	35
Motherese/Parantese	37
Contingency	38
Effects of the Family Context	39
Socioeconomic Status (SES)	39
Acoustic Features of the Environment	41
Noise	41
Television and Screen Media	43
Features of Technology Use and Screen Media Exposure	44
Infant Use	44
Background Television	46

Adult Use.....	47
Chapter 3: The Effects of Technology Use on Infant Language Development: Study of an Italian Cohort	51
Technology Use in the Italian Context	52
Some Methodological Concerns to Consider	54
The Relationship Between TV/ES and Language Development in Infancy	56
The Present Study	57
Chapter 4: A Continuation of The Effects of Technology Use on Infant Language Development: Study of an Italian Cohort.....	59
Method	59
Procedure	59
Participants	59
The Checklist:.....	60
The Language ENvironment Analysis (LENA) System.....	61
Statistical Analyses.....	63
Aim 1: Validating the Use of the LENA Device to Measure TV/ES Exposure	63
Aim 2: Understanding the Types of Sounds that Italian Infants Hear in a Typical Day	63
Aim 3: Understanding TV/ES Exposure Over the Course of Infant Language Development	64
Results	64
Aim 1: Validating the Use of the LENA Device to Measure TV/ES Exposure	64
Aim 2: Understanding the Types of Sounds that Italian Infants Hear in a Typical Day	65
Aim 3: Understanding TV/ES Exposure Over the Course of Infant Language Development	69
Discussion	74
Aim 1: Validating the Use of the LENA Device to Measure TV/ES Exposure	75
Aim 2: Understanding the Types of Sounds that Italian Infants Hear in a Typical Day	75
Aim 3: Understanding TV/ES Exposure Over the Course of Infant Language Development	77
Limitations	80
Conclusion.....	81

References	83
Appendix	105

Chapter 1: Infant Language Development

Language and Communication

Language is the systematic and conventional use of sounds, signs, or written symbols, for the purpose of communication or self expression (Crystal, 2018). The language system consists of basic conventional units and rules, or grammar, for its arrangement in order to convey information (Ervin-Tripp, 1966; Wright, 2008). Most human languages employ speech sounds to represent the basic units, and the written languages are derived from these oral counterparts (Ervin-Tripp, 1966). While all typically developing children in typical environments acquire language (Hoff-Ginsberg & Shatz, 1982), language itself is comprised of many facets that the child must learn in order to be considered as having linguistic competence (Gleason & Philips, 2024), which is not a simplistic task.

The multiple facets of language development include phonological, lexical, morphological, and syntactical development. Phonological development includes the ability to discriminate, recognize, and produce the sounds of the language (Gleason & Philips, 2024; Hoff, 2014) as well as to understand the systems of the language (Hoff, 2006). For example, children who learn English will learn approximately 44 different vowel and consonant sounds (Hoff, 2014), 30 in Italian (Paoli, 2016), 35 in French (Tranel, 1987), and 403 in standard Mandarin, Chinese (Greenhoe, 2017). Lexical, or semantic, development involves understanding the words of the language and their associated meanings (Gleason & Philips, 2024; Hoff, 2006). Children must learn what sounds can be combined into which possible words as well as the knowledge of each word's meaning, such as if the word maps onto a specific object, person, or concept (Hoff, 2006). In English, people will develop a vocabulary of tens of thousands of words by adulthood (Hoff, 2006). Morphological development involves the knowledge of morphemes, which are the smallest pieces of linguistic information that serve a grammatical function (Aronoff & Fudeman, 2022; Gleason & Philips, 2024). This requires the development of knowledge about the system that is used to combine units of meanings into words (Aronoff & Fudeman, 2022; Hoff, 2006) as well as the internal structure of the words themselves (Aronoff & Fudeman, 2022). Syntactical development is learning the system to combine words into sentences (Gleason & Philips, 2024; Hoff, 2006) and knowing how words can be systematically combined with other words to form sentences (Gleason & Philips, 2024; Hoff, 2006).

Other features of language may include the prosody and pragmatics, or context, of the language. Prosody involves the stress, rhythm, and intonation of the language (Speer & Blodgett, 2006), that allows the user to communicate different attitudes or meanings of the utterance (Levey, 2024). Listeners are able to use the overall prosodic structure of the utterance in order to aid with syntactic parsing (analyzing a sentence's grammatical structure to determine the relationships between words) and syntactic constituency relations (analyzing the sentence to figure out its grammatical structure, specifically, how words are grouped together and related to each other) (Speer & Blodgett, 2006). Infants demonstrate a preference for their native language over an unfamiliar language as early as four days old, even when only the prosodic information of the language is presented (Mehler et al., 1988). Adult caregivers are able to act as communicative partners, or models, to help the infant form a relationship between the environment and themselves (Fodor, 1996 as cited in Hoff-Ginsberg & Shatz, 1982), which can aid in the development of pragmatic skills and effectively use the components of the language (Hoff, 2006). Reaching linguistic competence in a language includes the ability to combine sentences into larger units of discourse, tell stories, hold a conversation, or communicate thoughts and feelings with others (Hoff, 2014).

Communication, on the other hand, involves more than the spoken components of language (Wright, 2008). It includes other forms of communication, including body language and facial expressions, as well as other forms of expression, including art, photography, and dance (Finnegan, 2002, as cited in Wright, 2008). According to Wright (2008), it involves the speaker, or sender, and the listener, or receiver. The speaker must first have an intention to initiate communication and encode information in the form of a message. Here, the emphasis is focused on the speaker's intention, to exclude any unintentional messages that would not be considered communication. Meanwhile, the receiver must possess the ability to interpret these communication signals, which places emphasis on the extent to which the sender and receiver share similar signs or coding systems for communication. Paralinguistic cues, such as intonation or facial expressions, accompany spoken language and aid in the receiver's ability to better understand the speaker's meaning (Levey, 2024). It should be noted that there are factors other than exposure to a language that shape our interpretations of non-verbal signals and paralinguistic cues, including education, culture, and socio-economic conditions (Hall, 1997 as cited in Wright, 2008). Domains of knowledge in both verbal and non-verbal language are developed concurrently, whereby

developing knowledge in one domain often aids or influences acquiring knowledge in others (Hoff, 2014).

Theories of Language Development

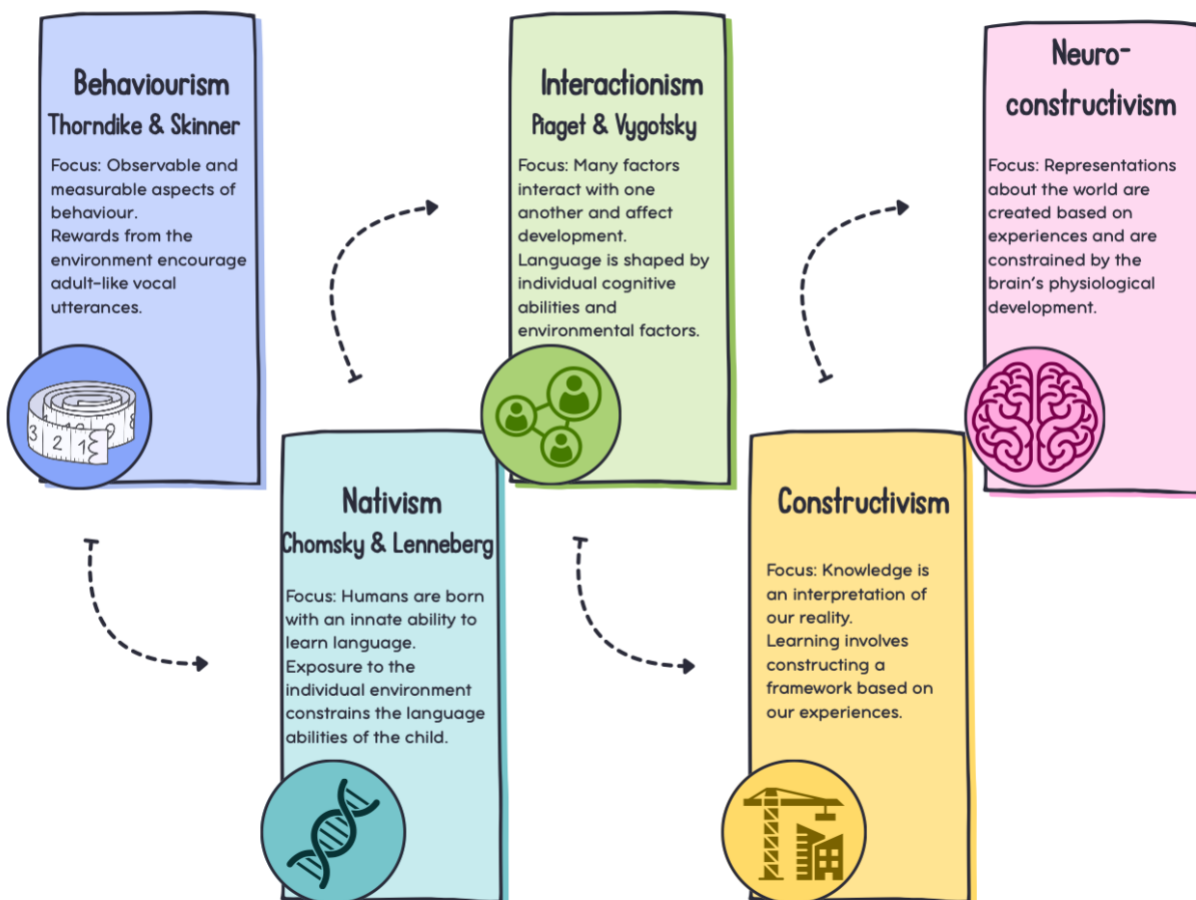


Figure 1: Progression of theories of language development.

Behaviourism

Theories for how children develop a language have been proposed in various approaches to psychology (Figure 1). One of the first approaches to apply a theory to and systematically study language development was behaviourism, which deals only with observable behaviours in studying the human psyche (Shahini, 2025; Stangor & Walinga, 2014). Behaviourism principles involve observable behaviours, conditioning, stimulus-response associations, shaping, and reinforcement (Shahini, 2025).

Focusing only on the observable and measurable aspects of language behaviour, Edward Thorndike proposed a basic reward model (Thorndike, 1911, as cited in Kymissis & Poulson, 1990). This theory was derived from the idea that infants instinctively produce sounds, and some of those sounds resemble words. Only those specific vocalizations are rewarded by the environment, such as through the caregiver's attention or verbal acknowledgement, leading to a feeling of satisfaction in the infant and an increase in the probability that the infant will make those, or similar, vocalizations again. Thorndike later refined some of the ideas from this basic model and put forward the theory of learning. The theory of learning states that behaviours followed by annoyance or that are unused will be weakened, while behaviours followed by satisfaction or that are used more often will be strengthened (Thorndike, 1932, as cited in Stangor & Walinga, 2014). Thorndike also pointed out that learning occurs automatically and incrementally over time. In other words, argument for infant language development from an early behaviourist perspective is that infants automatically produce vocalizations, some of which will sound like adult words, and only those will be rewarded, leading the infant to produce more of those exact or similar vocalizations (Figure 2). Slowly, over time, the infant will learn to only produce vocalizations that imitate adult words.

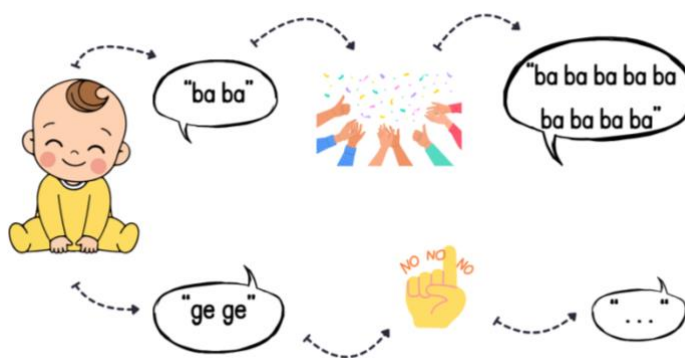


Figure 2: Progression of language production with adult feedback.

Other behaviourists developed models based on different types of conditioning behaviour. Classical conditioning is the idea that learning can occur when a neutral or conditioned stimulus becomes associated with a stimulus that naturally produces behaviour, or an unconditioned stimulus, eventually eliciting a similar response, also known as a conditioned response (Stangor & Walinga, 2014). As an example, used by Hamamouche & Bohannon III (2024), when an infant touches a hot stove (the unconditioned stimulus), they may experience physiological pain (an unconditioned response). If the caregiver shouts "Hot!" as this occurs, the infant may also associate the word "hot" (the conditioned stimulus) with the physiological pain they experienced (the

conditioned response). As the meaning of the word has been internalized, it can then be used as the unconditioned stimulus in other or similar circumstances, such as when looking at a campfire, in order to modify the infant's behaviour, or paired with other words, such as "fire" which the infant will come to associate with a similar internal meaning (Hamamouche & Bohannon III, 2024). Through classical conditioning, associations can be formed between arbitrary verbal stimuli and internal responses, creating a source of word meanings for infants (Staats, 1971, as cited in Hamamouche & Bohannon III, 2024)

Skinner, a behaviourist in the 1950s, however, argued that language cannot be looked at through individual linguistic units, as per the theory generated by Thorndike, and must instead be examined as a whole, similar to any other observable behaviour (Skinner, 1957, as cited in Hamamouche & Bohannon III, 2024). Skinner was known for his studies on reflexes, defined as a reliable production of a particular response by a specified stimulus (Catania, 1984, as cited in Kymissis & Poulson, 1990). According to Skinner, all responses are reflexes or chains of reflexes (Kymissis & Poulson, 1990). In this case, a stimulus, which is anything that can trigger a physical or behavioural change, is paired with a response, which is the observable behaviour that is exhibited by a reaction to that stimulus (Shahini, 2025). Skinner also placed emphasis on the idea that behaviours may be strengthened or weakened over time depending on the consequences of that behaviour (Skinner, 1938, as cited in Kymissis & Poulson, 1990), also known as operant conditioning. In this case, productive speech can be shaped by differential reinforces and punishments in the infant's environment (Hamamouche & Bohannon III, 2024). Verbal behaviour or speech, then, will only be rewarded when it closely approximates adult speech (Hamamouche & Bohannon III, 2024), and meaningless or inappropriate speech is "punished" (Kymissis & Poulson, 1990). As the child's utterances begin to more gradually approximate adult utterances, the rewards from the environment (e.g., caregivers and other adults) will change to encourage meaningful and adult-like utterances (Hamamouche & Bohannon III, 2024). Interaction between the infant's verbal production and the adult's responses are crucial in the language acquisition process (Bijou & Baer, 1965, as cited in Kymissis & Poulson, 1990).

Evidence to support the behaviourism perspective comes mainly from behaviour modification techniques that are applied to children with limited speech skills, primarily in the form of imitation and reinforcement techniques. However, language development during infancy is not shaped in

the same careful and structured way as behaviour modification strategies are administered (Hamamouche & Bohannon III, 2024). Young infants do not show similar patterns of reinforcement and imitation as do older children with limited speech skills (Hamamouche & Bohannon III, 2024). Sounds that adults produce have been shown to be different from sounds that infants produce, which may weaken properties of the caregiver's utterance, as the infant is unable to receive true contingent vocal behaviour (Risley, 1966; 1977, as cited in Hamamouche & Bohannon III, 2024). Caregivers are more likely to praise infants for any utterance or vocalization, regardless of correctness (Brown & Hanlon, 1970, as cited in Hamamouche & Bohannon III, 2024; Hirsh-Pasek et al., 1984, as cited in Hamamouche & Bohannon III, 2024). Imitative behaviours are bi-directional, as mothers imitate their children as much as children imitate their caregivers (Pelaez et al., 2011, as cited in Hamamouche & Bohannon III, 2024). And children may also have their own internal linguistic processes in operation regardless of caregiver reinforcement or response, as they often produce their own novel linguistic forms (Kymissis & Poulson, 1990). Instead, it may be a general rule that an environment that is responsive to a child is more likely to foster language development. For example, it's been shown that caregivers who demonstrate contingent responding encourage children to babble with more phonemes of their native language (Goldstein & Schwade, 2008), and that children of caregivers rated higher on levels of responsiveness demonstrate more rapid language growth (Hsu et al., 2000; Laakso et al., 1999; Tamis-LeMonda et al., 1996, 2014).

Nativism

Language learning happens to the child when they are placed in an appropriate environment (Chomsky, 2015). However, according to Chomsky (1981, as cited in Fernald & Marchman, 2006) speech from adults is full of hesitations, false-starts, mispronunciations, and ungrammaticalities. Learning from imitation or reinforcement alone may be hindered by the fact that linguistic input is limited in the scope and extent of syntactic information that it can provide (Gleitman et al., 1988, as cited in Fernald & Marchman, 2006). Therefore, the environment itself may not provide the regularities needed for children to acquire language. The nativist perspective, then, argues that children acquire language rapidly, effortlessly, and without direct instruction (Hoff, 2014). Language acquisition is dependent upon innate knowledge (Hamamouche & Bohannon III, 2024; Hoff, 2014) and all human beings are born with the ability to learn language to some degree

(Shahini, 2025). The idea that the structure or principles of language are universal and innate, and shared by all human individuals, is also known as the theory of universal grammar (Hoff-Ginsberg & Shatz, 1982; Shahini, 2025). Any cognitive or neural structures or behaviours that are related to the development of linguistic abilities are situated in the brain and waiting to develop (Shahini, 2025). These structures and behaviours may assist language acquisition by enabling children to deduce the specific parameters necessary for the language they are learning (Fernald & Marchman, 2006; Shahini, 2025). Although while nativists for the theory of universal grammar argue that the mechanism that exerts influence over these brain structures remains unspecified (Hoff-Ginsberg & Shatz, 1982), Chomsky argues for a language acquisition device (Chomsky, 1965, as cited in Hamamouche & Bohannon III, 2024). This device is a fixed, genetically determined system that constrains the forms of language that may appear (Chomsky, 1980, as cited in Cicourel, 1982). The exact structure of the language that is produced is set in motion through exposure to the infant's environment. In other words, while the rules of universal grammar are available to the infant from birth, it is the exposure to information in the environment that reveals the structures available to the infant for language development (Chomsky, 1980, as cited in Cicourel, 1982). Therefore, the language acquisition ability of the infant depends on the linguistic input available in the environment. Without these additional constraints, the presentation of language is not sufficient to guarantee language acquisition (Hoff-Ginsberg & Shatz, 1982). From this perspective, only the auditory information is important for language development. The communicative function of the language becomes irrelevant (Hoff, 2014).

Alongside the idea that children are born with an innate ability to learn a language is the idea that this ability is not available to the infant forever. Lenneberg (1967) proposed that there may be a critical period of development, where language learning may stop (as cited in Hoff, 2014). The three main claims of his critical period theory were that (1) adult native speakers can identify non-native speakers by their accent immediately and accurately, (2) there is a loss of brain plasticity around the age of puberty that may account for the emergence of foreign accents, and (3) that a critical period hypothesis only applies to productive abilities, such as speech and the native accent, and does not affect other areas of linguistic competence (Scovel, 2000). Support for the critical period theory comes from the fact that infants and young children are able to master a language system completely, while adults are unable to master, at least, the complete sound patterns of the language system (Scovel, 1969). There is considerable variation on the extent to which this may

be true, and discrepancies are often attributed to differences in the environment in which the language is learned (Scovel, 1969) and the language learning abilities of the adult, including the conscious versus unconscious awareness of the language (Wolfe, 1967, as cited in Scovel, 1969), or the type of processing that occurs during language acquisition, such as visual to auditory or auditory to auditory (Geschwind, 1966, as cited in Scovel, 1969). The cerebral dominance or left-hemisphere lateralization of language may also be a reason for the differences seen in the capacity to learn a language as an adult compared to as an infant (Scovel, 1969). This lateralization or transfer between hemispheres is proposed to stop around the time of puberty (Lenneberg, 1967, as cited in Hoff, 2014), meaning that the "cut-off" for learning a language in its entirety is around the age of 12 (Scovel, 1969). Evidence from neuroimaging studies demonstrates that younger, healthy children, are more likely to show an atypical pattern of language distribution than older, healthy children, suggesting that there is an increasing specialization of language functions to the left hemisphere with age (Holland et al., 2007; Szaflarski et al., 2006). However, the lateralization changes may also depend on the type of task, as lexical and semantic tasks yielded stronger patterns of lateralization, specifically in the anterior and posterior language regions, than did syntactic and prosodic processing tasks (Holland et al., 2007).

Further support for the theories of universal grammar and the critical period comes from observation and neuroimaging studies, as well as studies on atypically developing populations. Observations of classic errors in infant language development from Slobin (1986) demonstrated that regardless of the word order for subjects and objects in the native language, infants and young children tend to prefer subject-object word order (as cited in Hamamouche & Bohannon III, 2024). Meanwhile, observations from Brown and Bellugi (1964) noted that toddlers and young children show an overgeneralization or inappropriate application of grammar rules (as cited in Hamamouche & Bohannon III, 2024). Both studies suggest that children may converge on innate grammar rules, as these are not errors they would have been exposed to in adult speech. Neuroimaging studies also provide evidence that infants begin life with the neurological systems that allow them to acquire language through multiple modalities, while adults, instead, have a highly modularized system, dramatically decreasing the flexibility that is allowed to an infant (for a review see Kuhl, 2011). Research on children who are deaf or isolated show that children not exposed to verbal language from early infancy have a difficult time developing language abilities at later ages (Grimshaw et al., 1998; see also Hamamouche & Bohannon III, 2024). While there is

much evidence for these theories, there is also evidence to suggest that the nativist perspective is not sufficient to explain all language acquisition throughout infancy. Children acquire language, such as first words, and grammar abilities at differential rates across and within cultures (Hart et al., 1997; Hoff, 2006; Laakso et al., 1999; Rowe, 2012; Tamis-LeMonda et al., 1996; Weisleder & Fernald, 2013; see Dabrowska, 2015 for a review). In order to learn the language, children also require contingent interaction with capable adults (Tamis-LeMonda et al., 1996; see also Dabrowska et al., 2015 for a review). And neuroimaging studies have revealed that regions identified in the brain for language processing are neither highly localized nor are activated only during linguistic processing tasks (for a full review see Dabrowska, 2015).

Interactionism

As an attempt to compromise theories of behaviourist and nativist perspectives, interactionism assumes that many factors affect the course of an infant's development, and that these factors are mutually dependent upon, interact with, and modify one another (Hamamouche & Bohannon III, 2024). Both cognitive and social factors modify language development, and language development modifies the development of cognitive and social factors (Vygotsky, 1962, as cited in Hamamouche & Bohannon III, 2024). It is the interaction of innate characteristics of the child's mind and the child's language experience (Hoff, 2014). While there is still a high burden placed on the language-learning experience, which is heavily influenced by the cultural environment the child is placed in, the child in this perspective is a much more active participant in their development (Hoff, 2014). At the forefront of this perspective, are theories by Jean Piaget and Lev Vygotsky.

According to Piaget, language is structured, or constrained, by reason. Basic linguistic developments must be derived from even more basic, general changes in cognition, *implying* that linguistic development is driven by cognitive development (Hamamouche & Bohannon III, 2024). However, instead of an infant having innate cognitive structures, Piaget argued that infants are born with innate cognitive schemas of action that interpret perceptual experiences and create knowledge based on action (as cited in Cicourel, 1980). When an action is repeated or generalized through interaction with objects, a schema is created (Cicourel, 1980). The infant can then use that schema to assess further actions with new objects and assimilate the properties of various objects, which is when development occurs (Cicourel, 1980). The cognitive schemas would become

progressively better over time and never revert (Cicourel, 1980). As a result of the continuing interaction between the child's cognitive ability and their linguistic (and non-linguistic) environment, the structures for language development emerge (Hamamouche & Bohannon III, 2024). The ability to express a concept through language would depend upon the infant's ability to comprehend that concept (Hamamouche & Bohannon III, 2024). As well, the different ways in which the child reasons about the world may affect language acquisition (Hamamouche & Bohannon III, 2024). The final resulting structure may not resemble neither the structure of the environment nor the structure of the innate cognitive schemas from which the infant began exploring their environment (Bates & Syndar, 1987, as cited in Hamamouche & Bohannon III, 2024).

Vygotsky, on the other hand, explored more closely how learning was shaped by context, language, and cultural tools (Shahini, 2025). His sociocultural theory puts forward that language develops out of social interaction, and that the environment plays the largest role in development (Vygotsky, 1986). Vygotsky also stresses the importance of children's early social and communicative experiences for language development (Vygotsky, 1986; see also Elardo et al., 1977; Hoff, 2006; Hoff-Ginsberg & Shatz, 1982; Newport et al., 1977; Rowe, 2012, 2018), as these experiences assist in the development of higher order functions (Vygotsky, 1986). According to Lantolf (2000), these functions appear twice in the life of the individual. The first time is between the infant and the social partner, on an interpsychological level, where the infant can perform a task with help. This is also known as the zone of proximal development (Figure 3), or the difference between what an individual can achieve when acting alone as compared to acting with support from another (Lantolf, 2000). Later, the



Figure 3: Vygotsky's Zone of Proximal Development.

functions appear for a second time, when the individual is able to perform the task alone, often through the use of psychological mediation. In the case of language development, initially

caregivers may phonetically emphasize important content words, slow the rate of their speech, repeat themselves frequently, and supply critical nonverbal cues, such as pointing, to aide with the development of phonetic and syntactic skills and understanding (Hamamouche & Bohannon III, 2024).

Language can also be used to model or scaffold an action or ability, for which the caregiver will monitor the infant's performance of the task, increasing or decreasing the amount of monitoring and assistance given, until the infant is able to perform the task on their own (Bruner, 1985; see Mermelshtine, 2017 for a review). This strategy that can be employed to promote an infant's development within their zone of proximal development (Mermelshtine, 2017). In order for this process to be useful, the caregiver must be able to respond contingently to the infant's activities (Wood, Bruner, & Ross, 1967, as cited in Mermelshtine, 2017) and the guidance must be tied to the level of the infant's current understanding (Heckhausen, 1987) in order to promote an infant's problem-solving abilities and autonomy in goal-directed activities. This is a difficult task as the 'correct' steps may not be easily determined (Wertsch, 1979). In a study by Heckhausen (1987), mothers of 14-month-old infants were observed in a laboratory-based setting as their infant was instructed to complete a given task. No direct instructions to the mothers were provided, other than to behave as usual. During the observations, mothers seemed to intuitively provide a one-step-ahead strategy, where they refrained from offering task-related actions if the infants attempted a task, and provided task motivation when infants were not attempting the task. As the infants became more capable in their ability to perform the task, mothers became less explicit in their direction. In language development, caregiver-responsiveness, such as the ones described above, to their infants can be considered a form of scaffolding, as caregivers have been shown to modify their responses in line with infant development, which is linked to the promotion of communicative skills well before infants are able to produce conventional words (Tamis-LeMonda et al., 2014). Some other forms of scaffolding were described in a review by Tamis-LeMonda and colleagues (2013, as cited in Tamis-LeMonda et al., 2014), where mothers responded with mostly simple labels and descriptions to their one-year old infants, but increased responsive questions across the second year of life, as infants became more skilled at language (Bornstein et al., 2008), and increased referential responses to vocalizations and decreased responses to gestures of their infants between 14 and 24 months (Tamis-Lemonda et al., 2013, as cited in Tamis-LeMonda et al., 2014).

Constructivism and Neuroconstructivism

Although the theories of the interactionism have support, constructivism attempts to construct a model of how learning takes place, taking a more practical approach at infant cognitive development (Cobern, 1993). Constructivism argues that all knowledge lacks exactitude and comprehensiveness, so we can never know how close our knowledge approximates reality (Cobern, 1993). Knowledge, instead, is a meaningful interpretation of our experiences of reality (Cobern, 1993). We learn by making sense of what we experienced and constructing a conceptual framework (Cobern, 1993). Neuroconstructivism takes this a step further and places emphasis on the interrelation between brain and cognitive development (Sirois et al., 2008). Sirois and colleagues (2008) explained that, to the neuroconstructivist, development is the progressive increase in the complexity of representations, with the consequence that new competences can develop based on earlier, simpler ones. In other words, infants create representations about the world around them that have emerged within the context of cooccurring molecular and neural development. The physiological development of the brain places constraints on what an infant is able to experience, process, and do. These representations will increase in complexity throughout development and through interactions with the world as a natural process of a complex biological system through the processes of competition, cooperation, and chronotopy (Sirois et al., 2008). Competitions between neurons and brain regions means that only a subset of the initial contributors to an immature function will be involved in the mature function, to allow for minimal but stable representations. Cooperation integrates multiple contributors to a function to have efficiency in the coordination of the interrelated function. This process integrates multiple contributors to a function to have efficiency in the coordination of interrelated functions, and allows for enough redundancy to be relatively robust to damage. Chronotopy is the process that describes how some patterns of gene expressions are restricted to specific developmental times, and how some key aspects of neuronal development rely on a specific sequence of events. For example, gene expression is what drives progressive specialization, where the state of the system at any given time constrains the future state of that system. With regard to language development, newborns are born with the ability to distinguish speech sounds from all human languages (Werker & Tees, 1984), which decreases substantially after six months of age (Stager & Werker, 1997, as cited in Sirois et al., 2008; Trehub, 1976, as cited in Sirois et al., 2008). After which, older infants, toddlers, and even

adults, may no longer have the same capabilities to process speech sounds in languages other than their native language (Holland et al., 2007; Scovel, 1969; see also Hoff 2014).

The neuroconstructivist approach attempts to explain many processes in infant language development. As an example from phonological development, when young infants begin to babble, it creates a coupling between the perception and production of language (Sirios et al., 2008). Alongside the infant's cognitive maturation, the babbling becomes more articulate and progressively closer to the infant's native language (De Boysson-Bardies et al., 1989; Sirois et al., 2008; see also Sasaki et al., 1977; Stoel-Gammon et al., 2024). However, deaf infants babble later than hearing infants and produce different sounds, which are long-lasting and have negative effects on later speech (Wallace et al., 1998, as cited in Sirois et al., 2008). As auditory feedback is necessary for the coordination of the phonatory (larynx) and articulatory (vocal tract) systems for babbling and speech (Koopmans-van Beinum et al., 2001), deaf infants may not be able to build the auditory representation for the sounds of their native language. Their vocalization abilities are later constrained by the missing experiences that contribute to a hearing infant's language development. This is also supported by the mechanistic model from Westermann and Miranda (2004), whereby the researchers argue for the existence of two topographical maps, for articulation and perception, respectively. Neurons within each topographical map respond to input from their respective fields. Connections between the two maps are strengthened when there is high covariation between maps, and weakened when there is low covariation. This creates local competition within and cooperation between the two maps to couple the articulatory and perceptual processes. The synchronized associations of these processes become prototypes for the infant and are used as representations, or models, for future development when the infant is exposed to auditory input. Environmental context is therefore central to the infants' experiences and abilities to form accurate representations (Sirois et al., 2008). Auditory input may only lead to partial representations within these maps. Through babbling, the infant is also able to couple their perception and production of sounds with the initial model, making it progressively more selective to specific environmental input (Westermann & Miranda, 2004) and assisting in tuning the infant's own babbling to follow the sound patterns and prosody of the native language. Speech sounds then become activation patterns for both maps, and can no longer be isolated from auditory input (Sirois et al., 2008). This leads to enhanced auditory and productive performance within their native

language, at the expense of discrimination abilities in other languages (Westermann & Miranda, 2004).

Linguistic Milestones

Gestation & Birth: Perceptual Processes

Extremely young infants have speech perception capabilities that are relevant to language acquisition, and provide the infant with a framework for organizing linguistic input (Mehler et al., 1988). This framework is essential to the task of separating non-verbal from linguistic input, and non-native from native sounds (Mehler et al., 1988). The infant's brain, as young as a newborn, shows some specialization for speech processing, although these abilities are broad (Gervain, 2015). For example, although newborns have shown a preference for and selectively attend to speech over non-speech analogues (Vouloumanos & Werker, 2007, as cited in Gervain, 2015), this tuning ability also shows preference for primate vocalizations that are acoustically similar to human speech, until the infant reaches about 3 months of age (Vouloumanos et al., 2010, as cited in Gervain, 2015). Infants as young as 5 weeks have demonstrated the ability to discriminate phonemic contrasts used in their native, but also non-native languages, which adults are unable to perform (Werker, 1995). Newborns can also discriminate between rhythmically different languages as well as acoustic-signaling word and phrase boundaries (see Gervain 2015 for a review).

Although newborns come into the world with broad-based perceptual abilities, they do still show a preference for their mother's voice (DeCasper & Fifer, 1980) as well as for the language heard in utero (Moon & Cooper, 1993, as cited in Gervain, 2015). As hearing is functional between the 24th - 26th week of gestation (Eggermont & Moore, 2012, as cited in Gervain, 2015), although some researchers suggest that it is much earlier -at the 20th week of gestation- (Pujol et al., 1991, as cited in Gervain, 2018) infants are already receiving input from the auditory environment. This prenatal exposure to the native language can help to set the stage for further language development. However, auditory information gathered in utero does not encompass all aspects of language. The uterine lining and amniotic fluid act as a low-pass filter (Gervain, 2015), preserving the prosody of the language while the finer details and individual phonemes are lost (Gervain, 2015, 2018). Prosody provides important information of vocalic segmentation to the infant in utero, as it carries

most of the pitch, intensity, and durational information (Moon, Lagercantz, & Juhl, 2013, as cited in Gervain, 2015), and aids in the ability of the infant to break up a continuous speech signal into smaller units (Gervain, 2018). This early experience with prosody prepares newborns to identify and produce their native language in the acoustic environment, whereby before they know anything about the meanings of the sounds they hear, they are paying attention to the acoustic properties of them (Hoff, 2006).

Prelinguistic Production: The Cry, Coo, and Vocal Play Stages

Typical vocal development begins in the infant's first year of life. The first year is considered the prelinguistic or preverbal stage, as while the infant does not speak per se, the foundations of language acquisition for production are constructed (Gillis et al., 2016). Each stage in vocal development is

marked by the appearance of new vocal boundaries, although the previous vocal boundaries may still be present (Kent & Murray, 1982; Stoel-Gammon et al., 2024). Adapted from Oller, 2000, and Vihman, 2014, Stoel-Gammon and colleagues (2024) laid out six stages to vocal development (see also Gillis et al., 2016; Stark, 1980).

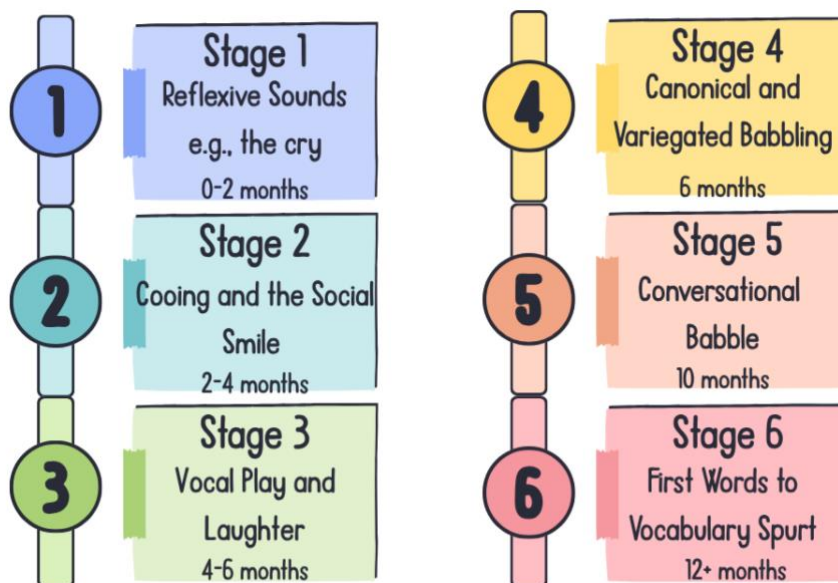


Figure 4: Stages of typical vocal development.

Stage One: How Reflexive Crying Paves Way for Language Production. Stage one begins at birth until the infant is about two months of age. This stage is characterized by reflexive vocalizations that may imitate vowel-like sounds, with the main form of reflexive vocalization being used by infants at this age is the cry (Stark, 1980). The cry is used to introduce pharyngeal and laryngeal movements to the vocal system (Gillis et al., 2016). From birth, the infant cry melody

demonstrates the features present in the native language, including its intensity, pitch, rhythmic patterning, and phrasing (Shinya et al., 2017; Stark, 1980). Newborns and young infants initially cry without communicative intent, to regulate their respiratory system (Gillis et al., 2016). From the caregiver responding to the infant's cry, it develops into a communicative tool for the infant (Gillis et al., 2016). At about six weeks of age, the cry is established as a sign of genuine discomfort for the infant (Gillis et al., 2016). The melody of the cry also changes with development, becoming increasingly complex as a reflection of the increasing voluntary vocal control from the cortex, which is responsible for controlling non-cry vocalizations related to language development (Wermke et al., 2007). By two months of age, the complexity of the infant cry develops from a single-arc melody structure to a multiple-arc melody structure, which is also linked to laryngeal control (Wermke et al., 2007). Shinya and colleagues (2017) investigated 77 preterm and 30 term infants on the variable and complex cry melodies at term with language outcomes at 18 months, in Kyoto, Japan. They discovered a significant positive correlation with transposition range variation, or pitch variation, across a 60-second crying episode and expressive language outcomes and overall cognitive outcomes at 18 months (as measured by an adapted version of the MCDI for Japanese infants and the Kyoto Scale of Psychological Development). These associations may reflect the integrity of a developing central nervous system in a similar manner to that of spontaneous physical movements. Furthermore, Wermke and colleagues (2007) recorded the cries of 34 infants over the first two months of life and compared them to language outcomes at 2.5 years. Infants who used complex melodies in more than 45% of their cries were more likely to undergo regular language development, whereas infants below that threshold had a 4.81 times higher risk of developing a language delay. The first two months of life see maturational changes within the infant's physiological development of their vocal system and their cognitive control over vocalizations that facilitate later non-cry vocalizations (Lieberman, 1984, as cited in Wermke et al., 2007).

Stage Two: Beginnings of Breath and Melody Control for Language Production. Stage two begins in the second month to the fourth month of life and is characterized by cooing and laughter sounds (Stark, 1980; Stoel-Gammon et al., 2024). There may be some continuity between the pitch of an infant's cry and their later coo and babbling vocalizations, as the same aspects of laryngeal regulation are used in all three processes (Stoel-Gammon et al., 2024). In this stage, phonation is more controlled, and the vocalizations are made during a single expiration of air

(Gillis et al., 2016). The development of vocalizations is continuous, but not necessarily linear (Hsu et al., 2000). These sounds may take the form of vocalic sounds (an utterance containing only vowel-sounds), which are more nasal-like in quality and have a shorter duration (Bloom et al., 1987; Hsu et al., 2000; Masataka & Bloom, 1994), or syllabic sounds (an utterance containing a consonant-vowel combination sound), which have more speech-like quality and the same duration as adult speech syllables (Kent & Murray, 1982). Syllabic sounds and vocalizations with more complex melodic contours were found to be longer than other vocalizations at three, four, and six months of age (Hsu et al., 2000). The complexity or type of melodic contour may vary based on the presence of a social partner. Infants at the age of three months demonstrated a preference for a falling intonation when alone, and a rising-falling contour in the presence of their mother or another social partner (Kent & Murray, 1982). As infants age, they come to produce more speech-like vocalizations, also known as protophones. Nathani and colleagues (2006) found that about 65% of an infant's total utterances between zero and two months were speech-like. This number increased to 94% of their total utterances by the time the infants reached between 16 and 20 months of age (Nathani et al., 2006). Non-speech sounds, such as cries and vegetative sounds demonstrated the opposite pattern, and decreased significantly with age (Nathani et al., 2006). Infants, as young as three months, who produce more speech-like sounds were found to be more favourable and perceived as more attractive, friendly, and sociable by adult social partners (Bloom et al., 1993), and received higher levels of responsiveness from their mothers (Masataka & Bloom, 1994). One explanation for this may be that speech-like sounds produced by infants allow mothers, caregivers, or other social partners to maintain conversation with the infants (Masataka & Bloom, 1994) which also served the evolutionary purpose of maintaining vocal contact with the infant when it was put down to allow the mother to forage (Masataka, 2009; see also Poliva, 2017 for an in-depth review).

In the fourth month of life there is a regressive transition, where the number of infant vocalizations produced in the context of socialization that had previously accelerated, drops substantially (Hsu et al., 2000). This may be a result of the reduction of simple vocalic sounds and melodic contours and an increase in more complex melodic contours (Hsu et al., 2000) that signifies the third stage of vocal development. Significant anatomical and physiological changes to the infant's vocal tract also begin in the fourth month, such as the separation between the oral and nasal cavities, as well as between the velum and epiglottis, (Sasaki et al., 1977) that allow the infant to begin producing non-nasal sounds (Kent, 1981, as cited in Hsu et al., 2000). Sound, instead, is produced when an

infant moves their articulators, for example, the tongue or lips, with an opened oral cavity (Gillis et al., 2016).

Using the Social Smile for Communication. Most infants in stage two will also begin to exhibit a true social smile, often elicited by social contact (Gillis et al., 2016). Smiles before typically occurred in the absence of known stimulation, and, most commonly, during sleep (Sroufe & Waters, 1976), or as a response to satiation following feeding (Wolf, 1963, as cited in Sroufe & Waters, 1976). According to Wolf (1963, as cited in Sroufe & Waters, 1976) the first alert smiles occur around three weeks old, whereby an infant will smile while awake and with focused attention, often at the caregiver's voice. This smile is elicited between four and five seconds after the stimulus, or voice, is presented to the infant. The progression of smiling continues into the fourth week, where vigorous stimulation, such as playing pat-a-cake, and the fifth week, where some visual stimulation, such as the nodding of a head, are also able to consistently elicit smiles. Between the eighth and tenth week a stationary face begins to consistently elicit a smile, reflecting a change in gaze patterns to include focus on the mouth (Messinger & Fogel, 2007), the first formation of a true visual schema (Sroufe & Waters, 1976). These smiles often reflect the involvement of cognitive developmental processes (Messinger & Fogel, 2007; Sroufe & Waters, 1976). The patterns described in a comprehensive overview of social smiling by Messinger and Fogel (2007) describe the smile beginning with a 3- to 20-second period of brow knitting and visual fixation on the mother's face, after which the brows relax indicating a cognitive recognition, and a smile appears. The social smile, along with infant cooing expressions, becomes sequentially and bidirectionally linked with maternal speech and smiling in the following months of development, suggesting a positive emotional attractor in the social communication system (Messinger & Fogel, 2007).

Stage Three: Vocal Play. The third stage of vocal development, then, is characterized by the production of some clear syllables and the beginning of vocal play, between four and six months (Stoel-Gammon et al., 2024). In vocal play, the infant plays with the various sounds they are able to produce, testing their vocal apparatus with different types of sounds, shapes, and intensities (Lee et al., 2018; Stark, 1980; Stoel-Gammon et al., 2024). In preparation for the babbling stage, they are able to produce a variety of vowel- and consonant-like sounds, exploring their articulatory organs in a sensory way (Gillis et al., 2016). This may include low and soft

sounds, such as yells and whispers, low and high sounds, such as squeals and growls, raspberries, tongue clicks, sustained vowels, and rudimentary consonant-plus-vowel syllables (Stark, 1980; Stoel-Gammon et al., 2024). The infant receives auditory feedback from their vocalizations and deliberately tries to reproduce the sounds (Gillis et al., 2016). Their articulatory movements gradually move from the back of the mouth to a more frontal position, such as the lips (Gillis et al., 2016), including early frontal consonants such as /m/, /b/, and /d/ (Lee et al., 2009, as cited in Stoel-Gammon et al., 2024). Velar activity, which is the positioning of the back of the tongue against the velum to produce /g/ and /k/ sounds, also increases (Gillis et al., 2016). Other commonly produced sounds are faint, low-pitched, grunt-like sounds with muffled resonance (Nathani et al., 2006). Some infants will also produce sounds not present in their maternal language as they attempt unknown and complex sounds throughout this stage (Gillis et al., 2016).

Laughter to Promote Adult-Infant Interactions. Stage three is also characterized by the reliable appearance of laughter (Stark, 1980; Stoel-Gammon et al., 2024), a highly stereotyped pattern of responding, (Ambrose, 1963, as cited in Sroufe & Wunsch, 1972), which increases over the developmental period (Sroufe & Wunsch, 1972). The first segments of laughter resemble certain portions of a cry episode, and are characterized by a rapid alternation of voicing (vibration of vocal cords) with voicelessness (air passes through with no vibration of the vocal cords) (Stark, 1980). This is followed by a voiced inspiratory segment that is also found in cry, and then a typical cooing sound (Stark, 1980). Laughter requires a large, rapid build up of tension, and usually occurs immediately or shortly after the stimulus is presented (Sroufe & Waters, 1976). The onset of laughter can be attributed to the changes in cognitive processes that infants experience, as laughter is produced mainly through interactions with adults and older children, although changes throughout the course of its and the infant's development (Sroufe & Wunsch, 1972; Stark, 1980). There is a progression from a passive to a more active role for the cognitively maturing infant to produce laughter, aiding in building the required tension for laughter to be elicited (Sroufe & Waters, 1976). For example, in its initial stages, between four and six months old, laughter is elicited mainly through physically vigorous stimulation (Sroufe & Waters, 1976). Between seven and nine months, tactile, social, and visual stimuli can produce laughter (Sroufe & Wunsch, 1972). Around ten to twelve months infants demonstrate an increase in laughter to visual stimuli, but a decrease towards tactile stimuli (Sroufe & Wunsch, 1972). At this age, infants will also begin to

laugh in anticipation of the stimuli and reach or lean forward to engage with the stimuli (Sroufe & Wunsch, 1972).

The Beginning of Language Comprehension. Prior to the onset of stage four, some infants may have already demonstrated comprehension of first, highly repetitive, words (Hoff, 2006), such as their name (Mandel, Jusczyk, & Pisoni, 1995, as cited in Hoff, 2006), and "mommy" and/or "daddy" (Tincoff & Jusczyk, 1999, as cited in Johnson, 2016). In order for infants to learn a word, it must be isolated, or segmented, from the stream of speech, and the child must develop a syntactic understanding, or an understanding of what that word means (Hoff, 2006). The characteristics of infant-directed speech, such as exaggerated stress patterns and high amounts of repetition, may assist infants in segmenting words from the streams of audio input (Hoff, 2006; Hoff-Ginsberg & Shatz, 1982).

Babbling

Stage Four: Babbling to Explore Sounds in the Native Language. Stage four begins around six months, and is characterized by the introduction of canonical babbling (Stark, 1980; Stoel-Gammon et al., 2024). As the majority of the world's languages are composed of canonical syllables, the ability to produce them must develop before the production of any substantial vocabulary (Lee et al., 2018). Infants are considered to have entered this stage when the ratio of canonical syllables to total syllables is at least equal to 0.15 (Lynch et al., 1995, as cited in Lee et al., 2018). Canonical babbling is the production of well-formed syllables containing at least one consonant and one vowel (Lee et al., 2018) which may also take on adult-like timing as well as the pitch and contour of the native language (De Boysson-Bardies & Vihman, 1991; Stoel-Gammon et al., 2024). Canonical babbling can take the form of reduplicated babbling, which are utterances that contain multiple repetitions of the consonant-vowel syllables, or variegated babbling, which is a vocalization that contains two or more different consonants (Hoff; 2014; Stark 1980; Stoel-Gammon et al., 2024). To investigate exactly when infants are more likely to produce canonical babbling and how it changes with age, Lee and colleagues (2018) used the Language ENvironment Analysis (LENA) system to record infants at the age of 6 and 11 months in home environments for up to 12-hour periods. They used a canonical babble ratio, which compares canonical vocalizations to total vocalizations. It was found that not only did American infants increase their canonical babbling ratio with age, but they exhibited a higher ratio of canonical babbling during

periods of adults speaking amongst one another with the infant in the room (adult-directed speech) as compared to adults speaking to the infant directly (infant-directed speech). This home-based study built upon many laboratory-based studies that had shown a higher number of overall vocalizations produced during infant-directed speech over adult-directed speech conditions (Hsu & Fogel, 2001), although some studies have also shown the opposite effect (Lee et al., 2018). Canonical babbling is an important milestone in overall language development, as infants build on the babbles produced for conversation babble and their first words. In fact, infants who did not produce canonical syllables by 10 to 12 months demonstrated a reduced expressive vocabulary at 18, 24, and 30 months of age (Oller et al., 1999, as cited in Lang et al., 2019).

The Growing Language Comprehension of Infants. Around eight months of age, often after the onset of canonical babbling, infants begin to understand short phrases (Fenson et al., 1994, as cited in Hoff, 2006). Their comprehension continues to grow, and by 11 months old, it ranges from 11 to 154 words (Hoff, 2006). The variability in the amount of language that children hear can effect their vocabulary development (Hoff, 2006). Exposure to higher levels of overheard speech (Hart et al., 1997) and to richer vocabularies using rare or sophisticated words (Hoff & Naigles, 2002, as cited in Hoff, 2006) are correlated with a faster rate of vocabulary development.

Stage Five: The Introduction of the Conversational Babble. In stage five, the infant begins to produce jargon, or conversational babble, around 10 months of age (Stark, 1980; Stoel-Gammon et al., 2024), with canonical syllables being the primary form of vocal production (Nathani et al., 2006). An abrupt increase in the production of canonical syllables is a key indicator in the advancement of prelinguistic vocal development (Stark, 1980). Vocalizations are often delivered with eye contact and gestures (such as reaching, pointing, or pushing away), and they may even imitate the outward form of adult conversations (Stark, 1980; Stoel-Gammon et al., 2024). During this stage, the strings of sounds and syllables are uttered with a rich variety of stress and intonational patterns (Stoel-Gammon et al., 2024). Infants in this stage often show a preference for 'stops' or /b/, /d/, and /g/ sounds over other forms of vocalizations (Vihman et al., 1986). Adults often engage in 'conversations' with infants in this stage using real words and phrases and some nonsense syllables (Stoel-Gammon et al., 2024). According to Vihman et al. (1986), when adults are present, infants show a preference for short vocalizations and monosyllables, perhaps to accommodate for the mother's presence and to use the adult-like words that they understand and

are able to produce. When infants are alone, however, they are more likely to engage in a greater proportion of long nonword vocalizations, perhaps in vocal play, practice, or exploration of their own capacities (Vihman et al., 1986).

Nonverbal Communication to Promote Language Comprehension and Production

Nonverbal communication refers to behaviours that are not verbal but still enable the child to understand and to be understood by the members of the child's social community (Mayo & Le France, 1978). Development of nonverbal communicative behaviours is shown by an increase in the infant's nonverbal repertoire, greater integration and coordination of verbal and nonverbal communication channels, and greater differentiation of nonverbal behaviour according to social context (Mayo & La France, 1978). Some infants have demonstrated vocal sensitivity and are able to differentially respond to changes in the tone of voice from as young as five months (Buhler & Hetzer, 1935, as cited in Masataka, 2009). Part of communication involves attending to others, responding to their social bids, and knowing how one's response may affect others (Mayo & Le France, 1978). Gaze is instrumental in these processes (Çetinçelik et al., 2021; Mayo & La France, 1978). During interactions, gaze can indicate social engagement, reflect the desire to communicate, reveal the speakers' or listeners' goals and feelings, and direct the attention of the listener to objects or parts of the environment (Kleinke, 1986, as cited in Çetinçelik et al., 2021).

From birth, newborns demonstrate a preference for open over closed eyes (Batki et al., 2000, as cited in Çetinçelik et al., 2021), and upright over inverted faces (Farroni et al., 2002, 2004). At birth, a newborn's eye has well-developed rods, but widely spaced and poorly developed cones. This means that they have low visual acuity in the fovea, and are sensitive only to images with high contrasts (Salomao & Ventura, as cited in Farroni et al., 2002). Over the first four months of life, their visual acuity develops, until about four months when they can see enough details in the face to be able to distinguish between happy and sad (Salomao & Ventura, as cited in Farroni et al., 2002), as well as between a direct or obviously averted gaze (Vecera & Johnson, 1995). Joint attention begins to emerge between six and eight months (Butterworth & Jarrett, 1991, as cited in Tenenbaum et al., 2015), alongside the onset of canonical babbling. Joint attention is the caregiver-infant dyad's shared attention to an object or event while maintaining their attention to each other (Çetinçelik et al., 2021). One form of joint attention, known as gaze following, requires the infant to first establish mutual eye contact with the adult, and then coordinate their attention with the

adult's gaze to the object(s) that the adult is referring to (Çetinçelik et al., 2021; Tenenbaum et al., 2015). According to a systematic review by Çetinçelik et al. (2021), gaze following enables the infant to engage in these joint attentional periods with a social partner and selectively attend to a single source (e.g., an object). The infant can then direct other attentional resources to the properties of the object, to better understand what the social partner is referring to. These processes may promote language development through word-object mapping, object processing, and speech processing. Infants who do not learn to reliably orient to the objects in the speaker's direction of gaze may fail to develop vocabulary at the same rate as children who respond to these visual cues more successfully (see also Tenenbaum et al., 2015). For example, differences in gaze following abilities at 6 months old contributed to differences seen in infant receptive (Morales et al., 2000, as cited in Çetinçelik et al., 2021) and expressive (Morales et al. 1998) vocabularies when measured at 12 months, and subsequent expressive vocabularies when measured at 18, 21, and 24 months (Morales et al., 1998). Gaze following abilities between 10 and 11 months of age also predicted receptive vocabulary scores at 14 months (Brooks & Meltzoff, 2005) while the ability to initiate these triadic interactions at 14 months was positively correlated with expressive and receptive language scores at 30 months (De Schuymer et al., 2011).

Infants may also use other forms of gestures, such as holding out, giving objects, reaching, and pointing from as young as nine months old to communicate prior to their first words (Cameron-Faulkner et al., 2021; Iverson & Goldin-Meadow, 2005). Some argue that gesture and speech form a single integrated system, with gesturing exploiting different representational resources than does speech (McNeil, 1992, as cited in Iverson & Goldin-Meadow, 2005) and facilitating language learning by providing a mode of communication when the infant is unable to use verbal expression (Iverson & Goldin-Meadow, 2005). Cameron-Faulkner and colleagues (2021) found a positive relationship between infant gesturing and contingent speech between 10 and 12 months and vocabulary production at 18 months. In a longitudinal study by Iverson and Goldin-Meadow (2005), infants from 10 months of age initially relied exclusively on using gestures to communicate. About three months before an infant was able to produce a word to represent an object of their desire, they increasingly used gestures to indicate the object. Before producing two-word utterances, infants had integrated word and gesture combinations to communicate. As their cognitive capacities for language increased, gestures became a less important modality for communication.

Protowords and First Words

The quality and complexity of an infant's babbling, such as the frequency of canonical syllables, correlates with early language development, including an earlier onset of word production, a larger productive vocabulary, and more accurate word productions (De Boysson-Bardies & Vihman, 1991). Throughout vocal development, infant babbling, which has a limited range of sounds, begins to be brought under more vocal control and take on the appearance of the native language (MacNeilage et al., 1997; Majorano & D'Odorico, 2011; Stoel-Gammon et al., 2024; Vihman et al., 1986). Prior to producing their first words, infants produce articulated word-like structures, known as protowords (Kent & Bauer, 1985, as cited in Laasko et al., 2010). Protowords consist of one articulatory movement, such as the closure of the airway with the tongue, during phonation, or the production of speech sounds (Menn, 1983, as cited in Laasko, 2010). Protowords are often syllable types that have been repeated many times over by the infant during the babbling stage to express the infant's attitudes and wishes. (Stark, 1980) These words do not have any linguistic meaning (Laasko, et al., 2010) but may become linked with specific objects (Nazzi & Bertoncini, 2003) and are often vocalized in conjunction with nonverbal behaviour, such as eye contact, facial expressions, or gesturing (Laasko et al., 2010; Stark, 1980). As the caregiver continues to respond to the child's protowords using the linguistic names, and, oftentimes, also incorporating gestures, between 7 and 12 months, the infant begins to integrate the correct linguistic names with the objects and their protoword production becomes increasingly similar to the referent word (Laasko et al., 2010; Levey, 2024).

Stage Six: Linking Sounds to Concepts. The first word is typically uttered shortly before an infant's first birthday (Johnson, 2016) and subsequent first words are acquired slowly over the course of the following several months (Goldfield & Reznick, 1990; Mervis & Bertrand, 1995), indicating the onset of the sixth and final stage of vocal development. The first word can be defined as a phonetically specified pattern of sound that is linked to a specific concept (Nazzi & Bertoncini, 2003). These sound patterns may be accurate, resemble the target word, or not resemble the target word at all, and are influenced by the words they hear or prefer to produce (MacNeilage et al., 1997; Majorano & D'Odorico, 2011; Stoel-Gammon et al., 2024; Vihman et al., 1986). According to previous literature from MacNeilage and colleagues (1997), some of the patterns present in babbling are also found in a slightly more reduced form in the first words of the same infants

(Davis & MacNeilage, 1995, as cited in MacNeilage et al., 1997). These specific patterns are 1) consonants that are constricted in the front region of the mouth, such as alveolar stops (e.g., /t/, /d/, and /s/) and nasals (e.g., /n/ and /m/), which are accompanied by frontal vowels (e.g., /i/ and /æ/), 2) consonants that are constricted in the back of the mouth, such as velar stops (e.g., /k/ and /g/), which are accompanied by back vowels (e.g., /u/, /oʊ/, and /ɑ/), and 3) labial consonants (e.g., /p/, /b/, and /m/) which are accompanied by central vowels (e.g., /ə/, /ʌ/, and /ɜ/). There is also evidence that the type of pattern that is favoured by infants is dependent on the native language. For example, while De Boysson-Bardies and Vihman (1991) found evidence of a preference for labials and dentals (e.g., /θ/ and /ð/) in English, French, Japanese, and Swedish infants, Majorano and D'Odorico (2011) found that Italian infants also showed a preference for labial and alveolar sounds. One explanation offered may be found for infant preference, or for the words that are targeted at the infant. However, MacNeilage and colleagues (1997) also found that, as compared to babbling, when infants produced their first words, they increased the proportion of labial sounds, decreased the reduplicated disyllabic pairs (attributed to an increase in vowel variegation), and decreased the amount of open vowels produced, in favour of high frontal and back vowels. One explanation for this may be a result of the tendency to have the last two vowels of the word differ, producing a non-reduplicative sequence. In terms of the actual words that an infant may produce as their first word or as part of their first words, a cross-cultural comparison between, English-, Potunghua-, and Cantonese-speaking infants by Tardif and colleagues (2008) found that cross-culturally, the common first words are a version of "mommy", "daddy", "bye-bye", "hi", "uh oh" and "woof". There was also evidence to show that cross-culturally, infants are more likely to produce terms for people, and that English-speaking, but not Pontunghua- or Cantonese-speaking infants, were also highly likely to produce terms for animals and objects. These early words may function initially as verbal components of specific situations, routines, or actions, but will eventually become a term that refer to objects or events more generally (Piaget, 1951, as cited in Goldfield & Reznick, 1990).

The Vocabulary Spurt

At some point in their lexical development, an infant's productive vocabulary begins to accumulate at a rapid pace within a short period of time (Goldfield & Reznick, 1990). This period has been coined the "vocabulary spurt" and is often defined as a rapid increase in the rate of word learning (Goldfield & Reznick, 1990), although some researchers have defined it as learning a pre-

determined number of words over a given period of time (e.g., Gopnik & Meltzoff, 1987; Lifter & Bloom, 1989), or reaching a specified vocabulary size (Mervis & Bertrand, 1995). Nonetheless, it appears that most typically developing children experience a vocabulary spurt at some point in their development. Oftentimes the types of words that children add to their lexicon over this period of rapid growth are nouns, which can encompass over 3/4 of their total productive lexicon post-spurt (Goldfield & Reznick, 1990; Mervis & Bertrand, 1995) and indicate significant cognitive changes that occur during this time (Goldfield & Reznick, 1990; Mervis & Bertrand, 1995; Nazzi & Bertoncini, 2003). Some researchers have found that the rate of vocabulary development varies from between 8 and 11 new words per month, to a rate of 22 to 37 words per month (Goldfield & Reznick, 1990; Hoff, 2014). Nazzi and Bertoncini (2003) proposed two possible interconnected mechanisms for this period of vocabulary development. The first links early words, or protowords, to concepts through repeated repetitions, with a second, referential mechanism, that immediately pairs a phonetically specified sound pattern to an unnamed object category. This would allow the infant to represent their world using language, and the referential mechanism would allow for the increase in the rate of learning new words, as the word-category pairing would only need to be presented once. The infant can then spend more time interpreting phonological information from the input. Evidence from these mechanisms comes from the infant's ability to demonstrate exhaustive categorization, which is the understanding that every object has a category and fits into only one category, which appears between 16 and 20 months. Infants are unable to demonstrate this ability prior to, but are able to demonstrate this ability following, the vocabulary spurt, regardless of whether this spurt was documented as earlier or later in their development (Mervis & Bertrand, 1995).

Some researchers have argued against the notion of a vocabulary spurt, as although infants show an increase in the rate of vocabulary learning, there is no true inflection point, or point where their rate of increase is greater than it is before, and sustained for a period of time (Ganger & Brent, 2004). Instead, infants have demonstrated a linear, or regular, increase of word acquisition (Goldfield & Reznick, 1990), or have undergone multiple vocabulary spurts over the course of their early development (Dandurand & Shultz, 2011). One explanation for the lack of a spurt being detected is that the acceleration of learning is an unavoidable by-product of variation in the difficulty of word learning, and should not be misconstrued as evidence for specialized learning (McMurray, 2007).

Summary

To summarize, infant language development begins before birth (DeCasper & Fifer, 1980), with evidence from newborn infants that show a preference for their mother's voice (DeCasper & Fifer, 1980) and for the language heard in utero (Moon & Cooper, 1993, as cited in Gervain, 2015). After birth, language development continues and progresses in stages (Gillis et al., 2016; Stark, 1980). It is shaped by the acoustic environment of the infant, from the caregiver responding to an infant crying (Gillis et al., 2016), to the caregiver engaging in 'conversations' with the infant (Stoel-Gammon et al., 2024). This environment is critical for any developing infant, and can positively or negatively impact the infant's language development. The following chapter will describe in-detail the role of the adult, and some sources of potential disruption during the first two years of infant language development.

Chapter 2: The Role of the Adult and Sources of Interruption in Infant Language Development

The Role of The Adult

Quantity and Quality of Linguistic Input

An infant's vocabulary growth is impacted by the language they are exposed to (Hart et al., 1997; Hoff, 2006; Laasko et al., 1999; Rowe, 2012; Tamis-LeMonda et al., 1996; Weisleder & Fernald, 2013). Both the quantity (Anderson et al., 2021; D'Odorico & Jacob, 2006; Hart et al., 1997; Huttenlocher et al., 1991; Rowe, 2012; Weisleder & Fernald, 2013) and quality (Anderson et al., 2021; D'Odorico & Jacob, 2006; Hart et al., 1997; Rowe, 2012, 2018) of language exposure can explain some differences seen in infant's vocabulary growth. For instance, in a 90-minute free-play session in a laboratory setting, the number of words that parents uttered to their 18-month-old infant, otherwise known as infant-directed speech, ranged from 360 to over 9200, and had a significant positive relationship with infant's receptive vocabulary at 30, 42, and 54 months (Rowe, 2012). Huttenlocher and colleagues (1991) followed infants from 14 to 26 months and measured the amount of infant-directed speech they received over the course of their development. Infants who received a higher amount of infant-directed speech demonstrated faster vocabulary growth, accounting for 29% of the variance in their final vocabularies. Weisleder and Fernald (2013) investigated both infant-directed speech and adult-directed speech on receptive and productive vocabulary size at 19 and 24 months of age using an all-day audio recording. Parents produced a range of 670 words to 12 000 words directed to their infant. Infant-directed speech at 19 months predicted vocabulary size at 24 months, as well as processing efficiency at 19 months as measured using a looking-while-listening task. Adult-directed speech, or the speech that infants may only overhear in their environment, did not correlate with their vocabulary development, indicating that it may be more than just listening to adult speech that explains language development. Hart and colleagues (1997) wanted to investigate if differences in parent speech towards their 7- to 9-month-old infants was a factor in the infants' productive vocabularies at 3 years old, using socioeconomic status as a measure of how much parents may speak to their infants. Infants from higher income families received significantly more speech than infants in middle and lower income families over spontaneous verbal interactions, having 2152, 1251, and 616 words per hour directed at the infant and reading productive vocabularies of 1100, 750, and 500 words, respectively, by the age of three.

Although there were substantial differences in the amount of speech present in the environment, it was the language experience of the infants that was a better predictor for the infant's productive abilities. This included measures of vocabulary diversity, symbolic emphasis, responsiveness, and style of guidance, as these infants also received significantly more affirmative forms of speech, to encourage speech or expand the infant's vocabulary, as opposed to prohibitive speech, used to stop an infant from doing something.

In a meta-analysis performed by Anderson and colleagues (2021), there were significant moderate to large associations between speech quality (defined as vocabulary diversity and syntax complexity) and quantity (defined as the number of words, tokens, or utterances) of parental linguistic input and children's language development. The effect size for the quality of linguistic input was more statistically robust than the quantity of input, and the effect of quality increased with age. In the aforementioned study by Rowe (2012) further analyses were completed on the quality of the language that parents uttered during the 90-minute free-play session. The quality of input that infants receive was defined as a higher variability in speech, a higher use of "rare" or uncommon vocabulary, and the use of decontextualized language. Infant vocabulary at 30 months was associated with the number of parental word tokens, different word types, and explanations used during play at 18 months. At 42 months, infant vocabulary was associated with the number of parental word tokens, word types, and rare word types used during play at 30 months. And at 54 months, infant vocabulary was associated with the number of parental word tokens, word types, rare word types, narrative utterances, and explanations used during play at 42 months. It seems that specific qualitative features of infant-directed speech are more beneficial during a given age range, as compared to others, for productive vocabulary development. The quantity of input seems to be most important from the second year of life, diversity or sophistication of vocabulary from the third year, and the use of decontextualized language (such as narrative language and explanations) from the fourth year (Rowe, 2012). It may be that using more challenging language as the infant develops may facilitate vocabulary development (Rowe, 2012). When investigating children who are considered to be "late-talkers", meaning these children have not reached a 50-word vocabulary and/or do not use word-combinations by the age of 2 years old, D'Odorico and Jacob (2006) noted significant differences in maternal speech to their infants. Mothers of late-talkers produced less words per minute, used a lower frequency of nouns, and used a lower frequency of different types of nouns in a 30-minute free-play session at 18-months old. Not only

did these mothers speak less and use less sophisticated language, but the utterances produced by these mothers were flatter and had less varied contours, characteristics atypical of speech directed to young infants, compared to the speech from mothers of typically developing infants.

Motherese/Parantese

Speech directed towards young infants exhibits a different speech pattern than speech directed towards an older child or adult (Cutler, 1994; Hoff, 2006). This type of speech is often coined “motherese” or “parantese”, is highly exaggerated and redundant, and often focused on the here and now (Hoff, 2006). Infant-directed speech often has a higher pitch (Cooper & Aslin, 1994; Green et al., 2010; Narayan & McDermott, 2016; Smith & Trainor, 2008), slower rate of speech or articulation (Cox et al., 2022; Green et al., 2010; Narayan & McDermott, 2016), shorter utterances (Martin et al., 2016), longer vowels (Albin & Echols, 1996; Cox et al., 2022; Green et al., 2010), as well as exaggerated prosody (Thiessen et al., 2005), stress on the final syllable (Albin & Echols, 1996; Wang et al., 2015), and lip movements (Green et al., 2010; for a review see Cox et al., 2022). These exaggerated speech patterns may aid infants with figuring out the structural components of the language, or separate the sequences of sound, into their respective lexical units (Plunkett, 1993), as word boundaries are not consistently marked by pauses (Thiessen et al., 2005). Infants have also been shown to prefer (Cooper & Aslin, 1994), or hold their attention longer to (Thiessen et al., 2005), infant-directed speech over adult-directed speech. Holding the infants attention may promote a higher level of learning in infants, as it increases their attention, and may also direct their interest towards those specific aspects of the input to better understand the speech stream (Thiessen et al., 2005).

Infants as young as one month are able to discriminate between infant-directed and adult-directed speech, even when the speech is filtered to remove the higher frequencies of infant-directed speech (Cooper & Aslin, 1994). When the filter was removed, the infants also showed a preference for infant-directed speech (Cooper & Aslin, 1994), indicating that infants are attuned to the significantly higher variation in and frequency of pitch that is characteristic in infant-directed speech. Caregivers are also sensitive to the feedback provided by infants regarding their pitch (Smith & Trainor, 2008). Mothers of 4-month-old infants produced more infant-directed speech when they received positive reactions from their infants (Smith & Trainor, 2008). As a response to the positive infant feedback, the mothers raised their pitch even higher (Smith & Trainor, 2008).

The pitch in infant-directed speech has been shown to not only be higher, but have more variability during a naturalistic in-home observation session of mothers and their 4- to 16-month-old infants (Narayan & McDermott, 2016). The words themselves have a greater lengthening in infant-directed speech, as demonstrated during a free-play session and labelling game with 6- to 9-month-old infants (Albin & Echols, 1996). Albin and Echols (1996) found that single-syllabic words had a longer overall duration in infant-directed speech, while multi-syllabic words with a final stress showed a longer duration for the overall word and the final syllable, and multi-syllabic word with an unstressed final syllable showed a longer overall duration of the word. The final syllable also accounted for a greater proportion of the duration of the word, when compared with adult-directed speech. While pitch, length, and prosody all play an important role in infant directed speech, Green and colleagues (2010) demonstrated that the lip movements during storytelling and reading to infants between 9 and 10 months of age were different than when retelling the same stories to an adult. The overall working spaces of the lips, as well as the vertical aperture, were significantly larger in the infant-directed speech condition.

Contingency

Contingency is the extent to which a communication is produced when the intended recipient is fully orienting towards receiving and processing it (Topping et al., 2013). A caregiver who talks a lot when the infant is not attending to their speech, or an infant offers a bid for communication that is not received or attended to, is a caregiver who is not behaving contingently (Topping et al., 2013). Parents who offer contingent replies to their infants' early vocalizations and bids have infants who progress faster along in their language development (Hsu & Fogel, 2001; Laasko et al., 1999; Tamis-LeMonda et al., 1996; see also Tamis-LeMonda et al., 2014). Infants as young as four weeks old can benefit from contingent communication (Hsu & Fogel, 2001). Hsu and Fogel (2001) compared the rate of infant non-distress, speech-like syllabic, and nonspeech-like vocalic vocalizations from infants between four and nine weeks old during free-play sessions in the laboratory. During symmetric communication, which is a mutual coordination of novel actions in an ongoing interaction, infants produced more speech-like syllabic and nonspeech-like vocalizations, than during asymmetric or unilateral communication. Furthermore, the rate of non-distress vocalizations made by the infant was positively correlated with the amount of symmetrical communication during the session. In another 15-minute laboratory-based free-play session with

older infants, at 13 months old, the infants had greater productive vocabularies at 13 and 20 months old, greater language comprehension at 13 months old, and achieved a productive vocabulary of 50 words and combinatorial speech sooner when mothers were coded to be more responsive to their infants (Tamis-LeMonda et al., 1996). In this study, the mother's responsiveness to their infant's language, such as the infant's exploratory vocal behaviour, was also correlated with the infant engaging in higher levels of symbolic play at 13 months old, highlighting the link between language and cognitive development.

Effects of the Family Context

Socioeconomic Status (SES)

Parental language input has been known to vary with socioeconomic differences (Hart et al., 1977; Hoff, 2003; Rowe, 2018). According to the American Psychological Association (2023), socioeconomic status (SES) is not only a measure of income, but also "educational attainment, occupational prestige, and subjective perceptions of social status and social class." Socioeconomic status is most commonly measured through familial income, parental educational attainment, and parental occupation (American Psychological Association, Task Force on Socioeconomic Status, 2007) although studies have yet to reach a consensus on the most effective measure of it (Bornstein & Bradley, 2003, as cited in Pace et al., 2017). Some variables adjacent to or commonly associated with SES can also have a direct impact on the infant and their subsequent language development, including food insecurity, abuse, neglect, limited educational resources, limited opportunities, low birth weight, and a higher risk of physical health problems (Poitras et al., 2017). Other variables may have an overall impact on the family unit and are also associated with variations in infant language development, such as parent-child interaction (Elardo et al., 1977) or parental education (Gilkerson et al., 2017), parental stress and depression levels (Der Nederlanden et al., 2025; Havron et al., 2022; Rochette & Bernier, 2014; Rowe, 2018), household chaos (Rochette & Bernier, 2014; Rowe, 2018), as well as neighbourhood safety (De Marco & Vernon-Feagans, 2013) and other contextual factors (De Marco & Vernon-Feagans, 2013; Lapointe et al., 2007). The mechanism by which socioeconomic status exerts influence over infant language development appears to be significantly mediated by the quantity and quality of infant-directed speech that the

child receives (Elardo et al., 1977; Evans et al., 1999; Hart et al., 1997; Rowe, 2018) and through the family's available resources (e.g., Lapointe et al., 2007).

On average, families from higher SES backgrounds talk more with their infants during the first two years than those from lower SES backgrounds (Hart et al., 1977; Rowe, 2018) with differences in input seen as young as 7 to 9 months old (Hart et al., 1977). These differences can lead to infants, between 10 months and 3 years of age (Hart et al., 1977), and between 2 months and 4 years of age (Gilkerson et al., 2017), from low-SES households hearing an estimated 30 million words less than children from high-SES households (Gilkerson et al., 2017; Hart et al., 1977). Meaningful differences arising from this have been seen as young as 18 months old. In a looking-while-listening task administered by Fernald and colleagues (2013), infants at 18 months old from lower SES households, measured using parental income and education, took longer than infants from higher SES households to direct their attention to the correct image, and looked at the correct image for less time. By the time children from lower SES households reached 24 months of age, their performance was similar to that of 18-month-old children from higher SES households, demonstrating a delay of almost six months in processing efficiency between the two groups. Differences in expressive vocabularies were also apparent as young as 18 months old, with a difference from around 75 words growing to almost 150 words by the time the children reached 24 months. Hart and colleagues (1997) reported that by the age of 3 years old, children from higher SES homes have produced twice as many words as children from lower SES households, also using familial income as the main form of SES measure, alongside the type of job the parents performed, classified as professional workers, working class, or on welfare assistance. The quality of interactions between caregivers and their infants have also been shown to vary with, and have differential effects depending on, levels of SES (Elardo et al., 1977; Rochette and Bernier, 2014; Tamis-LeMonda et al., 2014). Rochette and Bernier (2014) compared different aspects of maternal behaviour, including response to positive signals and distress, positive affect sharing, hostility and rejection, sensitivity and responsiveness, teaching orientation, and physical proximity, measured when the child was one year old, with the child's performance on executive functioning tasks at the age of three. While children from high-SES families, based on measures of maternal education and familial income, demonstrated no effects for maternal behaviour and executive functioning, children from low-SES families showed significantly higher performance on the tasks when maternal behaviour was of higher-quality. The differences in teaching approaches and responding

to infants that are observed may not vary directly with parental income and education levels, but with the specific knowledge that parents and caregivers have about infant and child development (Rochette & Bernier, 2014). Elardo and colleagues (1977) measured the HOME scores of even younger infants at six months old, and again at 24 months old, and compared them to the infants' psycholinguistic abilities at the age of 37 months. The Home Observation for Measurement of the Environment, or HOME, instrument measures the emotional and verbal responsibility of the mother, avoidance of restriction and punishment, organization of the environment, provision of appropriate play materials, maternal involvement with the child, and opportunities for variety in the daily routing. This measurement tool has demonstrated weak to moderate associations with variables that are also related to SES, such as welfare status, maternal education and occupation, paternal education, occupation and presence, and crowding in the home. The overall HOME scores at six months were correlated with auditory reception and association, visual reception and association, verbal expression, and grammatical closure. HOME scores demonstrated similar and even strong correlations at 24 months with auditory and visual reception, auditory and visual association, visual and grammatical closure, verbal expression, and manual expression, suggesting a cumulative effect on the home environment, parental responsiveness, and infant language development.

Acoustic Features of the Environment

Noise

Noise may be considered any unwanted or unattended to sound (Erickson & Newman, 2017). There are often multiple distinct noise types present in the child's environment, and noise in the background is a common feature of many settings experienced by children (Manlove et al., 2001, as cited in Erickson & Newman, 2017). According to a review by Erickson and Newman (2017), noise may "cover up" speech, as the target signals, or linguistic input, and background noise are funnelled down the ear canal and eardrum together, and excite the same auditory receptors and neural pathways. As a result, an incomplete representation of the speech sound pattern may be created, which may also impair the infant's ability to learn the language. Consistent interference from background noise can lead to significant impairments in the infant's ability to learn from speech signals, which may create long-lasting effects in their understanding of speech. Studies have shown that young infants struggle to learn new words when there is background noise

(McMillan & Saffran, 2016), as they require a higher signal-to-noise ratio than adults in order to successfully perceive speech (Trehub, Bull, & Schneider, 1981, as cited in Erickson & Newman, 2017). In non-ideal listening conditions, children may require higher levels of auditory and language input to develop robust representations of the words and sounds of the language (Newman, 2006, as cited in Erickson & Newman, 2017). McMillan and Saffran (2016) tested the ability for 22- to 24- and 28- to 30-month-old toddlers' abilities to learn novel word-object pairings when the speech is presented for the first time among background noise. Both younger and older toddlers were able to learn the novel word-object pairing with background speech present, when the background speech was, at most, 10dB quieter than the target speech. In another version of the same experiment, older toddlers were exposed to novel words within a stream of fluent speech with no background noise and no object-pairing. The toddlers were exposed to the novel words again, with only the word and image present for the pairing phase, among background noise that was only 5dB quieter than the target speech. The toddlers were able to successfully learn the word-object pairing in the noise level that had prevented learning in the previous experiments, suggesting that prior exposure to novel words may assist learning even among high levels of background noise. Newman & Hussain (2006) investigated infants preferences for infant-directed speech in quiet and noisy conditions, with infants at 4.5, 9, and 13 months of age. Infants in this study did show an overall preference for infant-directed speech, and speech under quiet conditions, but they did not show a preference for infant-directed speech in a quiet environment. Further analyses demonstrated that it was only the youngest infants, at 4.5 months old, that demonstrated these strong preferences, with the other two groups showing no significant preferences. Infant-directed speech and quiet settings may be more beneficial to younger infants, when they still have minimal exposure with words. However, the authors also pointed out that the control speech used as comparison in this study was recorded at a slower pace as compared to typical adult-directed speech, so older infants may not have shown a preference for either speech as they are still able to benefit from slower speech without exaggerated melodies. These studies combined suggest that younger infants may need more experience with speech at a slower speed with more exaggerated melodies to encourage language development, such as learning new words. Older infants, who may have previous experience with speech, may not require the same accommodations in speech to be able to successfully perceive speech and learn new words.

Importantly, other mechanisms, such as a decrease child vocalizations and parent speech, are also effected by noisy environments. While infants may be able to learn in these environments, they may not be given as much opportunity for their development when exposed to these conditions. Suarez-Rivera and colleagues (2024) investigated child vocalizations and parent speech in the presence of background noise when infants were 13, 18, and 23 months old using 2-hour long video recordings. The main sources of background noise in the home were found to be from traffic and construction, appliances, communication devices, television, and music, with music being the most prominent form of noise. The presence of noise ranged from only 7% to 97% of the total intervals recorded, and most infants experienced multiple sounds simultaneously. Both infants and mothers reduced vocalizations and speech, respectively, in the presence of all forms of background noise, with the exception of traffic and construction. Furthermore, when analyzed, the prevalence of different types of background noise varied with the family's socioeconomic status, whereby infants from higher-SES households, were exposed to music as the main source of background noise, and infants from lower-SES households were primarily exposed to television or screen media. What may be important to consider is the unique home environment of the infant, as there may be differential effects caused by SES, background noise, and the combination thereof.

Television and Screen Media

The use of television and screen media can be both harmful and beneficial to infants, depending on the context in which the viewing occurs (for a review see Guellai et al., 2022). While the use of television is fairly straightforward to investigate and define, the investigation into screen media may be a bit more challenging. Screen media is often referred to as a general category of overall device or technology use (Mannell et al., 2024). The specific types of screen media being measured has encompassed a combination of the television and DVDs, or primarily television, (Duch et al., 2013; Lin et al., 2015; Tomopoulos et al., 2010; Zimmerman et al., 2007a), the television, phone, tablet, or other screen (Krogh et al., 2021; Mendelsohn et al., 2008; Stockdale et al., 2020), primarily phone and small device use, (Beynon et al., 2024; Corkin et al., 2021; Garg, 2021); or a combination of all of the above (Asikainen et al., 2021; Durham et al., 2021). Although these categories are familiar and have some overlap between them, the types of use and the primary user may vary. For example, when studying primarily television, the primary user may be the infant, the adult, or both, depending on the type of show being played. It may also refer to television

that is being actively or passively attended to, or television being on in the background during other activities, which is often left undefined within the study. Studies which include the use of a cell phone or have a primary focus on cell phone and small device use, may capture a more adult-centred and active form of use. However, even very young infants now have access to cell phones and tablets (Mann et al., 2025). Questions surrounding the infants usage are often included, but then collapsed across a general category of "technology use" or "screen media use", although the impact on development may be different when the user is primarily an adult versus an infant. Including various forms of technology and use under the same category, such as those mentioned above, fails to differentiate entirely between the type of use (active or passive) and the primary user. Furthermore, while the primary form of measurement is often time exposed to screen media and technology (Mannell et al., 2024), some studies may consider context of use (Garg, 2021), such as interruptions or disruptions from technology (Corkin et al., 2021), but then fail to consider these factors separately within the analyses. Nonetheless, over the multitude of definitions and measurements of screen media and technology use, robust associations with various aspects of language and overall cognitive development have been observed.

Features of Technology Use and Screen Media Exposure

Infant Use

Studies have investigated the effects of infant exposure on language development, often defined as the amount of time infants spend watching television, videos, or DVDs (Chonchaiya & Pruksananonda, 2008, Lin et al., 2015; Tomopoulos et al., 2010; Zimmerman et al., 2007a), or using a mobile device or tablet (Alroqi et al., 2022; Supanitayanon et al., 2020). Although infants now have access to the use of handheld devices, through a cell phone or tablet, both inside and outside the home, (Mann et al., 2025) infants are primarily exposed to the television, followed by handheld devices as the second-most form of exposure (Durham et al., 2021; Supanitayanon et al., 2020). Parents have reported that screen use is the most prevalent activity that their infants partake in throughout the day, even when compared to playing indoors and/or outdoors and reading (Alroqi et al., 2022). Infants have demonstrated lower levels of overall language development (Alroqi et al., 2022; Chonchaiya & Pruksananonda, 2008; Lin et al., 2015; Sundqvist et al., 2023; Tomopoulos et al., 2010; Zimmerman et al., 2007a), as well as specific effects for language production (Alroqi et al., 2022) and comprehension (Tomopoulos et al., 2010) with high levels of

television exposure. Negative correlations between vocabulary development and screen time seen at the age of 5, can be traced back to exposure from 10 months, 2 years, and 5 years old (Sundqvist et al., 2023). Although it is also important to consider that the effects shown may be cumulative, as there are strong correlations between trends of television watching, whereby young infants exposed to high levels of television exposure are more likely to be exposed to high levels of television also at later ages (Krogh et al., 2021).

With television as the primary form of exposure, it is important to consider the specific aspects of television watching on infant language development. Lin and colleagues (2015) investigated these effects on Taiwanese infants under three years of age, using interviews with parents, the Bayley Scales of Infant Development (BSID-II), and the Peabody Developmental Motor Scales (PDMS-2). Compared with infants who had been exposed to low or moderate amounts of television, infants exposed to high amounts of television demonstrated delayed cognitive, language, and motor skills. For cognitive and motor delays, television, as well as lower maternal education and living in a double-income family, were associated with lower levels of cognitive and motor functioning. However, with regard to language delay, the only significant predictor was high television exposure, which had increased the risk of a delay by 3.3 times. Zimmerman and colleagues (2007a) collected data on infant television watching and reading habits using phone surveys, and correlated this with data on their linguistic development using the Communicative Development Inventory (CDI). For every hour that infants watching media, specifically, DVDs and videos that were aimed at infant use, between the ages of 8 and 20 months, it resulted in a 16.99 lower score on the CDI. Reading may help buffer some of the negative effects of television watching in infants. Reading once daily between 8 to 16 and 17 to 24 months, was associated with a 7.07 and 11.71 increase in scores on the CDI, respectively. Longitudinal investigations on even younger, American, infants, by Tomopoulos and colleagues (2010) measured the number of minutes of infant exposure to television, videos and DVDs, and movies, using a 24-hour recall diary, and their subsequent language development at 14 months, with the Preschool Language Scale-4 (PSL-4). A longer daily duration of television exposure at 6 months, predicted lower language development at 14 months, with associations for the expressive communication subscale, and even stronger associations with the auditory comprehension subscale. Thus, indicating that negative effects on language development can occur at very young ages, before they have even begun to babble (see Stark, 1980 for stages of infant vocal development). In a retrospective study on children already diagnosed

with language delays between 15 and 48 months old, by Chonchaiya & Pruksananonda (2008), parents reported that these infants began watching television at younger ages, and watched higher levels of television, than was reported by parents of typically developing infants. These parents also reported spending less time with their children, on average, than did the parents of typically developing infants. This helps to highlight the importance of delayed and minimal exposure to television, where possible, as well as finding other interactive activities for the infant-caregiver dyad to participate in. In-fact, the simple activity of co-viewing during exposure to television may help to mitigate the negative effects of television on development. Parents who reported to co-view screen media with their 20-month-old infants half the time or less, had infants with lower levels of vocabulary, as compared to parents who reported to co-view most or all of the time (Corkin et al., 2021). Longitudinal studies of Thai infants in the first two years of life, found that alongside a later onset of screen media use and less screen exposure overall, higher levels of verbal interaction between infants and caregivers during screen time was associated with higher language scores (Supanitayanon et al., 2020).

Background Television

Infants, especially extremely young infants, may not be actively using technology or screen media. However, television may be turned on in the background during various activities in the home. Background television can be defined as both television that is turned on in the background, and/or as pertains to the study of infants, television with adult-directed content (Anderson & Evans, 2001, as cited in Guellai et al., 2022).

The dangers of background television, such as disrupting play from as young as 6 to 26 months of age, have been studied immensely (Kirkorian et al., 2009; Schmidt et al., 2008; Setliff & Courage, 2011). Play is an important part of infant cognitive and language development (Barnett, 1990). According to Barnett (1990), play is known to aid in the development of problem-solving abilities (Barnett, 1976, as cited in Barnett, 1990), providing infants with a flexible approach to their environment and contributing to a generalized mode of cognitive approach. Play often precedes language development. It may help the child become more actively involved in what they are hearing (Piaget, 1926, as cited in Barnett, 1990) and may facilitate social interaction between the infant and their peers and/or caregivers. According to Valentine (1938), important developmental milestones also exist for play. For example, social play may appear as young as six months old,

playing with objects as early as one year, and make-believe play becoming prominent from the second year of life. Play being interrupted through the use of television and screen media may be one of the mechanisms by which screen media negatively effects language and cognitive development. Even toys that imitate forms of screen media have been shown to have negative effects, whereby infants produce less vocalizations, heard fewer words from parents, and experienced less conversational turns than infants who played with books and traditional wooden toys (Sosa, 2016).

Masur and colleagues (2015) obtained data from American infants using parent questionnaires about television exposure during play within the home from middle- and low-SES households, with infants from low-SES households separated into groups of younger (4 to 11 months old) and older (12 to 19 months old) infants. It was reported that background television accompanied mother-infant dyadic play half the time or more for 77% to 90% of infants from low-SES households, and for half the time or more for 57% of infants from middle-SES families. Television use during play was associated with more television use in the household, and the content was likely to be adult focused. Schmidt and colleagues (2008) found that infants at 12, 24, and 36 months of age who were exposed to adult-directed background television had less play overall, shorter play episodes, and shorter bouts of focused attention. In an experimental study, Setliff and Courage (2011) manipulated whether or not the television was on, as well as the type of program playing, during individual infant play. On average, play occupied the majority of the infants' time, although the range was relatively spread from 4 seconds of looking at toys to 552 seconds, out of a possible 600 seconds. Infants also engaged in longer episodes of focused attention when the television was turned off. Infants held their attention towards their toys for an average of 20.85 seconds with the television off, as opposed to an average of 13.91 seconds when the television was on. Masur and colleagues (2016) explored the speech behaviours of adults and 13- to 17-month-old infants during background television. It was found that greater background television during dyadic play at 13 months of age was negatively associated with their expressive, but not receptive, vocabularies at 17 months, as well as the number and variety of words that mothers produced at 17 months. Infants who are exposed to less speech, and produce less speech, demonstrate greater detriments in their later language development.

Adult Use

Adult use of different media may come in the form of directly using screen media, having the television or other media on in the background, or receiving distractible audible notifications. A literature review by Kildare & Middlemiss (2017) found that studies addressing the frequency of parental phone use revealed that 35% of adults report to use their phone while playing with their children (Qualcomm, 2013), spending between 0 to 5 hours on a device in the presence of children (Blackman, 2015), and engaging with their devices during mealtime (Radesky et al., 2014) or at the park (Hiniker et al., 2016).

Parent-infant interactions may be effected differentially through all forms of screen media, as there have been reports of less parent-infant interactions during adult- over child-directed programming (Mendelsohn et al., 2008). The brief interruptions to social interactions as a result of media use is known as technoferece (McDaniel, 2015). Stockdale and colleagues (2020) investigated this phenomenon in a laboratory-based interaction using a modified version of the still-face paradigm (for the original methodology of the still-face paradigm, see: Tronick et al., 1978). During the "still-face" portion of the interaction, parents of infants between 5 and 14 months were instructed to direct their attention towards their mobile devices and disengage from their infants. Infants responded with a decrease in their positive affect during this phase, with older infants displaying a more negative affect than younger infants, and an increase in escape behaviours. When considering parental beliefs about the appropriateness of technology use around infants, beliefs were associated with a higher number of escape attempts and self-comforting behaviours for younger infants, and a lower negative affect for all infants. Parental beliefs were found to be a stronger predictor of infant behaviours. The disruptive nature of technology may violate infant expectations leading to changes in infants behavioural organization in related to parents' mobile device use (Stockdale et al., 2020).

In the presence of background television, parents have been shown to talk less, talk more passively, and are less attentive than when the television is off (Kirkorian et al., 2009). A study on Finnish infants investigated the number of hours the child and their parents watched television and how long the television was kept on for, as well as shared reading daily, on infant language development (Asikainen et al., 2021). Using the Bayley Scales of Infant Development and parent reports of television use and infant vocabulary, increased screen time, was associated with a smaller vocabulary at 18 months (Asikainen et al., 2021). Parental screen time and shared reading less than

daily, was associated with a smaller vocabulary at 24 months. The ability to follow short and simple instructions was also at least somewhat deficient in 7.5% of children at 18 months and 3.4% of children at 24 months, which also related to the size of the child's expressive vocabulary and using one-word expressions. Not only does parental use have a negative effect on overall language production, but clear effects on language comprehension and cognitive development as well.

One aspect of cell phone use that may be overlooked, is the frequency of audible notifications that parents receive, while they are engaged with their infant and not their phone. Corkin and colleagues (2021) found that a higher frequency of audible notifications per hour was associated with higher levels of directiveness during play, and lower levels of scaffolding and responsiveness in a laboratory-based observation. Importantly, directiveness during play moderated the relationship between audible notifications and lower levels of infant vocabulary, whereby making the conscious decision to turn off or put away the mobile device was associated with a larger infant vocabulary (Corkin et al., 2021). Directiveness, such as giving commands or instructions, has been shown to negatively impact infant language development, specifically their use of auxiliary verbs between the ages of 15 and 27 months, such as *can*, *should*, and *would*, as well as noun inflections, such as adding "s" to use a plural (Newport et al., 1977).

It may also be important to consider other factors associated with high levels of parental screen media and technology use, such as socioeconomic factors (Krogh et al., 2021; Linder et al., 2021; Slobodin et al., 2024), parental health, (Beynon et al., 2024; Duch et al., 2013), and parental beliefs (Garg, 2021). For example, a study by Beynon and colleagues (2024) investigated maternal cell phone use and 12-month-old infant development across communication, motor (fine and gross), problem-solving, and personal-social domains using the Ages and Stages Questionnaire (ASQ-3). They found that a higher use of touch screens was associated with poorer infant fine and gross motor skills and personal-social skills, as well as overall lower ASQ-3 scores. Furthermore, poorer maternal and/or paternal mental health, as measured with the Depression, Anxiety, and Stress Scale was associated with lower scores across all domains except for an infant's gross motor skills. Poorer parental mental health was also associated with higher infant use of mobile touch screens, indicating that parental use and infant use may be influenced by the mental health state of the parent or caregiver. Garg (2022) performed semi-structured interviews and collected daily diaries of parental cell phone use. Parents who had access to private spaces and worked long hours away

from their children reported lower levels of use. Parental cultural backgrounds, emotional reactions of the children, and self- or mutually-decided norms also influenced their cell phone use in both directions. Some parents also believed that they could use their own use to showcase ideal behaviours for their children and teach their children technology-related skills and understanding. Parents often switched between different types of cell phone use (contented use, cautious and reluctant use, resistance and deferral, and contented non-use), indicating that most parents are aware of the cell phone use and make conscious decisions around it based on their workload and notions of appropriate levels of responsiveness to their infants. Above all, many parents reported that external recommendations for technology use around children are often contradictory and may not involve solutions that are easy or realistic to implement.

Chapter 3: The Effects of Technology Use on Infant Language Development: Study of an Italian Cohort

Guidelines set out by the World Health Organization (World Health Organization, 2019) highly discourage screen media for children under 12 months. More recently, the Canadian Pediatric Society (Ponti, 2023) set out recommendations that discourage screen media exposure to infants under two years old, apart from video chatting with caregivers. The idea is to avoid sedentary behaviour as becoming a part of a child's routine if they are under five years old. They also recommend maintaining daily screen-free times, such as during mealtimes or book reading, and to avoid screens altogether at least one hour before bedtime. The American Academy of Pediatrics updated their recommendations, focusing less on recommending "time limits" for children's viewing habits, and instead focusing on rules to create balance, observe content, encourage co-viewing, and facilitate communication between infants and their parents (Munzer et al., 2026). They claimed that these types of rules are often associated with better outcomes for the infants well-being. Emphasis has also been placed on parents and caregivers to model healthy screen use themselves, by participating in activities away from screens and turning off the television and background television when it is not in use (Ponti, 2023).

Despite many organizational recommendations and concerns for infant technology use and screen exposure, many studies have demonstrated that children are not meeting the recommended guidelines (Alroqi et al., 2022). For example, a census on American infants discovered that infants under two years old are exposed to over an hour of screen media per day, with 6% of families declaring that their infants under two own their own tablets (Mann et al., 2025). Studies on American infants have reported various averages of daily screen time, from as little as 40.2 minutes per day (Zimmerman et al., 2007b) to as much as 124.5 minutes per day (Tomopoulos et al., 2010), using parental reports of infant viewing time. Studies from countries outside of the United States have also attempted to understand and quantify the amount of infant technology use and screen exposure. In a study of French infants by Cristia and Seidl (2015), 30% of infants at 5 months old and 90% of infants at 2 years old used touchscreens, with 21% of the total sample reporting that their infants used touchscreens daily. In Germany, by 12 months of age, 33% of children had been exposed to television and 16.9% exposed to smartphones, with 45% of the infants not meeting the American Academy of Pediatrics' guidelines for no hours of screen time under the age of two (Durham et al., 2021). Furthermore, a study conducted on New Zealand infants found that, on

average, infants at the age of 20 months used screens for 55 minutes per day (Corkin et al., 2021), while another study of Thai infants between 15 and 35 months old reported 67.4 minutes of television exposure per day (Lin et al., 2015). Given the above reports, it is apparent that technology plays a large role in infants' daily lives, and, by default, in their development over the first two years of life. It is also evident that infants are exposed to high levels of technology and screen media, and are not meeting the recommendations set out for no screen exposure under the age of two.

Investigations into the effects of high levels of technology use and screen exposure on infant development have seen an uptake in recent years, in all areas of infant development. For example, a meta-analysis by Mallawaarachchi and colleagues (2024) on the effects of television on infant cognitive development reported that cognitive development of infants between the ages of 0 and 6.5 years old is negatively associated with television watching and exposure to background television. Furthermore, watching television and age-inappropriate content, as well as caregiver use of screens during daily routines were negatively associated with the infants' psychosocial outcomes. Of note, a study from Krogh and colleagues (2021) found moderate to strong associations between earlier and later amounts of screen time, in infants between the ages of 2 and 11 months old, with interruptions of mother-infant interactions due to maternal screen use between 5 and 6 times a day. Thus, suggesting that high levels of television exposure do not decrease with age, but are held constant across development. Sundqvist and colleagues (2023) extended these age ranges, and found that screen time from when infants are as young as 9 months old correlated with their screen time at 5 years of age. This brings to question the cumulative or longitudinal effects that repeated exposure to high levels of technology may have on developmental outcomes.

Technology Use in the Italian Context

Although there has been a recent uptake in research on infant technology use and screen exposure in non-English speaking countries, a scoping review by Mannell and colleagues (2024) argued that there was still an overrepresentation of English-speaking infants in the current available literature. In order to better understand whether technology impacts infant development similarly across various cultures, this study focused its investigation on Italian-speaking infants. For Italian-speaking infants, the Italian Society of Pediatric Research (SIRP) put forward multiple recommendations, similar to the guidelines and recommendations put forward by the

aforementioned pediatric societies and WHO. These guidelines included the recommendations that children should not have a personal smartphone until at least the age of 13 years old, that all devices should be prohibited during mealtimes, before bedtime, and within bedrooms, and that outdoor activities, sports, reading, and creative play should be encouraged (Bozzola et al., 2026). Despite these recommendations, questionnaires completed both online and in pediatricians offices reported that over 10% (in office) and 5% (online) of parents reported allowing their 12-24 month old child use a preferred device for more than one hour per day, which rose to over 25% (in office) and 15% (online) by the time the child was 24-36 months old (Balbinot et al., 2016). Furthermore, for infants between 12 and 24 months of age, over 50% (in office and online) of parents reported using a device to "keep the child quiet". Importantly, over 60% of parents (in office and online) reported always or almost always co-viewing with the child when using the device, which may help to offset the negative effects of using high amounts of technology and screen media.

Few studies have tried to gain a better understanding of the impact of technology use and screen exposure on Italian-speaking infants. Operto and colleagues (2020) performed a cross-sectional analysis of infants between the ages of 8 and 36 months, using parent-reported measures such as the Digital Device Questionnaire, and *Il primo vocabolario del bambino* (PVB). They found that, on average, parents reported children using digital devices for 2.13 ± 2.04 hours per day, suggesting a high variability of technology use and screen exposure within Italian infants, similar to previous studies on American infants (Tomopoulos et al., 2010; Zimmerman et al., 2007b). When dividing this group of infants into older and younger infants, the younger infants, aged 8 to 17 months, demonstrated a significant negative relationship between the total daily time of exposure to digital devices and the actions and gestures quotient derived from the PVB, but not for the lexical understanding or lexical production quotients. The older infants, aged 18 to 36 months, demonstrated a significant negative relationship between the total daily time of exposure to digital devices and both lexical quotients. These data suggest that the impacts of digital device use may be differential based on the age of the infant, whereby younger infants, or infants with lower levels of productive vocabulary, may experience negative effects on aspects of nonverbal communication, while older infants are more likely to have their language production and comprehension impacted. Paoletti and colleagues (2025) explicitly investigated infant use and maternal use of screen media devices. Infants were tested at 4 months and 8 months old through the use of online questionnaires and developmental assessments. Infant use at 4 months of age was

found to be negatively associated with word communication abilities at 8 months of age, while maternal use of digital media was negatively associated with infant gesture production and use during infant routines was negatively associated with infant language comprehension. In line with previous research, however, the time infants spent looking at traditional and electronic books, listening to adults reading a paper book (Asikainen et al., 2021; Zimmerman et al., 2007a) and listening to music (Houde & Narendran, 2018) were positively associated with infant gestures and with first word production. Thus, highlighting the importance of adult-infant interaction as well as joint-engagement in an activity, regardless of the medium.

Some Methodological Concerns to Consider

Of the two studies described above, the first study took advantage of a cross-sectional investigation, which allows for quicker data collection, the ability to gather data from more participants, and a general understanding of the effects across different ages and stages of development. On the other hand, the second study implemented a brief longitudinal design, which can investigate and map individual trajectories of the effects being seen, to allow for closer investigation into individual differences that may lead to some variability between subjects. What is missing from these studies that have investigated the Italian context, is a longitudinal design to understand how the change in technology use and screen exposure impacts infant development over a longer period of development. Both of the aforementioned studies also made use of parent questionnaires as a primary source for data collection. Concerns have been raised about the use of parent questionnaires due to the subjective nature of parent-reporting and the potential for parent bias to influence their responses, and, by extension, the results, as well as concerns related to the comparability across studies. Two of the main sources of bias from parents responding to questionnaires are social desirability and recall biases. Social desirability bias is the tendency for individuals to present themselves in a manner that will be viewed favourably by others (American Psychological Association, 2018). Socially desirable responding has been linked to parenting self-reports, whereby parents are less likely to report negative or harsher parenting as opposed to positive parenting (Bornstein et al., 2015), which have consequences for the interpretation of these results. This also brings into question how accurately parents are responding to digital media questionnaires, when they are aware of the recommendations and guidelines for infant technology use, or aware of the negative effects of high levels of technology use. For example, in the

aforementioned study from Zimmerman and colleagues (2007b), where parents had reported their own infants' daily viewing time at 40.2 minutes per day, they also reported that their infants' peers' viewing was considerably higher at 69.6 minutes per day. Meanwhile, recall bias is when retrospective reporting includes inaccurate aspects, such as the systematic over- or under-count of a frequency with which a behaviour occurred (American Psychological Association, 2018), which may lead to inaccurate information being reported by the parents, and inaccurate data being collected and analyzed by the researchers. In terms of comparability, questionnaires may include open- or closed-ended questions, and may or may not be standardized or used in previous research. These different types of questionnaires may yield different levels of validity in the parent reports (see Beynon et al., 2024). Diaries have been used as a helpful alternative or supplement to these questionnaires, as they require shorter recall periods (Beynon et al., 2024). Although concerns for the questions and style of reporting, as well as standardization for the diary may be a factor for its use, parents have also reported that the completion of a diary for more than three days is taxing (Tey et al., 2007).

In order to have an objective, unbiased estimate of technology use and screen exposure in the home environment, many researchers have turned to recording parent-infant interactions. For naturalistic observations, some studies have used video recordings to capture short interactions in the home (e.g., Suarez-Rivera et al., 2024), while others have taken advantage of audio recording technology (e.g., Brushe et al., 2023; Christakis et al., 2009; Ramírez et al., 2021; Weisleder & Fernald, 2013; Zimmerman et al., 2009). Audio recording technology can allow researchers to have day-long or multiple-day-long recordings of the infant's audio environment, without intrusion from the researcher themselves. Designed to provide insight into patterns of child vocal development and behaviour, and to generate simple, high-quality feedback on a child's natural language environment to promote adult behaviour change, the Language ENvironment Analysis (LENA) system was created (Gilkerson & Richards, 2020). The LENA device was also designed to be used in LENA clothing, which is a specialized vest that an infant can wear, with a pocket situated in the front centre of the vest for the LENA device (Gilkerson & Richards, 2020). The material for this vest was selected for its low friction properties, to allow for high-quality audio to pass through to the device (Gilkerson & Richards, 2020). While the main purpose of LENA is to record and count speech-related variables, including adult speech, child vocalizations, and conversational turns that occur between an adult and the child wearing the vest, or the target child (Gilkerson & Richards,

2020), it has also demonstrated strengths in recording, measuring, and reporting the variable of television and electronic sounds (TV/ES). The LENA device only reports TV/ES segments that are at least 1000ms in duration, which may also help to identify the effects of exposure to technology that are experienced in shorter increments throughout the day, for which parents may not consider in their reports. Christakis and colleagues (2009) reported that there was moderate to good agreement between LENA and human coders, on several measures, including adult speech, child vocalizations, conversational turns, and TV/ES with a Cohen's Kappa statistic ranging from 0.43 to 0.70.

The Relationship Between TV/ES and Language Development in Infancy

Studies utilizing the LENA device have demonstrated relationships between the LENA-reported variable of TV/ES and other reported variables from LENA, including adult speech and infant vocalizations (Brushe et al., 2023; Ramírez et al., 2021; Zimmerman et al., 2009). Brushe and colleagues (2023) reported an overall negative association with the amount of TV/ES infants were exposed to and overheard adult speech, in Australian infants from 12 to 36 months old. The associations for vocalizations were less clear, where TV/ES was associated with a decrease in vocalizations at 12 months, but an increase at 30 months. Ramírez and colleagues (2021) investigated the effects of TV/ES on adult and infant speech in American families with varying socioeconomic backgrounds. In this study, TV/ES was associated with SES which interacted with the age of the infant, whereby infants from high-SES families demonstrated a significant negative association between TV/ES exposure and age, but infants from low-SES backgrounds did not. Interestingly, neither of the LENA-reported variables of adult or infant speech were associated with TV/ES when looking at the overall daily amounts of TV/ES. However, when a closer analysis was completed to consider every 5-minute interval, exposure to TV/ES was negatively associated with adult speech, independent of age and SES. It is important to consider the potential for systematic effects that may be seen from families with varying degrees of SES, and to understand if effects seen from shorter interval periods, as opposed to the average over the course of the day, also have an effect on the overall language development of the infant. To investigate this line of inquiry, some studies have included measures of development in addition to the LENA-reported variables of infant and adult speech. For example, Zimmerman and colleagues (2009) tested infants using the PLS-4, in addition to the audio recordings from LENA, and found positive associations

between adult speech and conversational turns from LENA with language scores on the PLS-4, and negative associations between the TV/ES exposure reported by LENA and language scores from the PLS-4 in infants between 2 and 48 months of age. The effect of TV/ES, however, was shown to be mediated by infant-adult conversation, as the interaction between TV/ES and language development was no longer significant when the conversation variable was included in the analysis. This highlights the need for future research to also consider the various aspects involved in infant language development, including the potential mediating effects of the adult, and how they may influence one another.

The Present Study

The present study aims to rectify some of the above concerns in investigating the relationship between infant exposure to technology use and screen exposure and language development, by using the LENA device as an unbiased, objective measure of technology use and screen media, and investigating a sample of Italian-speaking infants. As this study will rely on the application of LENA to measure technology and media exposure, the first aim of this study was to validate the device in measuring TV/ES in an Italian-speaking sample (Aim 1). Based on previous validations of LENA for the variable of TV/ES (Christakis et al., 2009; Zimmerman et al., 2009), as well as previous validations of other LENA variables in Italian-speaking samples (Bastianello et al., 2023), we expect to see a moderate to strong correlation between the human-coder reported TV/ES exposure and the LENA-reported TV/ES exposure. The second aim attempts to understand and describe the day of a typical Italian infant, in relation to their TV/ES exposure and overheard adult speech (Aim 2). This includes understanding the types of TV/ES an infant may be exposed to, when peak periods of TV/ES may occur, and how infant-overheard adult speech evolves over the course of the day. Given the exploratory nature of this aim, no specific hypotheses were formulated.

The third and final aim of this study is to explain the relationship between TV/ES and its influence on language throughout infant development (Aim 3). Previous studies have focused solely on the age of the infant, and some have attempted to use the age of data collection as an approximation for the phase of language development the infant may be in (e.g., Ramírez et al., 2021). However, previous studies have not found associations between adult speech (Rowe, 2012) or technology use and age (Ramírez et al., 2021), suggesting high variability in what infants are exposed to across

the first two years of life. For a better understanding of how these factors interact with an infant's ability to produce and comprehend language, this study uses the phase of language development directly instead of relying on the infant's age, measuring four of its key phases: pre-babbling, babbling, first words, and vocabulary spurt. We expect to see that, across the phases of language development, infants who are exposed to high levels of TV/ES will have lower levels of language development (i.e., Brushe et al., 2023; Ramírez et al., 2021; Zimmerman et al., 2009). Although the primary goal of this aim is to understand TV/ES on language development, it is important to consider the other factors involved that may mitigate the risk of high levels of TV/ES, such as overheard adult speech. For this reason, the third aim will also investigate the role of adult speech on infant language development, and how it may be effected by TV/ES. In line with previous research, we expect to see that higher levels of TV/ES will be associated with lower levels of adult speech (Brushe et al., 2023; Kirkorian et al., 2009; Ramírez et al., 2021; Zimmerman et al., 2009), and that lower levels of adult speech will be associated with lower levels of infant language development (Hart et al., 1977; Huttenlocher et al., 1991; Rowe, 2012; Weisleder & Fernald, 2013).

Chapter 4: A Continuation of The Effects of Technology Use on Infant Language Development: Study of an Italian Cohort

Method

Procedure

This study is part of a greater longitudinal study within the IN-MIND ERC project lead by Prof. Benavides-Varela.

Each infant of the cohort was observed at four distinct time points. Each time point was defined by the acquisition of key linguistic milestones (pre-babbling, babbling, first words, and vocabulary-sport phases). Infants were recruited during their pre-babbling phase and followed through the remaining phases over the first two years of life. To identify the onset of a linguistic phase, a checklist was provided to families biweekly, with telephone interviews used to supplement the checklists as necessary. Upon entering the next stage, the Italian version of the MacArthur-Bates Communicative Developmental inventory short form (CDI) (Caselli, et al., 2015) was then administered, and the LENA device was delivered to the family for data collection.

Participants

Thirty-two families agreed to participate in the longitudinal study. Parents declared that their infants had no birth complications or critical health issues during the first two years of life. All infants were born full-term (gestation: 276.94 ± 8.46 days (approximately 39.5 ± 1.2 weeks); weight: 3363.79 ± 423.24 grams). Parents reported no familial history of language or developmental delays. Infant characteristics are reported in Table 1. As a result of the longitudinal nature of the study and associated attrition, some infants were not available for data collection at each time point in the study.

The average number of years spent in school for the mother was 17.22 years (equivalent to the completion of a bachelor's degree), and 16.38 years for the father (equivalent to the completion of a bachelor's degree), with one mother and three fathers completing a PhD, 18 mothers and 13 fathers completing a master's degree, and 4 mothers and 3 fathers completing post-graduate specialization or training, indicating a high level of parental education present within this sample.

As a result of practical constraints on recruitment capacity, time, and resources, this study adopts an exploratory approach to inform the design of future research.

Families were recruited through advertisements at social events, on social media, and on flyers. A 15 euro travel reimbursement was provided for participation. This project protocol was approved by the Institutional Ethics Committee of Psychology Research at the University of Padova (Protocol N. 3697).

Table 1. Sample size and age in months at each data collection point.

	Pre-Babbling	Babbling	First Words	Vocabulary Spurt
n	29	32	30	31
Age (months), mean (sd)	7.17 (2.30)	11.67 (3.35)	17.47 (3.28)	23.10 (3.00)
Age (months), min	5	6	11	17
Age (months), max	13	21	24	30
Gender	F 14; M: 15	F: 15; M: 17	F: 13; M: 17	F: 14; M: 17

Note. n indicates the number of participants at each data collection point. Age is summarized using the mean, standard deviation (sd), minimum (min), and maximum (max).

The Checklist:

An ad-hoc screening checklist was created by the researchers based on infant behaviours that are typically present within each phase of vocal development, to more clearly understand the linguistic development of each individual. The checklist targeted specific phonological and language skills that have been observed throughout infant development, that often indicate where the infant falls in their period of development (see Chapter 1 for information on infant behaviours during each stage of language development), including questions on the infant's language perception, comprehension, and production (see Appendix A).

Parents were asked to complete the checklist every two weeks throughout their involvement in the study. Parents reported whether or not an item was occurring, as well as how often it was occurring. When an infant was reported to have at least three behaviours of the next phase, the infant was considered as entering a new phase of linguistic development, and the LENA device given to the

family and the CDI was administered. If no entry into the next phase was reported, the checklist continued to be administered every two weeks until a notable shift in development occurred (e.g., if the infant is transitioning from the pre-babbling to the babbling phase, at least three items under the babbling phase must be checked off as occurring, to be considered as the infant entering into the babbling phase).

The Language ENvironment Analysis (LENA) System

A LENA device was provided to the families at each phase of vocal development to be turned on and recording for a full day. The LENA system is a wearable audio recording device that is placed into the front chest pocket of a specialized vest, for which the target child wears throughout the period of recording. Families were instructed to switch the device on in the morning when the infant wakes up, and keep it on until the infant goes to sleep at the end of the day. On average, the LENA device was recording for 14.05 hours across all phases. Upon inspection of the data, segments before 7:00 (7:00 a.m.) and after 22:00 (10:00 p.m.) included only silence, with none of the other LENA metrics being reported. This suggests that the devices were turned on early or left on while the infant was asleep for the evening, and so these segments were deleted, and is consistent with previous studies investigations into Italian-infants bedtimes being relatively stable throughout the first year of development at 22:00 (10:00 p.m.; Bruni et al., 2014). Using only the segments with other recorded audio information, LENA was recording for an average of 13.21 hours. The LENA device recorded for an average of 13.28 hours during the pre-babbling phase, 13.24 hours during the babbling phase, 13.45 hours during the first words phase, and 13.30 hours during the vocabulary spurt phase.

The LENA Online Platform uses the audio information from the speech signal (syllable count, consonant distribution, and segment duration) to detect and extract four key components of the target infant's auditory environment, including Adult Word Count (AWC), Child Vocalization Count (CVC), Conversational Turn Count (CTC), Television and Electronic Sounds (TV/ES), as well as other components of the auditory environment, such as noise, silence, vocalizations from other children, and when two categories overlap (Gilkerson & Richards, 2020). For the purpose of this study and to focus on the aims described above, we will analyze the AWC, CVC, and TV/ES variables that are automatically extracted by LENA. According to The LENA Foundation (Gilkerson and Richards, 2020), the CVC is an estimate of the number of times a child produced

communicative vocalizations, such as speech-related babbles, words, and sentences, but not vegetative or fixed, sounds. All vocalizations must have a duration of at least 600 milliseconds and be separated by 300 milliseconds of silence or other sounds. The AWC is an estimate of the number of words spoken by any post-pubescent (adult) male or female in the child's environment that have a minimum duration of 1000 milliseconds, or, one second. The TV/ES is an estimate of the amount of time that sounds originating from an electronic speaker were the dominant sound in the child's environment and have a minimum duration of 1000 milliseconds, or one second. These segments are not divided by the type of noise that is being emitted from the electronic device, and may include television, electronic toys, music, and cell phone or laptop speakers. Of note, segments may also be misclassified if there is a reverb, or echo, as this can effect the integrity of the speech signal reaching the device (Xu et al., 2009).

Although LENA has been validated mainly with American-English speaking infants and children, Bastianello and colleagues (2023) provided validation for use in an Italian-speaking sample with infants between the ages of one and two years old. Significant strong positive correlations were found for the AWC ($r = .78, p < .001$), with LENA slightly overestimating the number of adult words spoken, and significant but weaker correlations found for the CVC ($r = .47, p = 0.001$), with LENA slightly underestimating the number of vocalizations produced when compared to the human coder. Furthermore, as LENA was weakly correlated with CTC ($r = .33, p = .005$), and significant differences were found between LENA and the human coder ($p = .041$), this variable will not be analyzed for our research aims (see Bastianello et al., 2023 for a full review).

However, LENA has not been validated for its TV/ES metric when used with an Italian sample. In fidelity matrices reported by Christakis and colleagues (2009) and Zimmerman and colleagues (2009) from recordings of English-speaking infants, of the segments that humans coded as TV/ES, LENA correctly identified 70.5%, erroneously labelling 7.8% of the segments as adult speech, 0.5% of the segments as child vocalizations, and 21.2% of the segments as "other". The results indicated a high degree of agreement between the LENA device and the human coders. The authors also point out that when LENA did misidentify TV/ES, the segments were not often being mistaken for one of the key variables being used in the analyses, the adult speech and child vocalization variables.

Statistical Analyses

Data processing was completed using the LENA Online System, and all analyses were conducted using R (4.4.1), using the RStudio (Version 2026.04.0+526) development environment. Linear mixed effects models were run using the lme4 package in R (Bates, 2015), while the post-hoc comparisons were calculated using the emmeans package (Lenth, 2026), and plots were generated using the ggplot2 package (Wickham, 2016).

Aim 1: Validating the Use of the LENA Device to Measure TV/ES Exposure

To address the limitations described above, the first aim of this project was to validate LENA for the use of TV/ES in our Italian-speaking sample. A randomly selected portion of audio recordings, encompassing 35 hour-long audio segments from across 28 participants between 8 and 13 months, for a total of 2100 minutes. These segments were selected using audio data from both the longitudinal and cross-sectional LENA data that has been collected as a part of a larger study. A minute-by-minute analysis was performed whereby each minute was classified as having the presence or absence of an electronic sound. Electronic sounds were defined using a similar metric to LENA, to which it was recorded as such if there were any sounds that came through an electronic speaker. The coder also reported if the device was being used for a full or partial minute, whereby full minutes were calculated as sixty seconds and partial minutes were calculated as thirty seconds. Due to software constraints, the coder was unable to report with a less discrete measurement of time. As LENA produces reports only for 5-minute, hourly, or daily segments, the minute-by-minute data from the human coder was then transformed into segments of 5-minutes, by adding together the full or half minutes of TV/ES for the five consecutive minutes, to allow for cross-comparison. Comparisons of the total amount of TV/ES recorded by the human coder and LENA were reported and analyzed with Spearman correlations.

Aim 2: Understanding the Types of Sounds that Italian Infants Hear in a Typical Day

The second aim of this study was to understand a day of a typical Italian infant, the types of sounds they may hear, and how that may vary over time. First, we attempted to describe the types of TV/ES that infants are exposed to, on average. During the validation of the LENA device, the human coder was asked to record the types of electronic sounds present in the infants environment. A percentage of use for each type of device was calculated dividing the number of

minutes each device was present for by the total number of minutes of TV/ES exposure as recorded by the human coder. We then attempted to understand how TV/ES exposure evolved over the course of a 16-hour day (from 7:00 to 22:00, or 7:00 a.m. to 10:00 p.m.). We then investigated how CVC and AWC may fluctuate with TV/ES, as well as if AWC and CVC fluctuate with each other. For this, we calculated the average TV/ES exposure per hour across all infants, for each phase of development, and then conducted a repeated-measures correlation analysis (see Bakdash & Marusich, 2017; see Figure 6 for the average course of development) to account for each infant's individual trajectory within the course of the day.

Aim 3: Understanding TV/ES Exposure Over the Course of Infant Language Development

The third and final aim of this study was to understand the effects of TV/ES exposure on infant language development, using the LENA-reported metrics of AWC and CVC. For this analysis, the daily totals of TV/ES, AWC, and CVC were used. Linear mixed effects models were conducted to allow for each participant to be compared to their own baseline. Checks of singularity and convergence, and additional checks for collinearity were completed on models with interaction effects. Model fit was confirmed with visual inspection and testing of simulated residuals using the DHARMA package in R (Hartig, 2024). Planned post-hoc pairwise comparisons using the Bonferroni adjustment to control for multiple comparisons were run for significant models. If multiple models fit the data and demonstrated significance for the outcome variable, a likelihood-ratio chi-square test was conducted to determine the best fitting model.

Results

Aim 1: Validating the Use of the LENA Device to Measure TV/ES Exposure

After a visual inspection through the scatterplot analysis of both the human- and LENA-coded TV/ES data, it was evident that the data were not normally distributed. A Shapiro-Wilk's test was conducted and returned as significant, confirming the visual observations of the data (human-coded data: $W(35) = .73, p < .001$; LENA-coded data: $W(35) = .84, p < .001$). A Spearman correlation was then used to account for the non-normal distribution of the data, and was significant, confirming that the human- and LENA-correlated data were significantly positively correlated ($\rho = .50, p < .001$), indicating that as the human coder identified more minutes of TV/ES, so did the LENA device. On average, the human coder identified 1.94 minutes and LENA coded

2.23 minutes of TV/ES exposure for every 5-minute segment, with a 1.39 difference for the amount of TV/ES exposure minutes recorded by the LENA and the human coder. A Wilcoxon signed rank test was also used to confirm that these differences were not significant ($W = 24,922$, $p = .07$), so that LENA is able to be used as an objective measure of TV/ES in Italian-speaking infants.

Aim 2: Understanding the Types of Sounds that Italian Infants Hear in a Typical Day

Of the 2100 minutes analyzed, 815.49 minutes were coded as involving the exposure of TV/ES. In order to obtain an estimate of the type of device being used during this time, the human coder was asked to record the types of electronic sounds present in the infants environment, the number of devices used in each segment was totaled, and the segment (either a full-minute or half-minute) was divided by the number of devices present, to obtain an estimate of how long each device was used during that minute.

According to the data reported by the human coder, Italian infants were most commonly exposed to music, followed by live music, phones and computers, the television, and electronic toys. As live music or musical street performances do not necessarily represent the use of screen media or digital technology, as described in previous chapters, it was kept as a separate category titled "Live Music". Because the LENA device records all sounds coming through electronic speakers as TV/ES, this helps to understand that conclusions of TV/ES exposure using this device should be

Table 2: Types of TV/ES to which the Italian sample was exposed to.

Type of TV/ES	Amount of Minutes	Percent of Time
Live Music	192.00	23.54%
Music	339.91	41.68%
Toy	78.33	9.61%
TV	88.92	10.90%
Phone/Computer	100.58	12.33%
Other	15.75	1.93%

undertaken with caution. The category of "Music" included a home and car radio, Bluetooth speakers used at home, and car speakers. However, as no information on this data were collected from parents, and LENA records only audio information, the music heard may also have come from a television, phone, or laptop that was also playing a video in the background. The category of "Other"

included sounds that did not appear to come from any other category, but did appear to come through an electronic speaker, such as buttons being pressed on a cashier till and a navigation

system. See Table 2 for the exact minutes and percent values for the types of TV/ES exposure, and Figure 5 for a visual representation.

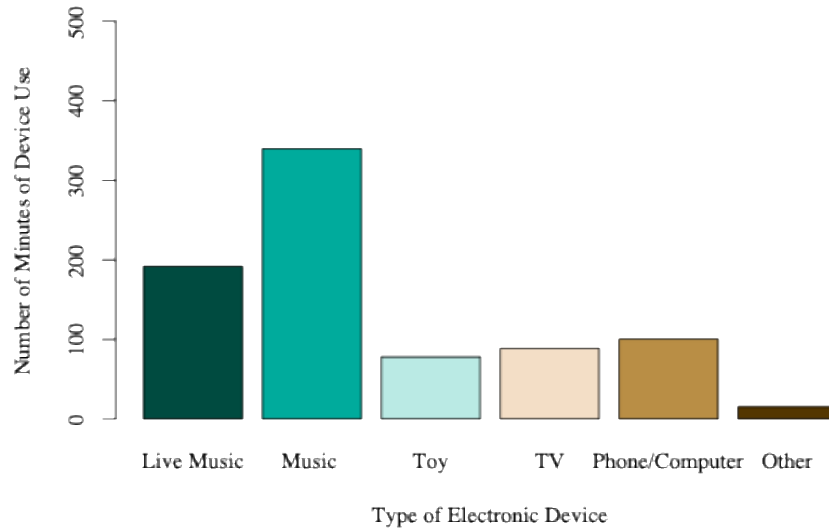


Figure 5: Types of TV/ES exposure in the typical day of an Italian infant.

Italian infants were exposed to a total average of 44.94 minutes of television and 24,932.69 adult words, as well as produced an average of 1794.97 linguistic vocalizations in a typical day. Table 3 holds information for the averages in each phase of development, and Figure 6 displays these trends with the variables transformed to fit a z-distribution for comparability and readability.

Table 3: Average TV/ES, AWC, CVC, and CTC in each phase of development.

LENA Descriptor	Pre-Babbling		Babbling		First Words		Vocabulary Spurt	
	Mean $\pm$ SD	Range	Mean $\pm$ SD	Range	Mean $\pm$ SD	Range	Mean $\pm$ SD	Range
TV/ES (in minutes)	45.45 $\pm$ 2.12 43.04	154.27 -	37.32 $\pm$ 1.66 46.10	202.97 -	53.01 $\pm$ 3.62 56.17	221.98 -	44.52 $\pm$ 0.82 43.19	172.30 -
AWC (count)	23,884.52 $\pm$ 9435.17	12,389 - 48,891	24,457.84 $\pm$ 9204.36	12,197 - 54,899	24,022.87 $\pm$ 8758.63	8324 - 42,258	27,283.87 $\pm$ 10,380.03	12811 - 47791
CVC (count)	1404.69 $\pm$ 479.54	643 - 3092	1493.06 $\pm$ 559.19	94 - 2906	1777.37 $\pm$ 845.52	73 - 3406	2488.74 $\pm$ 1204.70	80 - 4279

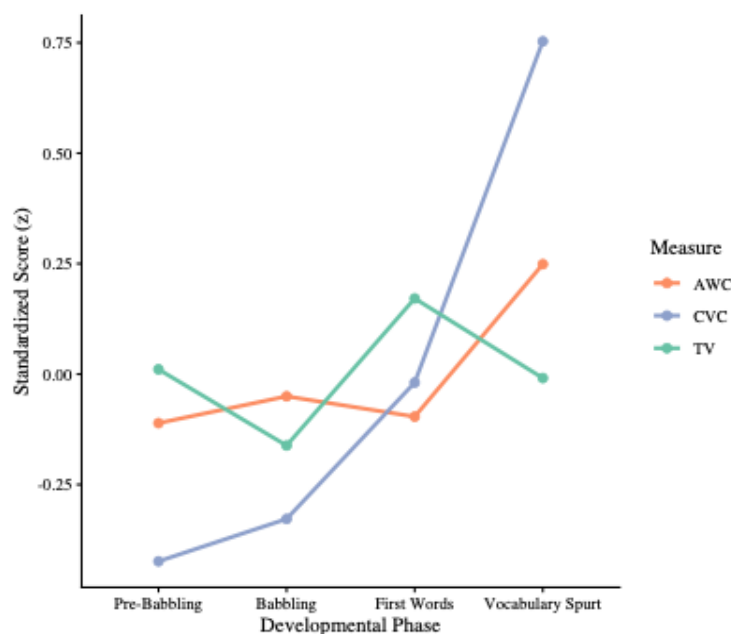


Figure 6: Average normed-scores of TV/ES exposure (in minutes), AWC, and CVC across phases of development.

To understand the audio environment of an Italian infant in a typical day, TV/ES, adult word count, and child vocalization count, were also graphically represented separately from each other. Visual inspection of these graphs revealed that there were some obvious peaks and depressions in specific times throughout the day (see Figures 7 and 8 for the z-transformed graphs, and Appendix B, C, D, and E for bar

charts representing exact averages). In the pre-babbling phase, TV/ES peaked between 13:00 and 15:00 (1:00 p.m. and 3:00 p.m.), with depressions around 11:00 (11:00 a.m.), 16:00 (4:00 p.m.), and after 19:00 (7:00 p.m.), which may be associated with infant daytime sleep. Adult word count and child vocalization counts did not demonstrate any peaks, but obvious depressions around 19:00 (7:00 p.m.). In the babbling phase, there were no clear peaks for TV/ES, AWC, or CVC, but obvious depressions for TV/ES at 20:00 (8:00 p.m.) and for AWC and CVC at 19:00 (7:00 p.m.). In the first words phase, there were obvious peaks for TV/ES, AWC, and CVC between 7:00 and

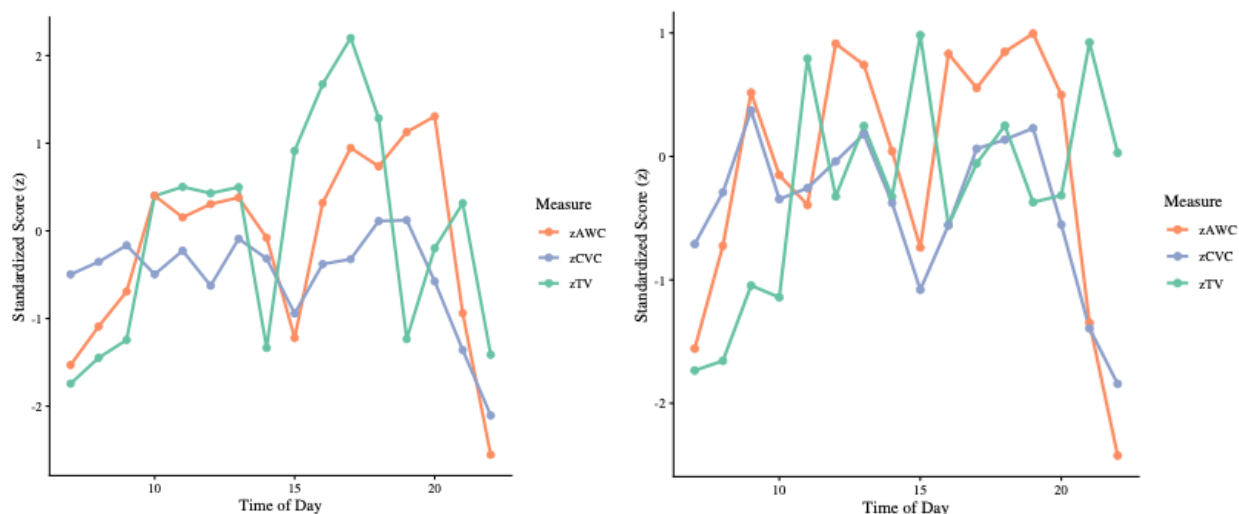


Figure 7: Pre-Babbling and Babbling Phases

9:00 (7:00 a.m. and 9:00 a.m.), 14:00 and 16:00 (2:00 p.m. and 4:00 p.m.), and 21:00 and 22:00 (9:00 p.m. and 10:00 p.m.). The obvious depressions in these phases may be a result of family routines, such as infant nighttime sleep routines. There was an obvious depression for TV/ES at 20:00 (8:00 p.m.) and for AWC and CVC at 19:00 (7:00 p.m.). In the vocabulary spurt phase, there were many peaks for TV/ES throughout the day, and obvious peaks for AWC and CVC from 7:00 to 10:00 (7:00 a.m. to 10:00 a.m.), 13:00 to 17:00 (1:00 p.m. to 5:00 p.m.), and from 20:00 to 22:00 (8:00 p.m. to 10:00 p.m.).

These patterns, while consistent with an infant napping twice daily, are interesting as the levels of TV/ES exposure are often peak during times of high adult speech and child vocalizations.

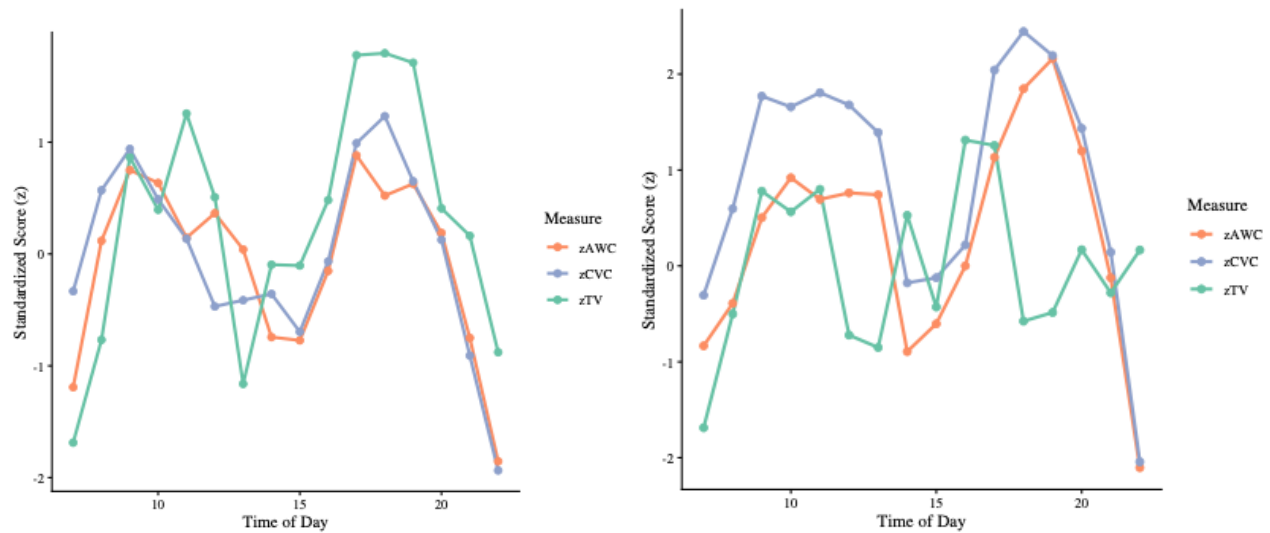


Figure 8: First Words and Vocabulary Spurt Phases

In order to understand how TV/ES exposure effects the LENA measures related to language development, repeated measures correlations were run (see Figures 9, 10, and 11; see Appendix F, and H for the repeated measures correlation plots separated by phases of development). The results revealed significant negative correlations between TV/ES exposure and adult word count ($r_{\text{rm}}(1694) = -0.13, p < .001, 95\% \text{ CI } [-.18, -.09]$), and child vocalization count ($r_{\text{rm}}(1694) = -.15, p < .001, 95\% \text{ CI } [-.19, -.10]$). Adult word count was also significantly positively correlated with CVC ($r_{\text{rm}}(1694) = .44, p < .001, 95\% \text{ CI } [.40, .47]$). Within each phase of development, there were also significant negative correlations with TV/ES exposure and adult word count (pre-babbling:

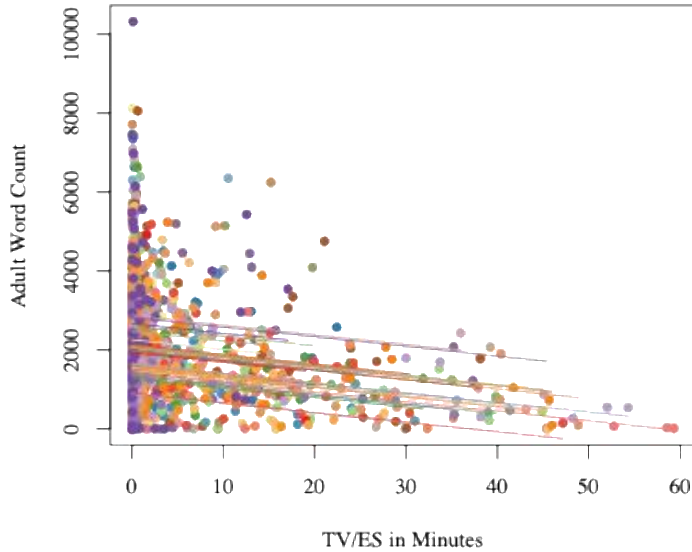


Figure 9: Repeated measures correlation plots for TV/ES (in minutes) and AWC across all phases of development.

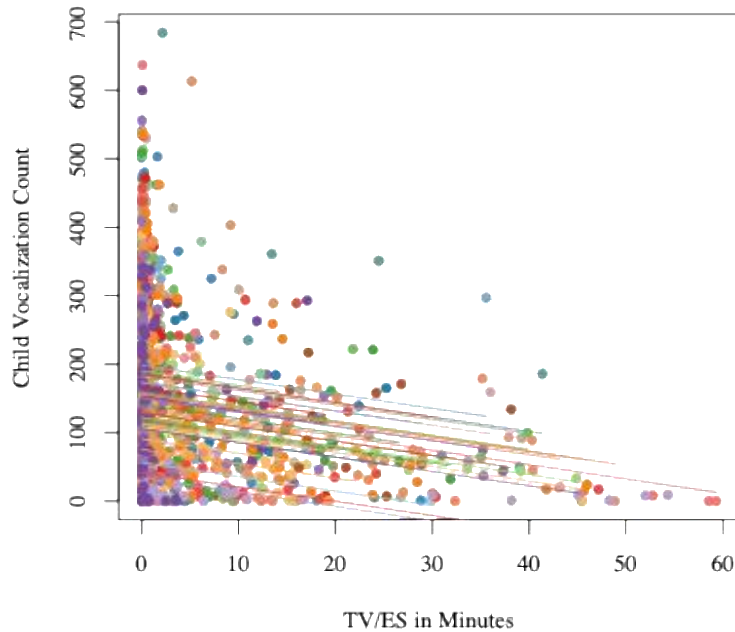


Figure 10: Repeated measures correlation plots for TV/ES (in minutes) and CVC across all phases of development.

$r_{rm}(380) = -.18, p < .001, 95\% \text{ CI } [-.27, -.08]$; babbling: $r_{rm}(421) = -.15, p = .003, 95\% \text{ CI } [-.24, -.05]$; first words: $r_{rm}(394) = -.11, p = .03, 95\% \text{ CI } [-.20, -.01]$; vocabulary spurt: $r_{rm}(406) = -.15, p = .002, 95\% \text{ CI } [-.24, -.05]$) and child vocalizations in the pre-babbling phase ($r_{rm}(380) = -.17, p < .001, 95\% \text{ CI } [-.27, -.07]$) babbling phase ($r_{rm}(421) = -.18, p < .001, 95\% \text{ CI } [-.27, -.08]$) and first words phase ($r_{rm}(394) = -.14, p = .006, 95\% \text{ CI } [-.23, -.04]$), but not the vocabulary spurt phase ($r_{rm}(406) = -.09, p = .06, 95\% \text{ CI } [-.19, .00]$). There were also significant positive correlations between AWC and CVC at all time points (pre-babbling: $r_{rm}(380) = -.19, p < .001, 95\% \text{ CI } [-.38, .19]$; babbling: $r_{rm}(421) = .42, p < .001, 95\% \text{ CI } [.34, .49]$; first words: $r_{rm}(394) = .51, p < .001, 95\% \text{ CI } [.44, .58]$; vocabulary spurt: $r_{rm}(406) = .55, p < .001, 95\% \text{ CI } [.48, .62]$).

Aim 3: Understanding TV/ES Exposure Over the Course of Infant Language Development

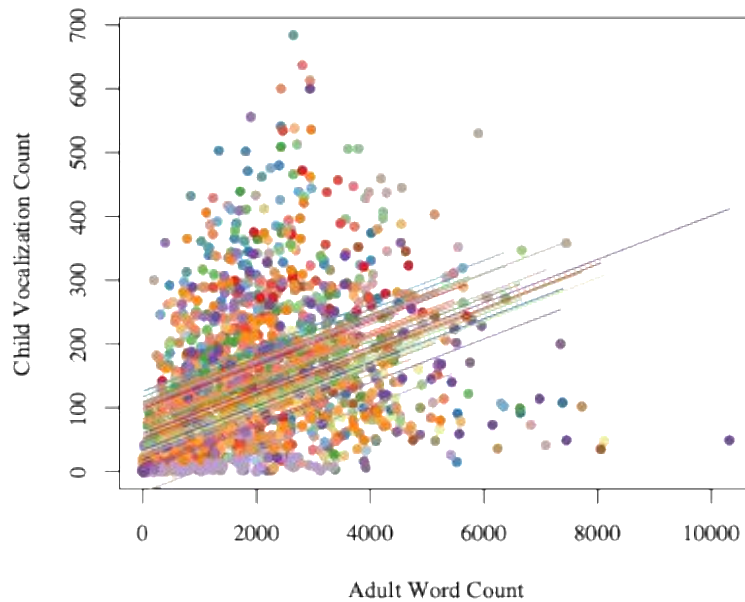


Figure 11: Repeated measures correlation plots for AWC and CVC across all phases of development.

then, in the case of AWC and CVC, with a developmental phase by TV/ES interaction as the predictor, to understand how the effects of TV/ES may change AWC or CVC over the course of development. Finally, an interaction model with an AWC and developmental phase interaction effect as the predictor and CVC as the outcome variable was run to investigate if the changing levels of AWC effected CVC. Because the TV/ES exposure (in minutes) was positively skewed and included zero values, it was transformed using the $\log_{1p}$ function in R for the first model, in which TV/ES was the outcome variable. This transformation calculates the natural log of TV/ES exposure +1, allowing zero values to be retained while reducing skew. For models in which TV/ES exposure and AWC were included as predictors, both TV/ES and AWC were standardized and mean-centered, such that a value of 0 represented the sample mean and a one-unit increase represented one-standard-deviation increase in the variable (also known as a z-transformation of the variable) to improve model fit.

Linear mixed effects models were run on the three main LENA-reported variables outlined above: TV/ES, AWC, CVC. All models included the individual participant as a random effect, to compare the changes throughout course of development to their unique baseline. The variables were tested with only developmental phase as a predictor, to understand if there were differences between each phase of language development, and

To understand if TV/ES changed over the course of development, the log of TV/ES was the outcome variable for the model, and developmental phase was the only predictor variable (Figure 12). There were no significant differences in the log of TV/ES over the course of development (all p 's > .05).

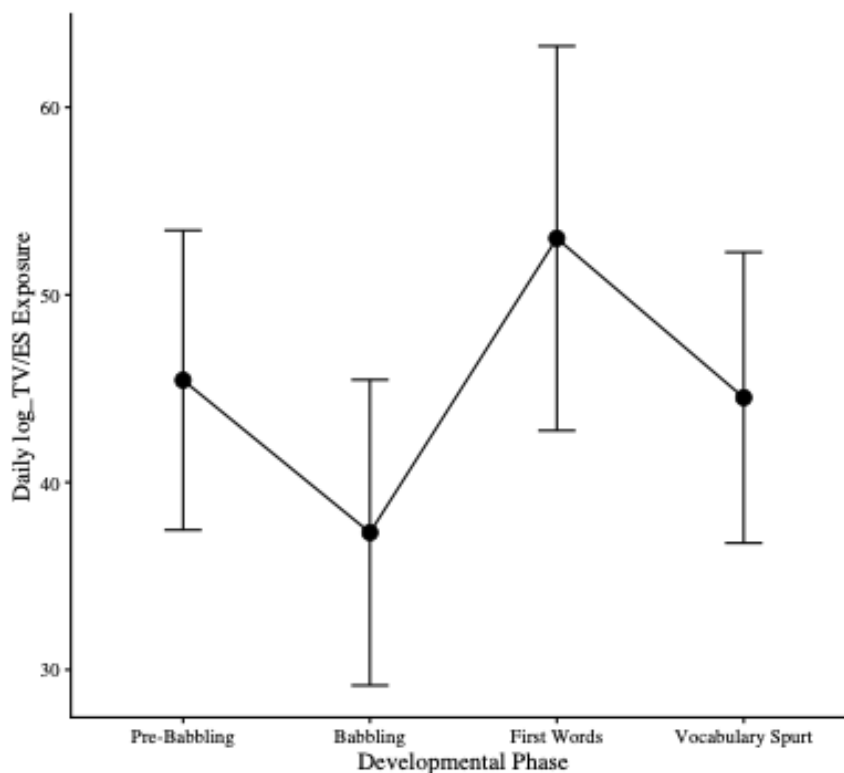


Figure 12: The log of TV/ES (in minutes) across phases of development.

To understand if AWC changed over the course of development, or was affected by the amount of TV/ES exposure, two models were run. The first included AWC as the outcome variable, and developmental phase as a single predictor (Figure 13). There were no significant differences found for AWC across phases of development

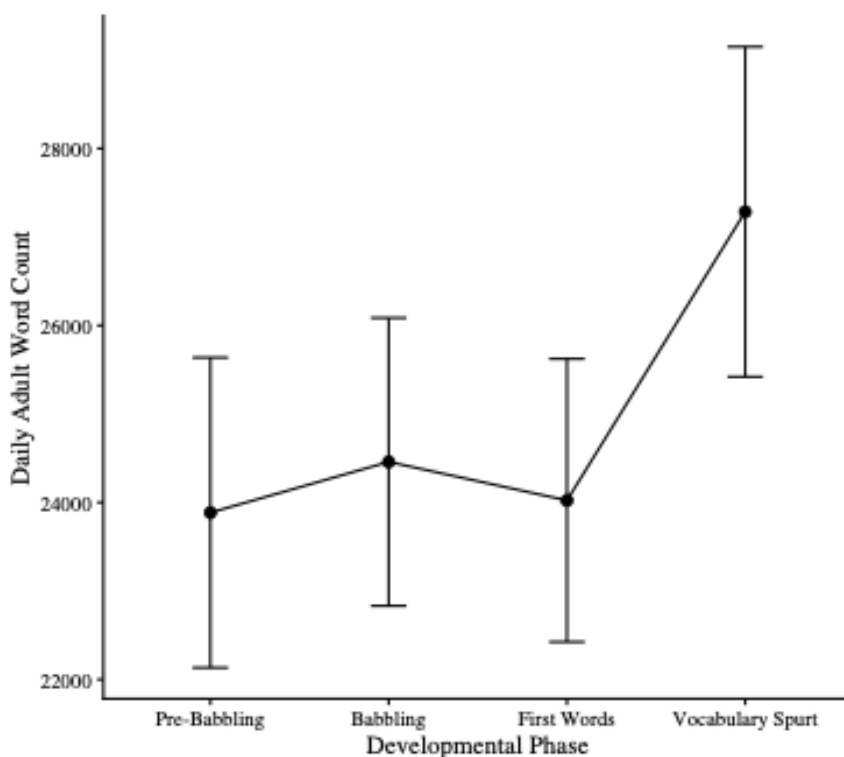


Figure 13: AWC across phases of development.

(all p 's > .05). The second model included TV/ES as a predictor variable (Figure 14). There were

no significant differences found for AWC when an interaction term between the phase of development and TV/ES was included in the model (all p 's > .05).

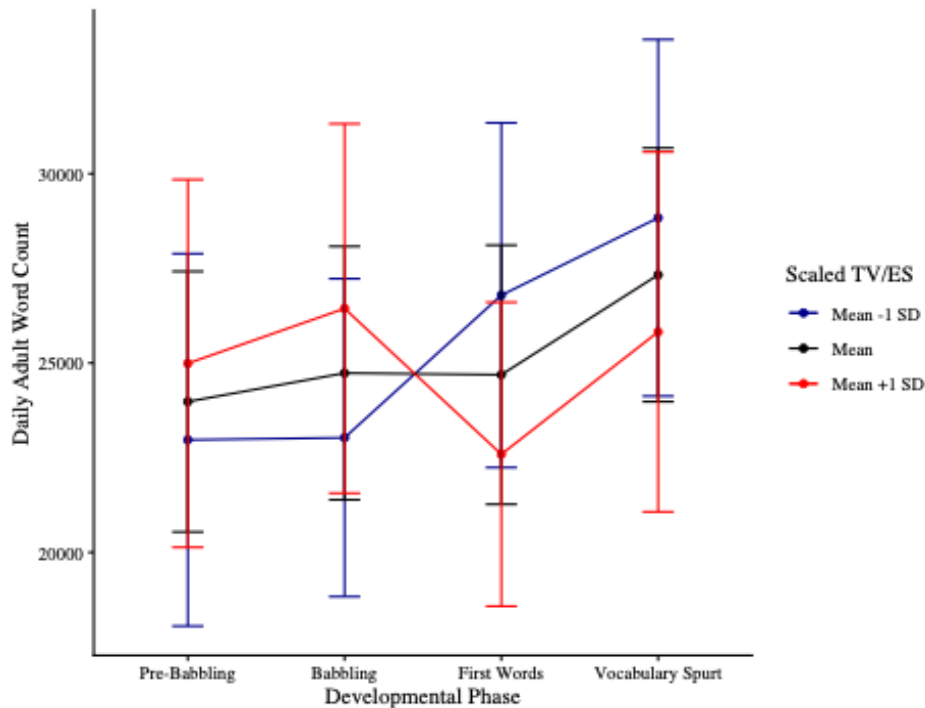


Figure 14: AWC across phases of development when scaled TV/ES exposure (in minutes) is included as an interaction term.

The first words phase ($b = 369.25$, $SE = 166.92$, $t(89.29) = 2.21$, $p = .03$) and the vocabulary spurt phase ($b = 1090.59$, $SE = 165.25$, $t(88.96) = 6.60$, $p < .001$) were significantly different from the pre-babbling phase, with an increase of approximately 369 vocalizations in the first words phase and 1090 vocalizations in the vocabulary spurt phase, as compared with the pre-babbling phase. Further analyses using the Bonferroni adjustment for post hoc pairwise comparisons showed that the vocabulary spurt phase remained significantly different from the pre-babbling phase, and was also shown to be significantly different from the babbling and first words phase (all p 's < .001). The first words phase no longer remained significant.

Considering both the developmental phase and TV/ES exposure (Figure 16), as well as the interaction between the two predictors, developmental phase remained a significant predictor in the first words ($b = 403.16$, $SE = 158.76$, $t(86.05) = 2.54$, $p = .02$) and vocabulary spurt phases ($b = 1082.22$, $SE = 155.98$, $t(85.40) = 6.94$, $p < .001$). The interaction between developmental phase and the scaled TV/ES exposure was also significant ($b = -526.94$, $SE = 191.03$, $t(98.51) = -2.76$, p

To understand if CVC changed over the course of development, or was affected by the amount of TV/ES exposure, two models were run. The first included CVC as the outcome variable, and developmental phase as a single predictor (Figure 15).

The first words phase ($b = 369.25$, $SE =$

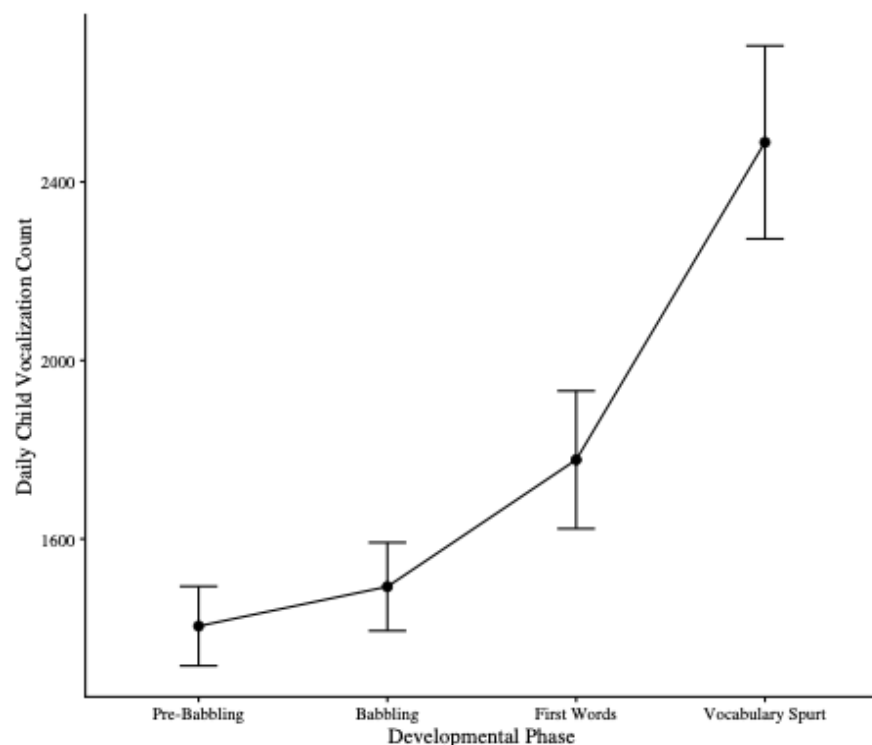


Figure 15: CVC across phases of development.

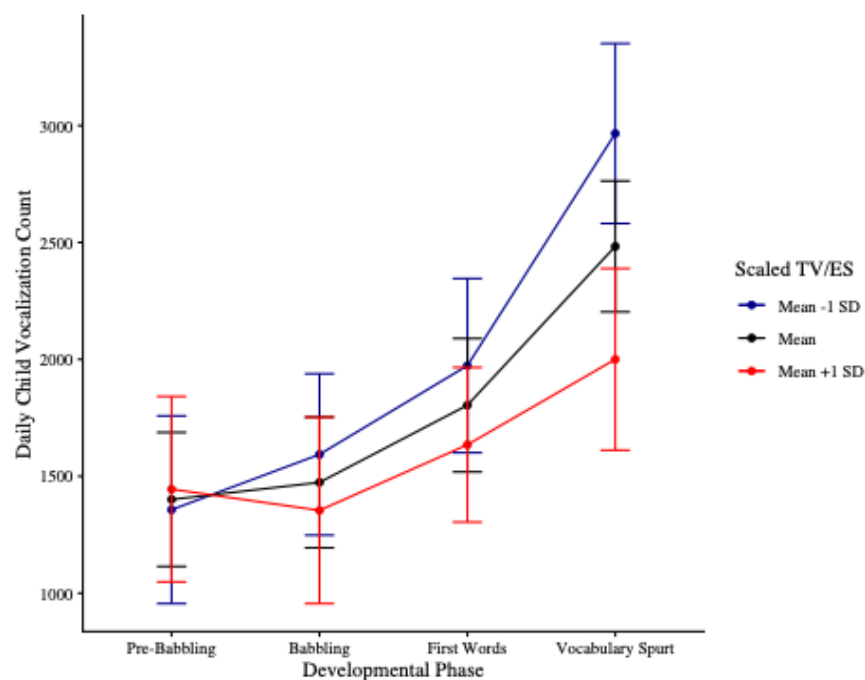


Figure 16: CVC across phases of development when scaled TV/ES exposure (in minutes) is included as an interaction term.

= .007), indicating that higher scaled TV/ES exposure was associated with a smaller increase in child vocalizations in the vocabulary spurt phase, relative to the pre-babbling phase. Upon further investigation using pairwise comparisons, with scaled TV/ES exposure, the vocabulary spurt phase remained significantly different from the pre-babbling, babbling, and first words phases (all p 's <.001). Model comparison indicated that the interaction provided a significantly better fit than the developmental-phase-only model, ($\chi^2(4) = 16.73$, $p = .002$) with a moderately lower AIC value ($\Delta AIC = 8.8$).

To understand if CVC also varied with AWC, CVC was used as the outcome variable for the model, and a scaled variable of AWC

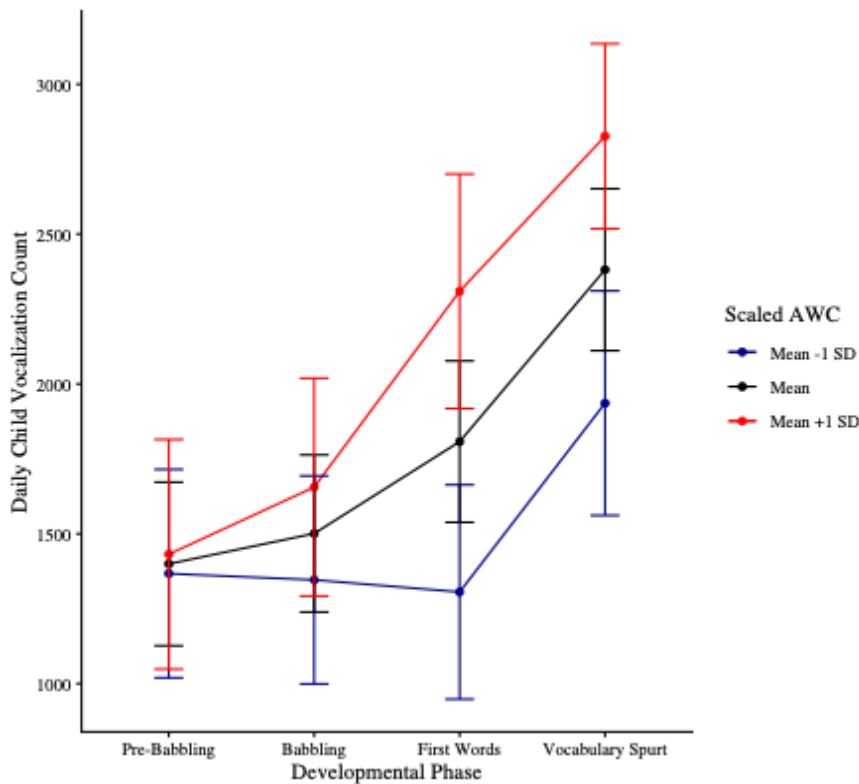


Figure 17: AWC across phases of development when scaled AWC is included as an interaction term.

was included as a predictor variable for the model, alongside developmental phase (Figure 17). There was a significant difference between the first words phase ($b = 408.11$, $SE = 152.11$, $t(85.92) = 2.68$, $p = .009$) and the vocabulary spurt phase ($b = 981.87$, $SE = 152.98$, $t(86.64) = 6.42$, $p < .001$), when compared to the pre-babbling phase. There was also a significant

difference in the interaction between the scaled AWC and first words phase ($b = 469.45$, $SE = 167.97$, $t(93.27) = 2.80$, $p = .006$) as well as the vocabulary spurt phase ($b = 413.12$, $SE = 156.76$, $t(95.55) = 2.64$, $p = .01$). Upon further investigation using pairwise comparisons, with scaled AWC, the vocabulary spurt phase remained significantly different from the pre-babbling phase, as well as the babbling and first words phases (all p 's $< .001$). Model comparison using the AIC index and the r-squared index to assess "goodness of fit" indicated that the model with the scaled AWC as a predictor provided a better overall fit for the data ($AIC = 1857.70$). This model explained 35% of the total variance that can be attributed to the fixed effects ($R^2 = .35$) and 62% of the variance when both the random and fixed effects were considered ($R^2 = .62$).

Discussion

The present exploratory study examined media exposure in a longitudinal sample of Italian-speaking infants. Ecologically valid recordings of infants' auditory environments were collected using LENA technology through the four main phases of language development: the pre-

babbling, babbling, first words, and vocabulary spurt phases. The primary goals of the analyses were to validate the use of LENA to measure TV/ES in an Italian sample, understand the acoustic environment during a typical day for an Italian infant, and to understand the longitudinal associations between TV/ES and infant language development.

Aim 1: Validating the Use of the LENA Device to Measure TV/ES Exposure

Our findings demonstrated that, when compared with a human coder, the LENA device has a weak to moderate association in its accuracy of recording TV/ES exposure. Compared to previous validation studies of the TV/ES variable (Christakis et al., 2009; Zimmerman et al., 2009), our study demonstrated relatively low correlations between the human coder and the LENA device. For example, fidelity matrices calculated between the human-coder-reported and LENA-reported TV/ES variable achieved 70.5%, indicating a high level of concordance between LENA and the human coder in previous studies (Christakis et al., 2009; Zimmerman et al., 2009).

One explanation for the weaker correlations seen in the Italian-speaking sample, as compared to previous samples of English-speaking infants, is likely due to the methodology in which the human coder performed the analysis. The human coder recorded, for an hour-long segment, whether or not there was technology present for a full or partial minute, assuming that all partial minutes were equivalent to 30 seconds of technology use, and that all full minutes had no overlapping sounds, adult speech, or child vocalizations in the foreground, that may have changed which sound was recorded by LENA. Each hour of recording was then transformed into a collection of 5-minute segments (as opposed to the coded 1-minute segment) to be comparable to the LENA device, which may have systematically over or under inflated the amount of technology present within the segment. Previous studies have comprised a fidelity matrix, to categorize TV/ES, AWC, and CVC, as well as "other", to consider the potential for overlap between AWC or CVC and TV/ES, and understand the mis-categorization of the TV/ES variable (Christakis et al., 2009; Zimmerman et al., 2009).

Aim 2: Understanding the Types of Sounds that Italian Infants Hear in a Typical Day

This study added to the literature highlighting the need for more encompassing descriptions of technology use and exposure, and analyses that include a more in-depth understanding of the types of devices that infants are exposed to (see a review by Mannell et al., 2024). Phones or

computers and television were logged by the human coder as being in the infant's auditory environment for 23.23% of the time that TV/ES was present. More interestingly, of the electronic devices recorded in the hours analyzed, the most often heard form of TV/ES was music, for 41.68% of the time. Although neither LENA nor the human coder were able to determine whether the music was coming from a radio, music-streaming application, or music video-streaming platform, it is well-established that music can promote development, learning, health, and well-being. A literature review from Houde and Narendran (2018) highlighted how music, such as the use of nursery rhymes, may promote early language acquisition in infants, through mechanisms of joint engagement, reciprocal interactions, and phonological processing. Based on these results, it is imperative that future research investigate the acoustic environment of the child, to better understand the source of electronic sounds as well as the use of the television, cell phones, and computers, as not all sources or types of use may impact development in the same way.

In terms of the amount of TV/ES exposure, Italian infants were exposed to an overall average of 44.94 minutes a day, with 45.45 minutes in the pre-babbling phase, 37.32 minutes in the babbling phase, 53.01 minutes in the first words phase, and 44.52 minutes in the vocabulary spurt phase. The amount of exposure time was on the lower end of technology use and media exposure when compared to studies on American infants (e.g., Mann et al., 2025; Tomopoulos et al., 2010; Zimmerman et al., 2007b) and other Italian samples (e.g., Balbinot et al., 2016; Operto et al., 2020). Importantly, regardless of amount of exposure, the LENA-reported TV/ES does not meet the guidelines set out by various organizations highlighting that infants under the age of two should have no exposure to electronic devices at all (Ponti, 2023; World Health Organization, 2019).

When hour-by-hour data of the average TV/ES, AWC, and CVC counts were plotted on a bar graph, there appeared to be prominent peaks and depressions in the TV/ES, AWC, and CVC within each phase of language development. In early language development, in the pre-babbling phase, there was an obvious peak of TV/ES in the early afternoon, followed by an obvious depression around 16:00 (4:00 p.m.). These times may be connected to an increase in TV/ES around lunchtime, which for Italian families is often between 13:00 (1:00 p.m.) and 14:30 (2:30 p.m.; Jhavar, 2019) while the immediate decline in TV/ES may be associated with other afternoon activities or naptimes. For the first words phase, there were clear peaks for TV/ES in the morning, early afternoon, and in the evening, and multiple peaks, but no obvious patterns in the vocabulary

spurt phase, implying that TV/ES is more present throughout the day as the infants develop. Importantly, as these data were averaged across infants, comparisons of the LENA reported variable of silence, or diaries for parent-reported variables of electronic use, daily habits, and daily naps, would be required to make any further conclusions.

Of note, correlations of these data revealed that higher amounts of TV/ES exposure was negatively associated with adult speech and child vocalizations throughout the day across all phases of development, except for TV/ES and child vocalizations in the vocabulary spurt phase. This is consistent with previous reports of digital device use (Kirkorian et al., 2009; Supanitayanon et al., 2020) and TV/ES exposure (Brushe et al., 2023; Ramírez et al., 2021; Zimmerman et al., 2009). It is well established in the literature that higher amounts of adult speech are associated with higher levels of infant receptive vocabulary (Rowe 2012), productive vocabulary (Hart et al., 1977), vocabulary growth (Huttenlocher et al., 1991; Weisleder & Fernald, 2013), and processing efficiency (Weisleder & Fernald, 2013), specifically when that speech is directed towards, and not just overheard by the infants. It is also established that infant vocalizations, including pre-linguistic vocalizations such as the cry and coo, as well as babbling, are imperative for the development of the vocal tract (Gillis et al., 2016; Stoel-Gammon et al., 2024) and help the infant to better explore their vocal capabilities (Gillis et al., 2016). Higher levels of TV/ES exposure may lead to interruption of the natural processes of infant language development, effecting both the infant themselves as well as the critical aspects of the acoustic environment, such as adult speech. These data also demonstrated that there were positive associations with adult speech and child vocalizations, which has shown to help buffer the negative effects of technology (Supanitayanon et al., 2020; Zimmerman et al., 2009).

While these relationships described above were statistically significant, many of the relationships were not particularly strong in relation to the variable of TV/ES exposure, with the strongest correlations being -0.18 between TV/ES exposure and adult speech in the pre-babbling phase and TV/ES exposure and child vocalizations in the babbling phase, and the highest correlation overall being 0.55 between adult speech and child vocalizations in the vocabulary spurt phase. This may be a result of the high degrees of freedom involved in these calculations, and further investigation into these relationships, possibly using Bayesian statistical modelling, may be required.

Aim 3: Understanding TV/ES Exposure Over the Course of Infant Language Development

The linear mixed effects models used to explore the relationships between TV/ES exposure, child vocalizations, and adult speech, had varying results (Figure 18).

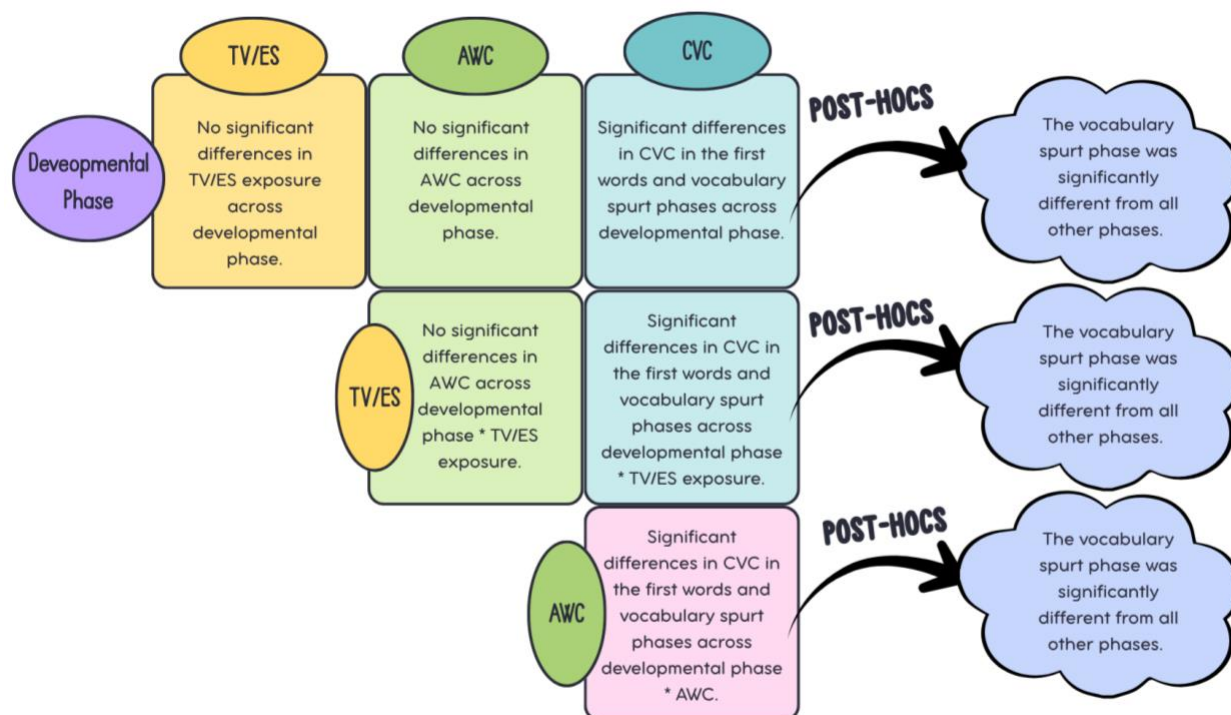


Figure 18: Summary of findings from the linear mixed effects models (Aim 3) and planned post-hoc comparisons.

The first model, exploring the trends TV/ES exposure across the phases of development with this Italian-speaking sample of infants, did not demonstrate significance. This indicated that, on average, TV/ES exposure does not differ between stages of language development, and may have further implications for the interpretations of the results. The effects of exposure of moderate to high TV/ES may be cumulative, such that daily exposure will have little to no effect on development initially, but have more severe effects in late infancy and/or throughout childhood and adolescence. The implications for future research are that each study should seek to understand the technology use and screen media habits of their specific sample, as there is high variability across studies, in order to draw accurate conclusions about technology use and language development. This finding was contrary to previous research that has shown positive relationships between screen media exposure and development, whereby the amount of exposure increased across the first year of development (Krogh et al., 2021). The findings, however, are in line with other research showing no viewing trends in TV/ES over the first two years of development (Ramírez et al., 2021).

The second model explored the trends in adult speech over the course of development, and how it may be affected by TV/ES. Neither the model with only developmental phase as a predictor nor the model with a developmental phase by TV/ES interaction were significant in our sample. This finding was contrary to previous research demonstrating that the presence of TV/ES is associated with a decrease in the quantity (Kirkorian et al., 2009; Masur et al., 2016; Mendelsohn et al., 2008) and quality (Corkin et al., 2021; Kirkorian et al., 2009; Masur et al., 2016) of adult speech. Associations may not have been found as a result of the LENA device having a lower predictive ability for adult speech in Italian samples (Bastianello et al., 2023), or a result of the infants in our sample potentially having a high exposure to music as opposed to the television, cell phones, or other digital devices compared to that reported in other studies.

The third model explored the trends in child vocalizations over the course of development, and how they may be affected by TV/ES. The models revealed that the first words phase and the vocabulary spurt phase differed significantly from the pre-babbling phase. However, when post hoc comparisons were conducted, it was shown that only the vocabulary spurt phase remained significantly different, and was significantly different from all other phases of infant language development. This is to be expected from previous research that has documented the existence of a period of rapid growth in language development (Goldfield & Reznick, 1990; Mervis & Bertrand, 1995). The models also revealed that, when considering the amount of TV/ES exposure, the first words phase and the vocabulary spurt phase differed significantly from the pre-babbling phase. However, when post hoc comparisons were conducted, it was shown that only the vocabulary spurt phase remained significantly different, and was significantly different from all other phases of infant language development. Infants exposed to higher levels of TV/ES had smaller increases in child vocalizations in the vocabulary spurt phase. This is in line with previous studies using LENA to measure child vocalizations (e.g., Ramírez et al., 2021). Infants as young as four months begin to play with various sounds, exploring the capabilities of their developing vocal tract (Stark, 1980). The auditory feedback they receive from these processes is critical to the development of language, as infants will try to reproduce the sounds they just made, that may or may not come from their native language (Gillis et al., 2016). The auditory intrusion from TV/ES may disrupt or prevent this process early on, and the long-term effects on the physiological or anatomical development required for language production are unknown. Although it has been

demonstrated that infants that do not go on to babble canonically before their first birthday, demonstrate reduced expressive vocabulary at later ages (Oller, 1999, as cited in Lang et al., 2019).

The fourth and final model explored how child vocalizations may be affected by adult speech. The models revealed that the first words phase and the vocabulary spurt phase differed significantly from the pre-babbling phase. However, when post hoc comparisons were conducted, it was shown that only the vocabulary spurt phase remained significantly different, and was significantly different from all other phases of infant language development. Infants exposed to higher levels of adult speech produced more vocalizations in the vocabulary spurt phase. Much of the literature exploring the role of the adult, or the role of adult speech, in infant development, has demonstrated that infant-direct speech, but not adult-directed speech, has a positive effect on infant language development (e.g., Weisleder & Fernald, 2013). Other research has demonstrated that while TV/ES may have a negative effect on language development, this effect may be mitigated with co-viewing (Corkin et al., 2021). As the LENA device does not differentiate between the two types of speech, and catalogues all speech coming from an adult within a given vicinity of the infant, it is difficult to draw any further conclusions based on this data. Furthermore, the model with adult speech was the strongest predictor of child vocalizations, as compared to the model with TV/ES and the model with only developmental phase as predictors. Further analyses on these data are needed to determine the potential mediating effects of adult speech in the presence of TV/ES on infant language development.

Limitations

Some limitations of the present study include the small, high-SES sample size, the limitations of the linear mixed-effects models, and the limitations of the LENA device. First, discussed in previous chapters, there may be differential effects from technology use and media exposure on language development (Ramírez et al., 2021), and the average amount of time spent using devices may systematically vary between different levels of SES (Masur et al., 2016). There is already evidence demonstrating that infants in low-SES households receive less linguistic input as compared to infants from mid- or high-SES backgrounds (Hart et al., 1977; Rowe, 2018). The combination of a lower baseline of input paired with higher average technology use may more strongly impact infants from low-SES backgrounds, and will need to be considered in future

studies, and studies considering potential interventions or guidelines for families with young infants.

Second, the linear mixed effects models used in the third aim of the present study demonstrated convergence. However, visual analyses of the residuals, alongside the QQ-plots, even after log-conversions and variable scaling, indicated that some of the models may be over- or under-dispersed. Furthermore, both of the models investigating child vocalizations violated assumptions of the DHARMA simulated residuals, the model with TV/ES as a predictor violated the assumption of homogeneity and the model with AWC as a predictor deviated from the predicted quantiles. In both cases, no other assumptions were violated, and so the models were acceptable to be used. However, future studies should consider using other modelling, such as Bayesian statistical modelling, to account for the fact that the data is not completely linear, and to ensure the effects are not being over-inflated. With a better statistical model for the data, it may also be possible, then, to look more closely at the mediating effects of adult speech, through the use of the AWC or CTC variables produced by LENA, which may help to explain more of and model better the variability in our data.

Finally, the use of the LENA device, while providing a more ecologically-valid measure of an infant's typical auditory environment, is unable to differentiate between types of technology use and media exposure. We are unable to determine if electronic devices are being used by the infant or by the adult, if it is on in the background, and whether or not the infant is attending to it. Furthermore, we are unable to identify the different sources of the electronic sounds being overheard by the infant directly from the data that LENA provides. As referenced earlier in the thesis, electronic sounds, such as music, may not be as harmful to the infant's language development.

Conclusion

The present study contributes to the ongoing literature on the impacts of technology use and screen media on infant language exposure. This study used an Italian-speaking sample to grow the understanding of how technology use impacts infants in non-Anglo-speaking regions. As this was the first study to use LENA on this sample to measure the impact of technology (TV/ES), this

study also provided a validation for the use of the LENA device in the sample, paving the way for research to implement this tool in Italian-speaking infants in future projects.

The exploration of the acoustic environment of an Italian-speaking infant provided novel insights into a typical day of these infants, highlighting periods of time within each phase of development where AWC, CVC, or TV/ES may be high, and times with which they may be low. These insights can lend themselves to interventions aiming to reduce infant exposure to TV/ES or increase adult speech in the home environment. Finally, the methodology used provided a unique view on language development, measuring the phase of language development directly, as opposed to approximating the phase of development based on the age of the infant. The results from this methodology demonstrated that the vocabulary spurt phase is when differences in infant linguistic vocalizations are most likely to be seen. These findings are critical for both future research and clinical interventions, as the age at which an infant may enter into this phase will vary. However, as neither TV/ES nor AWC varied significantly across the phases of development, this suggests that the effects of these variables may be cumulative, and interventions may need to be implemented at earlier ages, when smaller or more subtle differences are beginning to appear.

References

- Albin, D. D., & Echols, C. H. (1996). Stressed and word-final syllables in infant-directed speech. *Infant Behavior and Development*, 19(4), 401–418. [https://doi.org/10.1016/S0163-6383\(96\)90002-8](https://doi.org/10.1016/S0163-6383(96)90002-8)
- Alroqi, H., Serratrice, L., & Cameron-Faulkner, T. (2022). The association between screen media quantity, content, and context and language development. *Journal of Child Language*, 50(5), 1155–1183. <https://doi.org/10.1017/S0305000922000265>
- Alvarez Gutierrez, S., & Ventura, A. K. (2021). Associations between maternal technology use, perceptions of infant temperament, and indicators of mother-to-infant attachment quality. *Early Human Development*, 154, 105305. <https://doi.org/10.1016/j.earlhumdev.2021.105305>
- American Psychological Association. (2007). *Report of the APA Task Force on Socioeconomic Status*. American Psychological Association. <https://www.apa.org/pi/ses/resources/publications/task-force-2006.pdf>
- American Psychological Association. (2018). *Deferred imitation*. *APA Dictionary of Psychology*. <https://dictionary.apa.org/deferred-imitation>
- American Psychological Association. (2018). *Recall bias*. *APA Dictionary of Psychology*. <https://dictionary.apa.org/recall-bias>
- American Psychological Association. (2018). *Social desirability*. *APA Dictionary of Psychology*. <https://dictionary.apa.org/social-desirability>
- American Psychological Association. (2023, November 15). *Socioeconomic status (SES)*. *APA Dictionary of Psychology*. <https://dictionary.apa.org/socioeconomic-status>
- Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development*, 92(2), 484–501. <https://doi.org/10.1111/cdev.13508>
- Arabiat, D., Al Jabery, M., Robinson, S., Whitehead, L., & Mörelius, E. (2023). Interactive technology use and child development: A systematic review. *Child: Care, Health and Development*, 49(4), 679–715. <https://doi.org/10.1111/cch.13082>
- Aronoff, M., & Fudeman, K. (2022). Thinking about morphology and morphological analysis. In *What is morphology?* (3rd edition, p. 31). John Wiley & Sons, Ltd. https://media.wiley.com/product_data/excerpt/02/11197152/1119715202-15.pdf

- Asikainen, M., Kylliäinen, A., Mäkelä, T. E., Saarenpää-Heikkilä, O., & Paavonen, E. J. (2021). Exposure to electronic media was negatively associated with speech and language development at 18 and 24 months. *Acta Paediatrica*, 110(11), 3046–3053. <https://doi.org/10.1111/apa.16021>
- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, 8, 456. <https://doi.org/10.3389/fpsyg.2017.00456>
- Bakht, D., Yousaf, F., Alvi, Z., Buhadur Ali, M. K., Hadeed Khawar, M. M., Munir, L., Hussain Bokhari, S. F., Qureshi, M. S., Raza, M., & Qureshi, A. A. (2025). Assessing the impact of screen time on the motor development of children: A systematic review. *Pediatric Discovery*, 3(2), e70002. <https://doi.org/10.1002/pdi3.70002>
- Balbinot, V., Toffol, G., & Tamburlini, G. (2016). Digital technologies and young children: A survey on their use in the first years of life. *Medico e Bambino*, 35, 631–636.
- Barnett, L. A. (1990). Developmental benefits of play for children. *Journal of Leisure Research*, 22(2), 138–153. <https://doi.org/10.1080/00222216.1990.11969821>
- Barr, R. (2013). Memory constraints on infant learning from picture books, television, and touchscreens. *Child Development Perspectives*, 7(4), 205–210. <https://doi.org/10.1111/cdep.12041>
- Barr, R., Brito, N., & Simcock, G. (2013). Revisiting the effect of reminders on infants' media memories: Does the encoding format matter? *Developmental Psychology*, 49(11), 2112–2119. <https://doi.org/10.1037/a0031759>
- Bastianello, T., Lorenzini, I., Nazzi, T., & Majorano, M. (2023). The Language ENvironment Analysis system (LENA): A validation study with Italian-learning children. *Journal of Child Language*, 51(5), 1172–1192. <https://doi.org/10.1017/S0305000923000326>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bedford, R., Saez De Urabain, I. R., Cheung, C. H. M., Karmiloff-Smith, A., & Smith, T. J. (2016). Toddlers' fine motor milestone achievement is associated with early touchscreen scrolling. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01108>
- Beynon, A., Hendry, D., Lund Rasmussen, C., Rohl, A. L., Eynon, R., Thomas, G., Stearne, S., Campbell, A., Harris, C., Zabatiero, J., & Straker, L. (2024). Measurement method

- options to investigate digital screen technology use by children and adolescents: A narrative review. *Children*, 11(7), 754. <https://doi.org/10.3390/children11070754>
- Bloom, K., D'Odorico, L., & Beaumont, S. (1993). Adult preferences for syllabic vocalizations: Generalizations to parity and native language. *Infant Behavior and Development*, 16(1), 109–120. [https://doi.org/10.1016/0163-6383\(93\)80032-4](https://doi.org/10.1016/0163-6383(93)80032-4)
- Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, 14(2), 211–227. <https://doi.org/10.1017/S0305000900012897>
- Bornstein, M. H., Putnick, D. L., Lansford, J. E., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Uribe Tirado, L. M., Zelli, A., Alampay, L. P., Al-Hassan, S. M., Bacchini, D., Bombi, A. S., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Malone, P. S., & Oburu, P. (2015). Mother and father socially desirable responding in nine countries: Two kinds of agreement and relations to parenting self-reports: Agreement in social desirability. *International Journal of Psychology*, 50(3), 174–185. <https://doi.org/10.1002/ijop.12084>
- Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C.-S., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, 44(3), 867–874. <https://doi.org/10.1037/0012-1649.44.3.867>
- Bozzola, E., Irrera, M., Barni, S., Caruso, C., Leccese, B., Franzese, E., Fioretti, L., Bernardelli, L., Scarpato, E., Cupertino, V., Mazzone, T., Russo, R., Benevento, A., Strappato, B., Cervellini, M., Ferrara, P., & Agostiniani, R. (2026). Digital media exposure and pediatric health: The recommendations from the Italian Society of Pediatrics Digital Dependency Commission. *Italian Journal of Pediatrics*, 52(1), 30. <https://doi.org/10.1186/s13052-026-02198-6>
- Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8(6), 535–543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>
- Bruner, J. (1985). The role of interaction formats in language acquisition. In J. P. Forgas (Ed.), *Language and Social Situations* (pp. 31–46). Springer. https://doi.org/10.1007/978-1-4612-5074-6_2

- Bruni, O., Baumgartner, E., Sette, S., Ancona, M., Caso, G., Di Cosimo, M. E., Mannini, A., Ometto, M., Pasquini, A., Ulliana, A., & Ferri, R. (2014). Longitudinal study of sleep behavior in normal infants during the first year of life. *Journal of Clinical Sleep Medicine*, 10(10), 1119–1127. <https://doi.org/10.5664/jcsm.4114>
- Brushe, M. E., Lynch, J. W., Melhuish, E., Reilly, S., Mittinty, M. N., & Brinkman, S. A. (2023). Objectively measured infant and toddler screen time: Findings from a prospective study. *SSM - Population Health*, 22, 101395. <https://doi.org/10.1016/j.ssmph.2023.101395>
- Cameron-Faulkner, T., Malik, N., Steele, C., Coretta, S., Serratrice, L., & Lieven, E. (2021). A cross-cultural analysis of early prelinguistic gesture development and its relationship to language development. *Child Development*, 92(1), 273–290. <https://doi.org/10.1111/cdev.13406>
- Caselli, M. C., Bello, A., Pasqualetti, P., Rinaldi, P., & Stefanini, S. (2015). *Il primo vocabolario del bambino: Gesti, parole, e frasi. Forme lunghe e forme brevi del questionario e valori di riferimento per la fascia 8-36 mesi.*
- Çetinçelik, M., Rowland, C. F., & Snijders, T. M. (2021). Do the eyes have it? A systematic review on the role of eye gaze in infant language development. *Frontiers in Psychology*, 11, 589096. <https://doi.org/10.3389/fpsyg.2020.589096>
- Chomsky, N. (2015). *The minimalist program* (20th Anniversary Edition). The MIT Press.
- Chonchaiya, W., & Pruksananonda, C. (2008). Television viewing associates with delayed language development. *Acta Paediatrica*, 97(7), 977–982. <https://doi.org/10.1111/j.1651-2227.2008.00831.x>
- Christakis, D. A., Gilkerson, J., Richards, J. A., Zimmerman, F. J., Garrison, M. M., Xu, D., Gray, S., & Yipanel, U. (2009). Audible television and decreased adult words, infant vocalizations, and conversational turns: A population-based study. *Archives of Pediatrics & Adolescent Medicine*, 163(6), 554. <https://doi.org/10.1001/archpediatrics.2009.61>
- Cicourel, A. V. (1982). Constructivism versus innatism in theories of language and learning. *Contemporary Sociology*, 11(5), 526. <https://doi.org/10.2307/2068396>
- Cobern, W. W. (1993). Constructivism. *Journal of Educational and Psychological Consultation*, 4(1), 105–112. https://doi.org/10.1207/s1532768xjepc0401_8

- Cooper, R. P., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development*, 65(6), 1663.
<https://doi.org/10.2307/1131286>
- Corkin, M. T., Henderson, A. M. E., Peterson, E. R., Kennedy- Costantini, S., Sharplin, H. S., & Morrison, S. (2021). Associations between technofence, quality of parent-infant interactions, and infants' vocabulary development. *Infant Behavior and Development*, 64, 101611. <https://doi.org/10.1016/j.infbeh.2021.101611>
- Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, 7(1), 114–133.
<https://doi.org/10.1038/s41562-022-01452-1>
- Cristia, A., & Seidl, A. (2015). Parental reports on touch screen use in early childhood. *PLOS ONE*, 10(6), e0128338. <https://doi.org/10.1371/journal.pone.0128338>
- Crystal, D. (2018). *The Cambridge encyclopedia of the English language* (Third edition). Cambridge University Press. <https://doi.org/10.1017/9781108528931>
- Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, 92, 81–104.
- Dabrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00852>
- Dandurand, F., & Shultz, T. R. (2011). A fresh look at vocabulary spurts. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 33.
<https://escholarship.org/uc/item/36f265q7>
- De Boysson-Bardies, B. D., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, 67(2), 297–319.
<https://doi.org/10.2307/415108>
- De Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16(1), 1–17.
<https://doi.org/10.1017/S0305000900013404>
- De Marco, A., & Vernon-Feagans, L. (2013). Rural Neighborhood Context, Child Care Quality, and Relationship to Early Language Development. *Early Education & Development*, 24(6), 792–812. <https://doi.org/10.1080/10409289.2013.736036>

- De Schuymer, L., De Groote, I., Beyers, W., Striano, T., & Roeyers, H. (2011). Preverbal skills as mediators for language outcome in preterm and full term children. *Early Human Development*, 87(4), 265–272. <https://doi.org/10.1016/j.earlhumdev.2011.01.029>
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208(4448), 1174–1176. <https://doi.org/10.1126/science.7375928>
- Der Nederlanden, S. J., Schaeffer, J. C., Van Bakel, H. H. J. A., & Dirks, E. (2025). Socio-economic status and other potential risk factors for language development in the first year of life. *Journal of Child Language*, 52(1), 42–62. <https://doi.org/10.1017/S0305000923000478>
- D'Odorico, L., & Jacob, V. (2006). Prosodic and lexical aspects of maternal linguistic input to late-talking toddlers. *International Journal of Language & Communication Disorders*, 41(3), 293–311. <https://doi.org/10.1080/13682820500342976>
- Duch, H., Fisher, E. M., Ensari, I., & Harrington, A. (2013). Screen time use in children under 3 years old: A systematic review of correlates. *International Journal of Behavioral Nutrition and Physical Activity*, 10(1), 102. <https://doi.org/10.1186/1479-5868-10-102>
- Durham, K., Wethmar, D., Brandstetter, S., Seelbach-Göbel, B., Apfelbacher, C., Melter, M., Kabesch, M., Kerzel, S., & The KUNO Kids Study Group. (2021). Digital media exposure and predictors for screen time in 12-month-old-children: A cross-sectional analysis of data from a German birth cohort. *Frontiers in Psychiatry*, 12, 737178. <https://doi.org/10.3389/fpsyt.2021.737178>
- Elardo, R., Bradley, R., & Calwell, B. M. (1977). A longitudinal study of the relation of infants' home environments to language development at age three. *Child Development*, 48(2), 595–603.
- Erickson, L. C., & Newman, R. S. (2017). Influences of background noise on infants and children. *Current Directions in Psychological Science*, 26(5), 451–457. <https://doi.org/10.1177/0963721417709087>
- Ervin-Tripp, S. M. (1966). Language development. In L. W. Hoffman & M. L. Hoffman (Eds.), *Review of Child Development Research*. Russell Sage Foundation.
- Evans, G. W., Maxwell, L. E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology*, 35(4), 1020–1023. <https://doi.org/10.1037/0012-1649.35.4.1020>

- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, 99(14), 9602–9605. <https://doi.org/10.1073/pnas.152159999>
- Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, 16(8), 1320–1326. <https://doi.org/10.1162/0898929042304787>
- Fernald, A., & Marchman, V. A. (2006). Language learning in infancy. In M. J. Traxlor & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2nd ed., pp. 1027–1071). Elsevier.
- Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology*, 40(4), 621–632. <https://doi.org/10.1037/0012-1649.40.4.621>
- Garg, R. (2021). “Its changes so often”: Parental non-/use of mobile devices while caring for infants and toddlers at home. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–26. <https://doi.org/10.1145/3479513>
- Gervain, J. (2015). Plasticity in early language acquisition: The effects of prenatal and early childhood experience. *Current Opinion in Neurobiology*, 35, 13–20. <https://doi.org/10.1016/j.conb.2015.05.004>
- Gervain, J. (2018). The role of prenatal experience in language development. *Current Opinion in Behavioral Sciences*, 21, 62–67. <https://doi.org/10.1016/j.cobeha.2018.02.004>
- Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough Oller, D., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, 26(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169
- Gilkerson, J., & Richards, J. A. (2020). *A guide to understanding the design and purpose of the LENA® system (LTR-12)*. LENA Foundation. https://www.lena.org/wp-content/uploads/2020/07/LTR-12_How_LENA_Works.pdf
- Gillis, S., Koopmans-van Beinum, F., & van der Stelt, J. (2016). The prelinguistic stage. In P. Jordens & J. A. Lalleman (Eds.), *Language Development* (Reprint 2016). De Gruyter Mouton. <https://doi.org/10.1515/9783110854817>

- Gleason, J. B., & Philips, B. C. (2024). The development of language: An overview and a preview. In J. B. Gleason & N. B. Ratner (Eds.), *The development of language* (Tenth edition, pp. 209–254). Plural Publishing, Inc.
- Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, 17(1), 171–183.
<https://doi.org/10.1017/S0305000900013167>
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19(5), 515–523.
<https://doi.org/10.1111/j.1467-9280.2008.02117.x>
- Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development*, 58(6), 1523.
<https://doi.org/10.2307/1130692>
- Green, J. R., Nip, I. S. B., Wilson, E. M., Mefferd, A. S., & Yunusova, Y. (2010). Lip movement exaggerations during infant-directed speech. *Journal of Speech, Language, and Hearing Research*, 53(6), 1529–1542. [https://doi.org/10.1044/1092-4388\(2010/09-0005\)](https://doi.org/10.1044/1092-4388(2010/09-0005))
- Greenhoe, D. J. (2017). *The Sounds of Mandarin Chinese*.
<https://doi.org/10.13140/RG.2.2.20184.80643>
- Grimshaw, G. M., Adelstein, A., Bryden, M. P., & MacKinnon, G. E. (1998). First-language acquisition in adolescence: Evidence for a critical period of verbal language development. *Brain and Language*, 63(2), 237–255.
<https://doi.org/10.1006/brln.1997.1943>
- Guellai, B., Somogyi, E., Esseily, R., & Chopin, A. (2022). Effects of screen exposure on young children's cognitive development: A review. *Frontiers in Psychology*, 13, 923370.
<https://doi.org/10.3389/fpsyg.2022.923370>
- Hamamouche, K., & Bohannon III, J. N. (2024). Theoretical approaches to language acquisition. In J. B. Gleason & N. B. Ratner (Eds.), *The development of language* (Tenth edition, pp. 209–254). Plural Publishing, Inc.
- Hart, B., Risley, T. R., & Kirby, J. R. (1997). [Meaningful differences in the everyday experience of young american children]. *Canadian Journal of Education*, 22(3).
<https://www.proquest.com/scholarly-journals/meaningful-differences-everyday-experience-young/docview/215377356/se-2>

- Hartig, F. (2024). *DHARMa: Residual diagnostics for hierarchical (multi-Level / mixed) regression models* (R package version 0.4.7). <https://CRAN.R-project.org/package=DHARMa>
- Havron, N., Lovcevic, I., Kee, M. Z. L., Chen, H., Chong, Y. S., Daniel, M., Broekman, B. F. P., & Tsuji, S. (2022). The effect of older sibling, postnatal maternal stress, and household factors on language development in two- to four-year-old children. *Developmental Psychology*, 58(11), 2096–2113. <https://doi.org/10.1037/dev0001417>
- Heckhausen, J. (1987). Balancing for weaknesses and challenging developmental potential: A longitudinal study of mother-infant dyads in apprenticeship interactions. *Developmental Psychology*, 23(6), 762–770. <https://doi.org/10.1037/0012-1649.23.6.762>
- Hoff, E. (2006). Language Experience and Language Milestones During Early Childhood. In K. McCartney & D. Phillips (Eds.), *Blackwell Handbook of Early Childhood Development* (Vol. 26, pp. 233–251). <https://doi.org/10.1002/9780470757703>
- Hoff, E. (2014). *Language development* (Fifth edition). Wadsworth Cengage Learning.
- Hoff-Ginsberg, E., & Shatz, M. (1982). Linguistic input and the child's acquisition of language. *Psychological Bulletin*, 92(1). <https://doi.org/10.1037/0033-2909.92.1.3>
- Holland, S. K., Vannest, J., Mecoli, M., Jacola, L. M., Tillema, J.-M., Karunanayaka, P. R., Schmithorst, V. J., Yuan, W., Plante, E., & Byars, A. W. (2007). Functional MRI of language lateralization during development in children. *International Journal of Audiology*, 46(9), 533–551. <https://doi.org/10.1080/14992020701448994>
- Hood, R., Zabatiero, J., Zubrick, S. R., Silva, D., & Straker, L. (2021). The association of mobile touch screen device use with parent-child attachment: A systematic review. *Ergonomics*, 64(12), 1606–1622. <https://doi.org/10.1080/00140139.2021.1948617>
- Houde, M., & Narendran, N. (2018). A literature review of the influence of early childhood music education and music therapy on childhood development. [Une revue de la littérature consacrée à l'influence de l'éducation musicale et de la musicothérapie durant la petite enfance sur le développement de l'enfant]. *Canadian Journal of Music Therapy*, 24, 27–39.
- Hsu, H., & Fogel, A. (2001). Infant vocal development in a dynamic mother-infant communication system. *Infancy*, 2(1), 87–109. https://doi.org/10.1207/S15327078IN0201_6

- Hsu, H.-C., Fogel, A., & Cooper, R. B. (2000). Infant Vocal Development during the First 6 Months: Speech Quality and Melodic Complexity. *Infant and Child Development*, 9, 1–16.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27(2), 236–248.
<https://doi.org/10.1037/0012-1649.27.2.236>
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, 16(5), 367–371. <https://doi.org/10.1111/j.0956-7976.2005.01542.x>
- Jhawar, M. (2019, May 15). Eat like an Italian. *Italy Beyond the Obvious*.
<https://italybeyondtheobvious.com/eat-like-an-italian/>
- Johnson, E. K. (2016). Constructing a proto-lexicon: An integrative view of infant language development. *Annual Review of Linguistics*, 2(1), 391–412.
<https://doi.org/10.1146/annurev-linguistics-011415-040616>
- Jusienė, R., Urbonas, V., Laurinaitytė, I., Rakickienė, L., Breidokienė, R., Kuzminskaitė, M., & Praninskienė, R. (2019). Screen use during meals among young children: Exploration of associated variables. *Medicina*, 55(10), 688. <https://doi.org/10.3390/medicina55100688>
- Kahn, M., Barnett, N., Glazer, A., & Gradisar, M. (2021). Sleep and screen exposure across the beginning of life: Deciphering the links using big-data analytics. *Sleep*, 44(3), zsaa158.
<https://doi.org/10.1093/sleep/zsaa158>
- Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America*, 72(2), 353–365.
<https://doi.org/10.1121/1.388089>
- Kerai, S., Almas, A., Guhn, M., Forer, B., & Oberle, E. (2022). Screen time and developmental health: Results from an early childhood study in Canada. *BMC Public Health*, 22(1), 310.
<https://doi.org/10.1186/s12889-022-12701-3>
- Kildare, C. A., & Middlemiss, W. (2017). Impact of parents mobile device use on parent-child interaction: A literature review. *Computers in Human Behavior*, 75, 579–593.
<https://doi.org/10.1016/j.chb.2017.06.003>

- Kirkorian, H. L., Pempek, T. A., Murphy, L. A., Schmidt, M. E., & Anderson, D. R. (2009). The impact of background television on parent–child interaction. *Child Development*, 80(5), 1350–1359. <https://doi.org/10.1111/j.1467-8624.2009.01337.x>
- Koopmans-van Beinum, F. J., Clement, C. J., & Van Den Dikkenberg-Pot, I. (2001). Babbling and the lack of auditory speech perception: A matter of coordination? *Developmental Science*, 4(1), 61–70. <https://doi.org/10.1111/1467-7687.00149>
- Kostyrka-Allchorne, K., Stoilova, M., Bourgaize, J., Rahali, M., Livingstone, S., & Sonuga-Barke, E. (2023). Review: Digital experiences and their impact on the lives of adolescents with pre-existing anxiety, depression, eating and nonsuicidal self-injury conditions – a systematic review. *Child and Adolescent Mental Health*, 28(1), 22–32. <https://doi.org/10.1111/camh.12619>
- Krogh, M. T., Egmoose, I., Stuart, A. C., Madsen, E. B., Haase, T. W., & Væver, M. S. (2021). A longitudinal examination of daily amounts of screen time and technoferece in infants aged 2–11 months and associations with maternal sociodemographic factors. *Infant Behavior and Development*, 63, 101543. <https://doi.org/10.1016/j.infbeh.2021.101543>
- Kuhl, P. K. (2011). Brain mechanisms underlying the critical period for language: Linking theory and practice. In A. M. Battro, S. Dehaene, & W. J. Singer (Eds.), *Human Neuroplasticity and Education*. The Pontifical Academy of Sciences.
- Kymissis, E., & Poulson, C. L. (1990). The history of imitation in learning theory: The language acquisition process. *Journal of the Experimental Analysis of Behavior*, 54(2), 113–127. <https://doi.org/10.1901/jeab.1990.54-113>
- Laakso, M.-L., Poikkeus, A.-M., Katajamäki, J., & Lyytinen, P. (1999). Early intentional communication as a predictor of language development in young toddlers. *First Language*, 19(56), 207–231. <https://doi.org/10.1177/014272379901905604>
- Laasko, M., Helasvuo, M.-L., & Savinainen-Makkonen, T. (2010). Children’s early actions in learning language: A study of proto-words and pointing gestures in interaction between one-year-old child and parent. *SKY Journal of Linguistics*, 23, 199–226.
- Lang, S., Bartl-Pokorny, K. D., Pokorny, F. B., Garrido, D., Mani, N., Fox-Boyer, A. V., Zhang, D., & Marschik, P. B. (2019). Canonical Babbling: A Marker for Earlier Identification of Late Detected Developmental Disorders? *Current Developmental Disorders Reports*, 6(3), 111–118. <https://doi.org/10.1007/s40474-019-00166-w>

- Lantolf, J. P. (2000). Introducing sociocultural theory. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 1–26). Oxford University Press.
- Lapointe, V. R., Ford, L., & Zumbo, B. D. (2007). Examining the Relationship Between Neighborhood Environment and School Readiness for Kindergarten Children. *Early Education & Development, 18*(3), 473–495. <https://doi.org/10.1080/10409280701610846>
- LeBlanc, A. G., Spence, J. C., Carson, V., Connor Gorber, S., Dillman, C., Janssen, I., Kho, M. E., Stearns, J. A., Timmons, B. W., & Tremblay, M. S. (2012). Systematic review of sedentary behaviour and health indicators in the early years (aged 0–4 years). *Applied Physiology, Nutrition, and Metabolism, 37*(4), 753–772. <https://doi.org/10.1139/h2012-063>
- Lee, C.-C., Jhang, Y., Relyea, G., Chen, L., & Oller, D. K. (2018). Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings. *Infant Behavior and Development, 50*, 140–153. <https://doi.org/10.1016/j.infbeh.2017.12.002>
- Lenth, R. V. (2026). *emmeans: Estimated marginal means, aka least-squares means*. (R package version 2.0.3). <https://CRAN.R-project.org/package=emmeans>
- Levey, S. (2024). *Introduction to language development* (Third edition). Plural Publishing, Inc.
- Lifter, K., & Bloom, L. (1989). Object knowledge and the emergence of language. *Infant Behavior and Development, 12*(4), 395–423. [https://doi.org/10.1016/0163-6383\(89\)90023-4](https://doi.org/10.1016/0163-6383(89)90023-4)
- Lin, L.-Y., Cherng, R.-J., Chen, Y.-J., Chen, Y.-J., & Yang, H.-M. (2015). Effects of television exposure on developmental skills among young children. *Infant Behavior and Development, 38*, 20–26. <https://doi.org/10.1016/j.infbeh.2014.12.005>
- Linder, L. K., McDaniel, B. T., Stockdale, L., & Coyne, S. M. (2021). The impact of parent and child media use on early parent–infant attachment. *Infancy, 26*(4), 551–569. <https://doi.org/10.1111/infa.12400>
- MacNeilage, P. F., Davis, B. L., & Matyear, C. L. (1997). Babbling and first words: Phonetic similarities and differences. *Speech Communication, 22*(2–3), 269–277. [https://doi.org/10.1016/S0167-6393\(97\)00022-8](https://doi.org/10.1016/S0167-6393(97)00022-8)
- Majorano, M., & D’Odorico, L. (2011). The transition into ambient language: A longitudinal study of babbling and first word production of Italian children. *First Language, 31*(1), 47–66. <https://doi.org/10.1177/0142723709359239>

- Mallawaarachchi, S., Burley, J., Mavilidi, M., Howard, S. J., Straker, L., Kervin, L., Staton, S., Hayes, N., Machell, A., Torjinski, M., Brady, B., Thomas, G., Horwood, S., White, S. L. J., Zabatiero, J., Rivera, C., & Cliff, D. (2024). Early childhood screen use contexts and cognitive and psychosocial outcomes: A systematic review and meta-analysis. *JAMA Pediatrics*, 178(10), 1017. <https://doi.org/10.1001/jamapediatrics.2024.2620>
- Mann, S., Calvin, A., Lenhart, A., & Rob, M. B. (2025). The common sense census: Media use by kids zero to eight. *Common Sense Media*.
<https://www.common sense media.org/research/the-2025-common-sense-census-media-use-by-kids-zero-to-eight>
- Mannell, K., Bloul, S., Sefton-Green, J., & Willcox, M. (2024). Digital media and technology use by families with infants, toddlers, and young children: A scoping review and call for forward momentum. *Journal of Children and Media*, 18(4), 605–628.
<https://doi.org/10.1080/17482798.2024.2394939>
- Martin, A., Igarashi, Y., Jincho, N., & Mazuka, R. (2016). Utterances in infant-directed speech are shorter, not slower. *Cognition*, 156, 52–59.
<https://doi.org/10.1016/j.cognition.2016.07.015>
- Masataka, N. (2009). The origins of language and the evolution of music: A comparative perspective. *Physics of Life Reviews*, 6(1), 11–22.
<https://doi.org/10.1016/j.plrev.2008.08.003>
- Masataka, N., & Bloom, K. (1994). Acoustic properties that determine adults' preferences for 3-month-old infant vocalizations. *Infant Behavior and Development*, 17(4), 461–464.
[https://doi.org/10.1016/0163-6383\(94\)90038-8](https://doi.org/10.1016/0163-6383(94)90038-8)
- Masur, E. F., Flynn, V., & Olson, J. (2015). The presence of background television during young children's play in American homes. *Journal of Children and Media*, 9(3), 349–367.
<https://doi.org/10.1080/17482798.2015.1056818>
- Masur, E. F., Flynn, V., & Olson, J. (2016). Infants' background television exposure during play: Negative relations to the quantity and quality of mothers' speech and infants' vocabulary acquisition. *First Language*, 36(2), 109–123. <https://doi.org/10.1177/0142723716639499>
- Mayo, C., & La France, M. (1978). On the acquisition of nonverbal communication: A review. *Merrill-Palmer Quarterly of Behaviour and Development*, 24(4), 213–228.

- McDaniel, B. T. (2015). "Technoference": Everyday intrusions and interruptions of technology in couple and family relationships. In C. J. Bruess (Ed.), *Family Communication in the age of digital and social media*. New York: Peter Lang Publishing.
- McMillan, B. T. M., & Saffran, J. R. (2016). Learning in complex environments: The effects of background speech on early word learning. *Child Development*, 87(6), 1841–1855.
<https://doi.org/10.1111/cdev.12559>
- McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, 317(5838), 631–631. <https://doi.org/10.1126/science.1144073>
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
[https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- Mendelsohn, A. L., Berkule, S. B., Tomopoulos, S., Tamis-LeMonda, C. S., Huberman, H. S., Alvir, J., & Dreyer, B. P. (2008). Infant television and video exposure associated with limited parent-child verbal interactions in low socioeconomic status households. *Archives of Pediatrics & Adolescent Medicine*, 162(5), 411.
<https://doi.org/10.1001/archpedi.162.5.411>
- Mermelshtine, R. (2017). Parent-child learning interactions: A review of the literature on scaffolding. *British Journal of Educational Psychology*, 87(2), 241–254.
<https://doi.org/10.1111/bjep.12147>
- Mervis, C. B., & Bertrand, J. (1995). Early lexical acquisition and the vocabulary spurt: A response to Goldfield & Reznick. *Journal of Child Language*, 22(2), 461–468.
<https://doi.org/10.1017/S0305000900009880>
- Messinger, D., & Fogel, A. (2007). The interactive development of social smiling. In *Advances in Child Development and Behavior* (Vol. 35, pp. 327–366). Elsevier.
<https://doi.org/10.1016/B978-0-12-009735-7.50014-1>
- Morales, M., Mundy, P., & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, 21(2), 373–377.
[https://doi.org/10.1016/S0163-6383\(98\)90014-5](https://doi.org/10.1016/S0163-6383(98)90014-5)
- Munzer, T., Parga-Belinkie, J., Milkovich, L. M., Tomopoulos, S., Ajumobi, T., Cross, C., Gerwin, R., Madigan, S., Psych, R., Council on Communications and Media, Radesky, J., Evans, Y., Cahan, E., Milkovich, E., Bracho-Sanchez, E., Parga-Belinkie, J., Munzer, T.,

- Espinoza, J., Husain, A., ... Shapiro, I. (2026). Digital ecosystems, children, and adolescents: Policy statement. *Pediatrics*, 157(2), e2025075320.
<https://doi.org/10.1542/peds.2025-075320>
- Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America*, 139(3), 1272–1281. <https://doi.org/10.1121/1.4944634>
- Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics & Phonetics*, 20(5), 351–369.
<https://doi.org/10.1080/02699200500211451>
- Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, 6(2), 136–142. <https://doi.org/10.1111/1467-7687.00263>
- Neuman, S. B., & Celano, D. (2001). Access to print in low-income and middle-income communities: An ecological study of four neighborhoods. *Reading Research Quarterly*, 36(1), 8–26. <https://doi.org/10.1598/RRQ.36.1.1>
- Newman, R. S., & Hussain, I. (2006). Changes in preference for infant-directed speech in low and moderate noise by 4.5- to 13-month-olds. *Infancy*, 10(1), 61–76.
https://doi.org/10.1207/s15327078in1001_4
- Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, i'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to Children: Language Input and Acquisition* (pp. 109–149). Cambridge University Press.
- Operto, F. F., Pastorino, G. M. G., Marciano, J., De Simone, V., Volini, A. P., Olivieri, M., Buonaiuto, R., Vetri, L., Viggiano, A., & Coppola, G. (2020). Digital device use and language skills in children between 8 and 36 month. *Brain Sciences*, 10(9), 656.
<https://doi.org/10.3390/brainsci10090656>
- Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, 3(1), 285–308. <https://doi.org/10.1146/annurev-linguistics-011516-034226>
- Paoletti, M., Focaroli, V., Pecora, G., Barr, R., Chiarotti, F., Caravale, B., Gasparini, C., Sacco, G., Gastaldi, S., Addressi, E., & Bellagamba, F. (2025). Maternal media use and infant

- media exposure are associated with developmental outcomes at 4 and 8 months. *Parenting*, 25(1), 66–101. <https://doi.org/10.1080/15295192.2024.2437156>
- Paoli, S. (2016). *A short guide to Italian Phonetics and Phonology for students of Italian paper* (V). Oxford University. https://www.ling-phil.ox.ac.uk/romance-linguistics/assets/uploads/paoli/Booklet_P&P_right.pdf
- Plunkett, K. (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language*, 20(1), 43–60. <https://doi.org/10.1017/S0305000900009119>
- Poitras, V. J., Gray, C. E., Janssen, X., Aubert, S., Carson, V., Faulkner, G., Goldfield, G. S., Reilly, J. J., Sampson, M., & Tremblay, M. S. (2017). Systematic review of the relationships between sedentary behaviour and health indicators in the early years (0–4 years). *BMC Public Health*, 17(S5), 868. <https://doi.org/10.1186/s12889-017-4849-8>
- Poliva, O. (2017). From where to what: A neuroanatomically based evolutionary model of the emergence of speech in humans. *F1000Research*, 4, 67. <https://doi.org/10.12688/f1000research.6175.3>
- Ponti, M. (2023). Screen time and preschool children: Promoting health and development in a digital world. *Paediatrics & Child Health*, 28(3), 184–192. <https://doi.org/10.1093/pch/pxac125>
- Radesky, J. S., Silverstein, M., Zuckerman, B., & Christakis, D. A. (2014). Infant self-regulation and early childhood media exposure. *Pediatrics*, 133(5), e1172–e1178. <https://doi.org/10.1542/peds.2013-2367>
- Ramírez, N., Hippe, D. S., & Shapiro, N. T. (2021). Exposure to electronic media between 6 and 24 months of age: An exploratory study. *Infant Behavior and Development*, 63, 101549. <https://doi.org/10.1016/j.infbeh.2021.101549>
- Ricci, R. C., Paulo, A. S. C. D., Freitas, A. K. P. B. D., Ribeiro, I. C., Pires, L. S. A., Facina, M. E. L., Cabral, M. B., Parduci, N. V., Spegiorin, R. C., Bogado, S. S. G., Chociay Junior, S., Carachesti, T. N., & Larroque, M. M. (2023). Impacts of technology on children's health: A systematic review. *Revista Paulista de Pediatria*, 41, e2020504. <https://doi.org/10.1590/1984-0462/2023/41/2020504>
- Rochette & Bernier. (2014). Parenting, family socioeconomic status, and child executive functioning: A longitudinal study. *Merrill-Palmer Quarterly*, 60(4), 431. <https://doi.org/10.13110/merrpalmquar1982.60.4.0431>

- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, 83(5), 1762–1774. <https://doi.org/10.1111/j.1467-8624.2012.01805.x>
- Rowe, M. L. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives*, 12(2), 122–127. <https://doi.org/10.1111/cdep.12271>
- Sasaki, C. T., Levine, P. A., Laitman, J. T., & Crelin, E. S. (1977). Postnatal descent of the epiglottis in man: A preliminary report. *Archives of Otolaryngology - Head and Neck Surgery*, 103(3), 169–171. <https://doi.org/10.1001/archotol.1977.00780200095011>
- Schmidt, M. E., Pempek, T. A., Kirkorian, H. L., Lund, A. F., & Anderson, D. R. (2008). The effects of background television on the toy play behavior of very young children. *Child Development*, 79(4), 1137–1151. <https://doi.org/10.1111/j.1467-8624.2008.01180.x>
- Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19(3–4), 245–253. <https://doi.org/10.1111/j.1467-1770.1969.tb00466.x>
- Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics*, 20, 213–223. <https://doi.org/10.1017/S0267190500200135>
- Setliff, A. E., & Courage, M. L. (2011). Background television and infants' allocation of their attention during toy play: Background tv and attention to toys. *Infancy*, 16(6), 611–639. <https://doi.org/10.1111/j.1532-7078.2011.00070.x>
- Shahini, A. (2025). *Language Learning Theories: A Student's Guide* (1st ed. 2025). Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-92210-7>
- Shinya, Y., Kawai, M., Niwa, F., Imafuku, M., & Myowa, M. (2017). Fundamental frequency variation of neonatal spontaneous crying predicts language acquisition in preterm and term infants. *Frontiers in Psychology*, 8, 2195. <https://doi.org/10.3389/fpsyg.2017.02195>
- Sirois, S., Spratling, M., Thomas, M. S. C., Westermann, G., Mareschal, D., & Johnson, M. H. (2008). Précis of neuroconstructivism: How the brain constructs cognition. *Behavioral and Brain Sciences*, 31(3), 321–331. <https://doi.org/10.1017/S0140525X0800407X>
- Slobodin, O., Hetzroni, O. E., Mandel, M., Saad Nuttman, S., Gawri Damashi, Z., Machluf, E., & Davidovitch, M. (2024). Infant screen media and child development: A prospective community study. *Infancy*, 29(2), 155–174. <https://doi.org/10.1111/inf.12575>
- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13(4), 410–420. <https://doi.org/10.1080/15250000802188719>

- Sosa, A. V. (2016). Association of the type of toy used during play with the quantity and quality of parent-infant communication. *JAMA Pediatrics*, 170(2), 132.
<https://doi.org/10.1001/jamapediatrics.2015.3753>
- Speer, S., & Blodgett, A. (2006). Prosody. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (2nd ed.). Elsevier.
- Sroufe, L. A., & Waters, E. (1976). The ontogenesis of smiling and laughter: A perspective on the organization of development in infancy. *Psychological Review*, 83(3), 173–189.
<https://doi.org/10.1037/0033-295X.83.3.173>
- Sroufe, L. A., & Wunsch, J. P. (1972). The development of laughter in the first year of life. *Child Development*, 43(4), 1326. <https://doi.org/10.2307/1127519>
- Stangor, C., & Walinga, J. (2014). *Introduction to Psychology—1st Canadian Edition*.
<https://opentextbc.ca/introductiontopsychology/>
- Stark, R. E. (1980). Stages of speech development in the first year of life. In *Child Phonology* (pp. 73–92). Elsevier. <https://doi.org/10.1016/B978-0-12-770601-6.50010-3>
- Stockdale, L. A., Porter, C. L., Coyne, S. M., Essig, L. W., Booth, M., Keenan-Kroff, S., & Schvaneveldt, E. (2020). Infants’ response to a mobile phone modified still-face paradigm: Links to maternal behaviors and beliefs regarding technofence. *Infancy*, 25(5), 571–592. <https://doi.org/10.1111/inf.12342>
- Stoel-Gammon, C., Velleman, S., & Menn, L. (2024). Phonological development: Learning sounds and sound patterns. In J. B. Gleason & N. B. Ratner (Eds.), *The development of language* (Tenth edition, pp. 209–254). Plural Publishing, Inc.
- Suarez-Rivera, C., Fletcher, K. K., & Tamis-LeMonda, C. S. (2024). Infants’ home auditory environment: Background sounds shape language interactions. *Developmental Psychology*, 60(12), 2274–2289. <https://doi.org/10.1037/dev0001762>
- Sundqvist, A., Barr, R., Heimann, M., Birberg-Thornberg, U., & Koch, F. (2023). A longitudinal study of the relationship between children’s exposure to screen media and vocabulary development. *Acta Paediatrica*, 113(3), 517–522. <https://doi.org/10.1111/apa.17047>
- Supanitayanon, S., Trairatvorakul, P., & Chonchaiya, W. (2020). Screen media exposure in the first 2 years of life and preschool cognitive development: A longitudinal study. *Pediatric Research*, 88(6), 894–902. <https://doi.org/10.1038/s41390-020-0831-8>

- Szaflarski, J. P., Holland, S. K., Schmithorst, V. J., & Byars, A. W. (2006). fMRI study of language lateralization in children and adults. *Human Brain Mapping, 27*(3), 202–212. <https://doi.org/10.1002/hbm.20177>
- Tamis-LeMonda, C. S., Bornstein, M. H., Baumwell, L., & Damast, A. M. (1996). Responsiveness parenting in the second year: Specific influences on children's language and play. *Early Development and Parenting, 5*(4), 173–183. [https://doi.org/10.1002/\(SICI\)1099-0917\(199612\)5:4%3C173::AID-EDP131%3E3.0.CO;2-V](https://doi.org/10.1002/(SICI)1099-0917(199612)5:4%3C173::AID-EDP131%3E3.0.CO;2-V)
- Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Association for Psychological Science, 23*(2), 121–126. <https://doi.org/10.1177/0963721414522813>
- Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language, 42*(6), 1173–1190. <https://doi.org/10.1017/S0305000914000725>
- Tey, C., Wake, M., Campbell, M., Hampton, A., & Williams, J. (2007). The light time-use diary and preschool activity patterns: Exploratory study. *International Journal of Pediatric Obesity, 2*(3), 167–173. <https://doi.org/10.1080/17477160701369274>
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5
- Thompson, A., & Steinbeis, N. (2020). Sensitive periods in executive function development. *Current Opinion in Behavioral Sciences, 36*, 98–105. <https://doi.org/10.1016/j.cobeha.2020.08.001>
- Tomopoulos, S., Dreyer, B. P., Berkule, S., Fierman, A. H., Brockmeyer, C., & Mendelsohn, A. L. (2010). Infant media exposure and toddler development. *Archives of Pediatrics & Adolescent Medicine, 164*(12). <https://doi.org/10.1001/archpediatrics.2010.235>
- Topping, K., Dekhinet, R., & Zeedyk, S. (2013). Parent–infant interaction and children's language development. *Educational Psychology, 33*(4), 391–426. <https://doi.org/10.1080/01443410.2012.744159>
- Tranel, B. (1987). *The sounds of French: An introduction* (Transferred to digital reprinting). Cambridge Univ. Pr.

- Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, 17(1), 1–13. [https://doi.org/10.1016/S0002-7138\(09\)62273-1](https://doi.org/10.1016/S0002-7138(09)62273-1)
- Valentine, C. W. (1938). A study of the beginnings and significance of play in infancy. *British Journal of Educational Psychology*, 8(3), 285–306. <https://doi.org/10.1111/j.2044-8279.1938.tb03133.x>
- Vecera, S. P., & Johnson, M. H. (1995). Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition*, 2(1), 59–87. <https://doi.org/10.1080/13506289508401722>
- Vihman, M. M., Ferguson, C. A., & Elbert, M. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*, 7(1), 3–40. <https://doi.org/10.1017/S0142716400007165>
- Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Ed.). MIT Press.
- Walle, E. A., & Campos, J. J. (2014). Infant language development is related to the acquisition of walking. *Developmental Psychology*, 50(2), 336–348. <https://doi.org/10.1037/a0033238>
- Walter-Laager, C., Brandenburg, K., Tinguely, L., Schwarz, J., Pfiffner, M. R., & Moschner, B. (2017). Media-assisted language learning for young children: Effects of a word-learning app on the vocabulary acquisition of two-year-olds. *British Journal of Educational Technology*, 48(4), 1062–1072. <https://doi.org/10.1111/bjet.12472>
- Wang, Y., Seidl, A., & Cristia, A. (2015). Acoustic-phonetic differences between infant- and adult-directed speech: The role of stress and utterance position. *Journal of Child Language*, 42(4), 821–842. <https://doi.org/10.1017/S0305000914000439>
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143–2152. <https://doi.org/10.1177/0956797613488145>
- Werker, J. F. (1995). Exploring developmental changes in cross-language speech perception. In L. R. Gleitman & M. Liberman (Eds.), *Language: An Invitation to Cognitive Science* (2nd ed., Vol. 1, pp. 87–106). MIT Press.

- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49–63. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3)
- Wermke, K., Leising, D., & Stellzig-Eisenhauer, A. (2007). Relation of melody complexity in infants' cries to language outcome in the second year of life: A longitudinal study. *Clinical Linguistics & Phonetics*, 21(11–12), 961–973. <https://doi.org/10.1080/02699200701659243>
- Wertsch, J. V. (1979). From social interaction to higher psychological processes. A clarification and application of Vygotsky's theory. *Human Development*, 22(1), 1–22. <https://doi.org/10.1159/000272425>
- Westermann, G., & Miranda, E. R. (2004). A new model of sensorimotor coupling in the development of speech. *Brain and Language*, 89(2), 393–400. [https://doi.org/10.1016/S0093-934X\(03\)00345-6](https://doi.org/10.1016/S0093-934X(03)00345-6)
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://doi.org/10.1007/978-3-319-24277-4>
- World Health Organization. (2019). *Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age*. World Health Organization. <https://apps.who.int/iris/handle/10665/311664>
- Wright, S. (2008). Language, communication, and the public sphere: Definitions. In K. Knapp (Ed.), *Handbook of Communication in the Public Sphere* (Paperback ed). Mouton de Gruyter.
- Xie, H., Peng, J., Qin, M., Huang, X., Tian, F., & Zhou, Z. (2018). Can touchscreen devices be used to facilitate young children's learning? A meta-analysis of touchscreen learning effect. *Frontiers in Psychology*, 9, 2580. <https://doi.org/10.3389/fpsyg.2018.02580>
- Xu, D., Yapanel, U., & Gray, S. (2009). *Reliability of the LENA® language environment analysis system in young children's natural home environment*. LENA Foundation. https://www.lena.org/wp-content/uploads/2016/07/LTR-05-2_Reliability.pdf
- Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007a). Associations between media viewing and language development in children under age 2 years. *The Journal of Pediatrics*, 151(4), 364–368. <https://doi.org/10.1016/j.jpeds.2007.04.071>

- Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007b). Television and DVD/video viewing in children younger than 2 years. *Archives of Pediatrics & Adolescent Medicine*, 161(5), 473. <https://doi.org/10.1001/archpedi.161.5.473>
- Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yipanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, 124(1), 342–349. <https://doi.org/10.1542/peds.2008-2267>

Appendix

Appendix A: The Checklist

Checklist items organized by phase where items 1-10 represent the pre-babbling phase, 11-25 represent the babbling phase, 26-35 represent the first word phase, and 36-40 represent the vocabulary spurt.

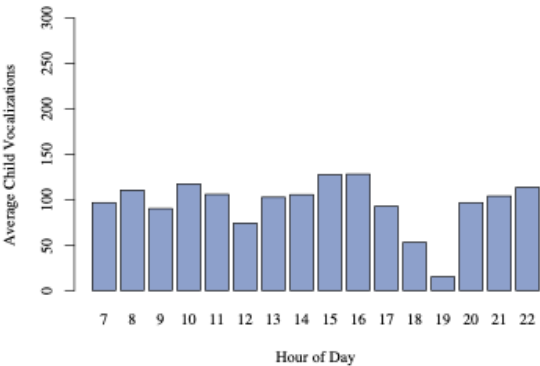
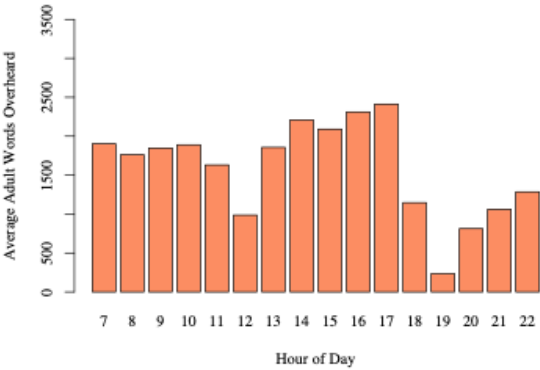
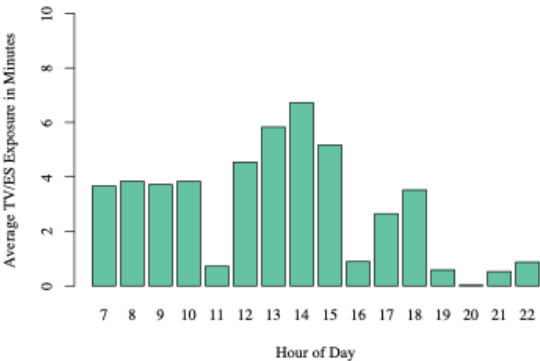
	Items
1	Does the child react when hearing sounds or noises from the surrounding environment (e.g., doorbell, telephone)?
2	Does the child react or turn when, in silence, they hear a person's voice (whether familiar or unfamiliar)?
3	Does the child recognize a familiar face (e.g., by smiling or making excited sounds)?
4	Does the child imitate facial expressions during close face-to-face interactions with an adult?
5	Does the child produce simple, guttural sounds or cooing noises?
6	Does the child produce simple vowel-like sounds (e.g., "ooh," "aah")?
7	Does the child smile spontaneously?
8	Does the child imitate simple vocalizations produced by a familiar adult?
9	Does the child extend their finger, without stretching out the arm, while looking at their hand?
10	Does the child produce the first isolated consonant sounds (e.g., m, n, p, d, t)?
11	Recognizes their own name; turns when called.
12	Produces isolated syllables (e.g., <i>pa</i> , <i>ba</i> , <i>ma</i> , <i>la</i>).
13	Produces repeated or reduplicated syllable sequences (e.g., <i>pa-pa-pa</i> , <i>ma-ma</i> , <i>la-la-la</i>).
14	Understands the meaning of some familiar, everyday objects.
15	Understands simple phrases.
16	Engages in proto-conversations with adults (characterized by some turn-taking).
17	Is able to share an object with an adult through mutual gaze.
18	Begins to use pointing gestures (with extended arm).
19	Communicates intentionally when attempting to express something.
20	Follows the adult's actions or gestures with their gaze.

21	Produces varied sequences of sounds (e.g., <i>ta-ti, ta-pa-pe</i>).
22	Uses different intonations and accents when producing sound sequences (e.g., varying rhythm or melody).
23	Uses gestures to request or express meanings.
24	Uses onomatopoeic sounds (e.g., animal or vehicle noises).
25	Understands the meaning of many words.
26	Produces specific sound sequences consistently associated with a particular referent (e.g., always saying <i>la-la</i> for a toy car).
27	Understands the word "no" when prohibited from doing something.
28	Uses conventional gestures to communicate meanings (e.g., waves goodbye, shakes head for "no").
29	Produces actual words for specific referents, even if not systematically (e.g., using the same word for the same object occasionally or only once).
30	Uses words in an overgeneralized way (e.g., <i>pappa</i> to refer to all foods) or undergeneralized way (e.g., <i>dog</i> only for the neighbor's dog).
31	Has said or used at least 10 different words.
32	Understands slightly more complex phrases and/or simple instructions (e.g., "put away the toys").
33	Is curious and uses newly learned words contextually.
34	Combines gestures with words to express meaning.
35	Uses a single word to express multiple meanings, often accompanied by gestures (e.g., says <i>doll</i> while placing hand on cheek and closing eyes to indicate the doll is sleeping; says <i>pappa</i> while circling a finger on their cheek to indicate food is tasty).
36	Says and uses more than 50 different words (mainly referring to objects/nouns).
37	Understands even very complex sentences.
38	Combines two words to form a phrase.
39	Engages in symbolic play (e.g., pretending to cook or playing with dolls).
40	Uses "yes" and "no" appropriately.
41	Constructs multi-word sentences.
42	Can tell a story while looking at the pictures in a book.
43	Recognizes colors (correctly identifies colors when asked).

- | | |
|----|--|
| 44 | Can name colors. |
| 45 | While playing, groups objects by shape, color, or size (any one of the three). |
| 46 | Correctly uses size-related words (e.g., <i>big/many</i> , <i>small/few</i>). |
| 47 | Can answer simple questions when told a story. |
| 48 | Understands pronouns (e.g., <i>he/she</i> , <i>my/your</i>). |
| 49 | Uses plural forms correctly. |
| 50 | Uses gestures while speaking to convey additional meaning |
-

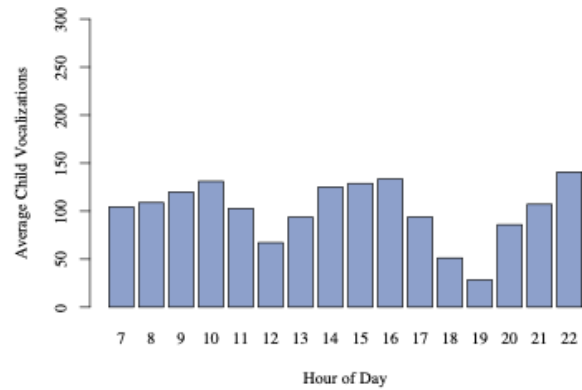
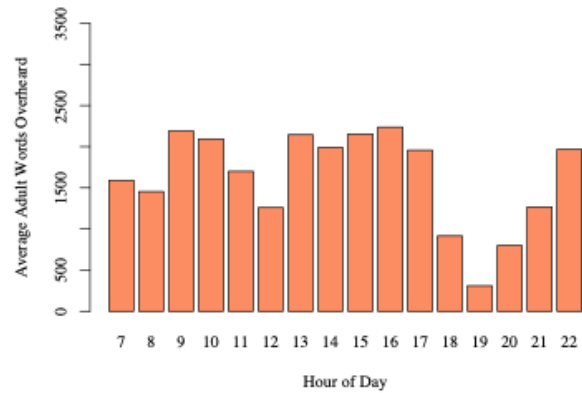
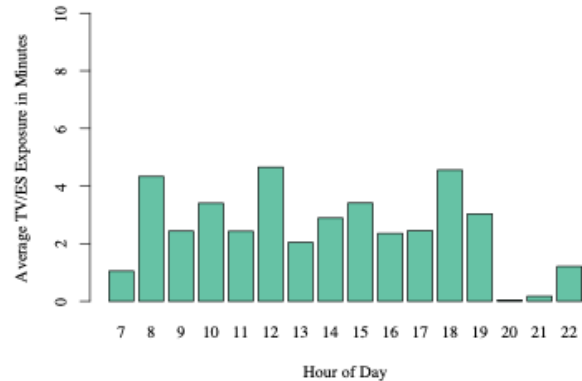
Appendix B

The Pre-Babbling Phase: Hourly Averages for TV/ES, AVW, and CVC



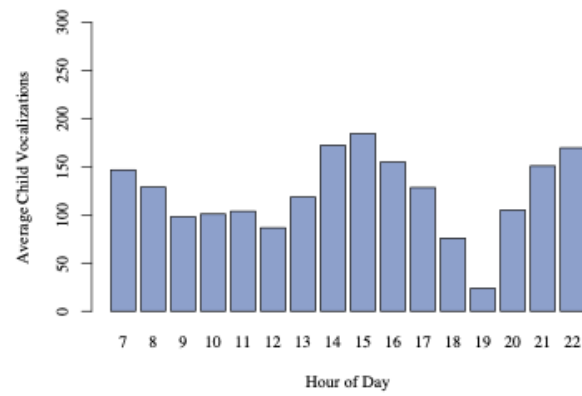
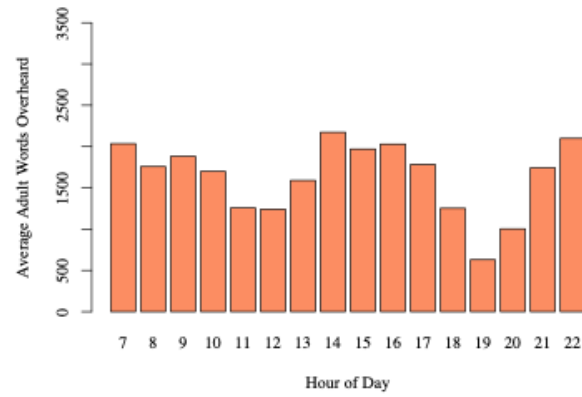
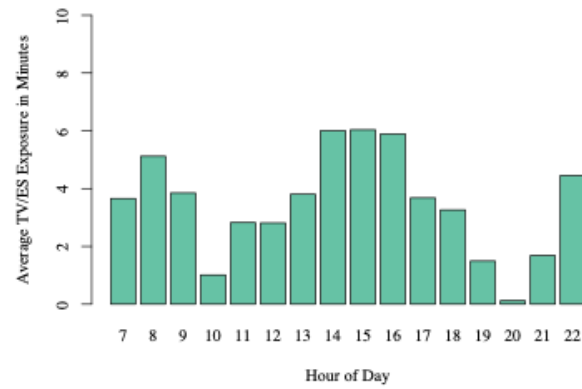
Appendix C

The Babbling Phase: Hourly Averages for TV/ES, AVW, and CVC



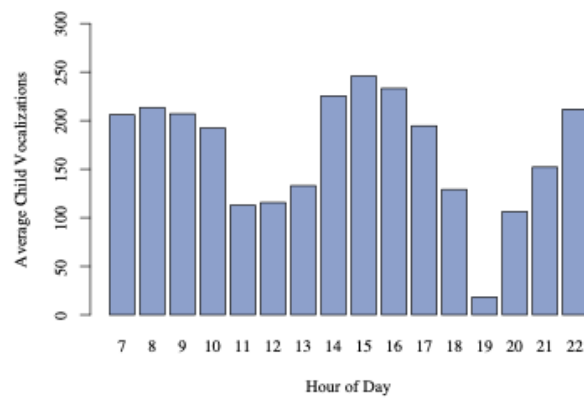
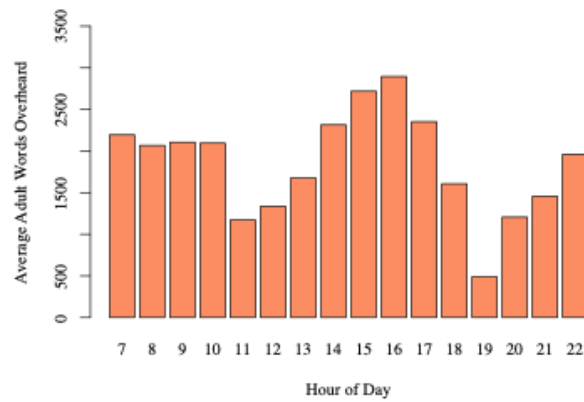
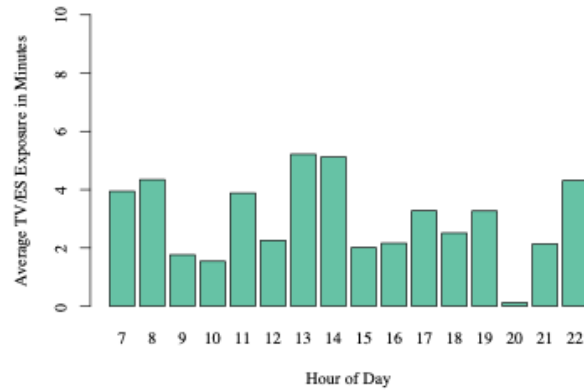
Appendix D

The First Words Phase: Hourly Averages for TV/ES, AVW, and CVC



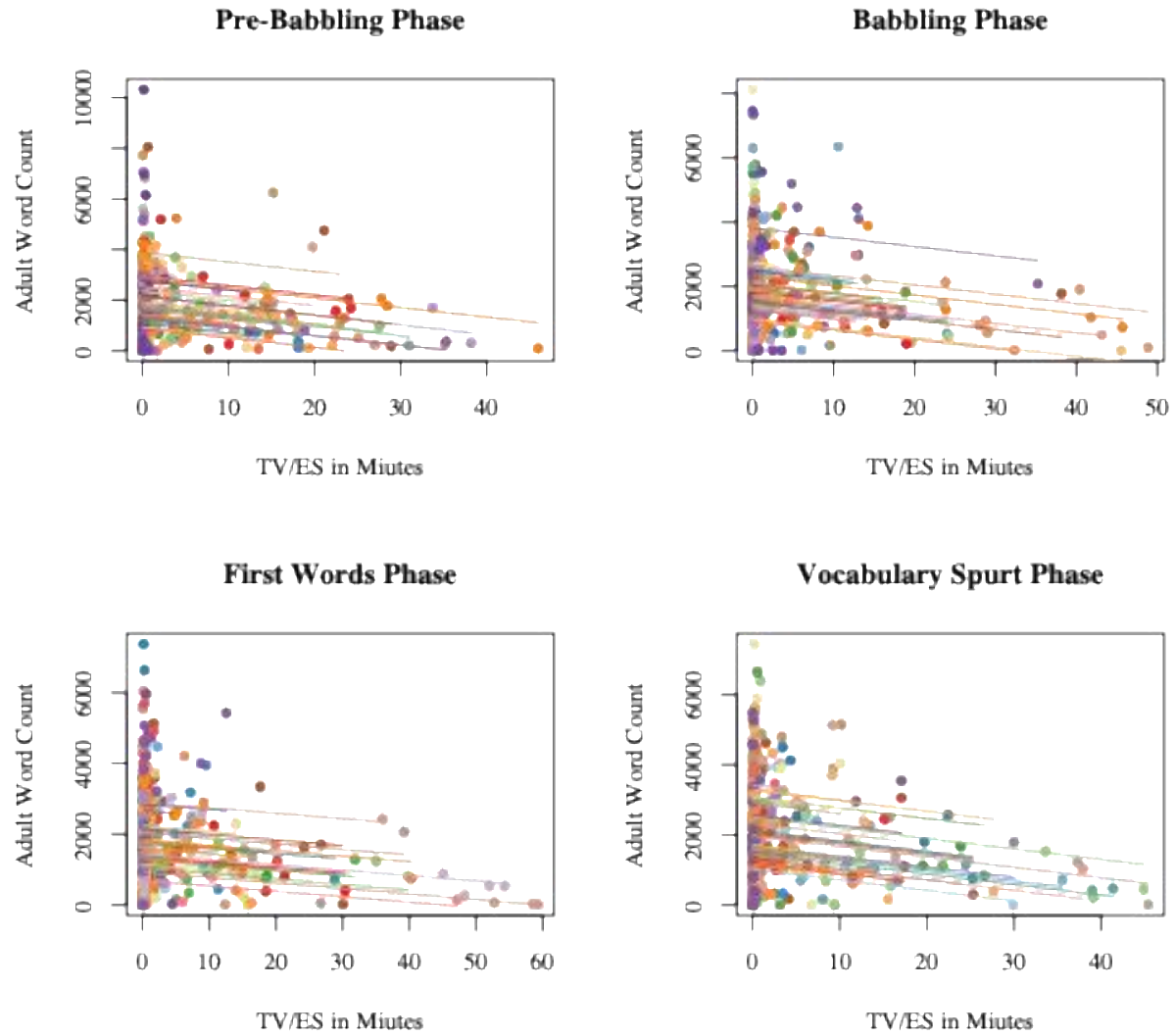
Appendix E

The Vocabulary Spurt Phase: Hourly Averages for TV/ES, AVW, and CVC



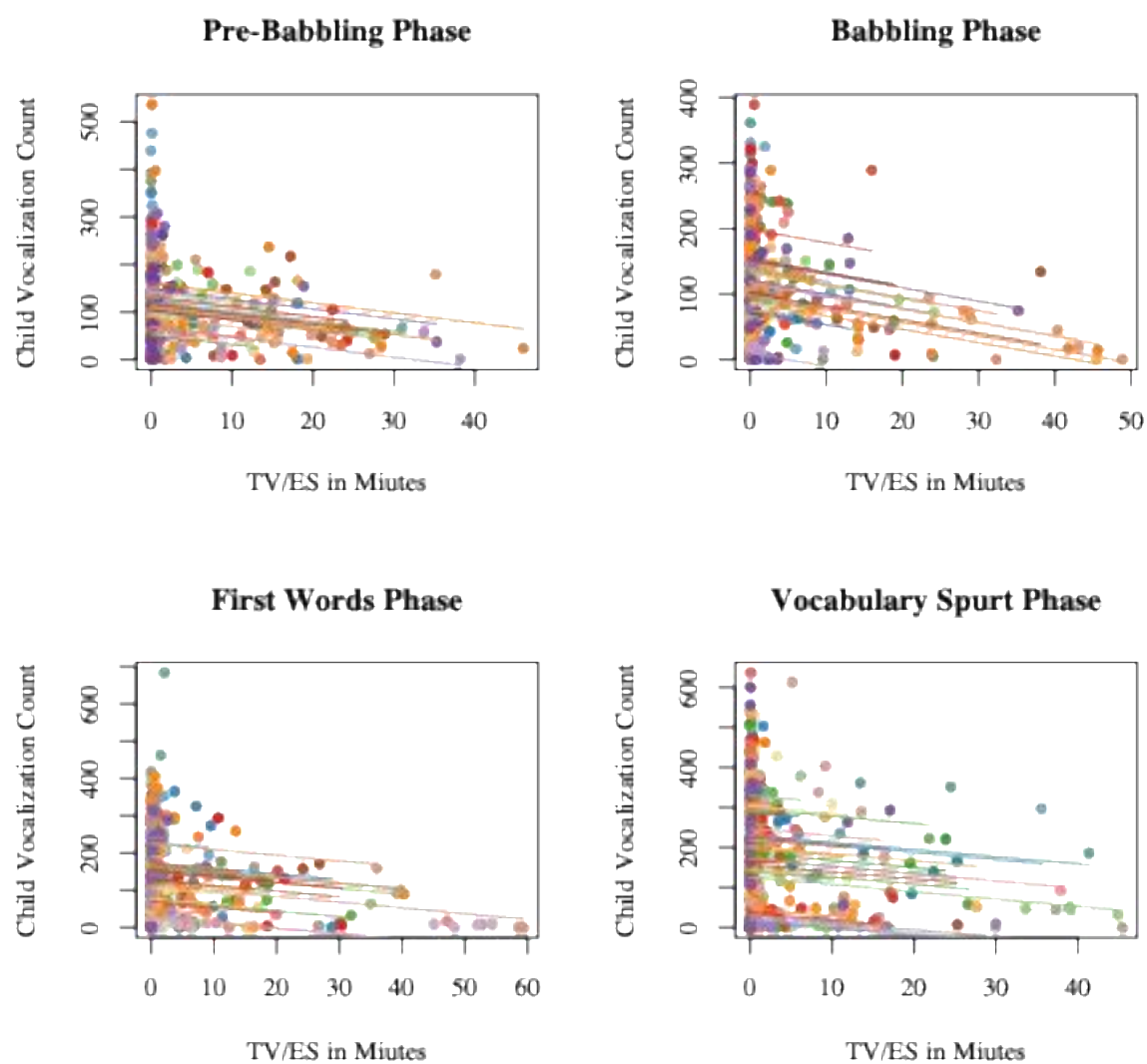
Appendix F

Repeated Measures Correlation Plots for TV/ES and AWC



Appendix G

Repeated Measures Correlation Plots for TV/ES and CVC



Appendix H

Repeated Measures Correlation Plots for AWC and CVC

