Processing Speech-in-Noise: The Effects of Children's Local and Global Familiarity with Naturalistic Background Noise

By

Brianna T. M. McMillan

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Psychology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2016

Date of final oral examination: 8/18/16

The dissertation is approved by the following members of the Final Oral Committee:

Jenny R. Saffran, Professor, Psychology
Maryellen MacDonald, Professor, Psychology
Vanessa Simmering, Assistant Professor, Psychology
Seth Pollak, Professor, Psychology
Jan R Edwards, Professor, Communication Sciences and Disorders

Table of Contents

Acknowledgements

During my graduate school journey, I have been fortunate to have an amazing group of people who helped, guided, and supported me. First and foremost, I must thank my adviser, Jenny Saffran. She has been everything I could have hoped for in an advisor, and I am forever grateful for her inviting me into her lab. Her encouragement when breaking new ground in the lab, as well as her patience and support throughout my entire graduate school career have been instrumental in helping me achieve all that I have accomplished while pursuing my Ph.D. Thank you Jenny for being such a genuinely awesome mentor – I have learned so many lessons that I will carry forward with me into my career.

I would also like to thank the members of my dissertation committee—Jan Edwards, Vanessa Simmering, Maryellen MacDonald, Seth Pollak, and Kristin Shutts—whose insight and suggestions have not only improved my dissertation, but have helped push me to become a better scientist. Additionally, I would like to thank Jan Edwards and Bianca Schroeder for their generosity in introducing me to the possibilities of LENA.

I am incredibly grateful for the members of the Infant Learning Lab. Thank you Rachel Reynders, Sarah Oakley, Kayla Gardner, and Lizzie Hoff for all of their help with recruiting, collecting data, and coding data. They are the wheels that kept my dissertation rolling. Chris Potter, Ron Pomper, Martin Zettersten, Viridiana Benitez, Erica Wojcik, and Tianlin Wang have been the best lab mates I could have asked for. Their generosity of time, skills, and ideas has been a fount for my growth through the years. I am incredibly thankful to have them as my lab mates and friends.

Finally, I would like to thank my family. My parents, Angela and Aaron, and my grandparents Larry and Vickie, have always been my biggest cheerleaders. I am grateful for all of their support and love throughout the years. To my husband, CJ, I owe a huge thank you for following me half-way across the country (and following me again the across the other half). Your encouragement pushes me to pursue my dreams, and your support makes sure I don't fall too hard when I stumble. Thank you for everything.

Abstract

How do children's day-to-day experiences with their home auditory environment affect their language processing? There are many ways to quantify the quality of a child's home auditory environment. Researchers interested in language development typically focus on the language directed towards the child. However, children live in environments that feature multiple sources of auditory information: near talkers, distant talkers, and electronic noise. The current set of studies combine observational and experimental methodologies to better understand how children's experiences with background noise affect their ability to process speech-in-noise. Recordings of children's home environments were collected, and samples of children's electronic noise were used as background noise in two experimental studies. Two studies examined whether a child's local familiarity with electronic background noise affects their ability to learn words (Study 1), and recognize familiar objects (Study 2) children heard either background noise from their own home or another child's home. In Study 1, children in both conditions were able to learn the novel words, despite the presence of background noise. In Study 2, children again showed a robust ability to overcome speech-in-noise in order to recognize familiar words. However, children who heard background noise from another child's home were better able to exploit the informativeness of the verb labeling phrase, shifting sooner and looking longer towards the correct image. In a series of exploratory analyses, individual differences in children's home environments were evaluated as possible mechanisms for explaining children's ability to process speech-in-noise. Neither the amount of overlapping speech nor the amount of electronic media children heard at home predicted children's ability to process speech-in-noise. By combining both experimental and observational methods, we have a fuller understanding of how children's home environments shapes speech-in-noise processing. More broadly, examining children's language development within a noisier auditory context allows us to examine the robustness of language development processes, and provides insights into how naturally occurring environments shape language development.

Introduction

The speech stream is filled with cues for the listener to exploit. During the first years of life, children must learn to extract information from the speech stream, using the phonological characteristics (Werker & Tees, 1984), the statistical properties (Saffran, Aslin & Newport, 1996), and the grammatical structure of their language (Fisher et al., 2010). To master their native language, children must be able to integrate these various aspects of language and in order to do so infants must attend to the relevant cues in their environment. The ability to use language is a fundamental skill that, barring any large cognitive deficits or extreme neglect, all children are able to acquire. While all children acquire language to varying proficiencies, how they acquire language may differ greatly. The environment in which a child grows up plays a large role in their language development, and variability in the environment may lead to variability in children's emergent language skills.

Parental Speech

One of the most well researched home environment variables that shape language development is the type of speech a child hears. Basil Bernstein began the investigation into how maternal speech varies as a function of social class (Bernstein, 1961), hypothesizing that children are differentially exposed to two distinct *language codes* that shape their language use. *Restricted code,* consisting of simple, shorter utterances, will often be found in informal, social speech. *Elaborated code*, consisting of structurally more complex utterances, are often more explicit, abstract, and nuanced as they serve to facilitate conversation about topics to which the listener may not have extensive background knowledge. According to Bernstein, *elaborated code* is the dominant form of communication among middle-class families, with *restricted code* being used when appropriate (e.g., in informal settings). He also theorized that working-class families rarely use *elaborated code*; instead*, restricted code* is frequently the sole mode of communication. Thus, children from middle-class families are primarily exposed to two types of language codes, and as such must learn to switch between the two codes when appropriate. Children from working-class families are exposed to only

one code type, thus not only do they not have experience with code switching, but the language input they are receiving is more sparse in nature.

To test Bernstein's hypotheses, Hess and Shipman (1965) observed Black mothers in four different socioeconomic status (SES) categories (i.e, college-educated, professional, executives; skilled-blue collar occupations with only high-school education; unskilled or semiskilled occupations with mostly elementary-school education; unskilled or semiskilled occupations with absent fathers and reliance on public assistance) interacting with their 4-year-old children during a structured task, to examine how mothers' teaching styles would differ as a function of SES, and how their use of restricted and elaborated codes would affect children's responses and behavior. Consistent with Bernstein's (1961) hypothesis, mothers in different SES groups differed in both the quantity and the quality of their language use, such that higher-SES mothers spoke more overall and used longer utterances, with more complex syntax and abstract concepts. Lower-SES mothers used more imperatives to control the child's behavior (e.g., "Be quiet!" versus "Would you keep quiet for a minute!"), and produced sentences that were overall less elaborate. Altogether, this series of findings establishes that the nature of the language input a child receives may vary as a function of their SES.

While a family's SES may impact the linguistic input, whether this actually impacts a child's language abilities was not established until Hart and Risley's (1995) seminal longitudinal study. Following 42 families with 7-month-old infants, Hart and Risley (1995) recorded interactions in the home every month for three years. By 36 months, the higher-SES children knew twice as many words as the lower-SES children. Additionally, children in higher-SES families were exposed to a greater amount of child-directed speech than children in lower-SES families, which correlated with children's vocabulary scores as well as later academic achievement (Walker, Greenwood, Hart, & Carta, 1994). Shockingly, Hart and Risley estimated that children in higher-SES families heard 30 million more words over the first 3 years of their lives than children in lower-SES families. Additionally, higher-SES mothers were more likely to use affirmatives and conversation-eliciting utterances, whereas

lower-SES mothers used more directives, consistent with observations made by Hess and Shipman (1965) three decades earlier.

Expanding Hart and Risley's (1995) findings, Huttenlocher, Waterfall, Vasilyeva, Vevea, and Hedges (2010) conducted a 32-month longitudinal study of children's syntactic development. Huttenlocher and colleagues (2010) found a similar pattern of results for children's acquisition of syntax as is found in their vocabulary development. Children from lower-SES families heard less complex syntactic utterances—frequently hearing sentences with fewer constituents (optional words or phrases, such as adverbs) and fewer clauses—than their higher-SES peers. In turn, from the age of 14-months to 46-months, children from lower-SES homes produced words and sentences that reflected the more simplistic speech of their environment. Huttenlocher and colleagues also found that a child's earlier vocabulary predicts their parents' future vocabulary, indicating a reciprocal relationship in which a parent tailors the words they use to their child's vocabulary level. Interestingly, this was not the case for syntax. Parents consistently produced the same structure of sentences, regardless of the structure of sentences produced by their child.

Language Processing

Language processing, the incremental processing of speech in order to understand language in real time, has been shown to be correlated with children's later vocabulary, such that children who were better at processing language have a larger vocabulary (Borovsky, Elman, & Fernald, 2012; Marchman & Fernald, 2008; Weisleder & Fernald, 2013). Rapid processing of speech facilitates listener's ability to integrate linguistic and non-linguistic stimuli. As a child's vocabulary develops, so does their speech processing ability (Fernald et al. 1998). As children are better able to incrementally process the speech stream, they are able to capitalize on the information in the speech stream in order to make predictions (Fernald, Perfors, Marchman, 2006). By the time a child is 21-months-old, they are able to process the speech stream with only partial target word information, for instance, looking towards a picture of a cat when they hear "Look at the ki". By their second year,

children are also able to use more semantically informative sentence frames, such as “Drink the juice” versus “Take the juice”, to initiate shifts to the target word earlier.

In a longitudinal study following children from 18 months to 24 months, Fernald, Marchman, and Weisleder (2013) found that a child’s ability to efficiently process verbal information is predicted by their SES. At 18 months, children in the lower-SES group were already observed to have lower vocabulary scores than children in the higher-SES group. By 24 months, both groups showed an increase in vocabulary size, but the gap between the two SES groups was found to have continued to increase. SES affects children’s processing abilities similarly to how it affects vocabulary; while both groups showed improvements in processing ability over time, children in the lower-SES group were slower to orient towards the correct familiar picture, and were less accurate overall than children in the higher-SES group. While these studies do not directly establish a causal link between a child’s language environment and their language development, they begin to explain a potential linguistic mechanism that is influenced by a child’s SES. According to these studies, the lower vocabulary levels of children in impoverished environments does not reflect just an imitation of the impoverished speech of their parents, but also arises from a greater difficulty in processing what language they do hear.

While differences in SES have been shown to lead to differences in the quality and quantity of input a child receives (Fernald, Marchman, & Weisleder, 2013), recent studies have shown that there is a great amount of variability in quality and quantity of the language input even within a socioeconomic group. In a study of low-SES Spanish speaking families, Weisleder and Fernald (2013) found that there was a large amount of variability in the quantity of both child-directed and overheard speech within this sample. Weisleder and Fernald found that variation in child directed speech in the home environment predicted children’s vocabulary scores, and this effect was mediated by children’s language processing ability. By expanding upon previous findings that established a relationship between children’s language environment and their subsequent

vocabulary, Weisleder and Fernald provide evidence that children's active processing of linguistic information may be directly influenced by environmental variability.

Indirect Auditory Information

While much is known about how the amount of speech a child hears affects their language development, children's home environments are frequently more complex than a single adult interacting one-on-one with the child. Many factors may contribute to the auditory complexity of the home environment. Three factors that may play an important role in shaping the auditory environment of the home, television, crowding, and noise pollution, may also have important implications for children's language development.

Television. Television is a ubiquitous part of the child's home environment; 97% of households in the US have a television (Common Sense Media & Rideout, 2013). The average two- to four- year-old child in the US spends over an hour of their day watching television (Common Sense Media, 2013). Children who live in heavy television viewing households, where the television is reported to be on "most of the time"—regardless of whether someone is watching the television, are more likely to watch television, and less likely to engage in enriching tasks, such as reading (Rideout, 2015; Vandewater et al., 2005). Even when a child is engaging in other tasks, the television is a pervasive force; the average child is exposed to 232 minutes of background television (Lapierre, Piotrowski, Linebarger, 2012). Excessive television during mealtime is particularly concerning for language development, as mealtime is an important activity for engaging children in discourse that encourages language development (Hoff, 2003).

By influencing how a child interacts with their environment, background television directly affects children's attention. In a cross-sectional experiment, Schmidt and colleagues (2008) examined whether background television would impact 12-, 24-, and 36-month-old's solitary toy play behavior. When the television was on during the free-play session, children played less, and were more distractible than when the television was off. Additionally, children who had more looks toward the television spent less time playing and had less focused attention. Attention in young infants is driven

by saliency, and the television is an incredibly salient object that might pull children's attention away from valuable learning opportunities.

In addition to affecting how children behave, television also shapes parent's behavior. Using the aforementioned free-play paradigm, Kirkorian and colleagues (2009) found that background television reduced parent-child interactions, particularly affecting parent's verbal responsiveness. Additionally, parents were more passive when interacting with their child while the television was on. Thus, the television impacts both the quantity and the quality of caregiver interactions.

While it seems clear that television affects children's and parent's attention, the effects of television on language development are less clear. In a longitudinal study of media exposure, Tomopoulos and colleagues (2010) found that, for low-SES mother-infant dyads, longer television exposure at 6-months predicts children's vocabulary at 14-months, such that children who are exposed to more television have lower vocabulary scores. The vocabulary development of very young children appears to be particularly susceptible to the effects of television. In a cross-sectional study of 8- to 24-month-olds, Zimmerman, Christakis, & Meltzoff (2007) found that parental reports of the amount of television a child is exposed to and their child's vocabulary are significantly, negatively correlated only for children under the age of 16-months. Additionally, the effect of television may be mitigated by providing age-appropriate educational content (Barr, Lauricella, Zack, & Calvert, 2010; Linebarger & Walker, 2005) or by providing verbal engagement while viewing educational television (Mendelsohn et al., 2010).

Rather than directly affecting children's language development, television may play an indirect role by shaping children's auditory environment. In a large observational study examining children's home auditory environment, Christakis and colleagues (2009) found that audible television in a child's home significantly reduced the amount of vocalizations a child produced, as well as the amount of speech a child hears; each additional hour of television exposure was associated with a 7% decrease in the amount of speech a child hears.

Crowding. Crowding in the home may also indirectly shape children's language development by shaping the home auditory environment. In particular, crowding affects language development by impacting the manner in which children are parented. Parents in crowded homes talk less (Wachs, 1979) and their speech is comprised of less complex utterances (Evans, Maxwell & Hart, 1999), beyond what is accounted for by SES. In examining the effects of residential crowding on children's language, Evans and colleagues found that the relationship between the complexity of parent's speech and crowding is mediated by parental responsiveness, indicating that the lack of speech in crowded homes is a result of parents not responding to their children. In consideration of the aforementioned work examining the importance of caregiver speech in children's language development, it stands to reason that crowding may be yet another important factor in shaping language development. Given that crowded homes are more prevalent in lower-SES families, it is possible that the previous connections found between maternal speech and language development are a byproduct of the number of people in the home.

While parents in crowded homes talk less, having extra people in the home creates an environment in which children could potentially hear a lot of speech that is not directly addressed to them. While assessing the effect of maternal speech on children's language development, Fernald, Marchman, and Weisleder (2013) found that the amount of child-directed speech in a child's home predicted the child's subsequent vocabulary, but the amount of non-child-directed speech did not predict the child's vocabulary development. Thus, it appears that children's learning is a reflection of direct interactions, as opposed to overheard conversation. Convergent evidence for the importance of child-directed input can be found in studies that look at language development in cultures that do not directly speak to children (Hoff, 2006).

In contrast, empirical studies directly measuring children's ability to learn novel nouns from overheard speech have found that two-year-olds are able to learn novel label-object associations from overhearing the object being labeled (Akhtar, Jipson, & Callanan, 2001; Floor & Akhtar 2006). Akhtar, Jipson, and Callanan (2001) contrasted a child's ability to learn words if they either a) were

directly told an object's name, or b) overheard the object's name when the experimenter was telling the assistant the name for the object. Children showed similar levels of learning in both conditions. While these results are promising in showing that children can learn without direct instruction, the pared down nature of the design may not accurately reflect a child's home environment. In this study design, the labeling sentences were not directed towards the child, but the child was in close proximity to the speakers and the object and the dialogue occurred at the child's eye level. The proximity of the learning moment to the child's location is entirely different than if a child was, for instance, sitting on the kitchen floor playing while their parents talk about what they are going to make for dinner. Additionally, the type of overheard conversation used in Akhtar and colleague's study is likely more captivating for a child than the other types of overheard speech in a child's natural environment. Thus, I am cautious in drawing strong conclusions on children's ability to learn words from overheard speech, particularly in light of research on noise pollution.

Noise Pollution in a Child's Home

Ecologically motivated studies on the role of environmental noise on children's cognition and reading have found that the presence of noise in the child's home environment has deleterious effects on their later performance in school (Cohen et al., 1973). Cohen and colleagues measured interior and exterior noise levels at the Bridge Apartments, an income-controlled housing project built over a highway in New York City, and examined the effects that noise had on children's reading ability and auditory discrimination skills. They found that the buildings' average external noise levels were relatively high, ranging from 75 dbA to 79 dbA. This is a stark contrast to the sound levels experienced in a typical quiet household (40 dbA) or in normal conversation (60 dbA). The measures of internal noise showed that as the floor level increased, the amount of noise exposure decreased.

Upon evaluating children's auditory discrimination abilities, as well as their standardized reading scores, Cohen and colleagues (1973) found that elementary school-aged children who had lived in the lower levels of the apartments for greater than four years showed poorer auditory

discrimination skills. Furthermore, children's auditory abilities were significantly correlated with their reading abilities. These findings indicate that chronic noise pollution may cause a child to become inattentive to acoustic cues, which in turn leads to a decreased ability to discriminate phonemic sounds, thus leading to poorer reading ability. Interestingly, additional research on the effects of chronic noise exposure have found that initially, children who are exposed to environmental noise may show an increased ability to tune out auditory distractors, but after just four years of exposure, the children lose this ability (Cohen et al., 1981). Thus noise exposure may initially confer an advantage to children; however, this advantage quickly diminishes when the exposure prevails for extended periods of time.

The presence of environmental noise outside of the home environment has also been found to have deleterious effects on children's learning (Maxwell & Evans, 2000). In the year prior to the installation of noise-dampening padding, preschool-aged children in a childcare cohort performed more poorly on measures of word recognition and showed a greater propensity towards the induction of helplessness in an unsolvable jigsaw puzzle task. Furthermore, after noise pollution abatement, teachers in this study rated children in the quieter conditions as being more likely to speak in full sentences, be more intelligible, and have a better understanding of what was being said to them.

While Cohen et al. (1973) found the strongest effects of noise impeding the reading and speech perception abilities in children who had been exposed to chronic noise pollution, it could also be the case that being exposed to noise pollution at a very young age, when a child is just beginning to form their phonological and semantic representations, can be even more detrimental and have cascading effects for their future language abilities. Additionally, co-occuring background noise during speech processing may increase the processing demands, making it even more difficult for children to process even familiar speech. In order to assess whether background noise changes how children learn, we must take the noise research back to the lab.

Speech-in-Noise Studies

The speech-in-noise paradigm is one methodology that has been employed in the laboratory setting to directly measure the effects of background noise. In these studies, children listen to a target speech stream while simultaneously listening to a distractor speech stream. This paradigm relies on the listener's ability to selectively attend to the target stimuli, while ignoring extraneous auditory information. This selective attention task can frequently be found in real life, and is common in complex home environments that include televisions, traffic noise, and multiple talkers.

Infants' abilities to recognize familiar words in noise have been extensively studied (Grieco-Calub, Litovsky, & Saffran, 2009; Newman, 2003; Newman, 2005; Newman, 2009; Newman & Jusczyk, 1996). Newman and Jusczyk (1996) found that infants were able to selectively attend to a specific talker in order to perceptually separate multiple speech streams. Infants are able to use the different intensity of talkers, as well as the different gender of talkers, as a cue to which speech stream was pertinent. In a study of the effects of background speech on highly familiar word recognition, Newman (2011) found that two-year-olds are able to recognize familiar objects when the target speech stream that directs their attention to an object is louder than the distractor speech stream, but not if the target speech stream is quieter than the distractor speech stream.

The familiarity of the target speaker's voice is also a useful cue for overcoming the deleterious affects of the background speech. Infants are better able to recognize familiar words with co-occurring background speech when the target words are spoken by a familiar voice, but not when they are spoken by an unfamiliar voice (Barker & Newman, 2004). Similarly, adults are better at identifying novel words in noise when they are produced by a familiar speaker, compared to a novel speaker (Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994). Additionally, adults underestimate the intensity of the background noise when listening to words spoken by a familiar talker, compared to words produced by a novel talker (Goldinger, Kleider, & Shelley, 1999; Jacoby, Allan, Collins, & Larwill, 1988). Thus, familiarity with a speaker falsely induces the perception of a lower signal-to-noise ratio (SNR), which enables speech processing.

The presence of noise in the environment also leads parents to alter their speech style (Newman, 2003). When teaching their two-year-old children novel words in a controlled play session, mothers frequently used child directed speech (CDS): shorter, hyperarticulated utterances, with longer pauses, and more fluctuating prosody (e.g., higher pitch, greater pitch variability, and more prosodic repetition; (Bernstein Ratner, 1984; Fernald & Simon, 1984; Fernald et al. 1989). Mothers' speech showed characteristics of CDS in both quiet and noisy environments. However, the prosody and syntax of utterances both changed when there was multitalker babble noise in the background during the play session—mothers spoke with both a higher pitch and a slower speaking rate, and used the target object name in a sentence medial position less frequently, than during the quiet play session. Although the context affected the mothers' teaching strategy, children failed to show consistent learning in either condition. While observational studies have shown that child directed speech is correlated with children's language development (Walker, Greenwood, Hart, & Carta, 1994), the interaction between parental input, the auditory environment and children's ability to learn remains unclear.

Speech-in-noise studies have also been useful in evaluating whether background noise affects children's ability to learn words. In a series of studies examining whether the language context directly affects whether children learn words, McMillan & Saffran (2016) found that, similar to word recognition, young children's word learning was sensitive to the intensity of the background speech. Two-year-olds are able to learn words concurrently with background speech at a +10 dB Signal-to-Noise ratio. If the SNR is decreased to +5 dB, the toddlers showed no evidence of word learning. However, hearing the target words prior to ostensive labeling of their referent object ameliorated the effect of background speech on word learning. From these studies we see that the environment in which children are learning words directly impacts whether or not children are able to learn words. Additional research is necessary in order to establish whether the environments simulated in the laboratory are truly reflective of children's home environments. The speech-in-noise paradigm has been helpful in examining whether background noise impacts speech processing and language

development, however, most of the studies rely on using contrived sources of noise, such as multitalker babble featuring 2-9 overlapping speakers. In order to fully assess whether background noise is detrimental to children's language development, it is important to consider the types of background noise children experience in their everyday lives. Children's home environments are highly variable, and different experiences with noise may differentially impact language development.

Current Study

From observational studies that have directly measured the auditory environments of children's homes, we know that there are multiple contributing factors that shape both children's home auditory environment and their language development. There are many ways to quantify the quality of a child's home auditory environment. With regards to language development, researchers have classically focused on the amount of language directed towards the child, finding that children's vocabulary and ability to process language is directly influenced by the amount of child directed speech in their environment. However, in addition to speech directed at them, children live in environments that feature other sources of auditory information, such as distant talkers, multiple talkers, and television.

We also know much about how the environment affects learning from empirical studies. Most studies examining how a child's environment directly influences their language development have relied on using multitalker babble consisting of overlapping utterances from up to nine different talkers (e.g., McMillan & Saffran, 2016; Newman, 2011). This research has been influential in establishing the limitations of children's perceptual systems. However, examining how more naturalistic background noise influences language development remains an open frontier. Additionally, by using naturalistic background noise, we are able to begin to look at potential mechanisms that drive children's ability to process speech-in-noise. For instance, children's prior experience with background noise may affect their ability to process speech-in-noise.

Children's familiarity with background noise on a local and global level may affect their ability to process-speech-in-noise. Locally, children's experiences with a specific source of background noise may impact their ability to process speech-in-noise. Young children and adults are both able to use their familiarity with the target speaker in order to more effectively process speech-in-noise (Barker & Newman, 2004; Nygaard & Pisoni, 1998). However, less is known about whether children are able to use their local familiarity with background noise when processing speech-in-noise. Children's familiarity with a specific source of noise may provide a useful cue to children, aiding them in separating the two auditory streams. Children live in homes that offer diverse sources of noise (e.g., multiple people talking, television, etc.), are children able to leverage their experiences with familiar sound sources in their home environment in order to process speech-in-familiar noise?

Globally, children's experiences with noise in their home environment may underlay their ability to process speech-in-noise. For school-aged children, experience with noise initially facilitates their ability to overcome distractions during a speech-in-noise task, however, after several years of exposure to noise, the facilitation effect disappears (Cohen, 1980). Do more moderate experiences with noise affect children's ability to process speech in noise? For instance, it is possible that given the ubiquity of television, children may become accustomed to having electronic sounds in their home environment. This experience may allow them to overcome the extra processing demands created by background television during a speech-in-noise task. Similarly, experience with hearing multiple overlapping sound sources may facilitate processing speech-in-noise.

We conducted a series of studies to test whether children's familiarity with background noise influences their language development. Specifically, three main research questions were investigated: 1) Does a child's familiarity with background noise impact their ability to learn new words? 2) Does a child's familiarity with background noise impact their ability to process familiar words? 3) How does the type of background noise a child experiences in their home environment influence their ability to process speech-in-noise? To address these questions, we combined observational and experimental methods, in order to test whether a child's familiarity with background noise affects their language

development. By using observational data to inform our experimental methods, we can begin to understand how naturalistic background noise affects children's language development.

Two experimental studies examined whether 26- to 30-month-old children's familiarity with electronic background noise affects their ability to learn new words (Study 1), and incrementally process speech (Study 2), using the looking-while-listening paradigm (LWL; Fernald et al., 2008). The target age of 26- to 30-month-olds was chosen for this study because children of this age are adept at recognizing familiar words in the context of background speech (Newman, 2011) and learning new words with background speech (McMillan & Saffran, 2016). In both studies, children were randomly assigned to either a familiar or an unfamiliar background noise condition. Children in the familiar background noise condition heard samples of background noise from their own home during the word recognition task. Children in the unfamiliar noise condition were yoked to a child in the familiar noise condition, and heard the background noise of the specific child to whom they were yoked. Additionally, in a series of exploratory analyses we examined how differences in children's environment affect their performance in learning new words (Study 1 results) and recognizing familiar words (Study 2 results) by looking at whether a child's experience with specific sources of noise in their home environment affects their performance in our experimental tasks.

Prior to participating in our experimental studies, daylong recordings were collected from each child's home environment and analyzed using the Language Environment Analysis system (LENA). The LENA system is comprised of two components: a digital language processor (DLP) that records up to 16-hours of the child's auditory experience, and computer software that analyzes the content of the recordings. The LENA software was used to quantify the overall amount of meaningful speech, electronic sounds, and overlapping noise a child heard in a day. The LENA software was also used to identify sections of the recording that were used as background noise stimuli; sections of the home audio recordings that had been identified as coming from electronic sources were used as background noise in the experimental tasks. The twenty-three sections used in the experimental tasks

were identified as coming from child television (14), child radio (2), adult television (1)[1], and adult radio (6).

Study 1: Word Learning

Does children's local and global familiarity with background noise affected their ability to learn words? To test whether children's local familiarity with background noise affects their ability to learn novel words, children were trained on novel object-referent mappings and then tested on whether they learned the object labels. In the *Referent Training* phase, children saw an object image appear on the screen and heard its corresponding label. During the *Test* phase, children were presented with a pair of objects on each trial. They then heard a target label that corresponds to one of the two objects. To test the effect of background noise familiarity on word learning, children heard background noise that was either sampled from their home environment, or sampled from another child's home during the referent training phase.

To examine whether children's global familiarity with noise affects their ability to process speech-in-noise in order to learn new words, we conducted a series of exploratory analyses looking at whether the proportion of electronic sounds a child heard throughout the day and the amount of overlapping noise a child heard predicts their speech-in-noise performance. We posit that individual differences in experience with noisy environments may lead to differences in children's ability to process speech-in-noise when learning new words. Given that the background noise used during word learning were from electronic sources, children's experiences with electronic sounds might facilitate their ability to process speech-in-noise. Additionally, the amount of overlapping noise a child hears might also help develop children's speech-in-noise processing abilities. Experience with these kinds of environments may provide children with more opportunities in which to learn to separate the relevant target speech directed at them from the irrelevant background noise.

---

[1] While the content of this sample was adult in nature, the parent of the child who was familiar with the background noise identified the sample as originating from the child's favorite movie.

Method

Participants: Participants were 46 monolingual English learners (22 male) with a mean age of 28.8 months (Range = 26.8-31.2 months). Two children were excluded from analyses for failing to complete the study and for providing a recording that was less than 10-hours long.

Children were randomly assigned to either a familiar (μ_{age} = 29, range = 26.8-31.2 months) or an unfamiliar (μ_{age} = 28.8, range = 27.1-30.6 months) background noise condition (Table 1). Children in the familiar background noise condition heard samples of background noise from their own home during the word recognition task. Children in the unfamiliar noise condition were yoked to a child of the same gender in the familiar noise condition, and heard the background noise of the specific child to whom they were yoked. Participants in the two conditions had comparable expressive vocabulary percentage scores (μ Familiar= .60, range = .18– 1; μ Unfamiliar = .67, range = .2 – 1), as measured by the MacArthur-Bates Communicative Development Inventory: Words and Sentences (Fenson et al., 2007) ($t(44) = .84$, $p = .40$).

	Familiar			**Unfamiliar**		
	n = 23			n = 23		
Gender	13 Female			13 Female		
Age in Months	29 (26.8-31.2)			28.8 (27.1-30.6)		
Vocabulary Percentile Scores	.60 (.18-1)			.67 (.2-1)		
Percent Overlapping Noise	.17 (.05-.29)			.15(.04-.26)		
Percent Electronic Sounds	.08 (.01-.13)			.05(.01-.11)		
Hours Meaningful Speech	3:04:58 (1:46:04 – 5:00:01)			3:03:08 (1:51:55 – 4:29:09)		
Parental Education	**Low** Less than College Degree	**Mid** College Degree	**Hi** Graduate Degree	**Low** Less than College Degree	**Mid** College Degree	**Hi** Graduate Degree
	5	9	9	1	12	10
Mean Parental Education (SD)	2.28 (.78)			2.48 (.59)		
Household Income	**Low** Less than $40,000	**Mid** $40,000-$100,000	**Hi** $100,000 +	**Low** Less than $40,000	**Mid** $40,000-$100,000	**Hi** $100,000+
	1	13	9	1	14	8
Mean Household Income (SD)	2.43 (.57)			2.38 (.56)		

Table 1. Characteristics of participants

Materials. Two different audio streams – target and background speech – were used in this experiment. Target speech stimuli were digitally recorded by a native-English speaking female in an infant-directed register and edited with Praat. All target stimuli were equalized to yield an average loudness of 65 dB for each phrase. Background noise was selected from recordings of the electronic noise in children's home auditory environment. LENA software was used to analyze the recordings

to identify the 1-hour section that had the most electronic noise. Within this hour segment, we selected a 2.25-minute segment of electronic noise to play during the training portion of the study. The electronic noise was equalized to yield a loudness of 55 dB (10 dB SNR). Background noise stimuli were edited to remove identifiable human sounds (e.g. coughing).

Referent Training. Children were taught the object-label pairings for three novel words (*blicket*, *coro*, and *tursey)*. Each novel label was paired with a single novel object. The labels were embedded in common naming phrases (Table 2). As noted above, the background noise played continuously throughout the training phase.

Object Image	Object Label	Training Phrase
	Tursey	"It's a tursey. Look at the tursey." "There's a tursey. See the tursey."
	Coro	"There's a coro. Look at at the coro." "This is a coro. See the coro."
	Blicket	"This is a blicket. Look at the blicket." "That's a blicket. See the blicket."

Table 2. Novel objects and labeling phrases.

Test Phase. During the test phase, two stationary images were positioned on the left and right sides of the screen. Target image locations were counterbalanced such that each object appeared as a target on the left and right side an equal number of times. Additionally, each object served as a distractor and a target object an equal amount of times. On each test trial, one object served as the target and one served as the distractor. To direct their attention to the target image, children heard either "*Find the* ____." or "*Where's the* ____?" paired with a generic attention getter phrase (e.g., "*Wow!", "Check that out.", "That's cool.", "Do you like it?"*). At the beginning of each trial, children saw the image pairs for 1500 ms in silence, before the auditory stimuli began to

play. At 1500 ms, the carrier phrase began, and was followed by the onset of the target word at 2467 ms into the trial. Each trial ended 6000 ms after picture onset. The test phase did not have background noise.

Procedure. The experiment took place in a double-walled sound-attenuated booth. Visual stimuli were displayed on a high definition, 55" Toshiba LCD monitor. Children were seated on a caregiver's lap 3 feet from screen. Caregivers wore blacked-out glasses to minimize their influence on the toddler's behavior. Background speech stimuli were presented in stereo via speakers located on either side of the chair. Target audio stimuli were presented through speakers located directly in front of the children. Children heard background speech stimuli simultaneously with target speech stimuli during *Referent Training* phase. During the *Test* phase children heard only the target speech. A digital camera mounted below the monitor captured the session, and children' looks to the monitor were coded off-line in 33ms frames.

The phases of the word learning study occurred sequentially while the children were seated on their caregiver's lap. During the *Referent Training* phase (2.25 min) children were taught the referents for three objects. During each naming event, a single object appeared on the left or right side of the screen, and briefly shook prior to being labeled. Trials were pseudo-randomly organized into four blocks of trials, with four trials per block. Each block began with one familiar-object naming trial (shoe, cookie, duck, or cat), followed by three novel-object naming trials. Children saw each novel object four times and heard the object labeled twice during each object presentation. Prior studies using background noise have found that this amount of exposure yields robust learning for words taught with a 10 dB SNR (McMillan & Saffran, 2016). A 7-second filler trial consisting of a cartoon was presented between each block to keep children engaged in the task. The order of the novel words was randomized within each block. The *Test* phase immediately followed the *Referent Training* phase. Two images appeared on the screen for 1.5 seconds prior to the onset of the auditory stimuli. Each trial lasted for 6 seconds. Familiar word trials (duck and cat or cookie and shoe) were used to add variety and keep the children engaged in the task. Each familiar word served as the target

4 times and each novel word served as the target 4 times, for a total of 16 trials. Trials were organized into 4 blocks, with each block starting with a familiar word trial, and featuring one trial for each novel word (Appendix A).

Results & Discussion

Data Analysis Overview. Videos were coded offline by trained coders using custom software (Swingley, Pinto & Fernald, 1999). Children's looking behavior was coded frame-by-frame as either being left, right, or off-screen. Coders were blind to participant condition, the target word, and the side of the target picture. The primary question was whether children's local and global familiarity with background noise would facilitate word learning. To address whether children's local familiarity affected their ability to learn words, data were analyzed using two methods: 1) by analyzing whether the groups differed in the proportion of time children looked towards the target object during a 1500 ms time window, and 2) by analyzing differences in the trajectories of children's looking behavior using growth curve analyses (GCA; Mirman, 2014).

The accuracy of each child's response on each trial was calculated as the proportion of time they looked at the target versus the distractor image (out of the total time spent fixating the target or distractor). The target window began 300 ms after noun onset in order to account for the planning of eye movements, and ended 1800 ms after noun onset (Fernald et al., 2008). Trials were excluded if the child was not fixating on either object for more than 33% of the analysis window. GCA with orthogonal polynomials was used to model how children's fixations to the target images changed over time, from 300 ms to 1800 ms after novel word onset (Mirman, 2014). Time was modeled using linear, quadratic, and cubic orthogonal polynomials.

Experimental and observational data were used to assess whether children's global familiarity with background noise affects their ability to process speech-in-noise in order to learn new words. GCA was used to analyze individual differences in children's ability to learn novel words with background noise. The percent of electronic sounds and the proportion of overlapping noise a child heard throughout the 16-hour recording obtained by LENA were used as measures of the home

environment. We used likelihood ratio testing to assess whether adding the individual differences measures improved model fit.

Mean Accuracy. Our central hypothesis for Study 1 was that children in the familiar background noise condition would be better able to learn the novel words, as indicated by higher mean accuracy scores. Children in the familiar background noise condition looked significantly longer at the target objects, when compared to chance (50%) [$M_{Familiar}$ = .62, *SD* = .13; t (22) = 4.40, *p* <.001, *d* = .93], suggesting that children in this condition successfully learned the novel words. Children in the unfamiliar background noise condition looked significantly longer at the target objects, when compared to chance (50%) [$M_{Unfamiliar}$ = .63, *SD* = .11; t (22) = 5.411, *p* <.001, *d* = 1.2], suggesting that children in this condition also successfully learned the novel words. An independent samples t-test found that the mean accuracy scores for children in the familiar and unfamiliar condition did not significantly differ from one another [t (44) = -.2, *p* = .84, *d* = .08] (Figure 1). These results suggest that children's familiarity with background noise did not affect their ability to learn new words.

While mean accuracy scores indicated that both groups looked towards the target object for an equal proportion of the target window, the time course of children's pattern of looking may have been different for each condition (Figure 2). If children in the unfamiliar background noise condition were affected by the familiarity of background noise, they may be slower to shift towards the target object and they might exhibit more shifting away from the target objects—behavior that would not be captured by the mean accuracy scores.

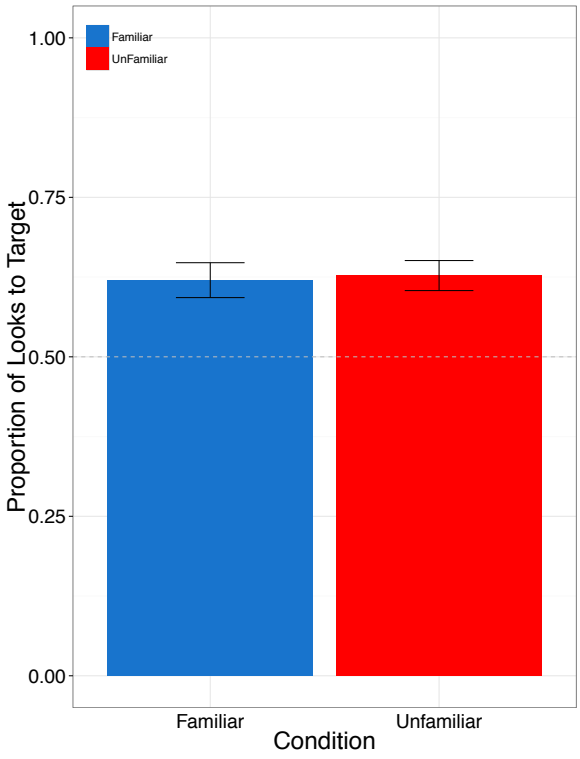

Figure 1. Children's mean accuracy by condition, calculated as the proportion of time looking towards the target object from 300-1800 ms after noun onset. Error bars represent +/- 1 SE.

GCA. GCA with orthogonal polynomials was used to evaluate whether children's familiarity with background noise affected their looking trajectories. We modeled the time-course of children's fixations to the target objects on the linear, quadratic, and cubic time terms from 300 ms to 1800 ms after novel word onset, with condition modeled as a fixed effect (Appendix C). All models were fit using Maximum Likelihood Estimation and likelihood ratio testing was used to compare the effect of condition on each time term. For both conditions, the linear term was significant. The quadratic, and cubic orthogonal time terms were not significant, indicating a curvilinear change in looks to the target image over time. There was a nonsignificant effect of condition on the intercept [b = .72, X^2 (1) = .50, p = .48], confirming that condition did not affect overall mean accuracy scores. There were also nonsignificant effects of condition on the linear time term [b = -.17, X^2 (1) = 1.03, p = .31], quadratic terms [b = .017, X^2 (1) = .03, p = .86], and cubic [b = .011, X^2 (1) = .02, p = .89] indicating that children's looking trajectories were not affected by condition.

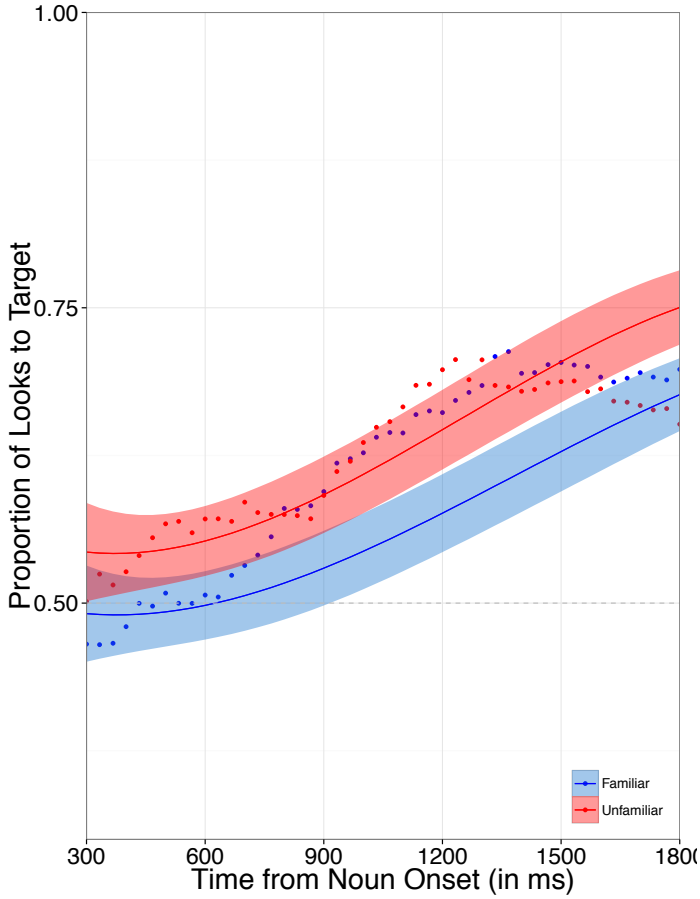

Figure 2. Time course of children's looking to the target object after noun onset. Data points are observed behavioral data and lines are the growth curve fit. Ribbons around the lines indicate +/- 1 SE.

Overlapping Noise. To assess whether the amount of overlapping speech a child hears affects their ability to overcome background noise when learning novel words, we added the proportion of overlapping speech to the original GCA model (Appendix D). Adding overlapping speech to the model did not improve the model fit [b = .20, X^2 (1) = 1.80, p = .15].

Electronic Sounds. To assess whether the amount of electronic media a child hears affects their ability to overcome background noise when learning novel words, we added the percent of television exposure to the original GCA model (Appendix E). Adding television exposure to the model did not improve the model fit [b = -.10, X^2 (1) = 1.70, p = .79].

In Study 1 the volume of the background noise was normalized to 55 dB, an intensity that produced robust learning in previous studies using two-talker background babble (McMillan & Saffran, 2016). Prior studies have found that children have greater difficulty contending with background speech featuring a single talker, compared to speech created with multiple, overlapping talkers (Newman, 2009). Given this, we predicted that the background noise abstracted from children's home environments would be particularly difficult, as it features only one individual speaking at a given time. We hypothesized that children in the familiar background noise condition would capitalize on their familiarity with the background noise, in order to push through the noise and learn words. We found evidence of robust learning for children in both conditions. Both groups of children showed very similar mean accuracy scores and looking trajectories, indicating that the background noise did not impede word learning for either group. Conflicting findings from word learning studies with two-year-old children indicate that the amount of overlapping talkers in background speech during word learning impacts whether or not children are able to learn. Two-year-olds are able to learn words when the background babble speech features 9 people talking simultaneously and is the same volume as the target speech (Dombroski & Newman, 2014). However, others have found that two-year-old children fail to learn novel words when listening to two-talker background babble that was 5 dB quieter than the target speech (McMillan & Saffran, 2016). Taken together, these findings, and the results from Study 1 indicate that two-year-old children have a prodigious ability to learn new words.

What experiences might underlay children's ability to overcome background noise when learning new words? In a series of exploratory analyses we examined potential factors that might explain this ability: experience with electronic sounds and experience with overlapping noise. Neither overlapping speech nor electronic noise significantly affected children's ability to process speech-in-noise when learning new words. Children in our study were very successful in overcoming the effect of background noise in order to successfully learn novel words. Our experimental task may not have elicited enough variability to find individual differences in word learning. Furthermore,

much of the research on noise in children's home environments has focused on children who experience chronic noise pollution (e.g., Cohen, 1973). Children in our study only experienced moderate amounts of noise in their home. It is possible that children who experience chronic background noise in their homes may process speech-in-noise differently than children who experience moderate amounts of background noise.

The ability to incrementally process speech underlies children's ability to learn novel words. Many studies have found that background noise hampers children's ability to recognize familiar objects when listening to background noise (e.g., Grieco-Calub, Litovsky, & Saffran, 2009; Newman, 2003; Newman, 2011). However, again these studies typically use multitalker babble speech that is unfamiliar to the child. In Study 2 we examine whether a child's familiarity with background noise impacts their ability to process familiar words.

Study 2: **Familiar Word Processing**

In study 2 we examine whether children's local and global familiarity with background noise affects their ability to process speech-in-noise when recognizing familiar objects. To test whether children's local familiarity with background noise affects their ability to process speech, children were tested on their recognition of familiar objects, using LWL (Fernald et al., 2008). Children saw pairs of familiar objects, and heard a sentence directing their attention towards one of the objects. Two-year-old children are particularly adept at recognizing objects using this paradigm. Thus, in order to have more individual variability in the task, children were prompted to look at the objects by a phrase that was either informative (e.g., "Drink the juice") or non-informative (e.g., "Find the juice") about the identity of the target object (Fernald, Perfors, Marchman, 2006). When hearing semantically informative verbs, adults and children are able to use their knowledge of the verb to shift their attention towards the target image sooner than if they hear a neutral verb (Altmann Kamide, 1999; Venker et al., 2016). Background noise played concurrently with each test trial. Familiarity with background noise was manipulated as in Study 1; children heard background noise that was either sampled from their home environment, or sampled from another child's home.

To examine whether children's global familiarity with noise affects their ability to process speech-in-noise, we conducted a series of exploratory analyses looking at whether the proportion of electronic sounds a child heard throughout the day and the amount of overlapping noise a child heard predicts their speech-in-noise performance. Again, we posit that individual differences in experience with noisy environments may lead to differences in children's ability to process speech-in-noise when learning new words.

Method

Participants. Same as Study 1

Material

Auditory Stimuli. As in Study 1, two audio streams – target and background speech – were used in this experiment. A different native-English speaking female speaking in an infant-directed register will digitally record target speech stimuli. Using Praat, all target stimuli were equalized to yield an average loudness of 65 dB for each phrase. Target speech stimuli consisted of informative (e.g., "Drink the juice") or neutral (e.g., "Find the juice") labeling phrases (Appendix B). The timing of the auditory stimuli were standardized in Praat, such that target noun onset began at 990ms after phrase onset, and the entire auditory stimulus was 4.5 seconds.

As in Study 1, LENA software was used to analyze the recordings to identify the 1-hour section that had the most amount of electronic noise. A different 3.5-minute subset of background noise was selected from recordings of the electronic noise in children's home auditory environment, and was equalized to yield a loudness of 65 dB (0 dB SNR). Background noise stimuli were from the same type of content as in Study 1 (i.e., children who heard a clip from a child television show in Study 1 heard another clip from a child television show).

Visual Stimuli. The visual stimuli were images of familiar objects: ball, door, cake, bike, juice, and book. These objects were different than the familiar objects used in Study 1 (i.e, cat, cookie, duck, and shoe). Each object was represented by four different images throughout the study.

Procedure. Study 2 was run immediately after Study 1; a brief, 1-minute Sesame Street video gave children a break between tasks. The design of the speech processing study is similar to the *Test* phase of the word learning study, but with familiar objects. Two familiar objects appeared on the screen for1.5 seconds prior to the onset of the labeling phrase. Children heard either a neutral verb or informative verb phrase directing their attention to one of the objects, and each trial lasted for 6 seconds. Each word served as the target twice with a verb phrase and twice with a noun phase, for a total of 24 trials. Trials were organized to minimize repetition in object, target location, image pairings, and carrier phrases (Appendix B).

Results and Discussion

Data Analysis Overview. The data were coded and analyzed in a similar manner as Study 1. In order to assess whether children were able to use the informativeness of the verb phrase to shift towards the target object earlier, we analyzed neutral verb trials and informative verb trials separately. For neutral verb trials, the analysis window began 300 ms after the noun onset and continued to 1800 ms after the noun onset (Fernald et al., 2008). For informative verb trials, the analysis window was extended, begining 300 ms after verb onset and continuing to 1800 ms after noun onset. The analysis window begins at 300 ms after verb onset in order to capture the effect of verb informativeness, and ends at 2500 ms after verb onset (which is 1800 ms after noun onset) to include the standard analysis window (Fernald et al., 2008). Data were analyzed using two methods, 1) by comparing differences in mean accuracy scores, and 2) using GCA to analyze the looking trajectories across the analysis windows.

As in Study 1, experimental and observational data were used to assess whether children's global familiarity with background noise affects their ability to process speech-in-noise in order to learn new words. GCA was used to analyze individual differences in children's ability to learn novel words with background noise. The percent of electronic sounds and the proportion of overlapping noise a child heard throughout the 16-hour recording obtained by LENA were used as measures of the

home environment. We used likelihood ratio testing to assess whether adding the individual differences measures improved model fit.

Mean Accuracy. Our central hypothesis for Study 2 was that children in the familiar background noise condition would be better able to recognize familiar words. Children in the familiar background noise condition looked significantly longer at the target objects, when compared to chance (50%) for both the neutral verb trials [$M_{Familiar}$ = .63, *SD* = .10; t (22) = 6.34, *p* <.001, *d* = .97], and the informative verb trials [$M_{Familiar}$ = .62, *SD* = .10; t (22) = 5.63, *p* <.001, *d* = .95] indicating that children were successful in recognizing the familiar objects. Children in the unfamiliar background noise condition looked significantly longer at the target objects, when compared to chance (50%) for both the neutral verb trials [$M_{unfamiliar}$ = .61, *SD* = .10; t (22) = 6.34, *p* <.001, *d* = .97] and the informative verb trials [$M_{Unfamiliar}$ = .66, *SD* = .07; t (22) = 5.63, *p* <.001, *d* = .95], indicating that children who heard unfamiliar background noise were also successful in recognizing the familiar objects (Figure 3). An independent samples t-test found that children in the unfamiliar background noise condition had marginally higher accuracy scores for the informative verb trials [t (44) = -1.77, *p* = .06, *d* = .5]. There were no significant group differences in means for neutral verb trials [t (44) = .51, p = .61, d = .32]. Taken together, these results indicate that children in both groups were able to recognize the familiar objects. Additionally, there is a trend for the informative verb condition that suggests that children in the unfamiliar background noise condition have higher mean accuracy scores than children in the familiar background noise condition.

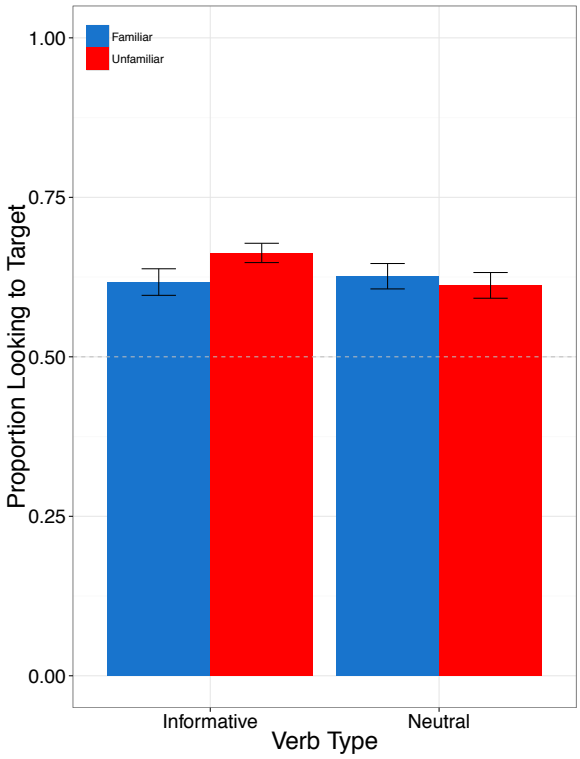

Figure 3. Children's overall mean accuracy scores for neutral and informative verb trials

GCA. GCA was used to further interrogate whether children in the unfamiliar background noise condition are exploiting the informativeness of the verb moreso than children in the familiar background noise condition. We used orthogonal polynomials to model the time-course of children's fixations to the target objects on the linear, quadratic, and cubic time terms from 300-1800 ms after noun onset for neutral verbs and 300 ms to 2500 ms after verb onset for informative verbs. Each type of verb trial was modeled separately, and condition was modeled as a fixed effect for both models. All models were fit using Maximum Likelihood Estimation and likelihood ratio testing was used to compare the effect of condition on each time term.

Neutral Verb GCA. For both conditions, the linear, quadratic, and cubic orthogonal time terms were significant, indicating a curvilinear change in looks to the target image over time (Appendix F). There was a nonsignificant effect of condition on the intercept [b = .01, X^2 (1) = .08, p = .78], confirming that condition did not affect overall mean accuracy scores. There were also nonsignificant effects of condition on the linear time term [b = .20, X^2 (1) = .72, p = .40], quadratic

terms [b = -.04, X^2 (1) = -.67, p = .69], and cubic [b = .05, X^2 (1) = -.61, p = .54] indicating that children's looking trajectories were not affected by condition (Figure 4).

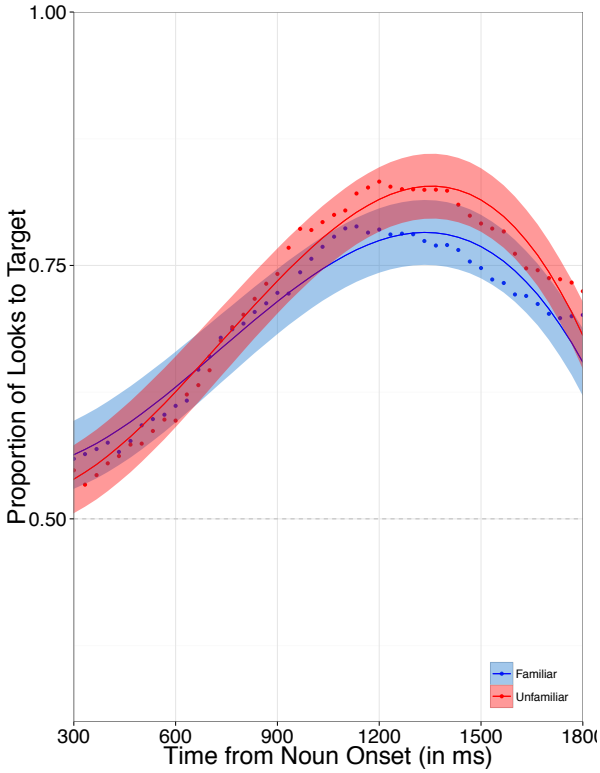

Figure 4. Time course of children's looking to the target object after 300-1800 ms after noun onset for neutral verb trials. Data points are observed behavioral data and lines are the growth curve fit. Ribbons around the lines indicate +/- 1 SE.

Informative Verb GCA. For both conditions, the linear time term was significant, and the quadratic, and cubic orthogonal time terms were non-significant, indicating a linear change in looks to the target image over time(Appendix G). There was a significant effect of condition on the intercept [b = .07, X^2 (1) = 4.33, p = .04], confirming that condition did affect overall mean accuracy scores. There were nonsignificant effects of condition on the linear time term [b = .05, X^2 (1) = .06, p = .80], quadratic terms [b = -.03, X^2 (1) = -.26, p = .79], and cubic [b = -.02, X^2 (1) = -.27, p = .79] indicating that children's looking trajectories were not affected by condition (Figure 5).

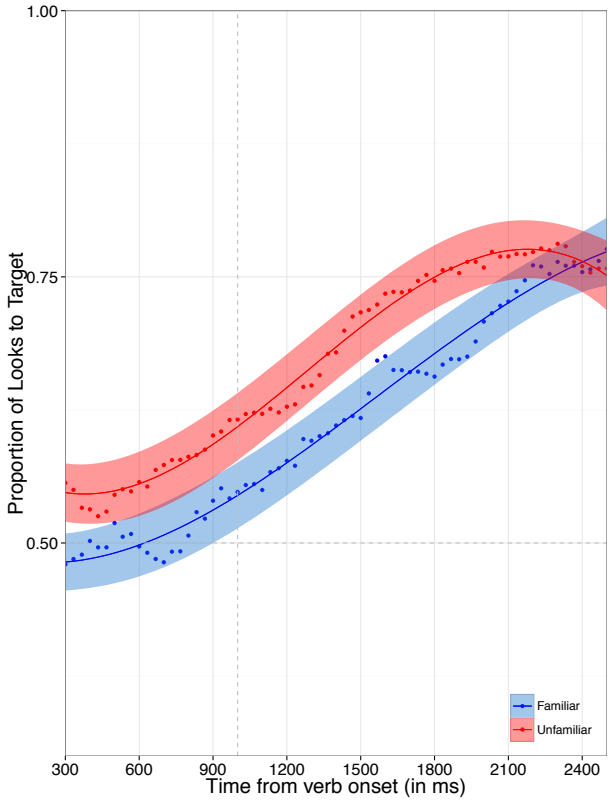

Figure 5. Time course of children's looking to the target object 300-2500 ms after informative verb onset. Vertical gray line indicates noun onset. Data points are observed behavioral data and lines are the growth curve fit. Ribbons around the lines indicate +/- 1 SE.

Overlapping Noise. To assess whether the amount of overlapping noise a child hears affects their ability to recognize familiar words, we added the proportion of overlapping speech to the original GCA models. For neutral verb trials, the effect of overlapping noise was non-significant [t (45) = -1.1, p = .28], and adding overlapping noise to the model did not improve the model fit [b = -.09, X^2 (1) = .46, p = .41] (Appendix H). Similarly, informative verb trials, the effect of overlapping noise was non-significant [t (45) = .98, p = .33], and adding overlapping noise to the model did not improve the model fit [b = .09, X^2 (1) = .36, p = .55] (Appendix I).

Electronic Sounds. To assess whether the amount of electronic media a child hears affects their ability to overcome background noise when looking at familiar objects, we added the percent of electronic sound exposure to the original GCA model. For neutral verb trials, the effect of electronic

sounds was non-significant [t (45) = .83, p = .41], and adding electronic sounds to the model did not improve the model fit [b = .38, X^2 (1) = .24, p = .63] (Appendix J). Similarly, informative verb trials, the effect of electronic sounds was non-significant [t (45) = -1.28, p = .21], and adding electronic sounds to the model did not improve the model fit [b = -.57, X^2 (1) = 1.29, p = .26] (Appendix K).

In Study 2 informative (e.g., "Eat the cake") and neutral (e.g., Find the cake") verb phrases directed children's attention towards images of highly familiar objects. While background noise played continuously, each trial began with displaying the objects on the screen for 1.5 seconds before the verb phrase would begin. If children were attending to the verb phrase, the informative verb trials would facilitate early shifting towards the target object, as children could exploit their knowledge of the verb in order to predict the noun. We hypothesized that children in the familiar background noise condition would be better able to ignore the background speech, and would shift towards the familiar objects sooner. Surprisingly, children in the unfamiliar condition disengaged from the distracting background noise and shifted their attention towards the target object more readily than their peers in the familiar background noise condition. While children's familiarity with background speech did not lead to differences in their overall accuracy in identifying the target objects, it did inhibit their ability to quickly take advantage of linguistic information. All together, these results indicate that children are proficient at processing speech-in-noise, but the increased processing load from the distractions has consequences for their ability to attend to other information. Other studies on children's ability to process speech-in-noise have found similar robust processing of speech-in-noise (e.g., Newman, 2011). The results from Study 2 add to the body of research and further our understanding of what experiences influence children's ability to contend with background noise.

In Study 2, as in Study 1, the proportion of overlapping noise or electronic sounds children heard in their home environment did not significantly affect their ability to process speech-in-noise. Children in our study were very successful in overcoming the effect of background noise in order to successfully learn novel words and recognize familiar words. Multilevel modeling techniques, such as GCA often rely on large sample sizes in order to reliably measure robust effects (Curran, Obeidat, &

Losardo, 2011). This is particularly necessary when studying individual differences, as the models require enough statistical power to adequately explain variability in the measures of interest. Given our relatively small sample size of 46 children, and children's robust ability to overcome background noise, it is likely that we did not have enough statistical power to find strong effects. Additionally, the overall proportion of overlapping noise and electronic sounds may be too gross of a measure of noise in children's home environments. Investigating the proportion of overlapping noise and electronic sounds relative to the amount of beneficial, informative speech (i.e., adult speech) may provide a more meaningful measure of the noise in children's home environments(Christakis et al., 2009). Overall, the results from Studies 1 and 2 indicate that children are adept at processing speech-in-noise in order to learn new words and recognize familiar objects.

General Discussion

In this series of studies we looked at the effect of children's local and global familiarity with background noise on their ability to process speech and learn new words. In Studies 1 and 2, children heard either familiar or unfamiliar background noise while learning new object-label mappings and viewing well-known objects. In Study 1, we found that children in both conditions were able to learn the words, indicating that children in this study were able to successfully release from any masking effects. In Study 2, the intensity of the background noise was increased to the same volume as the target speech, and we found that children in the unfamiliar background noise condition were better than children in the familiar background noise condition at exploiting the informativeness of the verb phrase in order to identify the target words. This indicates that children in the unfamiliar background noise condition were less encumbered by the background noise than their yoked partner in the familiar background noise condition. Children in both conditions received the same speech signal, which contained the same amount of energetic masking, but led to greater attentional demands for the familiar background noise group. Through a series of exploratory analyses, we examined two home environment variables as possible mechanisms for explaining individual differences in children's

ability to process speech-in-noise. Neither the amount of overlapping speech a child hears in their home, nor the percentage of television predicted children's abilities to process speech-in-noise.

Noise Familiarity

The aim of Studies 1 and 2 was to examine whether children's local familiarity with background noise affects their ability to segment overlapping sound sources in order to successfully process speech-in-noise. While previous studies have not looked at whether children's familiarity with background noise affects their ability to process speech-in-noise, studies with adults and children have looked at whether familiarity with the target speech affects performance. Seven-month-old infants are able to exploit the familiarity of their mother's voice in when it is presented with background speech (Barker & Newman, 2004). Infants listened longer to words spoken by their mothers than those produced by an unknown woman, indicating that they were able to use the familiarity of the target voice in order to separate the two speech streams. Studies 1 and 2 were a first attempt to understand whether children's familiarity with background noise would show similar facilitation of speech-in-noise processing. In Study 1, both groups of children were able to successfully learn words with background noise. In Study 2, children in the unfamiliar background noise condition were more attentive to the target images than children in the familiar background noise condition, for informative verb trials. These results indicate that familiar background noise is more salient than unfamiliar background noise, thus children might be better able to ignore unfamiliar background noise.

The aim of our exploratory studies throughout Studies 1 and 2 was to examine whether individual differences in children's global familiarity with background noise affects their ability to process speech-in-noise. Much of the focus of research on children's experiences with noise at home has focused on chronic noise pollution, and children in low income households. Studies of environmental noise have found that constant noise (e.g., traffic noise) may be less harmful than intermittent noise (e.g., airplane noise) (Matheson et al., 2010). Additionally, Matheson and colleagues (2010) suggest that the most deleterious effects of noise might be found for children who

encounter high levels of noise in multiple environments, such as at home and at school. Less is known about whether moderate amounts of background noise in the home environment affect language development. In classroom studies, children are better able to understand speech in quiet than in background noise, and younger children are particularly vulnerable to noise (Jamieson, Kranjc, Yu, & Hodgetts, 2004). While the home environment measures explored in Study 3 did not affect children's ability to process speech-in-noise, children in our study likely only experience moderate amounts of background noise. Exploring whether children who experience moderate amounts of noise are differentially affected by speech-in-noise than children who chronically experience background noise may help further our understanding of the deleterious effects of noise.

Why Language Development May Be Particularly Vulnerable

Background noise may be particularly problematic for language development as it affects the information children receive on both a local and global level. Locally, background speech can lead to energetic and informational masking of the speech stream, which makes it harder for children to understand the intended language (Mattys, Brooks, Cooke, 2009; Wightman & Kistler, 2005). The volume (McMillan & Saffran; Newman, 2012), content (Newman, 2009; Van Heugten, Krieger, & Johnson, 2014), and location (Litovsky, 2005; Johnstone & Litovsky, 2009) of background noise interact when masking the target speech stream. Background noise that is louder, produced by fewer talkers, and spatially in the same location as the target speech is more likely to mask the speech signal. The background noise used in our studies was presented 90° from the target speech, and featured a single person talking at any given time. The volume, location, and content of background noise in study 1 sufficiently differed from the target speech, so children were successful in learning new words, regardless of whether the background noise was familiar or not. However, by increasing the volume of background noise in Study 2 to the same volume as the target speech, the background noise distracted children in the familiar condition, depressing their ability to utilize the informativeness of the verb phrase when recognizing familiar objects. Taken together, these findings

suggest that children are vulnerable to specific conditions of background noise, and must recruit greater resources in order to overcome more difficult speech-in-noise.

Globally, background noise affects language development by suppressing the amount of beneficial information in the home environment. Having a television on in the home reduces both the amount of speech a child produces, and the amount of speech they hear throughout the day (Christakis et al., 2009). Additionally, both observational and experimental studies have found that the presence of television reduces parental responsiveness, and the amount of parent-child interactions (Christakis et al. 2009; Kirkorian, Pempek, Murphy, Schmidt, & Anderson, 2009; Nathanson & Rasmussen, 2011). Background noise generated by crowding has a similar effect as television – parents in crowded homes are less responsive to their children and produce less elaborative language. An abundance of studies have shown that the quality and quantity of speech a child hears during formative stages of language development has cascading consequences for their subsequent development (e.g., Hart & Risley, 1995; Ramirez-Esparza, Garcia-Sierra, Kuhl, 2014; Veronon-Feagans, Bratsch-Hines, & The Family Life Project Investigators, 2013; Weisleder & Fernald, 2013). In our individual difference analyses, we examined whether a child's global familiarity with noise, as measured by the amount of electronic media they heard and the amount of overlapping noise they heard, affects their ability to learn new words and recognize familiar words. We found that neither the amount of overlapping speech in a child's home, nor the amount of television they heard explain a child's ability to process speech-in-noise. Given the inconclusiveness of these findings, it is difficult to determine whether language development is most hindered on a global, compared to a local scale.

Limitations of the Current Design

While the results extend our understanding of how background noise affects children's language development, there are a few limitations of the current set of studies that should be considered. First, the LENA system uses a recording device that, while sophisticated enough to allow the LENA software to identify the different sources of auditory input a child hears throughout a 16-

hour time period, generated audio files that varied greatly in their quality. The LENA system is focused on capturing audio within 6-feet of the child, and while the microphone does this task adequately, frequently the quality of the audio samples extracted are lower than the original, live sample, would have been. This may have led to some samples being more difficult for children to recognize, as the audio quality was lower than what they were accustomed to. Additionally, since the recorders are worn in a vest, the resulting files frequently had friction noises whenever the child moved. When extracting audio samples for background noise in Studies 1 and 2, great care was taken to select audio samples that minimized the amount of friction noise, but this was not feasible for all samples. These potential audio quality issues may explain why the familiarity manipulation in studies 1 and 2 did not lead to greater effects – some children might have had a harder time recognizing the audio.

A second limitation of the LENA system is that the algorithms used for identifying the different sources of auditory input a child hears are not perfect. While LENA did an adequate job of identifying samples of electronic noise, it frequently overestimated the amount of electronic noise a child heard throughout the day due to misidentifying a sound source. One common occurrence of LENA misidentifying the sound source occurred when a child was in a vehicle; the LENA system would often confuse the road noise for electronic noise. The LENA also frequently identified adult singing as an electronic source, which again would lead to a child having a higher amount of electronic noise than they really experienced. Issues with the construct validity of the LENA system have direct implications for our individual difference analyses. If the LENA system has overestimated the amount of electronic noise for some children in our sample, our individual difference analyses may be less valid, and less generalizable.

Another key limitation of the current studies arises due to how the segments of the audio files were selected as background noise. Since the electronic audio was strategically selected from areas of the recording that had fewer additional auditory sources, it was impossible to know if the chosen audio would be easily identifiable to the child. Frequently these segments of the recording had less

talking from parents and children, which may indicate that children were very attentive to what they were watching—however, this is highly speculative. Despite this, 20-out-of-23 parents were able to correctly identify the background noise as originating from their home, when asked informally after the study. Relatedly, the study design also prevented us from being able to gauge how unfamiliar a sample of electronic noise would be. Indeed, when asked if they were able to recognize the background stimuli, 4-out-of-15 parents with children in the unfamiliar condition reported that the television show was familiar to their child, though the specific episode featured in the study may not have been familiar. Interestingly, all of the parents (8 total) whose children heard either adult or child radio content were able to correctly identify that the background noise was not from their home. In hindsight, formalizing this question, and asking parents to rate how confident they are with the source of the background noise, would have been helpful in determining the degree to which the background noise was familiar to the children. Thus, not all samples were equally familiar; hearing a sample of background noise from a child's favorite television show, versus hearing a sample of background noise from their parent's favorite radio station may elicit different responses. How familiar or unfamiliar a sample needs to be in order to affect language processing is an important empirical question that remains to be answered. In our studies, children in the familiar noise condition might need to hear background noise that was highly familiar (e.g. from television episode that they had watched many times, or a television show that was highly familiar to them) and children in the unfamiliar noise condition might need to hear background noise that was completely unlike what they hear in their home.

An additional limitation of the study design results from relying on only one recording of children's homes. By having only a single recording, it is unclear if the samples, and subsequent quantified data, are truly representative of children's experiences in their home. To address this issue, parents were asked to indicate if the recording occurred on a typical day for their child, and all of the parents selected "yes"—though many of them wrote-in exceptions, such as "dad was home late", "less TV than usual", or "we went camping that evening". Given that the majority of the data were

collected during the summer, many families' patterns and routines may be different than if the recordings were captured during another season. Furthermore, additional recordings would bolster the validity of the data obtained by the LENA, which would help to directly address any issues with the validity of the recordings. If the data collected by the LENA are skewed, this limits the robustness of our exploratory analysis findings, as any individual differences found in children's home experiences may actually be an artifact of recording on that specific day.

One final limitation of this study is the lack of economic diversity of the participants. 89% of the families who participated in these studies reported a household income that was over $60,000 a year. Given that the median household income in the United States is $53,842, our sample oversampled from households that are above the median income bracket (US Census, 2015). Furthermore, only 2 of the 46 families who participated in these studies reported income that was below the poverty level, which contrasts with 14.8% of the population who are living in poverty (US Census, 2015). The lack of economic diversity in our sample limits our ability to generalize our findings to other populations. Moreover, it is possible that testing children from a more diverse and representative sample may elicit different results, and a different pattern of findings. Not only would diversifying our sample potentially lead to differences in the amount of exposure a child has to electronic noise, but it may also lead to differences in speech-processing and word learning abilities.

Future Directions

Despite these limitations, the current set of studies takes a first step in understanding whether children's familiarity with background noise is a mechanism that underlies their ability to process speech-in-noise. The findings add to a growing literature that has found that children are sensitive to specific characteristics of the background noise stream. The results of the current study hint that children's local familiarity with background speech makes it increasingly difficult for children to selectively attend to the target speech stream. Additionally, it appears that the affect of local background noise familiarity on speech-in-noise processing is accumulative with the other characteristics of the distractor stream, such as the signal-to-noise ratio.

To further our understanding of how children's experiences with noise shapes their ability to process speech-in-noise, there are several future directions that will be taken. In extending these research findings, the first step I will explore is to examine whether a stronger manipulation of children's familiarity with background noise replicates the findings that children's sensitivity to noise familiarity interacts with other perceptually challenging characteristics of the distractor stream (i.e., SNR). By experimentally manipulating the amount of experience a child has, or prescreening a child's experiences with the background speech, I will be able to more clearly assess the magnitude to which children are familiar with the background noise.

Another area I would like to expand this line of research into is to begin to understand how attention affects a child's ability to process speech-in-noise. In order to successfully process speech-in-noise, children must be able to successfully selectively attend to the target speech stream, while ignoring the distracting stream (Newman, 2011). To recognize the target objects, children in our study had to disengage from the distractor stream. Children in the unfamiliar background noise condition were more successful at attending to the target stream than children in the familiar background noise condition. One unexplored factor that may have influenced how well children disengaged from the task is the salience of the background noise. Does the salience of background noise affect children's ability to selectively attend to the target speech stream? For children in the familiar condition, the background noise may have been more salient, as it was identifiable. Background noise salience can be manipulated in many different ways in order to understand how salience affects children's ability to allocate attention.

Furthermore, in our studies background noise requires children to choose which stream they would like to attend to, with the distractor stream being overtly less meaningful than the target stream. Outside of the lab, which stream a child should attend to may be less obvious. What are the consequences of maintaining this informational gating when learning language in situ? Family routines and rituals play an important role in children's language development (Fiese, Foley, Spagnola, 2006). Do these rituals and routines also support the development of children's selective

attention? By looking at the contexts in which children are hearing background noise, I will be able to begin to address how the home environment supports children's ability to select the relevant speech stream.

Finally, in order to broaden our understanding of how children process speech-in-noise, we must expand our participant base to include children from diverse backgrounds. Children who are from low-income homes may be particularly vulnerable to the effects of background noise and television (Mendelsohn et al., 2008). Much of the research with children living in low-income households has focused on the multiple environmental stressors that children experience in their home environment (Evans, 2013). Studying how children in low-income households process speech-in-noise will help us have a deeper understanding of how noise contributes to the cumulative risks children in poverty face.

Conclusion

How does background noise affect language development? The proposed research explored how elements of children's home auditory environments shape them into effective speech processors and word learners. Home environments are highly variable, and children's personal experience with this variability may have important consequences for their language development. By using naturalistic background noise, obtained from children's home environments, we examined whether children's local and global familiarity with background noise affects their ability to process speech-in-noise. Experimental studies of noise often compare children's performance in quiet to that in noise – establishing thresholds for children's ability to contend with speech-in-noise. However, children's homes are ever increasing in auditory complexity, making it evermore important to understand the limitations of children's abilities to adapt to adverse listening conditions.

References

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

Altmann, T.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73, 247-264.*

Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, *94*(2), B45-B53.

Barr, R., Lauricella, A., Zack, E., & Calvert, S. L. (2010). Infant and early childhood exposure to adult-directed and child-directed television programming: Relations with cognitive skills at age four. *Merrill-Palmer Quarterly*, *56*(1), 21-48.

Bernstein, B. (1964). Elaborated and restricted codes: Their social origins and some consequences. *American Anthropologist*, *66*(6), 55-69.

Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language, 11,* 557–578.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*(4), 417-436.

Christakis, D. A., Gilkerson, J., Richards, J. A., Zimmerman, F. J., Garrison, M. M., Xu, D., ... & Yapanel, U. (2009). Audible television and decreased adult words, infant vocalizations, and conversational turns: A population-based study. *Archives of Pediatrics & Adolescent Medicine*, *163*(6), 554-558.

Cohen, S., Glass, D.C., & Singer, J.E. (1973). Apartment noise, auditory discrimination and reading ability in children. *Journal of Experimental Social Psychology,* 9, 407-422.

Cohen, S., & Weinstein, N. (1981). Nonauditory effects of noise on behavior and health. *Journal of Social Issues*, *37*(1), 36-70.

Common Sense Media, & Rideout, V. (2013). *Zero to eight: Children's media use in America*. Common Sense Media.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121-136.

Evans, G. W., Maxwell, L. E., & Hart, B. (1999). Parental language and verbal responsiveness to children in crowded homes. *Developmental Psychology*, *35*(4), 1020.

Fenson L, Marchman VA, Thal D, Dale PS, Reznick JS, Bates E. MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual. 2. Baltimore, MD: Brookes; 2006.

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

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months.*Developmental Science,* 16(2)*,* 234-248.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98.

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year.*Psychological Science*, *9*(3), 228-231.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. *Developmental Psycholinguistics: On-line Methods in Children's Language Processing*, 113-132.

Fiese, B. H., Foley, K. P., & Spagnola, M. (2006). Routine and ritual elements in family mealtimes: Contexts for child well-being and family identity. *New Directions for Child and Adolescent Development*, *2006*(111), 67-89.

Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(2), 143-149.

Floor, P., & Akhtar, N. (2006). Can 18 - Month - Old Infants Learn Words by Listening In on Conversations?. *Infancy*, *9*(3), 327-339.

Goldinger, S. D., Kleider, H. M., & Shelley, E. (1999). The marriage of perception and memory: Creating two-way illusions with words and voices. *Memory & Cognition*, *27*(2), 328-338.

Grieco-Calub, T. M., Saffran, J. R., & Litovsky, R. Y. (2009). Spoken word recognition in toddlers who use cochlear implants. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1390-1400.

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

Hess, R. D., & Shipman, V. C. (1965). Early experience and the socialization of cognitive modes in children. *Child Development*, 36, 869-886.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. C*hild Development*, *74*(5), 1368-1378.

Hoff, E. (2006). Environmental Supports for Language Acquisition. *Handbook of Early Literacy Research*, *2*, 163.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343-365.

Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 240.

Johnstone, P. M., & Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical Society of America*, *120*(4), 2177-2189.

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.

Lapierre, M. A., Piotrowski, J. T., & Linebarger, D. L. (2012). Background television in the homes of US children. *Pediatrics*, *130*(5), 839-846.

Linebarger, D. L., & Walker, D. (2005). Infants' and toddlers' television viewing and language outcomes. *American Behavioral Scientist*, *48*(5), 624-645.

Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *The Journal of the Acoustical Society of America*, *117*(5), 3091-3099.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood.*Developmental Science*, *11*(3), F9-F16.

Matheson, M., Clark, C., Martin, R., van Kempen, E., Haines, M., Barrio, I. L., ... & Stansfeld, S. (2010). The effects of road traffic and aircraft noise exposure on children's episodic memory: The RANCH project. *Noise and Health*, *12*(49), 244.

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, *59*(3), 203-243.

Maxwell, L. E., & Evans, G. W. (2000). The effects of noise on pre-school children's pre-reading skills. *Journal of Environmental Psychology*, *20*(1), 91-97.

McMillan & Saffran (2016). Learning in complex environments: The effects of background speech on early word learning. *Child Development*.

Mendelsohn, A. L., Brockmeyer, C. A., Dreyer, B. P., Fierman, A. H., Berkule‐Silberman, S. B., & Tomopoulos, S. (2010). Do verbal interactions with infants during electronic media exposure mitigate adverse impacts on their language development as toddlers?. *Infant and Child Development*,*19*(6), 577-593.

Mirman, D. (2014). *Growth curve analysis and visualization using R*. CRC Press.

Nathanson, A. I., & Rasmussen, E. E. (2011). TV viewing compared to book reading and toy playing reduces responsive maternal communication with toddlers and preschoolers. *Human Communication Research*, *37*(4), 465-487.

Newman, R. S. (2003). Prosodic differences in mothers' speech to toddlers in quiet and noisy environments. *Applied Psycholinguistics*, *24*(04), 539-560.

Newman, R. S. (2005). The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology*, *41*(2), 352.

Newman, R. S. (2009). Infants' listening in multitalker environments: Effect of the number of background talkers. *Attention, Perception, & Psychophysics*, *71*(4), 822-836.

Newman, R. S. (2011). 2‐year‐olds' speech understanding in multitalker environments. *Infancy*, *16*(5), 447-470.

Newman, R. S., & Jusczyk, P. W. (1996). The cocktail party effect in infants. *Perception & Psychophysics*, *58*(8), 1145-1156.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, *60*(3), 355-376.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*(1), 42-46.

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

Rideout, V. (2014). Learning at home: Families' educational media use in America. *New York: Joan Ganz Cooney Center, January*.

Saffran, J.R., Aslin, R.N, & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science,* 274(5294), 1926-1928.

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.

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), 1105-1111.

U.S. Census Bureau. (2015). *Quick Facts*. Washington, D.C.: Government Printing Office. U.S. Census Bureau. (2015). Retrieved From https://www.census.gov/quickfacts/table/PST045215/00

Vandewater, E. A., Bickham, D. S., Lee, J. H., Cummings, H. M., Wartella, E. A., & Rideout, V. J. (2005). When the Television Is Always On Heavy Television Exposure and Young Children's Development. *American Behavioral Scientist*, *48*(5), 562-577.

Venker, C. E., Premo, E., Edwards, J., Saffran, J.R., & Ellis Weismer, S. (2016). Read the book vs. find the book: Use of informative verbs by children with autism spectrum disorder. *Poster presented at Symposium on Research in Child Language Disorders*.

Vernon-Feagans, L., Bratsch-Hines, M. E., & Family Life Project Key Investigators. (2013). Caregiver–child verbal interactions in child care: A buffer against poor language outcomes when maternal language input is less. *Early Childhood Research Quarterly*, *28*(4), 858-873.

Wachs, T. D. (1979). Proximal experience and early cognitive-intellectual development: The physical environment. *Merrill-Palmer Quarterly of Behavior and Development*, *25*(1), 3-41.

Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development*, *65*(2), 606-621.

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

Werker, J., & Tees, R.C, (1984) Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of Americas,* 75(6), 1866-1878.

Wightman, F. L., & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *The Journal of the Acoustical Society of America*, *118*(5), 3164-3176.

Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007). Associations between media viewing and language development in children under age 2 years. *The Journal of Pediatrics*, *151*(4), 364-368.

Appendix

Appendix A. Study 1 test trial order. 22-out-of 46 children received order B

Order A

Trial	Left Image	Right Image	Sound
1	Cookie	Shoe	Where's the shoe
2	Blicket	Tursey	Where's the tursey
3	Coro	Tursey	Find the coro
4	Blicket	Coro	Where's the blicket
5	Cat	Duck	Find the cat
6	Blicket	Coro	Where's the coro
7	Tursey	Blicket	Find the blicket
8	Tursey	Coro	Where's the tursey
9	Cookie	Shoe	Find the cookie
10	Coro	Tursey	Find the tursey
11	Blicket	Tursey	Find the blicket
12	Tursey	Coro	Where's the coro
13	Duck	Cat	Where's the duck
14	Coro	Blicket	Find the coro
15	Tursey	Blicket	Find the tursey
16	Coro	Blicket	Where's the blicket

Order B

Trial	Left Image	Right Image	Sound
1	Cookie	Shoe	Where's the shoe
2	Coro	Blicket	Where's the blicket
3	Tursey	Blicket	Find the tursey
4	Coro	Blicket	Find the coro
5	Cat	Duck	Find the cat
6	Tursey	Coro	Where's the coro
7	Blicket	Tursey	Find the blicket
8	Coro	Tursey	Find the tursey
9	Cookie	Shoe	Find the cookie
10	Tursey	Coro	Where's the tursey
11	Tursey	Blicket	Find the blicket
12	Blicket	Coro	Where's the coro
13	Duck	Cat	Where's the duck
14	Blicket	Coro	Where's the blicket
15	Coro	Tursey	Find the coro
16	Blicket	Tursey	Where's the tursey

Appendix B. Study 2 test trial order. 22-out-of 46 children received order B

Order A

Trial	Left Image	Right Image	Sound
1	ball3	door3	Look at the ball
2	bike3	juice3	Find the bike
3	ball4	cake4	Eat the cake
4	bike4	door4	Open the door
5	book3	cake3	Look at the book
6	juice2	bike2	Drink the juice
7	book1	cake1	Find the cake
8	door2	ball2	Open the door
9	juice4	bike4	Ride the bike
10	cake1	ball1	Find the ball
11	book2	juice2	Read the book
12	ball1	door1	Look at the door
13	juice3	book3	Find the juice
14	ball4	door4	Throw the ball
15	cake4	book4	Read the book
16	door3	bike3	Find the door
17	book4	juice4	Drink the juice
18	cake3	ball3	Look at the cake
19	bike2	door2	Ride the bike
20	juice1	book1	Find the book
21	cake2	cake2	Throw the ball
22	bike1	juice1	Look at the juice
23	cake2	book2	Eat the cake
24	door1	bike1	Look at the bike

Order B

Trial	Left Image	Right Image	Sound
1	ball1	door1	Look at the door
2	book1	cake1	Find the cake
3	juice2	bike2	Drink the juice
4	ball3	door3	Look at the ball
5	book2	juice2	Read the book
6	juice4	bike4	Ride the bike
7	cake3	ball3	Look at the cake
8	book4	juice4	Drink the juice
9	cake2	cake2	Throw the ball
10	bike3	juice3	Find the bike
11	cake2	book2	Eat the cake
12	door3	bike3	Find the door
13	cake4	book4	Read the book
14	cake1	ball1	Find the ball
15	door2	ball2	Open the door
16	book3	cake3	Look at the book
17	ball4	door4	Throw the ball
18	juice3	book3	Find the juice
19	door1	bike1	Look at the bike
20	ball4	cake4	Eat the cake
21	bike4	door4	Open the door
22	bike1	juice1	Look at the juice
23	juice1	book1	Find the book
24	bike2	door2	Ride the bike

Appendix C. Novel word learning GCA model

Accuracy ~ (Time + Time^2 + Time^3) * Condition + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.61	.03	23.69	<.001
Time	.54	.12	4.41	<.001
Time^2	-.14	.07	-1.85	.07
Time^3	-.09	.05	-1.56	.13
Unfamiliar Condition	.01	.04	.36	.72
Time*Unfamiliar Condition	-.18	.17	-1.04	.31
Time^2*Unfamiliar Condition	.003	.10	.03	.98
Time^3* Unfamiliar Condition	-.004	.077	-.056	.96

Random Effects		Variance	SD	Correlations		
Child	Intercept	.015	.12			
	Time	.35	.59	.21		
	Time^2	.12	.35	-.49	.02	
	Time^3	.065	.26	-.14	-.55	-.09
Residuals		.003	.05			

Appendix D. Novel word learning GCA model with overlapping noise

Accuracy ~ (Time + Time^2 + Time^3) * Condition*Noise + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.53	.05	11.66	<.001
Time	.06	.21	.29	.78
Time^2	-.019	.13	-.14	.88
Time^3	-.013	.10	-.13	.89
Unfamiliar Condition	.13	.07	1.83	.07
Overlapping Noise	.20	.09	2.15	.07
Time*Unfamiliar Condition	.34	.32	1.07	.29
Time^2*Unfamiliar Condition	-.16	.20	-.81	.42
Time^3* Unfamiliar Condition	-.03	.15	-.21	.83
Time*Overlapping Noise	.99	.44	1.69	.1
Time^2*Overlapping Noise	-.29	.28	-1.0	.31
Time^3* Overlapping Noise	-.18	.21	-.86	.40
Unfamiliar Condition*Overlapping Noise	-.17	.14	-.94	.16
Time*Unfamiliar Condition * Overlapping Noise	-.93	.66	-.95	.16
Time^2*Unfamiliar Condition * Overlapping Noise	.40	.41	.97	.34
Time^3* Unfamiliar Condition * Overlapping Noise	.07	.31	.23	.81

Random Effects		Variance	SD	Correlations		
Child	Intercept	.014	.12			
	Time	.30	.55	.11		
	Time^2	.12	.34	-.47	.09	
	Time^3	.064	.25	-.12	-.55	-.11
Residuals		.003	.058			

Appendix E. Novel word learning GCA model with electronic sounds

Accuracy ~ (Time + Time^2 + Time^3) * Condition*Electronic + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.63	.05	13.74	<.001
Time	.40	.22	1.86	.07
Time^2	-.16	.13	-1.23	.22
Time^3	-.07	.10	-.71	.48
Unfamiliar Condition	-.02	.07	-.35	.73
Electronic Sounds	-.20	.49	-.41	.69
Time*Unfamiliar Condition	.01	.31	.04	.97
Time^2*Unfamiliar Condition	.11	.18	.58	.57
Time^3* Unfamiliar Condition	-.06	.14	-.44	.66
Time*Electronic Sounds	1.80	2.3	.77	.44
Time^2*Electronic Sounds	.31	1.37	.22	.82
Time^3* Electronic Sounds	-.21	1.03	-.20	.84
Unfamiliar Condition*Electronic Sounds	.63	.94	.67	.51
Time*Unfamiliar Condition * Electronic Sounds	-2.91	4.50	-.65	.52
Time^2*Unfamiliar Condition * Electronic Sounds	-1.98	2.67	-.74	.46
Time^3* Unfamiliar Condition * Electronic Sounds	1.94	2.00	.52	.61

Random Effects		Variance	SD	Correlations		
Child	Intercept	.015	.12			
	Time	.35	.59	.23		
	Time^2	.12	.35	-.49	.01	
	Time^3	.065	.26	-.15	-.55	-.08
Residuals		.003	.006			

Appendix F. Familiar word recognition after neutral verb GCA model

Accuracy ~ (Time + Time^2 + Time^3) * Condition + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.67	.02	27.91	<.001
Time	.53	.13	4.19	<.001
Time^2	-.25	.08	-2.92	<.001
Time^3	-.22	.06	-3.59	<.001
Unfamiliar Condition	.01	.03	.38	.71
Time*Unfamiliar Condition	.19	.18	1.09	.28
Time^2*Unfamiliar Condition	-.05	.12	-.40	.69
Time^3* Unfamiliar Condition	-.05	.09	-.61	.54

Random Effects		Variance	SD	Correlations		
Child	Intercept	.013	.11			
	Time	.36	.60	-.07		
	Time^2	.17	.41	-.45	.04	
	Time^3	.09	.29	.08	-.54	-.22
Residuals		.003	.05			

Appendix G. Familiar word recognition after informative verb GCA model

Accuracy ~ (Time + Time^2 + Time^3) * Condition + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.58	.02	24.44	<.001
Time	.55	.12	4.71	<.001
Time^2	.06	.09	.344	.52
Time^3	-.05	.06	-.86	.39
Unfamiliar Condition	.07	.03	1.9	.05
Time*Unfamiliar Condition	.05	.17	.29	.77
Time^2*Unfamiliar Condition	-.03	.13	-.26	.79
Time^3* Unfamiliar Condition	-.02	.08	-.27	.79

Random Effects		Variance	SD	Correlations		
Child	Intercept	.013	.11			
	Time	.31	.56	-.03		
	Time^2	.18	.42	-.53	.06	
	Time^3	.07	.27	-.26	-.24	-.06
Residuals		.004	.06			

Appendix H. Familiar word recognition after neutral verb GCA model with overlapping noise

Accuracy ~ (Time + Time^2 + Time^3) * Condition*Noise + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.72	.04	16.61	<.001
Time	.11	.22	.49	.62
Time^2	-.24	.16	-1.54	.13
Time^3	-.09	.11	-.78	.44
Unfamiliar Condition	.04	.06	.65	.52
Overlapping Noise	-.10	.09	-1.1	.28
Time*Unfamiliar Condition	.60	.33	1.80	.08
Time^2*Unfamiliar Condition	-.01	.24	-.05	.96
Time^3* Unfamiliar Condition	-.32	.17	-1.91	.06
Time*Overlapping Noise	.63	.46	2.25	.30
Time^2*Overlapping Noise	-.02	.33	-.05	.96
Time^3* Overlapping Noise	-.33	.23	-1.46	.15
Unfamiliar Condition*Overlapping Noise	-.06	.13	-.46	.65
Time*Unfamiliar Condition * Overlapping Noise	-.10	.68	-1.46	.15
Time^2*Unfamiliar Condition * Overlapping Noise	-.08	.49	-.17	.87
Time^3* Unfamiliar Condition * Overlapping Noise	.64	.34	1.86	.07

Random Effects		Variance	SD	Correlations		
Child	Intercept	.012	.11			
	Time	.32	.57	-.02		
	Time^2	.17	.41	-.47	.04	
	Time^3	.08	.28	.10	-.52	-.23
Residuals		.003	.05			

Appendix I. Familiar word recognition after informative verb GCA model with overlapping noise

Accuracy ~ (Time + Time^2 + Time^3) * Condition*Noise + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.55	.04	12.50	<.001
Time	.95	.19	4.96	<.001
Time^2	.05	.16	.32	.76
Time^3	-.13	.11	-1.30	.21
Unfamiliar Condition	.10	.07	1.47	.15
Overlapping Noise	.09	.09	.98	.33
Time*Unfamiliar Condition	-.81	.29	-2.77	.008
Time^2*Unfamiliar Condition	.29	.24	1.21	.23
Time^3* Unfamiliar Condition	-.02	.16	-.12	.90
Time*Overlapping Noise	-.60	.40	-1.49	.16
Time^2*Overlapping Noise	.02	.33	.06	.96
Time^3* Overlapping Noise	.21	.22	.94	.35
Unfamiliar Condition*Overlapping Noise	-.08	.14	-.59	.56
Time*Unfamiliar Condition * Overlapping Noise	.90	.60	1.25	.12
Time^2*Unfamiliar Condition * Overlapping Noise	-.75	.49	-1.52	.14
Time^3* Unfamiliar Condition * Overlapping Noise	-.02	.33	-.07	.95

Random Effects		Variance	SD	Correlations		
Child	Intercept	.013	.11			
	Time	.24	.49	.01		
	Time^2	.16	.41	-.55	.19	
	Time^3	.07	.27	-.29	-.26	-.03
Residuals		.004	.06			

Appendix J. Familiar word recognition after neutral verb GCA model with electronic sounds

Accuracy ~ (Time + Time^2 + Time^3) * Condition*Electronic + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.65	.04	15.25	<.001
Time	.64	.22	2.89	.006
Time^2	.03	.14	.19	.85
Time^3	-.33	.11	-3.07	.004
Unfamiliar Condition	.01	.06	.22	.72
Electronic Sounds	.38	.45	.84	.41
Time*Unfamiliar Condition	-16	.32	.51	.31
Time^2*Unfamiliar Condition	-.41	.20	-2.02	.05
Time^3* Unfamiliar Condition	.05	.06	.34	.73
Time*Electronic Sounds	-1.47	2.35	-.63	.53
Time^2*Electronic Sounds	-3.55	1.52	-2.34	.03
Time^3* Electronic Sounds	1.43	1.15	1.24	.22
Unfamiliar Condition*Electronic Sounds	.22	.88	.26	.80
Time*Unfamiliar Condition * Electronic Sounds	-.21	4.57	-.05	96
Time^2*Unfamiliar Condition * Electronic Sounds	5.38	2.96	1.82	.08
Time^3* Unfamiliar Condition * Electronic Sounds	-1.34	2.25	-.60	.55

Random Effects		Variance	SD	Correlations		
Child	Intercept	.013	.11			
	Time	.35	.59	-.05		
	Time^2	.15	.38	-.45	.01	
	Time^3	.08	.29	.06	-.53	-.18
Residuals		.003	.05			

Appendix K. Familiar word recognition after informative verb GCA model with electronic sounds

Accuracy ~ (Time + Time^2 + Time^3) * Condition*Electronic + (Time + Time^2 + Time^3 | Child)

Fixed Effects	Estimate	SE	*t*	*p*
Intercept	.63	.04	15.07	<.001
Time	.56	.21	2.72	<.001
Time^2	.08	.16	.52	.61
Time^3	-.05	.10	-.50	.62
Unfamiliar Condition	.01	.06	.17	.87
Electronic Sounds	-.57	.44	-1.28	.21
Time*Unfamiliar Condition	.19	.30	.63	.53
Time^2*Unfamiliar Condition	-.19	.23	-.84	.98
Time^3* Unfamiliar Condition	.06	.15	.40	.69
Time*Electronic Sounds	-.11	2.18	-.05	.96
Time^2*Electronic Sounds	-.31	1.67	-.19	.99
Time^3* Electronic Sounds	.012	1.09	.011	.99
Unfamiliar Condition*Electronic Sounds	.80	.86	.93	.36
Time*Unfamiliar Condition * Electronic Sounds	-2.95	4.255	-.96	.49
Time^2*Unfamiliar Condition * Electronic Sounds	3.05	3.24	.94	.35
Time^3* Unfamiliar Condition * Electronic Sounds	-1.70	2.13	-.80	.43

Random Effects		Variance	SD	Correlations		
Child	Intercept	.013	.11			
	Time	.30	.55	-.03		
	Time^2	.18	.42	-.53	.08	
	Time^3	.07	.27	-.26	-.26	-.04
Residuals		.004	.06			