Title
Early auditory-semantic integration and organization: Behavioral and Electrophysiological Evidence

Author
Hendrickson, Kristi

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

SAN DIEGO STATE UNIVERSITY

Early auditory-semantic integration and organization: Behavioral and Electrophysiological Evidence

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in

Language and Communicative Disorders

by

Kristi Hendrickson

Committee in charge:

San Diego State University

Professor Margaret Friend, Chair
Professor Phillip Holcomb

University of California, San Diego

Professor Leslie Carver
Professor Sarah Creel
Professor Jeff Elman

2016

The dissertation of Kristi Hendrickson is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

San Diego State University

2016

DEDICATION

For my parents, Andy and Donna, and for my partner Jan.

This would not have been possible without your unconditional love and support.

EPIGRAPH

"There is no such thing as an empty space or an empty time. There is always something to see, something to hear. In fact, try as we may to make a silence, we cannot."

- John Cage

"There is no competition of sounds between a nightingale and a violin."

- Dejan Stojanovic

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deep appreciation and gratitude to my advisor, Dr. Margaret Friend, for the patient guidance and mentorship she provided to me. Dr. Friend's sincere concern for the training of her graduate students is unparalleled. Further, she has given me the freedom to explore new research areas and learn new skills related to the techniques used in this dissertation. I am truly fortunate to have had the opportunity to work with her.

I also would like to thank members of my committee, Drs. Leslie Carver, Sara Creel, Jeff Elman, and Phillip Holcomb, for their feedback and support throughout the dissertation processes. Your input has contributed greatly to this dissertation. I feel honored to have had the opportunity to work with each of you. I have also been especially privileged to learn from Drs. Tracy Love and Matthew Walenski who were central to the success of the ERP studies within this dissertation. Dr. Love, thank you for being there for me both as a mentor and as a supportive director. Dr. Walenski you went above and beyond what you were required to do. Your expertise and support was vital to the quality of studies 3 and 4 in this dissertation. Further, I would like to thank Dr. Alyson Abel for allowing me to conduct ERP research in her lab, and for her constant support, and to Dr. Faraz Farzin for inciting my passion for developmental research.

Thank you to members of the JDP for being by my side throughout this adventure. A very special thanks to Stephanie DeAnda, my partner in crime throughout my graduate career. I am indebted to Stephanie for several discussions that helped me to focus this dissertation. You know I could not have done this without you. Thank you to members of the Infant and Child Development Lab for being such a collaborative team. I am especially grateful to the team of research assistants who helped me gather data for this dissertation.

To my friends and family, thank you for always being there for me. Thank you Paulina and Blue Walichiewicz and Katie Wilson for adding a note of levity on this journey. To Barret, Victoria, Jaime, Ted, Grady, Charissa, Brent, Doug, Laurie, and the grandkids, for being a source of coziness when I needed a break from baby brains. My parents, Andy and Donna, receive my deepest gratitude and admiration. Words cannot even begin to express how much they were a part of this accomplishment. Thank you for instilling in me an unwillingness to give up. Your love and support has been absolute. You two are the tenderest of tunas. Finally Jan, you supported me in every way possible. You were a shoulder to cry on, a friend to laugh with, and a mentor to learn from. Thank you for holding up the fort while I was dissertating. Love you num.

Chapter 2, in full, is a reprint of material as it appears in Hendrickson, K., Mitsven, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2015). Looking and touching: What extant approaches reveal about the structure of early word knowledge. *Developmental Science*, *18*(5), 723-735. The dissertation author was the primary investigator and author of this paper. We are grateful to the participant families and also to Anya Mancillas, Stephanie DeAnda, Brooke Tompkins, Tara Rutter, and Kasey Palmer for assistance in data collection and coding. This research was supported by NIH #HD068458 to the senior author and NIH Training Grant 5T32DC007361 to the first author and does not necessarily represent the views of the National Institutes of Health.

Chapter 3, in full, is being prepared for submission for publication of the material, Hendrickson, K., Poulin-Dubois, D., Zesiger, P., & Friend, M. (in preparation). Assessing a continuum of lexical-semantic knowledge in the 2nd year: A multimodal approach The dissertation author was the primary investigator and primary author of this paper. We are grateful to the participant families and also to Kelly Kortright for assistance in data collection and coding. This research was supported by NIH #HD068458 to the senior authors and NIH Training Grant 5T32DC007361.

Chapter 4, in full, is a reprint of material as it appears in Hendrickson, K., Walenski, M., Friend, M., & Love, T. (2015). The organization of words and environmental sounds in memory. *Neuropsychologia*, *69*, 67-76. The dissertation author was the primary investigator and author of this paper. This research was supported by NIH R01DC009272 to the senior author, and the National Institutes of Health Training grant 5T32DC007361. We thank Zaira Flores, Kasey Palmer, and Jessica Onanian for assistance with data acquisition. Special thanks are extended to Dr. Phillip Holcomb and Dr. Jan R. Wessel for helpful comments on an earlier draft of this paper, and for their insights into data processing and analysis.

Chapter 5, in full, is being prepared for submission for publication of the material, Hendrickson, K., Love, T., Walenski, M., & Friend, M. (in preparation). The organization of words and environmental sounds in the 2nd year: Behavioral and Electrophysiological evidence. The dissertation author was the primary investigator and primary author of this paper. We gratefully acknowledge Alyson Abel for providing facilities for this research and Leslie Carver, Philip Holcomb, and Jan R. Wessel for their seasoned review of our ERP findings. We also thank Taylor Gums and Carly Horne for assistance with data acquisition. A special thanks is extended to the families who participated in this research. This research was supported by NIH Training Grant 5T32DC007361, NIH R01HD068458, and R01DC009272. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD, NIDCD, or the NIH.

VITA

EDUCATION

Ph.D. 2011 - 2016	Joint Doctoral Program in Language and Communicative Disorders, San Diego State University/University of California, San Diego.
B.A. 2003 - 2007	Psychology and Political Science, University of California, Davis.

RESEARCH EXPERIENCE

2011-present	Graduate Student, Department of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, CA.
2009-2011	Lab Manager, Language Acquisition Lab, Department of Linguistics, University of California, Los Angeles.
2007-2009	Junior Research Specialist/Lab Manager, Neurocognitive Development Lab, Center for Mind and Brain, University of California, Davis.

TEACHING EXPERIENCE

Fall 2014	Graduate Teaching Associate, San Diego State University, Speech, Language, and Hearing Sciences. Undergraduate course: Language Development and Disorders in Early Childhood,
Spring 2014	Teaching Assistant, University of California, San Diego, Human Development Program. Undergraduate course; Social Development.
Spring 2013	Teaching Assistant, University of California, San Diego, Human Development Program. Undergraduate course: Social Development,

MANUSCRIPTS

Hendrickson, K., & Sundara, M. (2016). Fourteen-month-olds' decontextualized understanding of words for absent objects. *Journal of child language*, 1-16.

Hendrickson, K., Mitsven, S., Poulin-Dubois, D., Zesiger, P., Friend, M. (2015). Looking and touching: What extant approaches reveal about the structure of early word knowledge. *Developmental Science,* 18(5), 723–*735*

Hendrickson, K., Friend, M., Wallenski, M., & Love, T. (2015). The organization of words and environmental sounds in memory. *Neuropsychologia, 69*, 67-76.

Hendrickson, K., & Friend, M. (2013). Quantifying the relationship between infants' haptic and visual response to word-object pairings. In *Proceedings of the 37th annual Boston University Conference on Language Development*. Sommerville, MA: Cascadilla Press.

AWARDS

2015	NIMH funding for the UC-Davis ERP Boot Camp Awarded to selected PhD-level researcher scientists / faculty
2014	Travel Award San Diego State University College of Health and Human Services
2014	Travel Award International Association for the Study of Child Language
2013	Research Award CSU Research Competition
2013	President's Award San Diego State University Student Research Symposium
2013	Library Award San Diego State University Student Research Symposium
2012	Travel Award Boston University Conference on Child Development
2011-2014	Training Grant NIH/NIDCD (T32DC007361) Neurocognitive Approaches to Communication Disorders. Awarded to selected Ph.D. students for predoctoral research

INVITED TALKS

Hendrickson, K., Love, T., Walenski, M., & Friend, M. (2016). Semantic organization of words and environmental sounds in 20-month-olds: An ERP study. Paper accepted at the International Conference on Infant Studies, New Orleans, LA.

Hendrickson, K., Poulin-Dubois, D., Zesiger, P., and Friend, M. (2014). The temporal dynamics of early behavioral measures of language. Paper presented at the International Association for the Study of Child Language, Amsterdam, NL.

Hendrickson, K.I., Friend, M. (2014). The impact of semantic relatedness on verbal and nonverbal sound processing: A Developmental ERP Study. Presented at the SDSU Developmental Science Forum, San Diego, CA.

Hendrickson, K.I., Friend, M. (2013). You can look but don't touch: The real-time dynamics between infants' visual and haptic behavior. Presented at the 28th Annual CSU Student Research Competition, San Diego, CA.

PROFESSIONAL ORGANIZATIONS

National Down Syndrome Society
National Fragile X Foundation
National Student Speech Language Hearing Association
Society for Neuroscience
Society for Research in Child Development
The International Society for Infant Studies

PROFESSIONAL SERVICES
Journal Reviewing
Child Development
NeuroImage

Graduate Student Representative

Fall 2013 – Winter 2016	University of California, San Diego, Graduate Student Representative
Fall 2012 – Winter 2016	San Diego State University, Graduate Student Representative

POSTER PRESENTATIONS

Hendrickson, K., Friend, M., Wallenski, M., & Love, T. (2014). The organization of words and environmental sounds in memory. Presented at the Neuroscience, Washington, DC.

Hendrickson, K., Friend, M., Wallenski, M., & Love-Geffen, T. (2014). The neural response to spoken words and environment sounds in toddlers and adults. Presented at the Boston University Conference on Language Development, Boston, Massachusetts.

Hendrickson, K., Friend, M., Wallenski, M., & Love-Geffen, T. (2014). The ontogeny of differential sound processing: ERP evidence from infants and adults. Presented at the International Conference on Infant Studies, Berlin.

Hendrickson, K.I., Mitsven, S., Poulin-Dubois, D., Zesiger, P., Friend, M. (April 2013). Quantifying the Relationship between Infant Visual Attention, Reaching, and Lexical Knowledge. Presented at the Biennial Society for Research in Child Development Meeting, Seattle, WA.

Hendrickson, K.I., & Friend, M. (November, 2012). Quantifying the relationship between infants' haptic and visual response to word-object pairings. Presented at the Boston University Conference on Language Development, Boston, Massachusetts.

Hendrickson, K.**I**., & Sundara, M. (November, 2010). Emerging referential understanding of words by 14-month-olds. Poster at the 2nd Pan American / Iberian Meeting on Acoustics, Cancun.

Hendrickson, K.I., Farzin, F., Hagerman, R.J., Rivera, S.M. (April, 2009). Changing developmental trajectories in young infants with fragile X syndrome. Presented at the Biennial Society for Research in Child Development Meeting, Denver, CO.

ABSTRACT OF THE DISSERTATION

Early auditory-semantic integration and organization: Behavioral and Electrophysiological Evidence

by

Kristi Hendrickson

Doctor of Philosophy in Language and Communicative Disorders

University of California, San Diego, 2016
San Diego State University, 2016

Margaret Friend, Chair

At its core, word learning and recognition concerns a relation between a low-level acoustic signal and a high-level semantic representation. Fundamental to the study of the early lexical-semantic system is the manner and degree to which the acoustic signal activates the semantic representation, and how meanings associated with words relate to one another. An implicit notion in the developmental literature is that upon hearing a word, young children activate the corresponding semantic representation in a dichotomous fashion – i.e., the semantic representation is either activated resulting in word recognition or not activated resulting in lack of recognition. Further, it is not

entirely known the degree to which very young children appreciate that word referents are meaningfully related.

To date, models of early auditory-semantic processing are primarily based on studies of meaningful relations between words and referents. However, lexical information (i.e., words) is not the only type of auditory input that is meaningful – environmental sounds (e.g., *dog barking* or *pen scribbling*) are nonverbal yet complex sounds that carry deep semantic associations with a corresponding referent. Therefore, a thorough investigation into the fundamental relation between acoustic signals and meaning requires an understanding of how meaning is associated with both lexical and non-lexical sounds.

A series of four studies with toddlers and adults are presented to investigate three related issues in the study of early auditory-semantic development: 1.) Is lexical knowledge dichotomous or continuous? 2.) How is lexical information semantically processed and organized? 3.) To what degree is such processing specific to language?

In the first set of studies, Chapters 2 and 3 present behavioral evidence that early word recognition is not binary, but instead graded, and partial knowledge plays a role in future learning. In the second set of studies, Chapters 4 and 5 replicate and extend findings that suggest the lexical-semantic system is organized as an interconnected semantic network both early in life and into adulthood. Further, these chapters present evidence that the electrophysiological markers of semantic processing are present during environmental sound processing in toddlers and adults, however words and environmental sounds appear to be organized somewhat differently, with a more consistent fine-grained structure for words compared to environmental sounds. Together, these studies provide behavioral and electrophysiological evidence to further our understanding of the nature of the early auditory-semantic system.

CHAPTER 1:

Introduction

Introduction

Lexical items are central to the processing of language. They play a role in our ability to perceive speech, parse syntactic structures, and processes and organize semantic information (Bloom, 2000; Fenson, Dale, Reznick, Bates, & Pethick, 1994; Mayor & Plunkett, 2010; McMurray, Horst, & Samuelson, 2013; Xu & Tenenbaum, 2007). Furthermore, children who demonstrate both delays in lexical comprehension and production are at the greatest risk for continued language delays and other developmental deficits (Desmarais, Sylvestre, Meyer, Bairati & Rouleau, 2008; Law, Boyle, Harris, Harkness & Nye, 2000). As a result, it is important for research to characterize the nature of the early lexical-semantic system.

How early lexical-semantic knowledge is formed and stored is still widely debated. The accounts that guide the discussion can be grouped into ‘domain-general’ and ‘domain-specific’. Accounts that adopt a domain-specific theoretical framework suggest that language acquisition is subserved by innate language-specific processes (Chomsky, 1980; Fodor, 1975; 1983; Gelman; 1990; Gleitman, 1986; Grodzinsky, 2000; Pinker & Jackendoff, 2005; Spelke, Breinlinger, Macomber, & Jaconson, 1992; Van der Lely, 2005). A key part of such accounts is the idea that words are phonological-conceptual chunks that are themselves meaningful. Words are said to reside in long-term memory in something akin to a mental dictionary – i.e., the mental lexicon. Due to the discrete, static nature in which words reside in the mental lexicon, word recognition is often explicitly or implicitly defined in an all-or-none stepwise fashion (Gallistel, Fairhurst, & Balsam, 2004; Jackendoff, 2002; Pinker & Jackendoff, 2005). What follows from the assumptions made by such models is words are synonymous with meaning, and therefore word retrieval and meaning retrieval are conflated. This results in a lexicon that is a passive repository of lexical entries (Elman, 1995; 2004).

A set of alternative, dynamic, domain-general accounts posit that the mechanisms involved in language processing are shared by other developing cognitive skills (Aslin & Newport, 2012; Aydelott, Kutas, & Federmeier, 2005; Kirkham, Slemmer, Johnson, 2002; Elman, 2004; 2009; 2014; Saffran & Thiessen, 2007; Smith, 2001). Based on this account, the specificity with which lexical knowledge is established is borne out of general learning mechanisms. Such learning mechanisms are accumulative, such that learning a word's meaning unfolds overtime as partial knowledge becomes more robust with experience (Yurovsky, Fricker, Yu, & Smith, 2014). Some accounts that fall under a more dynamic model suggest that words are clues to meaning but do not have meaning themselves, and therefore the "lexicon" as a tangible mental entity does not exist (Elman, 2004; 2009; 2014; Rumelhart, 1979).

Which set of accounts prove more accurate has implications for how we study early lexical-semantic understanding. For instance, accounts that support a view of the lexicon as a static entity are less concerned with how contextual influences relate to the depth and speed of lexical-semantic processing and how partial knowledge can be leveraged for future learning, whereas such investigations would be central to a dynamic account (Elman, 2014). Further, each account would generate differential predictions about whether a specialized semantic network subserves the processing of lexical items, or whether electrophysiological markers of semantic processing can be instantiated to meaningful auditory information that is not lexical.

Overview of Dissertation

The primary aim of this dissertation then is to investigate the nature – i.e., semantic integration and organizational structure – of the early lexical-semantic system, by evaluating three debated issues: 1.) Is lexical knowledge dichotomous or continuous? 2.) How is lexical information semantically processed and organized? 3.) To what degree is such processing specific to language? For this dissertation, I put forth a series of four studies to help answer each of these highly related questions.

Question 1: Is early lexical semantic knowledge all-or-none or continuous?

Historically, in the infant literature knowledge for several aspects of language processing (e.g., speech sound categorization, lexical-semantic understanding) is often implicitly defined as all-or-none. However, it has been documented that infant competence is highly task dependent, such that infants exhibit behavioral dissociations characterized by demonstrating knowledge in one modality (e.g., visual) but not the other (e.g., haptic) (Ahmed & Ruffman, 1998; Diamond, 1985; Hofstadter & Reznick, 1996; Ruffman, Garnham, Import & Connolly, 2001; Shinskey & Munakata, 2005). Discrepancies between results obtained visually and haptically were interpreted as evidence that haptic measures underestimate infant knowledge, as though knowledge were all-or-none. However, more recently, behavioral dissociations have been explained by the graded representations approach, which suggests that when two response modalities conflict, underlying knowledge may be partial (Morton & Munakata, 2002; Munakata 1998, 2001; Munakata & McClelland, 2003). The first study of this dissertation reported in Chapter 2, evaluates the relation between the primary paradigms in use for measuring early word knowledge, and applies the graded representations approach to behavioral dissociations to instantiate a continuum of lexical-semantic knowledge in the 2nd year of life.

The second study of this dissertation reported in Chapter 3, seeks to replicate results of behavioral dissociations and corresponding partial knowledge states obtained in Study 1, throughout the 2nd year, and investigate the role partial knowledge plays in future lexical-semantic processing. Accumulative theories of word recognition suggest that learning the associated visual referent for a word requires the accrual of partial knowledge states. From this view, partial lexical-semantic understanding has a pivotal role in future learning, as partial knowledge of a word-referent relation at an earlier time point can influence the degree to which that word-referent pairing is known at a later time point. Study 2 is a longitudinal follow-up study, which tested children from Study 1 6-months later (at 22- months) with three primary aims: 1. To re-evaluate the relation between the primary paradigms in use for measuring early word knowledge throughout the 2nd year of life. 2. To replicate the finding of visual and haptic behavioral dissociations and corresponding partial knowledge states in older children, and 3. To investigate the utility of accumulative learning theories in explaining the role of knowledge partiality in future word recognition.

Question 2: How is lexical information semantically processed and organized?

Lexical-semantic knowledge not only requires an appreciation of the relation between words and concepts, but also an understanding of how concepts relate to one another. The organization of lexical-semantic memory is based on several factors. Among the most important factors is featural similarity – the perceived likeness between concepts – which aids in categorization (Kay, 1971; Murphy, Hampton, & Milovanovic, 2012; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Sajin, & Connine, 2014). We know that feature information in the brain is structured such that neurons responding to similar features tend to be organized in a proximal fashion (Brugge & Merzenich, 1973; Hubel & Wiesel, 1972). Semantic representations associated with words can be represented as collection of features, and two words that share many features (e.g., dog and cat) will show similarities in their underlying neural activity compared to two words that share few semantic features (e.g., dog and pen) (Amunts & Zilles, 2012). The N400 – an ERP component closely tied to semantic processing – appears to follow this logic. That is, the amplitude of the N400 is incrementally sensitive to differences in the featural similarity of concepts to which words refer in adults and young children (Federmeier & Kutas 1999b; 2002; Federmeir, Mclennan, Ochoa, & Kutas, 2002; Ibanez, Lopez, & Cornejo, 2006; von Koss Torkildson et al., 2006).

The studies reported in Chapters 4 and 5 seek to replicate results of incremental or graded effects in N400 amplitude for words based on featural similarity in toddlers and adults using a visual-auditory match-mismatch paradigm (Federmeir & Kutas 1999; Federmeir et al., 2002).

Question 3: To what degree is semantic integration and organization specific to language?

Models of auditory-semantic processing are primarily based on studies of how meaning is associated with language. However, speech or verbal information is not the only type of auditory input that is meaningful – environmental sounds (e.g., *dog barking* or *pen scribbling*) are nonverbal yet complex sounds that carry deep semantic associations with a corresponding referent. Environmental sounds are ever-present, and identical across linguistic communities. Like language, it is likely that humans have developed capabilities in processing the raw acoustics of the sounds – spectral and temporal properties – and further, skills to integrate sounds and meaning (Ballas, 1993). In fact, it is logical to assume that because environmental sound processing preceded language processing, the neural resources that subserve recognition of environmental sounds precede language (Engelien et al., 2006). Therefore, to assess whether meaning integration of auditory information is qualitatively different for linguistic and nonlinguistic stimuli, researchers have investigated the relation between semantic processing for words vs. environmental sounds.

Work with non-lexical sounds without inherent meaning (e.g., tones), suggests that at 24-months a transition takes place such that non-lexical sounds formerly accepted as an object associate, are considered an unacceptable labels, whereas words maintain their symbolic status (Namy & Waxman, 1998; May & Werker, 2014). Thus it has been suggested that at 24-months children undergo a form of linguistic specialization in which they begin to understand the role that language – as opposed to other types of information – plays in organizing objects according to subtle differences in featural similarity (Namy, Campbell, & Tomasello, 2004). Conversely, a majority of behavioral and electrophysiological work demonstrates that semantic processing of words and inherently meaningful non-lexical sounds (i.e., environmental sounds) may be quite similar from toddlerhood to adulthood (Aramaki, Marie, Kronland-Martinet, Ystad, & Besson, 2010; Cummings et al.; 2006; 2008; 2009; 2010; Daltrozzo & Schön, 2009; Frey, Aramaki, & Besson, 2014; Orgs, Lange, Dombrowski, & Heil, 2008; Orgs et al., 2006; Plante, Van

Petten, & Senkfor, 2000; Schirmer, Soh, Penney, & Wyse, 2011; Schön, Ystad, Kronland- Martinet, & Besson, 2010; Van Petten & Rheinfelder, 1995)

Though words and environmental sounds appear to be conceptually processed similarly, work to date has compared word vs. environmental sound recognition using between-category distractors (e.g., cat and pen) for two alternative forced-choice tasks or priming paradigms, making the task of recognition substantially easier. Further, no work to date has examined whether there are differences in how words and environmental sounds are semantically organized in the brain at any point in development. In Chapters 4 and 5, we investigate how words and environmental sounds are organized in semantic memory in adults and toddlers. We assess whether the conceptualization of lexical information diverges from the conceptualization of meaningful non-lexical information before (20 months) and after (adults) a putative shift in linguistic specialization is said to occur (24 months).

References

Ahmed, A., & Ruffman, T. (1998). Why do infants make A not B errors in a search task, yet show memory for the location of hidden objects in a nonsearch task? *Developmental Psychology*, 34 (3), 441–453.

Amunts, K., Zilles, K. (2012). Architecture and organizational principles of Broca's area. *Trends in Cognitive Science*, *16*(8), 418–426.

Aramaki, M., Marie, C., Kronland-Martinet, R., Ystad, S., & Besson, M. (2010). Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. *Journal of Cognitive Neuroscience*, *22*(11), 2555-2569.

Aslin, R. N., & Newport, E. L. (2012). Statistical learning from acquiring specific items to forming general rules. *Current Directions in Psychological Science*, *21*(3), 170-176.

Aydelott, J., Kutas, M., & Federmeier, K. (2005). Perceptual and attentional factors in language comprehension: A domain-general approach. *Beyond nature-nurture: Essays in honor of Elizabeth Bates*, 281-314.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of experimental psychology: human perception and performance*, *19*(2), 250.

Bloom, L. (2000). The intentionality model of word learning: How to learn a word, any word. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), Becoming a word learner: A debate on lexical acquisition (pp. 19–50). Oxford, U.K.: Oxford University Press.

Chomsky, N. (1980). Rules and representations. *Behavioral and brain sciences*, *3*(01), 1-15.

Cummings, A., & Čeponienė, R. (2010). Verbal and nonverbal semantic processing in children with developmental language impairment. *Neuropsychologia*, *48*(1), 77-85

Cummings, A., Čeponienė, R., Dick, F., Saygin, A. P., & Townsend, J. (2008). A developmental ERP study of verbal and non-verbal semantic processing. *Brain research*, *1208*, 137-149.

Cummings, A., Čeponienė, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds.*Brain research*, *1115*(1), 92-107.

Cummings, A., Saygin, A. P., Bates, E., & Dick, F. (2009). Infants' recognition of meaningful verbal and nonverbal sounds. *Language Learning and Development*, *5*(3), 172-190.

Daltrozzo, J., & Schön, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience*, *21*(10), 1882-1892.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development*, 56, 868–883.

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders*, *43*(4), 361–389.

Elman, J. L. (1995). Language as a dynamical system. *Mind as motion: Explorations in* *the dynamics of cognition*, 195-223.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in cognitive sciences*, *8*(7), 301-306.

Elman, J. L. (2014). Systematicity in the Lexicon: On Having Your Cake and Eating It Too. In P. Calvo, J. Symons (Eds.), *The Architecture of Cognition: Rethinking Fodor and Pylyshyn's Systematicity Challenge*, (115-146) Cambridge, MA : MIT press.

Engelien, A., Tüscher, O., Hermans, W., Isenberg, N., Eidelberg, D., Frith, C., Stern, E. & Silbersweig, D. (2006). Functional neuroanatomy of non-verbal semantic sound processing in humans. *Journal of neural transmission*, *113*(5), 599-608.

Federmeier, K.D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*(4),469–495. ropsychologia.40,430–474.

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, *40*(7), 730-747.

Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133-146.

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

Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard University Press.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.

Frey, A., Aramaki, M., & Besson, M. (2014). Conceptual priming for realistic auditory scenes and for auditory words. *Brain and cognition*, *84*(1), 141-152.

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proceedings of the national academy of Sciences of the united States of america*, *101*(36), 13124-13131.

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate ‐ inanimate distinction as examples. *Cognitive Science*, *14*(1), 79-106.

Gelman, R. (1986). The child's understanding of number. Harvard University Press.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23-64.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and brain sciences*, *23*(01), 1-21.

Hofstadter, M.C., & Reznick, J.S. (1996). Response modality affects human infant delayed-response performance. *Child Development*, 67, 646–658.

Hubel, D. H., & Wiesel, T. N. (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *Journal of Comparative Neurology*, *146*(4), 421-450.

Ibáñez, A., López, V., & Cornejo, C. (2006). ERPs and contextual semantic discrimination: degrees of congruence in wakefulness and sleep. *Brain and language*, *98*(3), 264-275.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), 35-42.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.

Kay, P. (1971). Taxonomy and semantic contrast. *Language*, *47*,866–887.

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language & Communication Disorders*, *35*(2), 165–188.

May, L., & Werker, J. F. (2014). Can a Click be a Word?: Infants' Learning of Non-Native Words. *Infancy*, *19*(3), 281-300.

Mayor, J., & Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning. *Psychological review*, *117*(1), 1.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological review*, *119*(4), 831.

Merzenich, M. M., & Brugge, J. F. (1973). Representation of the cochlear partition on the superior temporal plane of the macaque monkey. *Brain research*, *50*(2), 275-296.

Morton, J.B., & Munakata, Y. (2002). Active versus latent representations: a neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, 40 (3), 255–265.

Munakata, Y. (1998). Infant perseverative and implications for object permanence theories: A PDP model of the A-not-B task. *Developmental Science*, 1, 161–84.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, 5 (7), 309–315.

Munakata, Y., & McClelland, J.L. (2003). Connectionist models of development. *Developmental Science*, 6, 413–429.

Murphy, G. L., Hampton, J. A., & Milovanovic, G. S. (2012). Semantic memory redux: An experimental test of hierarchical category representation.*Journal of memory and language*, *67*(4), 521-539.

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child development*, *69*(2), 295-308.

Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, *5*(1), 37-57.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2006). Conceptual priming for environmental sounds and words: An ERP study. *Brain and cognition*, *62*(3), 267-272.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2008). N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. *Neuroscience letters*, *436* (2), 133-137.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition*, *95*(2), 201-236.

Plante, E., Van Petten, C., & Senkfor, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia*, *38*(13), 1669-1684.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, *8*(3), 382-439.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief ? Charting transitions in knowledge. *Journal of Experimental Child Psychology*, 80 (3), 201–224.

Rumelhart, D. E. (1979). Some problems with the notion of literal meanings. *Metaphor and thought*, *72*.

Saffran, J. R., & Thiessen, E. D. (2007). Domain-General Learning Capacities. *Blackwell handbook of language development*, 68-86.

Sajin, S. M., & Connine, C. M. (2014). Semantic richness: The role of semantic features in processing spoken words. *Journal of Memory and Language*, *70*, 13-35.

Schön, D., Ystad, S., Kronland-Martinet, R., & Besson, M. (2010). The evocative power of sounds: Conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience*, *22*(5), 1026-1035.

Schirmer, A., Soh, Y. H., Penney, T. B., & Wyse, L. (2011). Perceptual and conceptual priming of environmental sounds. *Journal of cognitive neuroscience*, *23*(11), 3241-3253.

Shinskey, J.L., & Munakata, Y. (2005). Familiarity breeds searching: infants reverse their novelty preferences when reaching for hidden objects. *Psychological Science*, 16 (8), 596–600.

Smith, L. B. (2001). How domain-general processes may create domain-specific biases. *Language acquisition and conceptual development*, 101-131.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological review*, *99*(4), 605.

Van der Lely, H. K. (2005). Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends in cognitive sciences*, *9*(2), 53-59.

von Koss Torkildsen, J., Sannerud, T., Syversen, G., Thormodsen, R., Simonsen, H. G., Moen, I., Smith, L., & Lindgren, M. (2006). Semantic organization of basic-level words in 20-month-olds: An ERP study. *Journal of Neurolinguistics*, *19*(6), 431-454.

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia*, *33*(4), 485-508.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological review*, *114*(2), 245.

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

CHAPTER 2:

Behavioral Dissociations and Partial Knowledge

Developmental Science

Developmental Science (2014), pp 1–13 DOI: 10.1111/desc.12250

PAPER

Looking and touching: what extant approaches reveal about the structure of early word knowledge

Kristi Hendrickson,[1,2] Samantha Mitsven,[2] Diane Poulin-Dubois,[3] Pascal Zesiger[4] and Margaret Friend[2]

1. Department of Psychology, University of California, San Diego, USA
2. Department of Psychology, San Diego State University, USA
3. Department of Psychology, Concordia University, Canada
4. Department of Psychology, Université de Genève, Switzerland

Abstract

The goal of the current study is to assess the temporal dynamics of vision and action to evaluate the underlying word representations that guide infants' responses. Sixteen-month-old infants participated in a two-alternative forced-choice word–picture matching task. We conducted a moment-by-moment analysis of looking and reaching behaviors as they occurred in tandem to assess the speed with which a prompted word was processed (visual reaction time) as a function of the type of haptic response: Target, Distractor, or No Touch. Visual reaction times (visual RTs) were significantly slower during No Touches compared to Distractor and Target Touches, which were statistically indistinguishable. The finding that visual RTs were significantly faster during Distractor Touches compared to No Touches suggests that incorrect and absent haptic responses appear to index distinct knowledge states: incorrect responses are associated with partial knowledge whereas absent responses appear to reflect a true failure to map lexical items to their target referents. Further, we found that those children who were faster at processing words were also those children who exhibited better haptic performance. This research provides a methodological clarification on knowledge measured by the visual and haptic modalities and new evidence for a continuum of word knowledge in the second year of life.

Research highlights

- Moment-by-moment analysis of 16–18-month-olds' looking and reaching behavior as measures of early knowledge in a two-alternative forced-choice word–picture matching task.
- Assessed speed with which a word was processed (visual reaction time) as a function of the type of haptic response: Target, Distractor, or No Touch.
- Participants are significantly slower at processing a word during No Touches compared to Distractor and Target Touches, which were statistically indistinguishable.
- Results suggest that incorrect and absent haptic responses appear to index distinct knowledge states.
- Further, children who were faster at processing words were also those children who exhibited better haptic performance (i.e. more Target Touches).

Introduction

The use of visual and haptic measures to estimate underlying cognitive abilities has a rich history in research on infant development of spatial concepts, object knowledge, and early vocabulary comprehension among others. However, it has been documented that infant competence is highly task dependent, such that infants exhibit behavioral dissociations characterized by demonstrating knowledge in one modality but not the other (Ahmed & Ruffman, 1998; Diamond, 1985; Hofstadter & Reznick, 1996; Ruffman, Garnham, Import & Connolly, 2001; Shinskey & Munakata, 2005). One problem that remains to be addressed in such behavioral tasks is the differential interpretation of incorrect relative to absent responses. To date, there are limited empirical data to disambiguate these two response classes.

Address for correspondence: Kristi Hendrickson, 9242 Regents Road, Apt F, La Jolla, CA 92037, USA; e-mail: krhendrickson@ucsd.edu

Few attempts have been made to assess early knowledge in infants by measuring the visual and haptic response modalities as they occur within the same task (Diamond, 1985; Hofstadler & Reznick, 1996; Ruffman *et al*., 2001; Gurteen, Horne & Erjavec, 2011). Further, no study to date has measured the moment-by-moment relation between visual and haptic responses as measures of early knowledge. The benefits of such an examination are threefold: (1) to guide the interpretation of behavioral responses and non-responses, (2) to clarify the relation between volitional (e.g. haptic, verbal) and spontaneous (e.g. visual, orienting) responses more generally, and (3) to facilitate discussion concerning the underlying knowledge educed in paradigms employing visual and haptic response modalities.

The study of early language comprehension presents a particularly ripe area within which to investigate the broader dynamics between visual and haptic measures of early knowledge. At present there are three primary paradigms in use for the assessment of early comprehension vocabulary: parent report, visual attention, and haptic response. Most of what we currently know about visual and haptic responses as measures of early language abilities are from studies that have been conducted in a piecemeal fashion, in which investigators selectively use either looking time (Behrend, 1988; Fernald, McRoberts & Herrara, 1991; Fernald, Zangl, Portillo & Marchman, 2008; Golinkoff, Hirsh-Pasek, Cauley & Gordon, 1987; Hirsh-Pasek & Golinkoff, 1996; Houston-Price, Mather & Sakkalou, 2007; Naigles & Gelman, 1995; Reznick, 1990; Robinson, Shore, Hull Smith & Martinelli, 2000; Schafer & Plunkett, 1998; Styles & Plunkett, 2009; Thomas, Campos, Shucard, Ramsay & Shucard, 1981) or haptic response (Bates, Bretherton, Snyder, Beeghly, Shore, McNew, Carlson, Williamson, Garrison & O'Connell, 1988; Snyder, Bates & Bretherton, 1981; Woodward, Markman & Fitzsimmons, 1994; Friend & Keplinger, 2003, 2008) but not both.

Micro-level measures of looking behavior that assess speed of processing, and pattern of visual attention (Aslin, 2007) have gained prominence in the infant literature within the last decade and have offered interesting insights into underlying cognitive processes. The 'looking-while-listening' paradigm first outlined in Fernald, Pinto, Swingley, Weinberg and McRoberts (1998) has evolved from the well-documented Intermodal Preferential Looking (IPL) paradigm to an on-line measure of saccades in response to speech. Eye movements are monitored by digital camcorders and saccades are coded frame-by-frame to determine infants' speed in processing words. These continuous data yield a richer, more nuanced picture of language processing than do dichotomous measures obtained by parent report, or macro-level looking time measures. In addition, it has been shown that the speed with which words are processed and the size of children's lexicons at 25 months are predictive of intellectual functioning and language skills at 8 years of age (Marchman & Fernald, 2008).

Researchers utilizing haptic response measures of early language have obtained comparable findings to visually based measures (Friend & Keplinger, 2003, 2008; Ring & Fenson, 2000; Woodward *et al*., 1994). Friend and colleagues conducted a series of studies investigating the psychometric properties of the Computerized Comprehension Task (CCT), a measure that uses touch responses to gauge early word comprehension. The score on the CCT (proportion of correct touches to a named visual referent) was found to be a reliable and valid measure of word comprehension in the 2nd year of life, and a significant predictor of productive language abilities in the 3rd year. Additionally, performance on the CCT was significantly correlated with parent report on the MCDI: WG (Friend & Keplinger, 2003; 2008). Despite the predictive value of this measure, it suffers from a quandary that exists for all measures that require a volitional response, that is, does one interpret both incorrect and absent responses equivalently, or do these two response types systematically index different levels of understanding?

To our knowledge there has been only one study that has used both looking and touching as measures of early word knowledge. Using an interactive modification of the IPL paradigm, Gurteen *et al*. (2011) investigated 13- and 17-month-olds' familiar and novel word comprehension. For both familiar and novel words, infants participated in two types of test trials: one requiring a looking response and one a touching response. There was no significant relation between MCDI: WG (comprehension or production) and target looking on the familiar or novel preferential looking tasks. The relation between infants' touching responses and MCDI: WG scores was not reported. For both novel and familiar words, 13- and 17-month-olds looked significantly longer to the target referent. However, across age, infants reached toward the target at a level greater than chance only for familiar words. These discrepant findings for visual and haptic behavior when measuring familiar versus novel word knowledge bring into question whether spontaneous and volitional measures more generally should be thought of as analogous. However, in Gurteen *et al*. (2011) looking and touching behaviors were assessed separately, thus the concurrent relationship between the modalities is still largely unknown.

One reason for the dearth of research on the synchronous relation between response modalities is that the task design must take into account the natural dependencies between response modalities. For example, there is evidence of cortical movement preparation in adults as early as 500 ms prior to voluntary hand movement and the decision to respond occurs ~200 ms before execution (Trevena & Miller, 2002; Gold & Shadlen, 2007; Romo & Salinas, 1999; Schall & Thompson, 1999; Shadlen & Newsome, 1996; Libet, Gleason, Wright & Pearl, 1983; VanRullen & Thorpe, 2001). Thus, looking to the target pre-reach onset may be influenced by neural activity related to movement anticipation. It has been shown that haptic responses can take up to ~7 secs post-stimulus to execute (Friend, Schmitt & Simpson, 2012). Looking responses can, and should, be captured within ~2 secs post-stimulus; the further from stimulus onset that looks occur, the less likely they are to be influenced by stimulus parameters, and the more likely they are to reflect processes other than comprehension of the target word (Aslin, 2007; Fernald, Perfors & Marchman, 2006; Swingley & Fernald, 2002). Therefore, due to differences in the relative timing of the visual and haptic modalities to respond, it is possible to acquire visual fixation data sufficiently early in the trial to minimize the effect of motor planning.

For the purpose of the present investigation then, a traditional measure of the macrostructure of looking time (e.g. proportion looking to target) would be confounded with information processing at every level in the task from the recognition of a word–referent relation to the preparation for and execution of the reach (Aslin, 2007). Given these considerations, we employed a micro-level measure of looking in the early post-stimulus period to maximize the stimulus–response contingency and minimize any influence by the decision to act.

The overarching goal of the current study is to assess the simultaneous moment-by-moment bidirectional relation between visual and haptic responses as measures of early word comprehension, and to evaluate the implications of our findings for the structure of early lexical-semantic knowledge. Children participated in a modified combination of the CCT (Friend & Keplinger, 2003, 2008) and the looking-while-listening procedure (Fernald *et al*., 1998, 2006, 2008; Fernald, McRoberts & Swingley, 2001). Participants were presented with within-category pairs of images (e.g. dog and cat) on a touch sensitive monitor and prompted to touch one of the images (e.g. 'dog'; Target), while their visual behavior and haptic responses were recorded concurrently on video. We analyzed infants' looking behavior at every 40 ms interval from image onset during the presentation of the target word on distractor-initial trials (i.e. those trials for which infants first fixated the distractor image upon hearing the target word). We calculated visual reaction time (RT) operationalized as the latency to shift from the distractor to the target image once both the target word in the first sentence prompt and the visual stimuli were presented. It is necessary to use distractor-initial trials in order to calculate the speed with which children shift fixation to the target. Visual RT calculated in this way is a measure of the efficiency of word processing and predicts subsequent development (Fernald *et al*., 1998, 2001, 2008). By definition, this measure cannot be obtained when children fixate on the target initially (Fernald *et al*., 2008). For each participant, trials were grouped by haptic response (Target, Distractor, No Touch). In the current study we ask whether there are differences in visual RT across Target, Distractor, and No Touches. There are two primary patterns of interest.

We predict that visual RTs and haptic responses will converge when word knowledge is most robust (i.e. visual RTs will be fastest during Target Touches). Of particular interest is the comparison between visual RTs during Distractor Touches and No Touches. One possibility is that both incorrect and absent responses reflect weak or nonexistent lexical access. From this view we would expect visual RTs during Distractor and No Touches to be indistinguishable, and slower than Target Touches. Another possibility is that these two response types gauge different capabilities in lexical access. Here, we would expect visual RTs for Distractor and No Touches to be significantly different, and likely, slower than Target Touches.

Moreover, no study to date has examined the relation between haptic, visual, and parent report measures within the same cohort of infants. Therefore, a coextending goal of the current study is to examine the correlations between children's vocabulary knowledge indexed by the haptic modality and children's speed of lexical access indexed by visual RT, and the well-documented MCDI: WG (Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly, 1993).

Method

Participants

Participants were drawn from a larger, multi-institutional longitudinal project assessing language comprehension in the second year of life. Infants were obtained through a database of parent volunteers recruited through birth records, internet resources, and community events in a large metropolitan area. All infants were full-term and had no diagnosed impairments in hearing or vision.

Seven infants were excluded from the study because of excessive fussiness ($n = 4$), experimenter error ($n = 1$), and technical error ($n = 2$). The final sample included 61 monolingual English-speaking infants (33 females, 28 males), ranging in age from 15.5 to 18.2 months (M = 16.6 months). Infant language exposure was assessed using an electronic version of the Language Exposure Questionnaire (Bosch & Sebastián-Gallés, 2001). Estimates of daily language exposure were derived from parent reports of the number of hours of language input by parents, relatives and other caregivers in contact with the infant. Only those infants with at least 80% language exposure to English were included in the study.

Apparatus

The study was conducted in a sound attenuated room. A 51 cm 3M SCT3250EX touch capacitive monitor was attached to an adjustable wall-mounted bracket that was hidden behind blackout curtains and between two portable partitions. Two HD video cameras were used to record participants' visual and haptic responses. The eye-tracking camera was mounted directly above the touch monitor and recorded visual fixations through a small opening in the curtains. The haptic-tracking camera was mounted on the wall above and behind the touch monitor to capture both the infants' haptic response and the stimulus pair presented on the touch monitor. Two audio speakers were positioned to the right and left of the touch monitor behind the blackout curtains for the presentation of auditory reinforcers to maintain interest and compliance.

Procedure and measures

Upon entering the testing room, infants were seated on their caregiver's lap centered at approximately 30 cm from the touch-sensitive monitor with the experimenter seated just to the right. Parents wore blackout glasses and noise-cancelling headphones to mitigate parental influence during the task. The assessment followed the protocol for the Computerized Comprehension Task (CCT; Friend & Keplinger, 2003, 2008). The CCT is an experimenter-controlled assessment that uses infants' haptic response to measure early decontextualized word knowledge. A previous attempt has been made to automate the procedure, such that verbal prompts come from the audio speakers positioned behind the touch screen instead of the experimenter seated to the right of the child. Pilot data using the automated version showed that children's interest in the task waned to such an extent that attrition rates approached 85% (attrition rates using the experimenter-controlled CCT are between 5 and 10%; M. Friend, personal communication, 17 June 2014; P. Zesiger, personal communication, 21 May 2014). Therefore, to collect a sufficient amount of data to yield effects we used the well-documented protocol of the CCT (Friend & Keplinger, 2003, 2008). Previous studies have reported that the CCT has strong internal consistency (Form A α = .836; Form B α = .839), converges with parent report (*partial r controlling for age* = .361, $p < .01$), and predicts subsequent language production (Friend *et al.*, 2012). In addition, responses on the CCT are nonrandom (Friend & Keplinger, 2008) and this finding replicates across languages (Friend & Zesiger, 2011) and monolinguals and bilinguals (Poulin-Dubois, Bialystok, Blaye, Polonia & Yott, 2013).

For this procedure, infants are prompted to touch images on the monitor by an experimenter seated to their right (e.g. 'Where's the *dog?* Touch *dog!*'). Target touches (e.g. touching the image of the dog) elicit congruous auditory feedback over the audio speakers (e.g. the sound of a dog barking). Infants were presented with four training trials, 41 test trials, and 13 reliability trials in a two-alternative forced-choice procedure. For a given trial, two images appeared simultaneously on the right and left side of the touch monitor. The side on which the target image appeared was presented in pseudo-random order across trials such that target images could not appear on the same side on more than two consecutive trials, and the target was presented with equal frequency on both sides of the screen (Hirsh-Pasek & Golinkoff, 1996). The item that served as the target was counterbalanced across participants such that there were two forms of the procedure. All image pairs presented during training, testing, and reliability were matched for word difficulty (easy, medium, hard) based on MCDI: WG norms (Dale & Fenson, 1996), part of speech (noun, adjective, verb), category (animal, human, object), and visual salience (color, size, luminance). The design of the study relied on participants producing Target, Distractor, and No Touches in sufficient numbers, and at roughly similar rates in order to address our hypotheses. Thus including words at varying degrees of difficulty was crucial.

The study began with a training phase to ensure that participants understood the nature of the task. During the training phase, participants were presented with early-acquired noun pairs (known by at least 80% of 16-month-olds; Dale & Fenson, 1996) and prompted by the experimenter to touch the target. If the infant failed to touch the screen after repeated prompts, the experimenter touched the target image for them. If a participant failed to touch during training, the four training trials were repeated once. Only participants who executed at least one correct touch during the training phase proceeded to the testing phase.

During testing, each trial lasted until the infant touched the screen or until 7 seconds had elapsed at which point the image pair disappeared. When the infant's gaze was directed toward the touch monitor, the experimenter delivered the prompt in infant-directed speech and advanced each trial as they uttered the target word in the first sentence prompt such that the onset of the target word occurred just prior to the onset of the visual stimuli (average interval = 238 ms).

Nouns: *Where is the _____? Touch _____.*
Verbs: *Who is _____? Touch _____.*
Adjectives: *Which one is _____? Touch _____.*

The criterion for ending testing was a failure to touch on two consecutive trials with two attempts by the experimenter to re-engage without success. If the attempts to re-engage were unsuccessful and the child was fussy, the task was terminated and the responses up to that point were taken as the final score. However, if the child did not touch for two or more consecutive trials but was not fussy, testing continued. Those participants who remained quiet and alert for the full 41 test trials (n = 34) also participated in a reliability phase in which 13 of the test trial image pairs were re-presented in opposite left–right orientation.

Parent report of infant word comprehension was measured using the MCDI: WG, a parent report checklist of language comprehension and production developed by Fenson *et al*. (1993), which has demonstrated good test–retest reliability and significant convergent validity with an object selection task (Fenson, Dale, Reznick, Bates, Thal, Pethick, Tomasello, Mervis & Stiles, 1994). Of interest in the current study was the 396-item vocabulary checklist for comparison with the infants' behavioral data.

Coding

A waveform of the experimenter's prompts was extracted from the eye-tracking video –positioned approximately 30 cm from the experimenter – using Audacity® software (http://audacity.sourceforge.net/). Subsequently, the eye-tracking video, haptic-tracking video, and a waveform of the experimenter's prompts were all synced using Eudico Linguistics Annotator (ELAN) (<http://tla.mpi.nl/tools/tla-tools/elan/>, Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands; Lausberg & Sloetjes, 2009). ELAN is a multi-media annotation tool specifically designed for the analysis of language. It is particularly useful for integrating coding across modalities and media sources because it allows for the synchronous playing of multiple audio tracks and videos. Only distractor-initial trials – those trials for which infants first fixated the distractor image upon hearing the target word – were included in the analyses of looking behavior.

Coders completed extensive training to identify the characteristics of speech sounds within a waveform, both in isolation and in the presence of coarticulation. Because a finite set of target words always followed the same carrier phrases (e.g. 'Where is the ____', 'Who is ___', or 'Which one is ____'?), training included identifying different vowel and consonant onsets after the words 'the' and 'is'. Coders were also trained to demarcate the onset of vowel-initial and nasal-initial words after a vowel-final word in continuous speech, which can be difficult using acoustic waveforms in isolation.

Coders were required to practice on a set of files previously coded by the first author with supervision and then to code one video independently until correspondence with previously coded data was reached. Two coders completed each pass, each coding ~50% of the data.

Trials with short latencies (200–400 ms) likely reflect eye movements that were planned prior to hearing the target word (Fernald *et al*., 2008; Bailey & Plunkett, 2002; Ballem & Plunkett, 2005). For this reason trials were included in subsequent analyses if the participant looked at the screen for at least 400 ms. In addition, looking responses were coded during the first 2000 ms of each trial. As previously mentioned, looking responses that are further from the stimulus onset are less likely to be driven by stimulus parameters (Aslin 2007; Fernald *et al*., 2006; Swingley & Fernald, 2002). Finally, by coding the first 2000 ms we are largely restricting our analysis to the period prior to the decision to touch.

Coding occurred in two passes. Coder 1 annotated the frame onset and offset of the target word as it occurred in the first sentence prompt using the waveform of the experimenter's speech. First the coder listened to the audio and zoomed in on the portion of the waveform that contained the target word in the first sentence prompt (e.g. Where is the DOG?). Once that section was magnified, the coder listened to the word several times precisely demarcating the onset and offset of speech information within the larger waveform. Coder 1 also marked the frame in which the visual stimuli appeared on the screen and the side of the target referent (note: side of the target referent was hidden from Coder 2). Coder 2 coded visual and haptic responses with no audio to ensure that she remained blind to the image that constituted the target. Coding began at image onset, roughly 238 ms after target word onset, and prior to target word offset in the first sentence prompt (see Figure 1). For the visual behavior, Coder 2 advanced the video and coded each time a change in looking behavior occurred using three event codes: right look, left look, and away look. For sustained visual fixations, Coder 2 advanced the video in 40 ms coding frames, and, because shifts in looking are crucial for deriving measures of reaction time, Coder 2 advanced the video during gaze shifts at a finer level of resolution (3 ms).

Haptic responses were coded over the course of the entire trial (7 secs). Only initial haptic responses were coded. The haptic response was coded categorically: Left Touch (unambiguous touch to the left image), Right Touch (unambiguous touch to the right image) or No Touch (no haptic response executed). Identifying touches as Target or Distractor was done post hoc, to preserve coders' blindness to target image and location.

Inter-rater reliability coding was conducted for both visual and haptic responses by a third, reliability coder. For looking responses a random sample of 11 videos (~ 25% of the data) was selected. Because our dependent variable (visual RT) relies on millisecond precision in determining when a shift in looking behavior occurred, only those frames in which shifts occurred were considered for the reliability score. This score is more stringent than including all possible coding frames because the likelihood of the two coders agreeing is considerably higher during sustained fixations compared to gaze shifts (Fernald *et al.*, 1998). Using this shift-specific reliability calculation, we found that on 90% of trials coders were within one frame (40 ms) of each other, and on 94% of the trials coders were within two frames (80 ms) of each other.

All haptic response coding was compared to offline coding of haptic touch location completed for the larger longitudinal project. Inter-rater agreement for the haptic responses was 95%. All haptic coding was completed blind to target image, location, and visual fixations.

Results

Calculating reaction time by including only distractor-initial trials and a narrow time window restricts the number of usable trials per condition. Consequently not all children contributed data to all experimental conditions and thus were removed from further analysis. Of the 61 infants originally included, 16 participants were excluded from subsequent analyses for not contributing data to all three haptic type conditions. The remaining 45 infants completed an average of 36 out of a possible 41 trials and their average MCDI: WG comprehension vocabulary was 188 words out of a possible 396 and ranged from 62 to 342 words (percentile range = 1st to

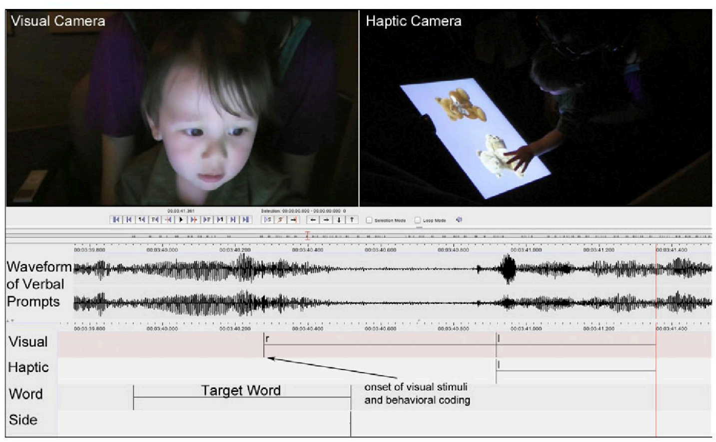

Figure 1 *Eudico Linguistic Annotator coding setup. The waveform of the experimenter's prompt is extracted from the video camera recording the visual behavior. The waveform is then synced with the video from the visual and haptic cameras. Coding was done using four tiers. On the Visual tier the onset of the visual stimuli was coded, and looking behavior was coded: right look (r), left look (l), or away look (a). On the Haptic tier the onset and offset of the haptic response and the direction of the touch (r or l) was coded. On the Word tier the onset and offset of the target word in the first sentence prompt was marked by viewing the waveform and using frame-by-frame auditory analysis. Finally, on the Side tier, the side (left or right) the target word appeared was coded and hidden from view. Behavioral coding began at the onset of the visual stimulus, which occurred ~ 238 ms after the target word in the first sentence prompt was uttered.*

91st). The average time to execute a haptic response was 3896.25 ms post image onset (< 14% of trials included a haptic response prior to 2000 ms). The average visual RT to shift to the target across haptic types was 862.43 ms, comparable to the mean visual RT found in similarly aged participants in previous research (827 ms; Fernald *et al*., 1998). Consistent with the literature, immediate test–retest reliability was strong for participants who completed reliability in the larger longitudinal project [$r(41) = .74$, $p < .0001$], and in the subset of data used for the current project [$r(32) = .67$, $p < .0001$]. Finally, internal consistency (Form A $\alpha = .931$ and Form B $\alpha = .940$) was excellent, indicating consistency within the test and between test and reliability phases.

Infants chose the target image on 11.78 trials ($SD = 6.76$), the distractor on 10.08 trials ($SD = 4.30$), and provided no haptic response on 13.03 trials ($SD = 7.78$). Thus Target, Distractor, and No Touches were elicited at roughly equal rates. This pattern of findings is expected given the task design. There are equal numbers of easy (comprehension ≥ 66%), moderately difficult (comprehension = 33–66%), and difficult words (comprehension < 33%) based on normative data at 16 months of age (Dale & Fenson, 1996). Therefore, if children correctly completed all 41 trials, and identified all words that most 16-month-olds are reported by parents to know, they would earn a score of roughly 14, indicating good knowledge of the easiest items.

To test the notion that children are more likely to execute target touches for highly known words, we analyzed haptic responses for those words for which there was a high probability of a target response (proportion of 16-month-olds expected to know each item appears in parentheses): ball (96.4%), juice (91.7%), dog (90.5%), and bottle (89.3%). This subset of words (59 trials in all) elicited a total of 35 Target Touches, 15 Distractor Touches, and 9 No Touches, indicating that, as expected, when children touched the screen they made a correct haptic response significantly more often than chance by Binomial Test (exact), $p = .007$. This, in conjunction with excellent internal consistency and strong test–retest reliability, reveals that children's responses were nonrandom.

To assess the potential contribution of side-bias effects to performance, we conducted two-sample t-tests and found no significant difference in number of touches, $t(88) = 1.8$, $p = .08$, or amount of looking time, $t(88) = 1.5$, $p = .13$, to images presented on the left relative to the right. To summarize, given the structure of the test and the average number of trials completed, the average number of target touches reported here is in line with expectations for performance at this age and is consistent with previous reports on the CCT (Friend & Keplinger, 2008; Friend & Zesiger, 2011). In addition, both internal consistency and test–retest reliability were strong and we found no evidence of side-bias effects. However, a two-sample t-test revealed a significant difference in visual RTs across Forms, $t(88) = 4.11$, $p = .0002$, but no difference in the number of target haptic responses. To determine whether this difference influenced our findings, we first analyzed visual RTs as a function of haptic response type by Form. The pattern of results was identical for both Forms and we report our findings collapsed across Form below.

Concurrent analyses of visual and haptic responses

The time-course of eye movements across the different haptic types for the first 2000 ms from visual onset can be seen in the onset-contingency plot (see Figure 2). As predicted, during Target Touches, infants shifted their gaze toward the target image rapidly following the target word. In contrast, during No Touches, infants were slower to fixate the target image. Distractor Touches appear to have an intermediate rising slope, however, roughly 1300 ms post image onset looking to the target image plateaus, and shifts towards the distractor.

We compared speed of processing across Haptic Types (Target, Distractor, No Touch) using visual RT. Average visual RTs were calculated for distractor-initial trials in which a shift in gaze occurred between 400 and 2000 ms post-visual onset. Visual RT was averaged for each participant by Haptic Type and subjected to a one-way

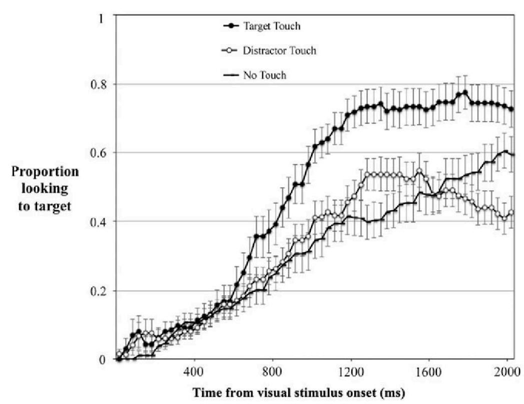

Figure 2 *Time-course analysis. Each data point represents the mean proportion looking to the target location at every 40 ms interval from the onset of the visual stimulus for each Haptic Type (Target, Distractor, No Touch); error bars show the standard error across participants.*

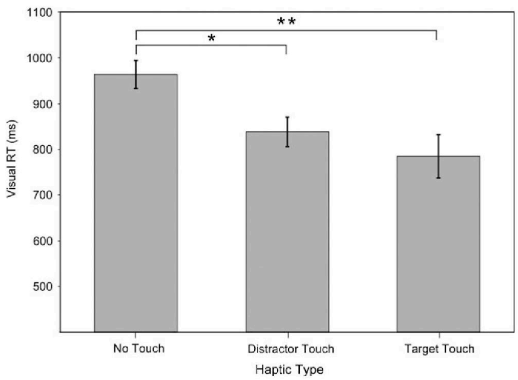

Figure 3 *Visual RT analysis. Mean visual RT to shift gaze from the distractor to the target image following the onset of the visual stimulus on distractor-initial trials.Note: Error bars show the standard error across participants. * p < .04; ** p < .01.*

ANOVA. There was a main effect of Haptic Type, $F(2, 43) = 6.8$, $p = .003$ (see Figure 3). Planned pairwise comparisons using a Bonferonni correction were conducted on the three levels of Haptic Type. As expected, infants processed the target word significantly faster during Target Touches (M = 784.76, SD = 30.87) compared to No Touches (M = 964.02, SD = 38.14). Interestingly, visual RTs were also significantly faster during Distractor Touches (M = 838.50, SD = 32.72) compared to No Touches. Finally, although visual RTs were faster for Target Touches than for Distractor Touches, this difference did not reach significance. Thus, statistically, infants shifted their gaze equally rapidly on Target and Distractor Touches. To ensure that this pattern of effects held for words that children this age had a high probability of knowing, and thus more closely mimic the majority of data collected using looking-while-listening procedure (Fernald *et al.*, 1998, 2008), we calculated the visual RT for highly familiar words (Dale & Fenson, 1996) in our task (ball, juice, dog, bottle). Consistent with the pattern observed for the full data set, visual RTs for Target Touches were fastest (M = 748 ms), followed by Distractor Touches (M = 787 ms), and finally, No Touches (M = 992 ms).

Relation of visual, haptic, and parent report measures

A series of Pearson's product-moment correlations was performed to analyze the relation between each of our response measures (visual RT and haptic) and MCDI: WG comprehension scores (the parent reported number of words understood by the child). For these analyses, the haptic measure was calculated in two ways: (1) as the number of Target Touches executed by the child, and (2) as a proportion of Target Touches (i.e. number of Target Touches divided by the total number of trials completed). Further, the visual RT measure was calculated across haptic response type in two ways: (1) for only those words reported by parents as 'known' and (2) for all words. There was no significant relation between visual RT for 'known' words and MCDI comprehension, and although the direction of the relation between visual RT for all words and MCDI comprehension was in the expected negative direction, the correlation was not significant ($r = -.16$, $p = .29$). The correlations for both haptic measures and MCDI comprehension were significant: the proportion of Target Touches ($r = .32$, $p = .03$), and the number of Target Touches ($r = .32$, $p = .03$). Finally, although the correlations comparing visual RT for all words and both haptic measures were not significant, the correlations between visual RT for 'known' words and each haptic measure were significant: proportion of Target Touches ($r = -.41$, $p = .009$), and number of Target Touches ($r = -.40$, $p = .01$) (see Table 1 for a summary of the correlation results).

Table 1 *Correlation (r) between visual, haptic, and parent report measures*

	Visual RT		
	All Words	"Known" Words	MCDI Comprehension
Visual RT			
All Words			-.16
"Known" words			.004
Haptic			
Number of Target Touches	-.12	.41**	.31*
Proportion of Target Touches	-.11	-.40**	.32*

*p < .05; **p < .01.

Discussion

In the current study we measured the dynamics of visual attention vis-à-vis haptic responses to examine the relation between two widely accepted measures of young children's language abilities and to facilitate the interpretation of incorrect in contrast to absent volitional responses. There are several metrics (average fixation duration, total looks, proportion looking, etc.) available to operationalize infant looking behavior. For the current investigation we were interested in linking looking behavior to infants' underlying lexical-semantic access by using a measure that is largely independent of subsequent action. Therefore we utilized a micro-level measure of visual attention (visual RT) known to gauge the speed with which infants process word–visual referent pairings. We found that visual RTs to shift gaze from the distractor to the target image varied as a function of whether a reach was subsequently executed (Touch vs. No Touch), but not where the reach was executed (Target vs. Distractor Touch). That is, infants quickly shifted visual attention on trials on which they touched either the target or the distractor image.

Of interest are the implications of these results for the structure of early lexical-semantic knowledge, particularly with respect to whether incorrect and absent volitional responses can be collectively bundled as representing lack of knowledge, or whether each indexes different capabilities in lexical access. A useful first step in understanding potential differences between these response types is to interpret the existence of behavioral dissociations during Distractor Touches (i.e. rapid visual RTs during incorrect haptic responses), but not during No Touches (i.e. visual and haptic behavior converge: slow visual RTs and absent haptic responses).

Traditionally discrepancies between results obtained visually and haptically have been interpreted as evidence that tasks requiring a haptic response underestimate infant knowledge. Thus, one explanation for why visual RTs are relatively quick during Distractor Touches is that visual measures are more sensitive than haptic measures and therefore more accurate at gauging what infants know. Haptic measures on the other hand may systematically underestimate knowledge because of the additional demands of executing an action, which may cause infants to perseverate on a prepotent response (Diamond, 1985; Baillargeon, DeVos & Graber, 1989; Hofstadter & Reznick, 1996; Gurteen *et al*., 2011).

Although several researchers have shown that haptic perseveration is common in infant participants (Clearfield, Diedrich, Smith & Thelen, 2006; Thelen, Schöner, Scheier & Smith, 2001; Smith, Thelen, Titzer & McLin, 1999; Munakata, 1998), flexible, goal-directed actions do occur when the input is salient and infants can execute actions without an imposed delay (Clearfield, Dineva, Smith, Diedrich & Thelen, 2009). Indeed some researchers have successfully used infants' haptic responses to gauge and predict language abilities (Friend *et al*., 2003, 2008; Ring & Fenson, 2000; Woodward *et al*., 1994), suggesting that haptic responses are a valid measure of early language.

Another interpretation for the conflicting results across modalities during Distractor Touches is based on the graded representations approach, which suggests that when two response modalities conflict, underlying knowledge may be partial (Morton & Munakata, 2002; Munakata 1998, 2001; Munakata & McClelland, 2003). Here the notion is that word knowledge is not all-or-none, but exists on a continuum from absence of knowledge, to partial knowledge, to robust knowledge (Durso & Shore 1991; Frishkoff, Perfetti & Westbury, 2009; Ince & Christman, 2002; Schwanenflugel, Stahl & McFalls, 1997; Steele, 2012; Stein & Shore, 2012; Whitmore, Shore & Smith, 2004; Zareva, 2012). Identifying measures that can gauge word knowledge across this continuum is vital because it has been well documented that infants who demonstrate both delayed language comprehension and production are at the greatest risk for continued language delay, and later development deficits (Karmiloff-Smith, 1992; Desmarais, Sylvestre, Meyer, Bairati & Rouleau, 2008; Law, Boyle, Harris, Harkness & Nye, 2000). From a graded representations account, haptic responses might be in a unique position to gauge these different levels of knowledge. Specifically, the convergence across modalities during correct touches and absent touches reveals the most and least robust levels of comprehension, respectively. From this view, behavioral dissociations emerge during distractor touches because knowledge is partial: knowledge is strong enough to support rapid visual RTs, but too fragile to overcome a prepotent haptic response to the first image fixated (the Distractor).

Indeed, adult work suggests that incorrect responses can be a good proxy for partial knowledge. For example, in a word-learning paradigm, Yurovsky, Fricker, Yu and Smith (2014) found that when word–object pairs to which adults executed an incorrect haptic response in the first block of testing were reencountered in a subsequent block, word–object identification dramatically improved when compared to a group of novel word–object pairs. Thus, while adults failed to encode enough information to support a correct haptic response in the initial test, they encoded partial knowledge, which increased subsequent word learning.

The haptic modality may be particularly susceptible to incorrect responses as a result of partial knowledge because to execute a correct haptic response activation from alternative responses must be inhibited (Woolley, 2006). Studies have shown that increases in processing load lead to greater distractor interference when target and distractor stimuli are presented visually (Fockert, 2013; Guy, Rogers & Cornish, 2012). In the current task paired images were from within the same category, making competition for activation between the incorrect and correct response particularly strong. Accordingly, this may have fostered greater distractor interference when knowledge of the target word was weak.

Crucially, these two interpretations rely on the notion that quick shifts from the distractor to the target image reflect the speed with which the target word was processed. It is possible that some shifts from the distractor to the target image simply reflect random eye movements that are not guided by speech. We attempted to mitigate the influence of such spurious orienting responses by removing gaze shifts that occurred < 400 ms from coding onset. Further, although it may be argued that using both familiar and more difficult words in the same task reduces the link from visual RT to speed of lexical access, there is evidence that visual RT can measure changes in speed of lexical access in words reported by parents as unknown (Fernald *et al*., 2006). Specifically, Fernald and colleagues found that visual RTs significantly decreased with age (15 to 25 months) at nearly identical rates for both words reported by parents as ‘known’ and ‘unknown’. Fernald *et al*. therefore concluded that the measure of visual RT is able to tap into children’s emerging word knowledge that was presumed to be unknown. Indeed, in the current study we calculated visual RTs on a subset of highly known words as reported by parents and obtained the same pattern of results observed in the full stimulus set.

Finally, it should be considered that although visual RTs were slower when children failed to make a haptic response, this does not exclude the possibility that children ‘knew’ the word, but failed to make a haptic response for reasons unknown (e.g. lack of cooperation). We tried to limit the influence of compliance by including only those children who completed the training phase and by utilizing criteria for ending the task when necessary so that we could be confident of the responses that contributed to the final dataset. So, one must wonder if No Touches were a result of noncompliance on the part of the participant, why would they fail to cooperate on some trials (No Touch), but not on others (Target and Distractor Touch). This, in addition to the visual RT evidence, suggests that absent responses, on the whole, reflect a true inability to successfully discriminate the target from the distractor: all children evinced compliance during the training phase, produced some of each response type (Target, Distractor, and No Touch) during test, and were slowest to shift their gaze to the target on No Touches.

A secondary goal of the present research was to assess the relation between visual, haptic and parent report measures of early vocabulary. We found no significant relation between comprehension on the MCDI and speed of lexical access as measured by visual RT. These findings are in line with results from Fernald *et al*. (2006) who found significant correlations between visual RT for ‘known’ words and vocabulary and grammar measures at older ages (25 months), but no significant correlation between visual RT for ‘known’ words and MCDI scores ($r = -.21$) with similarly aged participants (18 months). These results are consistent with a growing literature suggesting that the relation between visual and parent report measures of early language is highly variable (Fernald *et al*., 2006; Houston-Price *et al*., 2007; Marchman & Fernald, 2008; Styles & Plunkett, 2009). Consistent with previous findings from Friend and Keplinger (2008), we find a significant relation between haptic performance (proportion and number of Target Touches) and MCDI comprehension scores.

The finding that haptic performance but not visual performance was significantly correlated with parent report of early word comprehension is somewhat intuitive if we think about the information upon which parent judgments are based. It is likely that explicit kinds of behavioral responses are taken as evidence of comprehension. Indeed, it has been argued that, for this reason, parent reports more accurately estimate children’s productive lexicons (Killing & Bishop, 2008). However the present research suggests that parents provide a reasonably accurate assessment of robust early vocabulary.

Finally, we conducted a series of comparisons between the visual and haptic measures at the child level to examine whether those children who are faster at processing words are also those children who exhibit better haptic performance. Although we did not find a significant relation between visual RT across all words for the haptic measures (proportion or number of Target Touches), visual RT for ‘known’ words correlated significantly with both haptic measures. This suggests that although visual RT may be a more sensitive measure than haptic performance, which requires more robust understanding, the two measures potentially give us a similar picture about children’s level of lexical skill overall. This is supported by the findings that both visual RT for ‘known’ words and haptic performance are significant predictors of later language abilities. An interesting future question then is which measure is more predictive.

Conclusion

The ability to recognize and access the meaning of familiar words gradually increases over the second year of life. It has been suggested that learning the correct referent for a word involves the accumulation of partial knowledge across multiple exposures (Yurovsky *et al.*, 2014). Initially, word recognition may require supporting contextual cues. Eventually stronger, more symbolic representations of word–referent pairings must develop. Consequently, the early lexicon likely consists of both weak (i.e. contextually dependent) and strong (abstract) word representations (Tomasello, 2003). To investigate language acquisition in a developmentally minded way, researchers need to tease partial from fully formed knowledge and latent from active representations. The present results suggest that by implementing testing methods that exclusively measure complete knowledge, or lump partial with absent knowledge, we may not get the full picture of a developing lexicon. Obtaining a rich understanding of the nature of early vocabulary development may necessitate the use of multiple methodologies and modalities (Woolley, 2006). The present results help to clarify the relation between modalities in indexing early knowledge and contribute both to the literature on early language as well as to the broader developmental literature on cognition in the second year of life, especially with respect to the graded structure of early knowledge.

In future research, neuroimaging studies using methods such as event-related potentials (ERPs) may be valuable for exploring the strength of word representations that provoke different types of behavioral responses. An interesting question for future work is whether evidence for graded early knowledge obtains neurophysiologically. As Dale and Goodman wrote, 'Advances in observational and measurement techniques have often directly stimulated theoretical advances, because they do not simply lead to more precise measurement of what is already studied, but to the observation and measurement of new entities or quantities' (Dale & Goodman, 2005, p. 44). This nascent evidence for a graded structure in the developing lexicon has implications for connections between language and cognition early in life and motivates new research extending these findings more broadly both behaviorally and neurophysiologically.

Acknowledgements

We are grateful to the participant families and also to Anya Mancillas, Stephanie DeAnda, Brooke Tompkins, Tara Rutter, and Kasey Palmer for assistance in data collection and coding. This research was supported by NIH #HD068458 to the senior author and NIH Training Grant 5T32DC007361 to the first author and does not necessarily represent the views of the National Institutes of Health.

References

Ahmed, A., & Ruffman, T. (1998). Why do infants make A not B errors in a search task, yet show memory for the location of hidden objects in a nonsearch task? *Developmental Psychology*, **34** (3), 441–453.

Aslin, R.N. (2007). What's in a look? *Developmental Science*, **10** (1), 48–53.

Bailey, T.M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, **17** (2), 1265–1282.

Baillargeon, R., DeVos, J., & Graber, M. (1989). Location memory in 8-month-old infants in a non-search AB task: further evidence. *Cognitive Development*, **4** (4), 345–367.

Ballem, K.D., & Plunkett, K. (2005). Phonological specificity in children at 1; 2. *Journal of Child Language*, **32** (1), 159–173.

Bates, E., Bretherton, I., Snyder, L., Beeghly, M., Shore, C., McNew, S., Carlson, V., Williamson, C., Garrison, A., & O'Connell, B. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* New York: Cambridge University Press.

Behrend, D.A. (1988). Overextensions in early language comprehension: evidence from a signal detection approach. *Journal of Child Language*, **15** (1), 63–75.

Bosch, L., & Sebastián-Gallés, N. (2001). Early language differentiation in bilingual infants. In J. Cenoz & F. Genesee (Eds.), *Trends in bilingual acquisition* (pp. 71–93). Amsterdam: John Benjamins.

Clearfield, M.W., Diedrich, F.J., Smith, L.B., & Thelen, E. (2006). Young infants reach correctly in A-not-B tasks: on the development of stability and perseveration. *Infant Behavior and Development*, **29** (3), 435–444.

Clearfield, M.W., Dineva, E., Smith, L.B., Diedrich, F.J., & Thelen, E. (2009). Cue salience and infant perseverative reaching: tests of the dynamic field theory. *Developmental Science*, **12** (1), 26–40.

Dale, P.S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, **28** (1), 125–127.

Dale, P.S., & Goodman, J.C. (2005). Commonality and individual differences in vocabulary growth. In M. Tomasello & D. Slobin (Eds.), *Beyond nature–nurture: Essays in honor of Elizabeth Bates* (pp. 41–80). Mahwah, NJ: Erlbaum.

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders*, **43** (4), 361–389.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development*, **56**, 868–883.

Durso, F.T., & Shore, W.J. (1991). Partial knowledge of word meanings. *Journal of Experimental Psychology: General*, **120** (2), 190–202.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., Pethick, S.J., Tomasello, M., Mervis, C.B., & Stiles, J. (1994). Variability in early communicative development. *Monographs*

of the Society for Research in Child Development, **59** (5, Serial No. 242).
Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). *MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego, CA: Singular Publishing Group.
Fernald, A., McRoberts, G., & Herrara, C. (1991). Prosody in early lexical comprehension. Paper presented at the meeting of the Society for Research in Child Development, Seattle, WA.
Fernald, A., McRoberts, G.W., & Swingley, D. (2001). Infants' developing competence in understanding and recognizing words in fluent speech. In J. Weissenborn & B. Hoehle (Eds.), *Approaches to bootstrapping in early language acquisition* (pp. 97–123). Amsterdam: John Benjamins.
Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, **42** (1), 98–116.
Fernald, A., Pinto, J.P., Swingley, D., Weinberg, A., & McRoberts, G.W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, **9** (3), 228–231.
Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V.A. (2008). Looking while listening: using eye movements to monitor spoken language. In I.A. Sekerina, E.M. Fernández & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins.
Fockert, J.W. (2013). Beyond perceptual load and dilution: a review of the role of working memory in selective attention. *Frontiers in Psychology*, **4**.
Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, and Computers*, **35** (2), 302–309.
Friend, M., & Keplinger, M. (2008). Reliability and validity of the Computerized Comprehension Task (CCT): data from American English and Mexican Spanish infants. *Journal of Child Language*, **35**, 77–98.
Friend, M., Schmitt, S.A., & Simpson, A.M. (2012). Evaluating the predictive validity of the Computerized Comprehension Task: comprehension predicts production. *Developmental Psychology*, **48** (1), 136–148.
Friend, M., & Zesiger, P. (2011). Une réplication systématique des propriétés psychométriques du Computerized Comprehension Task dans trois langues. *Enfance*, **2011** (3), 329–344.
Frishkoff, G.A., Perfetti, C.A., & Westbury, C. (2009). ERP measures of partial semantic knowledge: left temporal indices of skill differences and lexical quality. *Biological Psychology*, **80** (1), 130–147.
Gold, J.I., & Shadlen, M.N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, **30**, 535–574.
Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, **14** (1), 23–45.
Guy, J., Rogers, M., & Cornish, K. (2012). Developmental changes in visual and auditory inhibition in early childhood. *Infant and Child Development*, **21** (5), 521–536.
Gurteen, P.M., Horne, P.J., & Erjavec, M. (2011). Rapid word learning in 13- and 17-month-olds in a naturalistic two-word procedure: looking versus reaching measures. *Journal of Experimental Child Psychology*, **109** (2), 201–217.
Hirsh-Pasek, K., & Golinkoff, R.M. (1996). The intermodal preferential looking paradigm: a window onto emerging language comprehension. In D. McDaniel, C. McKee & H.S. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 105–124). Cambridge, MA: MIT Press.
Hofstadter, M.C., & Reznick, J.S. (1996). Response modality affects human infant delayed-response performance. *Child Development*, **67**, 646–658.
Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language*, **34** (04), 701–724.
Ince, E., & Christman, S.D. (2002). Semantic representations of word meanings by the cerebral hemispheres. *Brain and Language*, **80** (3), 393–420.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Killing, S.E., & Bishop, D.V. (2008). Move it! Visual feedback enhances validity of preferential looking as a measure of individual differences in vocabulary in toddlers. *Developmental Science*, **11** (4), 525–530.
Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, **41** (3), 841–849.
Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language & Communication Disorders*, **35** (2), 165–188.
Libet, B., Gleason, C.A., Wright, E.W., & Pearl, D.K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain*, **106** (3), 623–642.
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.
Morton, J.B., & Munakata, Y. (2002). Active versus latent representations: a neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, **40** (3), 255–265.
Munakata, Y. (1998). Infant perseverative and implications for object permanence theories: A PDP model of the A-not-B task. *Developmental Science*, **1**, 161–84.
Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, **5** (7), 309–315.
Munakata, Y., & McClelland, J.L. (2003). Connectionist models of development. *Developmental Science*, **6**, 413–429.

Naigles, L.G., & Gelman, S.A. (1995). Overextensions in comprehension and production revisited: preferential-looking in a study of dog, cat and cow. *Journal of Child Language*, **22**, 19–46.

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, **17** (1), 57–70.

Reznick, J.S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics*, **11** (02), 145–166.

Ring, E.D., & Fenson, L. (2000). The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language*, **20** (59), 141–159.

Robinson, C.W., Shore, W.J., Hull Smith, P., & Martinelli, L. (2000). Developmental differences in language comprehension: what 22-month-olds know when their parents are not sure. Poster presented at the International Conference on Infant studies, Brighton, July.

Romo, R., & Salinas, E. (1999). Sensing and deciding in the somatosensory system. *Current Opinion in Neurobiology*, **9** (4), 487–493.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology*, **80** (3), 201–224.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, **69** (2), 309–320.

Schall, J.D., & Thompson, K.G. (1999). Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience*, **22**, 241–259.

Schwanenflugel, P.J., Stahl, S.A., & McFalls, E.L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, **29** (4), 531–553.

Shadlen, M.N., & Newsome, W.T. (1996). Motion perception: seeing and deciding. *Proceedings of the National Academy of Sciences, USA*, **93** (2), 628–633.

Shinskey, J.L., & Munakata, Y. (2005). Familiarity breeds searching: infants reverse their novelty preferences when reaching for hidden objects. *Psychological Science*, **16** (8), 596–600.

Smith, L.B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: the task dynamics of the A-not-B error. *Psychological Review*, **106** (2), 235–260.

Snyder, L.S., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language*, **8** (3), 565–82.

Steele, S.C. (2012). Oral definitions of newly learned words: an error analysis. *Communication Disorders Quarterly*, **33** (3), 157–168.

Stein, J.M., & Shore, W.J. (2012). What do we know when we claim to know nothing? Partial knowledge of word meanings may be ontological, but not hierarchical. *Language and Cognition*, **4** (3), 144–166.

Styles, S., & Plunkett, K. (2009). What is 'word understanding' for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. *Journal of Child Language*, **36** (04), 895–908.

Swingley, D., & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language*, **46** (1), 39–56.

Thelen, E., Schöner, G., Scheier, C., & Smith, L.B. (2001). The dynamics of embodiment: a field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, **24** (01), 1–34.

Thomas, D.G., Campos, J.J., Shucard, D.W., Ramsay, D.S., & Shucard, J. (1981). Semantic comprehension in infancy: a signal detection analysis. *Child Development*, **52**, 798–803.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Trevena, J.A., & Miller, J. (2002). Cortical movement preparation before and after a conscious decision to move. *Consciousness and Cognition*, **11** (2), 162–190.

VanRullen, R., & Thorpe, S.J. (2001). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience*, **13** (4), 454–461.

Whitmore, J.M., Shore, W.J., & Smith, P.H. (2004). Partial knowledge of word meanings: thematic and taxonomic representations. *Journal of Psycholinguistic Research*, **33** (2), 137–164.

Woodward, A., Markman, E.M., & Fitzsimmons, C. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, **30**, 553–566.

Woolley, J.D. (2006). Verbal–behavioral dissociations in development. *Child Development*, **77** (6), 1539–1553.

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

Zareva, A. (2012). Partial word knowledge: frontier words in the L2 mental lexicon. *International Review of Applied Linguistics in Language Teaching*, **50** (4), 277–301.

Received: 4 March 2014
Accepted: 20 August 2014

Acknowledgments

Chapter 2, in full, is a reprint of material as it appears in Hendrickson, K., Mitsven, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2015). Looking and touching: What extant approaches reveal about the structure of early word knowledge. *Developmental Science*, *18*(5), 723-735. The dissertation author was the primary investigator and author of this paper.

CHAPTER 3:

Role of partial knowledge

Abstract

Behavioral dissociations in young children's visual and haptic responses have been taken as evidence that word knowledge is not all-or-none, but exists on a continuum from absence of knowledge, to partial knowledge, to robust knowledge (Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015; Munakata, 2001). This longitudinal follow-up study tested a group of 16-months-olds, 6-months after their initial visit, to replicate results of partial understanding during behavioral dissociations, and determine if partial knowledge of word-referent relations can be leveraged for future word recognition. Results show that like 16-month-olds, 22-month-olds demonstrate behavioral dissociations exhibited by rapid visual reaction times to a named referent but incorrect haptic responses. Further, results suggest partial word knowledge at one time influences the degree to which that word will be understood in the future.

Introduction

Traditionally, investigations into the developing lexical-semantic system were mainly concerned with measuring the number of words children comprehend and produce. Indeed, much of what is known about early lexical knowledge is gathered from diverse measurement techniques (e.g., parent report, visual fixation, haptic response) that implicitly rely on the assumption that lexical knowledge is all-or-none. As a result, many discussions of word comprehension imply a form of abrupt acquisition, in which the child goes through stages of unknown to known in a stage-like fashion (Carey & Bartlett, 1978; Heibeck & Markman, 1987; Houston-Price, Plunkett, & Harris, 2005; Markson & Bloom, 1997; Woodward, Markman, & Fitzsimmons, 1994).

Despite the utility of measuring the number of words children comprehend at different points in development to obtain a general understanding of early lexical development, a central question that is often overlooked is not how many words children know, but what it means to "know" a word. Although researchers have attempted to determine a recognition-point in which a unique word is retrieved from the "mental lexicon" (Marslen-Wilson, 1980), the issue with such an enterprise is there is empirical and computational evidence that word knowledge is not dichotomous, but exists on a continuum from absence of knowledge, to partial knowledge, to robust knowledge (Shore & Durso, 1991; Frishkoff, Perfetti, & Westbury, 2009; Hendrickson, et al., 2015; Ince & Christman, 2002; McClelland & Elman, 1986; McMurray, 2007; Schwanenflugel, Stahl & McFalls, 1997; Steele, 2012; Stein & Shore, 2012; Whitmore, Shore, & Smith, 2004; Zareva, 2012).

In contrast to theories that suggest that word learning is all-or-none (Gallistel, Fairhurst, & Balsam, 2004), accumulative learning theories of word comprehension rely on the assumption that word knowledge is incremental and unfolds overtime as partial knowledge becomes more robust with experience (McMurray, Horst, & Samuelson, 2012; Rogers & McClelland, 2004; Siskind, 1996; Yu, 2008; Yu & Smith, 2007; Yurovsky, Fricker, Yu, & Smith, 2014). Recent connectionist models have corroborated the view of partial knowledge as a central component in characterizing lexical development (McMurray, 2007; McMurray et al., 2012; Yu, 2008). For instance, dynamic accounts suggest that the "lexicon" is active, and that competition exists between lexically related competitors (Elman, 1995). The competition develops dynamically overtime, which has been shown to result in unforeseen outcomes – e.g., partially activated lexical-semantic representations trump more active lexical-semantic representations (McMurray et al., 2012).

Behavioral evidence of such a phenomenon has recently been observed through dissociations of visual (looking) and haptic (touching) responses during a word comprehension task (Hendrickson, et al., 2015). Using a forced choice paradigm, a moment-by-moment analysis of looking and touching behaviors – measures of word processing and comprehension respectively – was conducted to assess the speed with which a prompted word was processed (visual reaction time) as a function of haptic response: Target (touched the picture of the word referent), Distractor (touched the picture of the unprompted word), or No Touch (failed to touch either image). Sixteen-month-olds' visual reaction times (visual RTs) to fixate a prompted image were significantly slower during No Touches compared to Distractor and Target Touches, which were statistically indistinguishable. Therefore, in the case of distractor touches, the visual and haptic response modalities conflict – i.e., children are quick to fixate the target image, but touch the distractor image. So, although evidence within the same study demonstrates a significant relation between different measures of word comprehension (visual RT, haptic response, parent report), word knowledge is highly task dependent; children can demonstrate knowledge in one modality (e.g., visual) but not the other (e.g., haptic) (Hendrickson, et al., 2015).

A recent computational model by Munakata and colleagues suggests that underlying knowledge may be partial when response modalities conflict in this way (Morton & Munakata, 2002; Munakata 1998, 2001; Munakata & McClelland, 2003). According to this view, incorrect (distractor touches) and absent haptic responses (no touches) may index different knowledge states: incorrect responses are associated with partial knowledge, whereas absent responses appear to reflect a true failure to map words to their target referents. Thus, utilizing behavioral dissociations as a measure of partial knowledge states has the potential to aid in our understanding of the continuum of word meaning and the role of partial knowledge in word learning and recognition.

A subset of incremental learning theories suggests that partial knowledge of a word–object mapping at an earlier time point, influences the degree to which that word-object mapping is recognized at a later time (Yurovsky et al., 2014). Consistent with this hypothesis, Yurovsky et al. (2014) found that when word-object pairs to which *adults* executed an incorrect haptic response in the first block of testing were reencountered in a subsequent block, word-object identification dramatically improved when compared to a group of novel word-object pairs. Whereas adults failed to encode enough information to support a correct haptic response in the initial test, they encoded partial knowledge, which increased subsequent word learning. While studies on adults suggest that partial knowledge plays a key role in word learning and recognition (Billman & Knutson, 1996; Rosch & Mervis, 1975; Trabasso & Bower, 1966; Yurovksy, et al. 2014; Yurovsky & Frank, 2015), there is controversy surrounding this question in the developmental word-learning literature. More specifically, it is unknown whether partial word knowledge demonstrates a similar influence over later knowledge states in early development.

Current Study

Most of what we currently know about the primary measures of early language comes from studies that have been conducted in a piecemeal fashion, in which investigators selectively use one, possibly two measures (DeAnda, Arias-Trejo, Poulin-

Dubois, Zesiger, & Friend, 2016; Houston-Price, Mather, & Sakkalou, 2007; Legacy, Zesiger, Friend & Poulin-Dubois, 2015; Marchman & Fernald, 2008; Fernald, Perfors, & Marchman, 2006; Hurtado, Marchman, Fernald, 2008; Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2012). Indeed, no study to date has examined the relation between haptic, visual, and parent report measures of word knowledge within the same cohort of children overtime Therefore, the current study has three primary aims. The first aim is to examine developmental changes in the speed and accuracy of word recognition across these three measures in the same cohort of children throughout the 2nd year. The second aim is to replicate and extend the finding from Hendrickson et al., (2015) of visual and haptic behavioral dissociations and corresponding partial knowledge states across the 2nd year. Finally, the third aim investigates the role partial knowledge plays in early word comprehension over this same time period.

For the analyses related to Aims 1 & 2 we included a group of children for whom we had visual response (visual reaction time), haptic response, and parent report measures of vocabulary knowledge at 16 months and 22 months (*N* = 39). For Aim 1, we examine the correlations between children's vocabulary knowledge indexed by the haptic modality and children's speed of lexical access indexed by visual reaction time, and the well-documented MacArthur-Bates Communicative Development Inventories (MCDI; Fenson et al., 1993) at 16 and 22 months. We anticipate these correlations to reveal stability from 16 to 22 months for all measures. For Aim 2, we conduct a moment-by-moment analysis of looking and reaching behaviors as they occurred in tandem to assess the speed with which a prompted word is processed (visual RT) as a function of the type of haptic response: Target (touched the picture of the prompted image), Distractor (touched the picture of the unprompted image), or No Touch (failed to touch either image). In line with results from Hendrickson et al., 2015, we predict that visual RT will vary as a function of haptic response. Specifically, we predict that visual RT's will be fastest for target touches, and slowest for no touches with an intermediate speed of processing for visual RT's associated with distractor touches.

For analyses related to Aim 3 we include a larger sample (n = 62) for whom we have haptic (but not visual RT) performance data at 16 and 22 months. We assess the claim that partial knowledge of a word–object mapping at an earlier time point, influences the degree to which that word-object mapping is "known" at a later time. Recall that correct (target touches), incorrect (distractor touches) and absent haptic responses (no touches) have been shown to index distinct knowledge states: correct responses represent the most robust levels of understanding demonstrated across modalities, incorrect responses appear to be associated with partial knowledge with evidence of knowledge in the visual but not haptic modality, and absent responses appear to reflect a true failure to map lexical items to their target referents (Hendrickson, et al., 2015; Yurovsky, et al., 2014). At 16 months, we assessed the participants' comprehension of 41 words. In contrast to previous studies, that focus on correctly selected referents to gauge vocabulary knowledge and size, for this analysis we instead focus on the words for which participants give incorrect answers: distractor touches (partially known) and no touches (unknown). Critically, participants are tested again on the same set of 41 words at 22 months. If word learning is accumulative such that partial knowledge is leveraged for future learning, then performance should improve for partially known words (distractor touches) compared to "unknown" words (no touches) (see Figure 1).

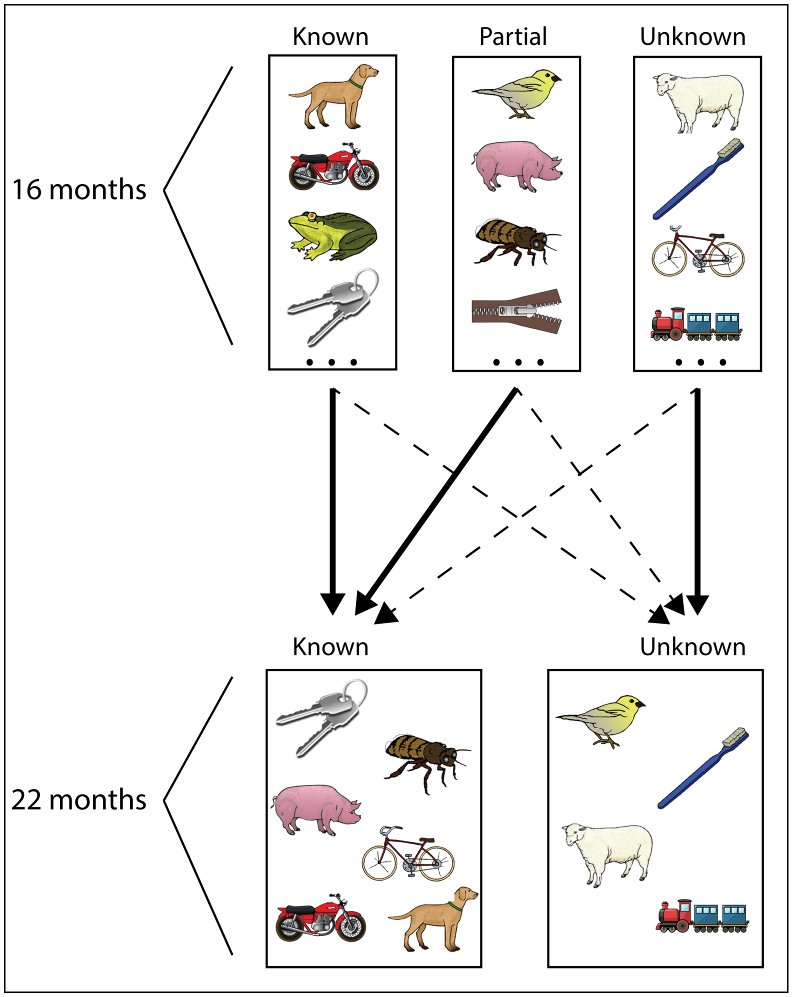

Figure 3-1. Schematic of the predicted influence of knowledge level at 16 months on knowledge level at 22 months. Schematic displays leveraging learning of partially known words compared to unknown words. Lines represent probability of an event (dotted line = low probability, solid line = high probability). Words that are known or partially known at 16 months have a higher probability of being known, and relatedly a lower probability of becoming unknown at 22 months. Conversely, unknown words at 16 months have a lower probability of becoming known, and higher probability of remaining unknown at 22 months.

Method

Participants

In this study, we brought back toddlers who participated in a larger, multi-institutional longitudinal project assessing language comprehension in the 2nd year of life. Participants were obtained through a database of parent volunteers recruited through birth records, internet resources, and community events in a large metropolitan area. Estimates of daily language exposure were derived from parent reports of the number of hours of language input by parents, relatives and other caregivers in contact with the infant. Only those infants with at least 80% language exposure to English were included in the study (Bosch & Sebastian-Galles, 2001).

Apparatus

The study was conducted in a room with sound attenuation paneling. A 51 cm 3M SCT3250EX touch capacitive monitor was attached to an adjustable wall mounted bracket that was hidden behind blackout curtains and between two portable partitions. Two HD video cameras were used to record participants' visual and haptic responses. The eye-tracking camera was mounted directly above the touch monitor and recorded visual fixations through a small opening in the curtains. The haptic-tracking camera was mounted on the wall above and behind the touch monitor to capture both the infants' haptic response and the stimulus pair presented on the touch monitor. Two audio speakers were positioned to the right and left of the touch monitor behind the blackout curtains for the presentation of auditory reinforcers to maintain interest and compliance.

Procedure and Measures

Upon entering the testing room, infants were seated on their caregiver's lap centered at approximately 30 cm from the touch sensitive monitor with the experimenter seated just to the right. Parents wore blackout glasses and noise-cancelling headphones to mitigate parental influence during the task. The assessment followed the protocol for the Computerized Comprehension Task (CCT; Friend & Keplinger, 2003; Friend, Schmitt, & Simpson, 2012). The CCT is an experimenter-controlled assessment that uses infants' haptic response to measure early decontextualized word knowledge. There are two between-subjects forms of the procedure such that distractors on one form serve as targets on the other. Targets and distractors are part of speech (noun, adjective, verb), category (animal, human, object), and visual salience (color, size, luminance) and difficulty as determined by the proportion of children reported to know the word on the MCDI at 16 months (Dale & Fenson, 1996; Jorgenson, Dale, Bleses, & Fenson, 2009). A previous attempt has been made to automate the procedure, such that verbal prompts come from the audio speakers positioned behind the touch screen instead of the experimenter seated to the right of the child. Pilot data using the automated version showed that children's interest in the task waned to such an extent that attrition rates approached 85% (attrition rates using the experimenter-controlled CCT are between 5 – 10%; M. Friend, personal communication, June 17, 2014, P. Zesiger, personal communication, May 21, 2014). Therefore, to collect a sufficient amount of data to yield effects we used the well-documented protocol of the CCT (Friend et al., 2003; 2012). Previous studies have reported that the CCT has strong internal consistency (Form A α =.836; Form B α =.839), converges with parent report (partial r controlling for age = .361, $p < .01$), and predicts subsequent language production (Friend et al., 2012). Additionally, responses on the CCT are nonrandom (Friend & Keplinger, 2008) and this finding replicates across languages (Friend & Zesiger, 2011) and monolinguals and bilinguals (Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2013).

For this procedure, infants are prompted to touch images on the monitor by an experimenter seated to their right (e.g. "Where's the dog? Touch dog!"). Target touches (e.g. touching the image of the dog) elicit congruous auditory feedback over the audio speakers (e.g., the sound of a dog barking). Infants were presented with four training trials, 41 test trials, and 13 reliability trials in a two-alternative forced-choice procedure. For a given trial, two images appeared simultaneously on the right and left side of the touch monitor. The side on which the target image appeared was presented in pseudo-random order across trials such that target images could not appear on the same side on more than two consecutive trials, and the target was presented with equal frequency on both sides of the screen (Hirsh-Pasek & Golinkoff, 1996). The design of the study relied on the successful performance of both 16- and 22-month-olds. That is, the task needed to be easy enough for 16-month-olds to complete the task, but hard enough such that children at 22 months did not perform at ceiling. To insure this outcome there are equal numbers of easy (comprehension = >66%), moderately difficult (comprehension = 33-66%), and difficult words (comprehension < 33%) based on normative data at 16 months of age (Dale & Fenson, 1996; Jorgenson et al., 2009).

The study began with a training phase to insure that participants understood the nature of the task. During the training phase, participants were presented with early-acquired noun pairs (known by at least 80% of 16-month-olds; Dale & Fenson, 1996; Jorgenson, et al., 2009) and prompted by the experimenter to touch the target. If the infant failed to touch the screen after repeated prompts, the experimenter touched the target image for them. If a participant failed to touch during training, the four training trials were repeated once. Only participants who executed at least one correct touch during the training phase proceeded to the testing phase.

During testing, each trial lasted until the infant touched the screen or until seven seconds elapsed at which point the image pair disappeared. When the infant's gaze was directed toward the touch monitor, the experimenter delivered the prompt in infant-directed speech and advanced each trial as she uttered the target word in the first sentence prompt such that the onset of the target word occurred just prior to the onset of the visual stimuli (average interval = 238 ms).

Nouns; Where is the _____? Touch _____.

Verbs; Who is _____? Touch _____.

Adjectives; Which one is _____? Touch _____.

The criterion for ending testing was a failure to touch on two consecutive trials with two attempts by the experimenter to re-engage without success. If the attempts to re-engage were unsuccessful and the child was fussy the task was terminated and the responses up to that point were taken as the final score. However, if the child did not touch for two or more consecutive trials but was not fussy, testing continued. Those participants who remained quiet and alert for the full 41 test trials (16 months, N = 21; 22 months, N = 34), also participated in a reliability phase in which 13 of the test trial image pairs were re-presented in opposite left-right orientation.

Parent report of infant word comprehension was measured at 16 months using the MacArthur-Bates Communicative Development Inventory: Words & Gestures (MCDI: WG) – a parent report checklist of language comprehension and production – and at 22 months using the MacArthur-Bates Communicative Development Inventory: Words &

Sentences (MCDI: WS – a parent report checklist of language production – developed by Fenson et al. (1993). Both inventories have good test-retest reliability and significant convergent validity with an object selection task (Fenson et al., 1994). Of interest in the current study was a comparison between vocabulary checklist and infants' behavioral data.

Coding

A waveform of the experimenter's prompts was extracted from the eye-tracking video –positioned approximately 30 cm from the experimenter – using Audacity® software (http://audacity.sourceforge.net/). Subsequently, the eye-tracking video, haptic-tracking video, and a waveform of the experiment's prompts were all synced using Eudico Linguistics Annotator (ELAN) (<http://tla.mpi.nl/tools/tla-tools/elan/>, Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands; Lausberg & Sloetjes, 2009). ELAN is a multi-media annotation tool specifically designed for the analysis of language. It is particularly useful for integrating coding across modalities and media sources because it allows for the synchronous playing of multiple audio tracks and videos. Only distractor-initial trials – those trials for which infants first fixated the distractor image upon hearing the target word – were included in the analyses of looking behavior.

Coders completed extensive training to identify the characteristics of speech sounds within a waveform, both in isolation and in the presence of coarticulation. Because a finite set of target words always followed the same carrier phrases (e.g., "Where is the ____", "Who is ___", or "Which one is ____"?), training included identifying different vowel and consonant onsets after the words "the" and "is". Coders were also trained to demarcate the onset of vowel-initial and nasal-initial words after a vowel-final word in continuous speech, which can be difficult using acoustic waveforms in isolation. Coders were required to practice on a set of files previously coded by the first author with supervision and then to code one video independently until correspondence with previously coded data was reached. Two coders completed each pass, each coding ~50% of the data.

Trials with short latencies (200 – 400 ms) likely reflect eye movements that were planned prior to hearing the target word (Fernald et al., 2008; Bailey & Plunkett, 2002; Ballem & Plunkett, 2005). For this reason, trials were included in subsequent analyses if the participant looked at the screen for at least 400 ms. Additionally, looking responses were coded during the first 2000 ms of each trial. As previously mentioned, looking responses that take place further from the stimulus onset are less likely to be driven by stimulus parameters (Aslin 2007; Fernald, Perfors & Marchman, 2006; Swingley & Fernald, 2002). Finally, by coding the first 2000 ms we are largely restricting our analysis to the period prior to the decision to touch.

Coding occurred in two passes. Coder 1 annotated the frame onset and offset of the target word as it occurred in the first sentence prompt using the waveform of the experimenter's speech. First, the coder listened to the audio and zoomed in on the portion of the waveform that contained the target word in the first sentence prompt (e.g. Where is the **DOG**?). Once that section was magnified, the coder listened to the word several times precisely demarcating the onset and offset of speech information within the larger waveform. Coder 1 also marked the frame in which the visual stimuli appeared on the screen and the side of the target referent (note: side of the target referent was hidden from Coder 2). Coder 2 coded visual and haptic responses with no audio to insure that she remained blind to the image that constituted the target. Coding began at image onset, roughly 238 ms after target word onset, and prior to target word offset in the first sentence prompt. For the visual behavior, Coder 2 advanced the video and coded each time a change in looking behavior occurred using three event codes: right look, left look, and away look. For sustained visual fixations, Coder 2 advanced the video in 40 ms coding frames and because shifts in looking are crucial for deriving measures of reaction time, she advanced the video during gaze shifts at a finer level of resolution (3 ms).

Participants' initial haptic response was coded categorically: Left Touch (unambiguous touch to the left image), Right Touch (unambiguous touch to the right image) or No Touch (no haptic response executed). Identifying touches as Target or Distractor was done post-hoc, to preserve coders' blindness to target image and location.

Inter-rater reliability coding was conducted for both visual and haptic responses by a third, reliability coder. For looking responses, a random sample of 11 videos (~ 25% of the data) was selected for each age group. Because our dependent variable (visual RT) relies on millisecond precision in determining when a shift in looking behavior occurred, only those frames in which shifts occurred were considered for the reliability score. This score is more stringent than including all possible coding frames because the likelihood of the two coders agreeing is considerably higher during sustained fixations compared to gaze shifts (Fernald, Pinto, Swingley, Weinbergy, & McRoberts, 1998). Using this shift-specific reliability calculation, we found that for 90% of trials coders were within one frame (40 ms) of each other, and on 94% of the trials coders were within two frames (80 ms) of each other.

All haptic response coding was compared to offline coding of haptic touch location completed for the larger longitudinal project. Inter-rater agreement for the haptic responses was 95%. All haptic coding was completed blind to target image, location, and visual fixations.

Results

The average time to execute a haptic response post image onset was 3896.25 ms for 16-month-olds and 2639.89 for 22-month-olds. The average visual RT to shift to the target across haptic types was 862.43 ms for 16-month-olds and 762.18 ms for 22-month-olds, comparable to the mean visual RTs found in similarly aged participants in previous research (Fernald et al., 1998). Consistent with the literature, immediate test–retest reliability on the CCT was strong for participants who completed reliability in the larger, 62 participant sample [$r(41) = .74$, $p < .0001$], and in the subset of data used for the analyses related to Aim 1 [$r(32) = .67$, $p < .0001$]. Finally, internal consistency on the CCT was excellent (Form A $\alpha = -.931$ and Form B $\alpha = -.940$)

At the age of 16 months, children executed target touches on 11.78 trials, distractor touches on 10.08 trials, and provided no haptic response on 13.03 trials. At 22 months, children executed target touches on 26.6 trials, distractor touches on 5.66 trials, and no touches on 4.29 trials. This pattern of findings was expected. As previously mentioned, to compare performance on the same set of words over a 6 month period, the word stimuli selected needed to be easy enough to keep 16-month-olds engaged in the task, whilst being difficult enough so that 22-month-olds did not perform at ceiling.

Aim 1: Relation of Visual, Haptic, and Parent Report Measures in 2nd year

Speed of processing and word recognition from 16 to 22 months.

One goal of this research was to extend the cross-sectional findings of Hendrickson et al. (2015) with a longitudinal design. To compare speed of word processing and word recognition in the same group of children at different ages, we performed correlations between visual RT, haptic response, and parent report at 16 and 22 months (see Figure 2 for summary of results).

The average visual RT was calculated for each child at 16 months and 22 months on distractor-initial trials in which a correct shift in gaze occurred between 400 and 2000 ms post-stimulus onset. The correlation between the average visual RT at 16 months (M = 862.43, SE = 23.37) and 22 months (M = 762.18, SE = 17.51) was significant (r =.39, p = .014). The haptic measure was calculated as the number of Target Touches executed by participants at 16 months and 22 months. Following our findings for visual RT, the correlation between the number of target touches executed at 16 and 22 months (M = 12.8, SE = .98 and M = 26.6, SE = 1.18, respectively) was significant (r = .39, p = .014). This finding extends previous research showing stability in performance on the CCT from 16 to 20 months of age (Friend & Keplinger, 2008; Friend & Zesiger, 2011; Legacy et al., 2016)

Finally, parent reported vocabulary comprehension and production was obtained by using the MacArthur-Bates Communicative Development Inventory Words and Gestures at 16 months, and Words and Sentences at 22 months. Parent reported vocabulary comprehension at 16 months (M = 188, SE = 10.40) was significantly correlated (r = .58, $p < .0001$) with reported vocabulary production at 22 months (M = 247, SE = 24.41).

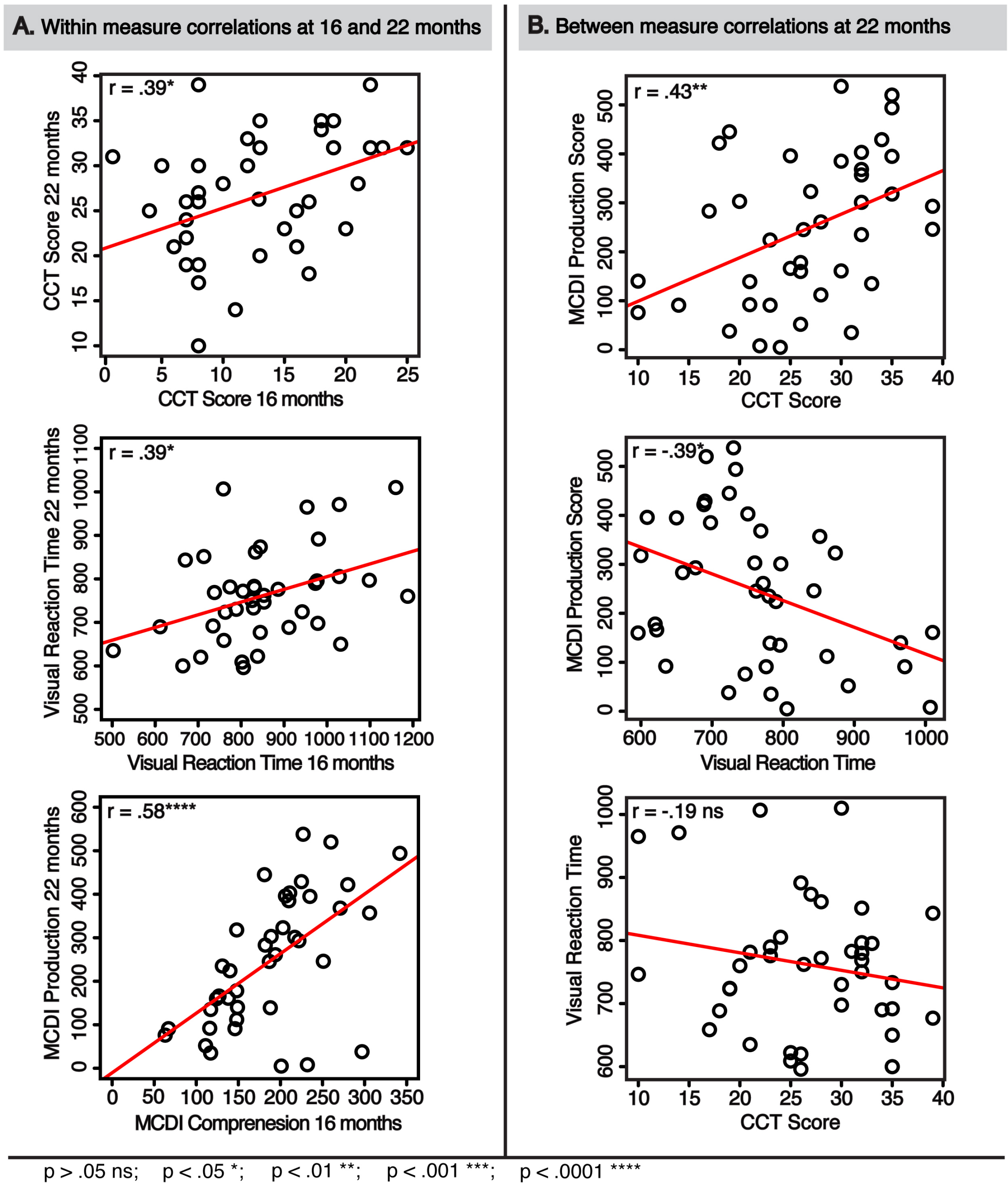

Figure 3-2. Correlations within each measure from 16 to 22 months (A), and inter-correlations between parent report, haptic response, and visual reaction time at 22 months (B).

Inter-correlation between Visual RT, Haptic Response, and Parent Report at 22 months.

A series of Pearson's product–moment correlations was performed to analyze the relation between each of our behavioral measures (visual RT and haptic) and MCDI: WS production score (the parent reported number of words produced by the child) at 22 months. There was a significant negative correlation between visual RT and MCDI production (r = -.39, p = .014), such that the faster a child processed words the more words they were reported to produce. Additionally, there was a significant positive correlation between the haptic measure and MCDI production (r = .43. p = .007), such that the more words children correctly identified on the haptic measure, the more words their parent reported they produced. Finally, although the correlation between visual RT and the haptic measure was in the expected direction it was not significant (r = -.19, p = .25).

Aim 2: Concurrent Analyses of Visual and Haptic Responses at 22 months

Calculating reaction time by including only distractor-initial trials and a narrow time window restricts the number of usable trials per condition. Consequently not all children contributed data to all experimental conditions and thus were removed from further analysis. Of the 39 infants originally included, 16 participants contributed visual (RT) data for all three haptic types (target, distractor, and no touch) and thus were included in analyses regarding the concurrent relation between visual RT and haptic response. The high rate of exclusion was due primarily to the lack of data in the No Touch condition indicating that roughly two-thirds of participants at 22 months executed a touch on every trial. This level of exclusion for the concurrent analysis of visual and haptic responses at 22-month-olds was expected since we anticipated growth in receptive vocabulary from 16 to 22 months. The remaining 16 participants who contributed visual RT data performed generated at least one Target, Distractor, and No Touch trial and completed an average of 37 out of a possible 41 trials.

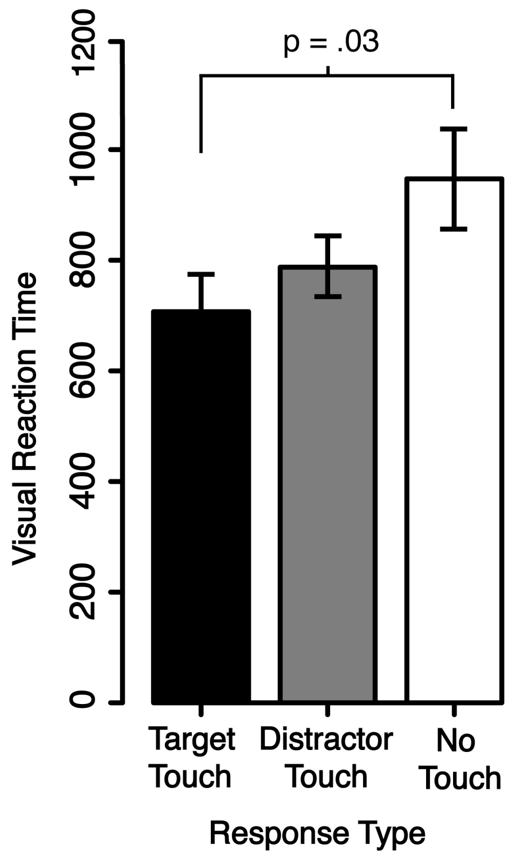

Figure 3-3. Visual RT analysis. Mean visual RT to shift gaze from the distractor to the target image by response type. Note: Error bars show the standard error across participants.

We compared speed of processing across the Haptic Types (Target, Distractor, No Touch) using visual RT. As previously mentioned, average visual RTs were calculated for distractor-initial trials in which a shift in gaze occurred between 400 – 2000 ms post-visual onset. Visual RT was averaged for each participant by Response Type and planned pairwise comparisons were conducted on the three levels of Response Type. Participants demonstrated a nearly identical pattern of looking times across the three response types as reported previously at 16-months (Removed for blinding, 2015; see Figure 3), such that infants processed words significantly faster ($t(14) = 2.33$, $p = .03$) during Target Touch (M = 707.06, SE = 31.45) compared to No Touch trials (M = 947.17, SE = 91.77), with Visual RT on Distractor Touches demonstrating an intermediate speed of processing (M = 787.45, SE = 57.41) that was not significantly different from that of Target Touches ($t(14) = 1.32$, $p = .19$). Finally, although visual RTs were faster for Distractor Touches than No Touches*,* this difference did not reach significance ($t(14) = 1.54$, $p = .14$) (see Figure X). To replicate the effects of Target and Distractor Touch on visual RT with a larger data set we again calculated visual RT by response type for participants who executed both target and distractor touches (N = 35). Results revealed a similar influence of response type on visual RT, such that visual RT for Distractor Touches (M = 790.62, SE = 42.96) and Target Touches (M = 732.29, SE = 118.10) were not significantly different ($t(33) = 1.26$, $p = .21$).

Aim 3: The Influence of Partial Knowledge on Word Recognition.

In order to address Aim 3, we evaluated the role partial knowledge plays in early word comprehension in the 2nd year of life. For analyses related to Aim 3 we include a larger sample ($N = 62$) for whom we have haptic performance data at 16 and 22 months. Recall that toddlers were tested on the same list of 41 words at 16 and 22 months. Therefore, for each participant we have a list of words at 16 months for which they executed a Target Touch, that we refer to as *Known Words* (e.g., dog, shoe, running etc.), a list for which they executed a Distractor Touch, referred to as *Partially Known* (e.g., cat, red, hand), and a list for which they made no haptic response, referred to as *Unknown Words* (No Touch, e.g., truck, bubbles, jumping.)

We used a similar analysis to that reported by Yurovsky et al., 2014, to evaluate partial knowledge states in cross-situational word recognition. We conducted an item level analysis to determine what proportion of Known, Partially Known, and Unknown words at 16 months are Known, Partially Known, and Unknown at 22 months creating a 3x3 design. For this analysis we were particularly interested in whether Partially Known words at 16 months, relative to Unknown words, were likely to become Known at 22 months.

We ran a 3 (proportion Known, Partially Known, and Unknown at 16 months) x 3 (proportion Known, Partially Known, and Unknown at 22 months) repeated measures ANOVA (see Figure 4). There was a significant main effect of response type at 22 months ($F(2, 60) = 109.6$, $p < .001$), suggesting that regardless of knowledge level at 16 months, participants had a greater proportion of Known Words compared to Partially Known and Unknown Words at 22 months. However, this main effect is qualified by the predicted interaction of age and response type ($F(4, 58) = 6.60$, $p < .001$), which suggests that there is a change with age in the proportions of different response types.

Planned pairwise comparisons were conducted to determine how proportion of each response type at 22 months was influenced by response type at 16 months. Of particular interest was whether Partially Known words at 16 months were more likely to be Known (and relatedly less likely to be in Unknown) at 22 months compared to Unknown words at 16 months. There was a significant difference in the proportion of Unknown (M = .64, SE = .03) versus Known words (M = .70, SE = .03) at 16 months

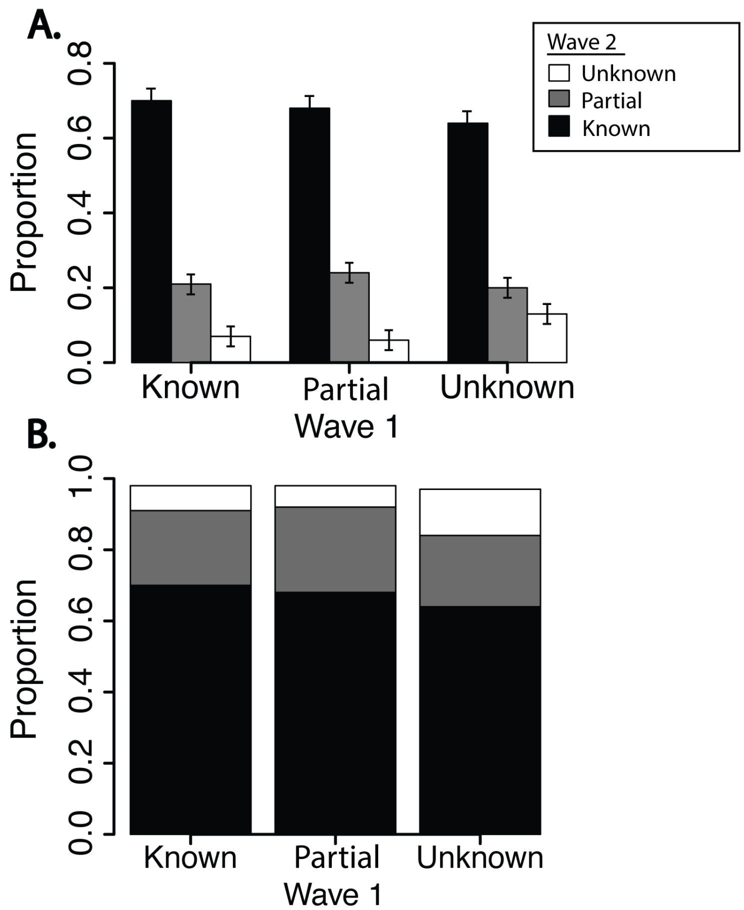

Figure 3-4. Proportion of Known, Partially Known, and Unknown Words at 22 months based on knowledge level at 16 months (Known, Partially Known, Unknown) displayed as clustered (A) and stacked (B) bar graphs. Error bars represent ± SE. Partially known words at 16 months (i.e. Distractor Touches) demonstrated more accurate performance (greater proportion of Target Touches, smaller proportion of No Touches) than "unknown" (i.e., No Touches) 6 months later.

that were Known by 22 months (t(61) = 2.13, p = .037). Further, words that were Partially Known (M = .059, SE = .020) were significantly less likely than Unknown at 16 months (M = .133, SE = .023) to be Unknown at 22 months (t(61) = 4.48, $p < .0001$). Thus, it was highly unlikely that a partially known word at 16 months would be unknown at 22 months. Finally, although the effect was in the expected direction there was no significant difference between the proportion of Unknown Words versus Partially Known at 16 months that were Known by 22 months (M = .64, SE = .030, and M = .68, SE = .027, respectively, t(61) = 1.23, p = .22). Nevertheless, in total, these findings suggest that partial knowledge provides a basis for developing more robust word representations. Partially Known words at 16 months either remained Partially Known or moved into the Known category by 22 months whereas Unknown words at 16 months were more likely to remain Unknown at 22 months.

Discussion

The overarching goal of the current study was threefold: 1. To evaluate the relation between visual RT, haptic response, and parent report as measures of early word knowledge throughout the 2^{nd} year of life, and 2. To assess whether there is a continuum of word knowledge over this same period, and 3. To examine the role of partial knowledge in future word recognition. This longitudinal study provides the first data on convergent and predictive associations between the three primary paradigms in use for measuring word processing and word comprehension across the 2nd year. Further, this research provides new evidence of a continuum of word knowledge and the role knowledge states play in future word comprehension in early development.

Relation of Visual, Haptic, and Parent Report Measures in the 2^{nd} year

The first aim of this research was to examine developmental changes in the speed and accuracy of word recognition across three measures of word knowledge (visual RT, haptic response, parent report) in the same cohort of children throughout the 2nd year. Findings revealed robust relations within each measure from 16 to 22 months. Children's speed of word processing (visual RT) decreased and word comprehension (haptic response, parent report) increased significantly over this period, consistent with earlier research, which examined this trend for each measure separately (Houston-Price, Mather, & Sakkalou, 2007; Legacy, Zesiger, Friend & Poulin-Dubois, 2015; Marchman & Fernald, 2008; Fernald, Perfors, & Marchman, 2006; Hurtado, Marchman, Fernald, 2008; Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2012).

We also conducted a series of comparisons between the visual, haptic, and parent report measures at 22 months. Consistent with previous results at 16 months, we continue to find a significant relation between haptic performance and parent reported vocabulary. Contrary to the results obtained at 16 months from Hendrickson et al., 2015, the current results revealed a significant relation between visual RT and parent report, such that children who are faster at processing words are also those children who exhibit more parent reported vocabulary knowledge at 22 months. That visual RT and parent reported vocabulary correlate at 22 months but not at 16 months is likely due to the substantial variability in mean visual RT in younger children. Indeed, it has been shown previously, and within the current study, that visual RT as a measure of processing speed may be less stable early in the 2nd year of life as variance in visual RT decreases with age (Fernald, et al., 2006).

Finally, visual RT and haptic response were not significantly related at 22 months. It has been previously suggested that although visual RT and haptic response potentially give us a similar picture of children' s level of lexical skill overall, each may be differentially sensitive to knowledge across a hypothetical continuum due to differences in task demands (Hendrickson et al., 2015; Munakata, 2001). Direct haptic measures of vocabulary are relatively demanding and therefore capture decontextualized or robust knowledge. Indeed, the effort involved in executing a looking response, in contrast, is minimal during visually based measures (e.g., looking time, first fixation). For instance, a weak understanding of a word-object pair may be enough to prompt a saccade away from a distractor image and to a matching referent, yet be insufficient to elicit an accurate haptic response due to the additional effort involved in executing an action and inhibiting a prepotent response to the first image fixated. The low-cost nature of executing a visual saccade may cause the visual RT measure to be geared toward measuring more fragile levels of understanding compared to haptic responses. Due to the design of the task, 22-month-old's understanding of the words tested was rather robust – i.e., these words were chosen to be highly familiar to children of this age. Therefore, it is possible that the lack of a significant correlation between the visual RT and haptic measure was due to the fact that certain tasks tap weaker representations (visual RT), while other tasks require stronger representations (haptic response), leading to dissociations in behavior, and somewhat discordant findings between response modalities (Munakata, 2001).

Concurrent Analyses of Visual and Haptic Responses at 22 months

The second aim sought to replicate and extend the finding from Hendrickson et al., (2015) of visual and haptic behavioral dissociations and corresponding partial knowledge states across the 2nd year. Indeed, we found that speed of processing differed as a function of haptic response, such that children were fastest at processing words for correct haptic responses (target touch), followed by incorrect haptic responses (distractor touch), and slowest to shift their gaze when they failed to make a haptic response (no touches).

Therefore, and in line with previous research, we found evidence for behavioral dissociations during Distractor Touches – i.e., rapid visual RTs during incorrect haptic responses – but not during No Touches – i.e., visual and haptic behavior converge: slow visual RTs and absent haptic responses. Based on the graded representations approach, when two response modalities conflict, underlying knowledge may be partial (Morton & Munakata, 2002; Munakata 1998, 2001; Munakata & McClelland, 2003). From this view, behavioral dissociations arise during distractor touches because knowledge is partial: knowledge is robust enough to catalyze rapid visual RTs, but too weak to surmount a predominant response to touch the first image fixated (the distractor). This replicates and extends findings from 16-months-olds that word knowledge is not all-or-none, but exists on a continuum from absence of knowledge, to partial knowledge, to robust knowledge (Hendrickson et al., 2015). These results also provide further evidence that incorrect and absent responses can not be grouped as representing lack of knowledge, but instead, each measure represents different abilities in lexical access that meaningfully measure knowledge across a hypothetical continuum.

One limitation of this analysis was the relatively small number of children who produced all three haptic response types at 22 months (N = 16). This is not surprising as few participants failed to produce a haptic response at this age, which corresponds to an increase in Target responses from 16 to 22 months. Thirty-five of the thirty-nine participants produced at least one Target Touch and one Distractor Touch. Thus, to insure that the pattern of observed results held for Target and Distractor Touches, we reran the analysis on this larger data set. Indeed results from this analysis confirmed the observed effect that visual RT during Target and Distractor Touches is relatively rapid, and is statistically indistinguishable.

The Influence of Partial Knowledge on Word Recognition.

The third aim investigated the role that partial knowledge plays in early word comprehension throughout the 2nd year. Accumulative theories of word learning suggest that partial understanding of words influences future word comprehension. Specifically, word learning is graded as knowledge moves from states of unfamiliarity through partial understanding to robust understanding. If participants had partial knowledge of words when they executed a distractor touch at 16 months, it could contribute to future learning by helping the participant to demonstrate knowledge of the same word at 22 months. That is, even when participants fail to demonstrate knowledge of the word and touch the distractor referent, they may have partial information about the word-referent pair, which may increase the probability of recognizing the same word, and decrease the probability of demonstrating a lack of knowledge (i.e., lower probability of executing no touch). However, if participants have very little to no knowledge of the word-object relation – as we argue is the case in no touches trials – we may expect a decreased probability of recognizing that same word and an increased probability of continued lack of knowledge (i.e., higher probability of executing a no touch again).

Consistent with this prediction, we found that words were more likely to be Unknown at 22 months if they were Unknown, as opposed to Partially Known, at 16 months. Therefore, demonstrating a lack of knowledge of a word at one point in development increased the probability of continuing to demonstrate a lack of understanding of a word at a later point in development. This suggests that even though participants fail to correctly identify a word's referent and touch the distractor, they may nevertheless have encoded, and continue to represent, partial information of the word-object relation.

However, we found that although words were more likely to be known at 22 months if they were Partially Known compared to Unknown at 16 months, this effect did not reach significance. One reason why we do not see significant effects here may be due to the time period between testing. In Yurovsky et al., 2014 – the study on which the current analysis was based – adults were presented with partially known and novel word-object pairs within the same day they were first tested. In the current study, both words that were unknown and partially known at 16 months had a 6 month period between testing occurrences. This long period between the initial testing of the words at 16 months and the follow-up testing at 22 months may have attenuated the effects. Specifically, there was significant vocabulary growth and our findings indicate that words that were previously partially known and unknown became known during this period. The time frame over which partial knowledge becomes robust is not known and may indeed be rather short given the rapid growth in vocabulary in the 2nd year. Considering the current results in this light, the influence of partial knowledge on learning early in development is swift and there is evidence that partially known words were more likely to become known over a 6 month period than were previously unknown words.

How then is knowledge for known or partially known words strengthened and made more robust over time? Classic approaches to word learning assume an all-or-none form of learning, such that once a word becomes learned, there is no change in knowledge state. An accumulative view of word learning suggests that even knowledge for familiar “known” words becomes stronger over development. It has been shown in a recent computational model that this improvement, which appears as gains in the efficiency of processing familiar words, is a result of changes in connection weights (McMurray, et al., 2012). Within the current two-alternative forced choice task, participants could arrive at the right answer in one of two ways: As a result of strengthening connection weights between the word and the object or by pruning unnecessary connections that may interfere by either slowing down the selection process, or by causing an incorrect selection of the referent due to increased competition from spurious connections. Of particular importance for the current results, the model showed that the majority of word learning and recognition of familiar words is in not forming new connections, but rather pruning irrelevant connections – i.e., identifying those words and referents that are *not* associated. From this view, partially known words at one time point can influence future learning not by helping the participant to recognize that same word, but instead, helping the participate to recognize what referents are *not* related to the word.

Conclusion

Classic theories of word learning suggest a form of all-or-none understanding. The current study provides further behavioral evidence of a continuum of knowledge that includes levels of partial knowledge states. Further, this work demonstrates that partial knowledge influences future word understanding, and offers evidence through accumulative learning theories, concerning the role partial knowledge plays in supporting later, adult-like levels of understanding. Finally, this research demonstrates that different measures for early word comprehension may provide a similar picture of children's level of lexical skill overall, however each may be differentially sensitive to knowledge across a hypothetical continuum due to differences in task demands.

Acknowledgments

Chapter 3, in full, is being prepared for submission for publication of the material, Hendrickson, K., Poulin-Dubois, Zesiger, & Friend (in preparation). Assessing a continuum of lexical-semantic knowledge in the 2nd year: A multimodal approach The dissertation author was the primary investigator and primary author of this paper.

We are grateful to the participant families and also to Kelly Kortright for assistance in data collection and coding. This research was supported by NIH #HD068458 to the senior authors and NIH Training Grant 5T32DC007361. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD or the NIH.

References

Aslin, R. N. (2007). What's in a look?. *Developmental Science*, *10*(1), 48-53.

Billman, D., & Knutson, J. (1996). Unsupervised concept learning and value systematicitiy: A complex whole aids learning the parts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(2), 458.

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*(1), 29-49.

Carey, S., & Bartlett, E. (1978). Acquiring a single word. *Papers and Reports on Child Language Development, 15*, 17–29.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125-127.

DeAnda, S., Arias-Trejo, N., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). Minimal second language exposure, SES, and early word comprehension: New evidence from a direct assessment. *Bilingualism: Language and Cognition*, *19*(1), 162-180.

Elman, J.L. (1995) Language as a dynamical system. In Mind as Motion (Port, R.F. and van Gelder, T., eds), pp. 195–225, MIT Press.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185.

Fenson, L., Dale, P.S., Reznick, J.S., Thal, D., Bates, E., Hartung, J.P., Pethick, S., & Reilly, J.S. (1993). MacArthur Communicative Development Inventories: User's guide and technical manual. San Diego, CA: Singular Publishing Group.

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.

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 302-

309.

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the Computerized Comprehension Task: Comprehension predicts production. *Developmental Psychology*, *48*(1), 136.

Friend, M., & Zesiger, P. (2011). A systematic replication of the psychometric properties of the CCT in three languages: English, Spanish, and French. *Enfance*, *3*, 329-344.

Frishkoff, G. A., Perfetti, C. A., & Westbury, C. (2009). ERP measures of partial semantic knowledge: Left temporal indices of skill differences and lexical quality. *Biological Psychology*, *80*(1), 130-147.

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(36), 13124-13131.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development*, 1021-1034.

Hendrickson, K., Mitsven, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2015). Looking and touching: What extant approaches reveal about the structure of early word knowledge. *Developmental Science*, *18*(5), 723-735.

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

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language*, *34*(4), 701-724.

Houston-Price, C., Plunkett, K. I. M., & Harris, P. (2005). 'Word-learning wizardry'at 1; 6. *Journal of Child Language*, *32*(1), 175-189.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31-F39.

Ince, E., & Christman, S. D. (2002). Semantic representations of word meanings by the cerebral hemispheres. *Brain and Language*, *80*(3), 393-420.

Jorgensen, R. N. Dale, P. S., Blesses, D., and Fenson, K. (2010). CLEX: A cross-linguistic lexical norms database. *Journal of Child Language*, *37*, 419-428.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*(3), 841-849.

Legacy, J., Zeisger, P., Friend, M., & Poulin-Dubois, D. (2015). Vocabulary size, translation equivalents, and efficiency in word recognition in very young bilinguals. *Journal of Child Language*, 1-24.

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.

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*(6619), 813-815.

Marslen-Wilson, W. D. (1980). Speech understanding as a psychological process. In J.C. Simon (Ed.) *Spoken language generation and understanding* (pp. 39-67). Dordrecht: Reidel

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*(5838), 631-631.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, *119*(4), 831.

Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, *40*(3), 255-265.

Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the AB task. *Developmental Science*,*1*(2), 161-184.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, *5*(7), 309-315.

Munakata, Y., & McClelland, J. L. (2003). Connectionist models of development. Developmental Science, 6(4), 413-429.

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, *17*(1), 57-70.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573-605.

Rogers, T., & McClelland, J. (2004). Semantic cognition: A parallel distributed processing approach. Cambridge, MA: MIT Press.

Schwanenflugel, P. J., Stahl, S. A., & Mcfalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, *29*(4), 531-553.

Shore, W. J., & Durso, F. T. (1991). Partial knowledge of word meaning.*Journal of Experimental Psychology*, *2*, 190-202.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1), 39-91.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558-1568.

Steele, S. C. (2012). Oral Definitions of Newly Learned Words An Error Analysis. *Communication Disorders Quarterly*, *33*(3), 157-168.

Stein, J. M., & Shore, W. J. (2012). What do we know when we claim to know nothing? Partial knowledge of word meanings may be ontological, but not hierarchical. *Language and Cognition, 4*(3), 144-166

Swingley, D., & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language*, *46*(1), 39-56.

Trabasso, T., & Bower, G. (1966). Presolution dimensional shifts in concept identification: A test of the sampling with replacement axiom in all-or-none models. *Journal of Mathematical Psychology*, *3*(1), 163-173.

Whitmore, J. M., Shore, W. J., & Smith, P. H. (2004). Partial knowledge of word meanings: Thematic and taxonomic representations. *Journal of Psycholinguistic Research*, *33*(2), 137-164.

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, *30*(4), 553.

Yu, C. (2008). A statistical associative account of vocabulary growth in early word learning. *Language learning and Development*, *4*(1), 32-62.

Yurovsky, D., & Frank, M. C. (2015). Beyond naïve cue combination: salience and social cues in early word learning. *Developmental science*, *145,* 53-62.

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

Zareva, A. (2012). Partial word knowledge: Frontier words in the L2 mental lexicon. *International Review of Applied Linguistics in Language Teaching*, *50*(4), 277–301

CHAPTER 4:

Organization of words and environmental sounds in adults

Neuropsychologia 69 (2015) 67–76

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

The organization of words and environmental sounds in memory ☆

Kristi Hendrickson [a,c,d,*], Matthew Walenski [c], Margaret Friend [b], Tracy Love [a,c]

[a] Center for Research in Language, University of California, San Diego, USA
[b] Department of Psychology, San Diego State University, USA
[c] School of Speech, Language, and Hearing Sciences, San Diego State University, USA
[d] Joint Doctoral Program in Language and Communicative Disorders, San Diego State University, USA

ARTICLE INFO

Article history:
Received 5 August 2014
Received in revised form
21 November 2014
Accepted 23 January 2015
Available online 24 January 2015

Keywords:
ERP
N400
Word
Environmental sound
Semantic memory

ABSTRACT

In the present study we used event-related potentials to compare the organization of linguistic and meaningful nonlinguistic sounds in memory. We examined N400 amplitudes as adults viewed pictures presented with words or environmental sounds that matched the picture (Match), that shared semantic features with the expected match (Near Violation), and that shared relatively few semantic features with the expected match (Far Violation). Words demonstrated incremental N400 amplitudes based on featural similarity from 300–700 ms, such that both Near and Far Violations exhibited significant N400 effects, however Far Violations exhibited greater N400 effects than Near Violations. For environmental sounds, Far Violations but not Near Violations elicited significant N400 effects, in both early (300–400 ms) and late (500–700 ms) time windows, though a graded pattern similar to that of words was seen in the mid-latency time window (400–500 ms). These results indicate that the organization of words and environmental sounds in memory is differentially influenced by featural similarity, with a consistently fine-grained graded structure for words but not sounds.

Published by Elsevier Ltd.

1. Introduction

Our ability to interpret the world around us crucially depends on how the brain organizes meaningful auditory information in memory. The organization of semantic memory for one form of meaningful information, linguistic items (e.g. words), has been well investigated, and is based on several factors. Among the most important is featural similarity (i.e. the perceived likeness between concepts), which aids in categorization (Kay, 1971; Murphy et al., 2012; Paczynski and Kuperberg, 2012; Rosch et al., 1976; Sajin and Connine, 2014). Far less is known about how the brain processes and organizes meaningful auditory information that is not linguistic (e.g. environmental sounds). The current paper examines whether semantic information is organized similarly in memory for words and environmental sounds, and specifically whether featural similarity is useful for the organization of environmental sounds in memory. Uncovering how the brain organizes meaning associated with diverse forms of referential auditory information is vital for understanding the relation between language and cognition.

1.1. The processing of words and environmental sounds

Words and environmental sounds share many spectral and temporal characteristics (Gygi, 2001; Shafiro and Gygi, 2004) and are modulated by contextual cues (Ballas and Howard, 1987), item familiarity (Ballas, 1993; Cycowicz and Friedman, 1998), and frequency of occurrence (Ballas, 1993; Cycowicz and Friedman, 1998). Like words, environmental sounds carry deep semantic associations with a corresponding referent (Ballas, 1993). Multiple lines of evidence suggest that words and environmental sounds are processed similarly. It has been shown with behavioral measures (accuracy, response time) that semantically congruent words or pictures can prime environmental sounds, and it has likewise been shown that environmental sounds can prime words or pictures (Ballas, 1993; Chen and Spence, 2011; Özcan and Egmond, 2009; Schneide et al., 2008; Stuart and Jones, 1995). Electrophysiological measures reveal a similar effect. N400 (described in detail below) priming effects (attenuated N400 amplitudes to semantically related compared to unrelated primes) have been found for word or picture targets primed by environmental sounds (Schön et al., 2010; Daltrozzo and Schön, 2009; Frey et al., 2014; Koelsch et al.,

☆This research was supported by NIH Training Grant 5T32DC007361 to the first author, NIH No. HD068458 (to Dr. Friend), and NIH R01DC009272 and R01DC03681 (to Dr. Love) The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD or the NIH.

* Correspondence to: San Diego State University, 6505 Alvarado Road, Suite 101, San Diego, CA 92120, USA.
E-mail addresses: krhendricson@ucsd.edu (K. Hendrickson), mwalenski@ucsd.edu (M. Walenski), mfriend@mail.sdsu.edu (M. Friend), tlove@mail.sdsu.edu (T. Love).

http://dx.doi.org/10.1016/j.neuropsychologia.2015.01.035
0028-3932/Published by Elsevier Ltd.

2004; Van Petten and Rheinfelder, 1995) and for environmental sound targets primed by words, pictures, or other environmental sounds (Aramak et al., 2010; Cummings et al., 2006, 2008; Cummings and Čeponienė, 2010; Daltrozzo and Schön, 2009; Orgs et al., 2008; Orgs et al., 2006; Plante et al., 2000; Schirmer et al., 2011; Schön et al., 2010; Van Petten and Rheinfelder, 1995). Indeed several studies of N400 priming effects using bimodal (visual/auditory) stimulus presentation have found similar scalp distributions for the N400 priming effects to words and environmental sounds across multiple ages (Cummings et al., 2006, 2008; Cummings and Čeponienė, 2010; Orgs et al., 2007) Finally, functional imaging results have shown activation to both word and environmental sound stimuli in areas commonly thought of as language specific: left inferior frontal and superior temporal regions (Binder et al., 2000; Leech and Saygin, 2011; Price et al., 2005; Thierry et al., 2003; Tranel et al., 2003) and similar neural networks being implicated in the semantic processing of speech and musical sounds (Koelsch, 2005; Koelsch et al., 2004; Steinbeis and Koelsch, 2008).

Despite these similarities, there are some important differences between words and environmental sounds. These differences exist on multiple dimensions since environmental sounds are non-linguistic. Whereas words have an arbitrary linkage to the items to which they refer, environmental sounds obtain meaning through the causal relation with the event or object that produces them (Ballas and Howard, 1987). Thus, the “lexicon” of environmental sounds is rather small, and tends to converge on a limited number of referents (Ballas, 1993).

Consequently, there is also some empirical support for the notion that distinct mechanisms underlie the processing of each sound type. Behavioral evidence suggests that environmental sound recognition is more susceptible to interference from semantically related competitors (e.g. cow and horse) than word recognition is (Saygin et al., 2005). Additionally, there is evidence from dichotic listening studies that environmental sounds are processed more efficiently in the right hemisphere, whereas words are processed more efficiently in the left hemisphere (Knox and Kimura, 1970; Kimura, 2011). While seemingly at odds with the dichotic listening research, ERP studies (using uni-modal auditory presentation) have found words and environmental sounds exhibit different scalp distributions for N400 priming effects: words showing a larger effect over the right hemisphere, whereas environmental sounds show a larger effect over the left hemisphere (Van Petten and Rheinfelder, 1995; but see above). We note here that the scalp topography of an ERP component does not correspond in any straightforward way to the location of its underlying neural generators, but reflects the summed activity of all generators, which vary in location, strength, and orientation with respect to the scalp. Therefore a right-sided asymmetry at the scalp does not implicate right-hemisphere generation, and results from dichotic listening tasks and ERP scalp topography are not necessarily at odds.

To further bolster the idea that words and environmental sounds indeed call upon different processing routines, and hence different neural networks, functional imaging research has revealed differential intra- and inter- hemispheric activation patterns for words (left angular gyrus, and left anterior and posterior temporal areas) and for environmental sounds (left superior and middle temporal gyri and right superior temporal cortex) (Noppeney et al., 2008; Thierry et al., 2003; Humphries et al., 2001). Finally, using electrophysiological and hemodynamic measures concurrently, Renvall et al. (2012) found that adding background noise affected the recognizability, timing, and location of cortical responses differently for each sound class.

1.2. The effect of featural similarity on N400 responses

All stimuli with referential meaning, whether auditory, visual, orthographic, or pictorial, elicit an N400 component, which is a negative voltage deflection peaking approximately 400 ms post-stimulus onset (Kutas and Federmeier 2011; Kutas and Hillyard, 1980; Kutas and Hillyard, 1983). The prototypical (visual) N400 semantic incongruity effect – the relative amplitude of the waveform compared to another experimental condition (e.g. unprimed target minus primed target) – is typically maximal over right parietal, posterior temporal, and occipital sites. However auditory N400s tend to begin earlier, last longer, and have a somewhat more frontal and less right-biased scalp distribution than visual N400s (Holcomb and Neville,1990; and reviewed in Kutas and Van Petten, 1994). It has been shown that N400 amplitude (to visual or auditory stimuli) is sensitive to category membership. For instance, following a series of prime words from the same taxonomic category (e.g. flower), N400 amplitudes are larger for target words that belonged to a different category (e.g. apple) than target words belonging to the primed category (e.g. tulip) (Polich, 1985; review in Kutas and Van Petten, 1988).

N400 amplitude for words is not only sensitive to gross category membership (member vs. non-member) but is also incrementally sensitive to differences in featural similarity (i.e. the perceived likeness between concepts as measured by the degree of overlap in their semantic features) (Federmeier and Kutas, 1999, 2002; Federmeier et al., 2002; Ibáñez et al., 2006; Torkildsen et al., 2006). We know that the brain often represents feature information in a structured fashion such that neurons responding to similar features are physically close to one another (Hubel and Wiesel, 1972; Tanaka, 1996). If we characterize neural representations of words as a collection of features, then two words that share many features will show similarities in their underlying neural activity (Amuntz and Zilles, 2012; Federmeier et al., 1999). Indeed an incremental or graded effect in N400 amplitude for words based on featural similarity was first found by investigating the effects of sentential context on semantic memory organization (Federmeier and Kutas 1999; Federmeier et al., 2002). Federmeir and colleagues defined feature likeness in terms of taxonomic semantic categories (e.g. bears and pandas are within the same taxonomic category and therefore share more features than do bears and zebras). Participants were visually presented with sentences that ended in three types of words: expected exemplars (e.g. panda), unexpected exemplars from the same category (e.g. bear), and unexpected exemplars from a different category (e.g. zebra). Both within- and between-category violations exhibited significant N400 effects; however between-category violations (e.g. ‘zebra’ instead of ‘panda’) exhibited greater N400 amplitudes than within-category violations (e.g. ‘bear’ instead of ‘zebra’) (Federmeier and Kutas, 1999). What's more, the graded effect in N400 amplitudes based on featural similarity has been replicated in visual-auditory priming paradigms across the lifespan (Federmeier et al., 2002; Ibáñez et al., 2006; Torkildson et al., 2006).

Active listening paradigms (i.e. paradigms in which participants are given a concurrent behavioral task to maintain attention), like those mentioned just above, are not mandatory for eliciting the N400 priming effect; the effect has been repeatedly found during attentional blink tasks in which a visual stimuli is not detected due to rapid presentation (Luck et al., 1996; Maki et al., 1997; Vogel et al., 1998), and has even been found for participants who were presented with stimuli while asleep (Ibáñez et al., 2006). Thus active listening does not appear necessary to elicit N400 priming effects, and effects found with passive listening paradigms appear to be comparable to those found during active listening paradigms.

Together these works demonstrate that featural similarities between concepts in the world influence the neural organization of lexical items, and further, that the N400 is sensitive to the organization of lexical-semantic categories in memory. Organizing representations of words based on subtle differences in feature overlap creates a processing benefit for items that are related; transitions between the pattern of activation corresponding to one word and that corresponding to a different but related word are likely easier (Federmeier et al., 1999). However, no previous evidence we are aware of addresses whether featural similarity between concepts acts on the organization and subsequent processing of environmental sounds similarly as for words. Thus this is a crucial question for the current paper to address.

1.3. Current study

In this study, we examine the effect of featural similarity on the processing of words and environmental sounds by using a cross-modal sound-picture match/mismatch ERP paradigm, in which we varied feature likeness among within-category pictures and auditory stimuli. Participants viewed pictures (e.g. dog) presented with words or environmental sounds at three levels of featural similarity: those that match the picture (Match: e.g. "dog" or *barking*), those that share semantic features with the expected match (Near Violations: e.g. "cat" or *meowing*), and those that share few semantic features with the expected match (Far Violations: e.g. "lion" or *roaring*).

To elucidate similarities and differences in the structure of words and environmental sounds in memory, our primary interest is how the relative N400 amplitude of Near Violations, a condition not examined in previous studies, compares to the Match condition and the Far Violations for each sound type. For words, we expect to replicate findings that show a graded organization: both Near and Far Violations will exhibit significant N400 effects, however Far Violations will exhibit greater N400 effects than Near Violations.

For environmental sounds we expect Far Violations to exhibit significant N400 effects, consistent with previous work. However, predictions for the Near Violations are not as certain. One possibility, based on behavioral, ERP, and fMRI evidence of similar processing for words and environmental sounds, is that words and environmental sounds are similar with respect to their organization in memory. On this view, environmental sounds should show a similar graded pattern as words, with both Near Violations and Far Violations exhibiting significant N400 effects, and Far Violations exhibiting larger N400 effects than Near Violations. This outcome would suggest that featural similarities between concepts influence the neurocognitive organization of environmental sounds in a graded fashion, similar to that of words.

An alternate possibility, consistent with fMRI findings that pinpoint distinct neural activation for words vs. environmental sounds and behavioral findings that people are less accurate at differentiating within-category members for environmental sounds than words (Saygin et al., 2005), is that words and environmental sounds will not show similar effects of featural similarity. Thus the processing of environmental sounds may not be dependent upon a strict feature match between sounds and concepts. In this case, environmental sounds, unlike words, will exhibit a non-graded effect: Far Violations will show a significant N400 effect but Near Violations will not.

2. Materials and methods

2.1. Participants

This study is part of a larger project examining the neural response to words vs. environmental sounds in infants, toddlers, and adults. Here, we report data for 22 adults (15 women, 7 men, 18–38 years of age, mean age 24). All participants were monolingual native English speakers and right-handed. Participants were recruited from the undergraduate and graduate populations at San Diego State University and either volunteered to participate or participated for course credit.

2.2. Stimuli

Stimuli for the cross-modal picture-word match/mismatch paradigm were line drawings, auditory words, and environmental sounds of 44 highly familiar concepts such as "dog" (all of which were nouns). A female native English speaker produced the word stimuli (mean duration = 876 ms, SD = 197 ms), which were recorded in a single session in a sound-attenuating booth (sampling at 44.1 Hz, in 16-bit stereo). Environmental sound stimuli were obtained from several online sources (www.soundbible.com, www.soundboard.com, and www.findsounds.com) and from a freely downloadable database of normed environmental sounds (Hocking et al., 2013). Environmental sounds were standardized for sound quality (44.1 kHz, 16 bit, stereo). The duration of words (876 ms, SD = 197 ms) and environmental sounds (1045 ms, SD: 616 ms) did not significantly differ ($t(86) = 1.73$, $p = 0.09$).Visual stimuli were black and white line drawings (600×600 pixels) depicting typical exemplars of each concept. The images were digitally edited to remove backgrounds and distracting features.

In order to ensure that the concepts were associated with easily identifiable environmental sounds for the match condition, a Likert scale pretest was conducted. Ten native English-speaking SDSU undergraduates (independent of those who participated in the EEG portion of the study) were presented with 51 images of prototypical members of highly familiar concepts (e.g., dog) paired with an associated environmental sound (e.g., barking). Each image was presented twice, though in a randomized order, each time with a different exemplar of the associated environmental sound. Therefore, 102 presentations of image/sound pairs were presented one at a time with participants asked to rate, on a 1–5 scale (1 = not related and 5 = very/highly related), how well the picture and sound went together. Only those sounds that received a mean rating of 3.5 or higher were included as stimuli (7 of the 51 concepts were excluded because they did not receive a score of 3.5 or higher for either example sound, resulting in a final set of 44 concepts). If both sounds for the same image were above 3.5, we chose the sound with the higher score; if both sounds obtained the same score we chose the sound we thought was more stereotypical (see Appendix).

To create stimulus sets based on feature similarity we used similarity scores from existing semantic feature production norms (McRae et al., 2005). McCrae and colleagues asked adults to list different types of features, such as physical (perceptual) properties (how it looks, sounds, smells, feels, and tastes), functional properties (what it is used for and where and when it is used), and other facts (such as where it is from). Concept similarity scores were derived by calculating the cosine between each pair of concepts (on the basis of feature production frequencies). The scores ranged from 1, indicating perfect correspondence (i.e. the cosine between a concept and itself) to 0 indicating complete concept independence (i.e. the concepts have no features in common). To be considered a violation, concept similarity scores had to be between 0.20 and 0.70 for Near Violations and 0.02–0.20 for Far Violations. Moreover, the difference between Near and Far Violations scores for the same concept had to be greater than 0.3. For both the word and environmental sound conditions, pictures of each concept were paired with one of three types of auditory stimuli: Matches (average concept similarity score = 1); Near Violations (average concept similarity score = 0.4); and Far Violations (average concept similarity score = 0.1). Six of our 44 stimuli (bunny, bug, bee, bird, monkey, and fish) were not included in the McRae norms, so we used scores from the closest prototypical member of that category for which similarity score were available (rabbit, ant, wasp, robin, gorilla, and goldfish, respectively).

Near and Far Violation conditions were created to be only within category: animals, vehicle, tools and household/outside objects. We used solely within-category stimuli for two reasons: firstly, between and within category items can differ not only based on featural similarity, but also by broader animacy distinctions. For example, it has been shown that animate (e.g. dog) and inanimate (e.g. pen) environmental sounds activate distinct neural substrates (Lewis et al., 2005). Thus, differences between violation conditions could be due to the activation of different underlying neural networks and not to changes in feature likeness per se. Secondly, by using only within category members we were able to control for feature likeness in a more systematic way than if we used between category violations, which often times resulted in a similarity score of zero in the McRae et al., 2005 semantic feature production norms.

The same 44 items were used to make Match, Near Violation, and Far Violation featural similarity conditions (Appendix A). Thus the three conditions were largely identical (e.g., *Ringing* is an environmental sound present in all three conditions – Match, Near, Far – as is the corresponding word "telephone"). Note, however, while all 44 items were used in the Match condition, not all 44 items had suitable scores for the Near and Far violation conditions, per the McRae et al., norms. As a result, a portion of the 44 items were used as Near Violations (32 unique pictures) or Far Violations (30 unique pictures) more than once (but no more than 3 times) while others were not included at all (see Appendix for list of stimuli). Therefore the final stimulus list consisted of 44 Match trials, 44 Near Violation trials, and 41 Far Violation trials. Nevertheless, as the three featural similarity conditions for both words and environmental sounds consisted almost entirely of the same stimuli, the conditions were very well controlled for word frequency, imageability, concreteness, phonology, and other properties of the stimuli.

2.3. Design

We used a 2 (sound class: Word, Environmental Sound)x3 (feature similarity: Match, Near and Far violation) within subjects design; sound class was presented in a blocked fashion, so that we had two back-to-back runs that were conducted in a single experimental session. Each of the runs was composed of a single presentation list with 129 trials: 44 in the Match condition, 44 in the Near Violation condition, and 41 in the Far violation condition. Repeated pictures were always at least 10 trials apart in the presentation list. Two versions of each run list for each sound condition were created, the second of each in reverse order of the first.

2.4. Procedure

Participants were seated in a comfortable chair at a distance of roughly 140 cm from a LCD computer monitor in a dimly lit, electrically shielded and sound-attenuated room. Each subject participated in two back-to-back runs, one for each sound type, each lasting approximately 12 min. The only differences between the runs were the type of sound (Word or Environmental) and the trial order (reversed). The order of the runs was counterbalanced such that half the participants received the Word run first. On each trial participants were presented with a line drawing of a familiar concept. The pictures were centered on screen and relatively small, so that they could be identified by central fixation (subtending a visual angle of 4.95° on average). After 1500 ms

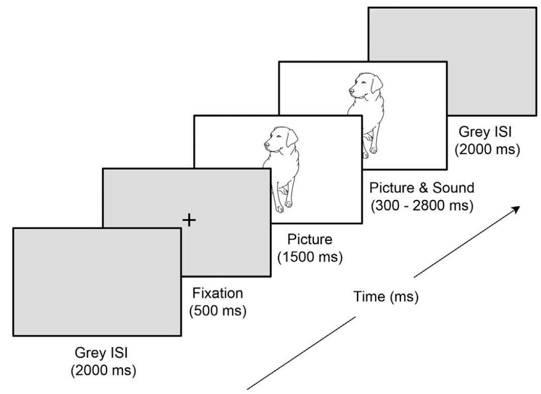

Fig. 1. Schematic of a single trial. For each sound type (Word and Environmental Sound), participants were presented with a line drawing of a familiar concept for 1500 ms before hearing a sound (300–2800 ms duration) from one of three conditions (Matches, Near Violation, Far Violation). The picture disappeared at the offset of the sound.

participants heard a sound from one of three levels of featural similarity (Match, Near Violation, Far Violation). As mentioned above, the current study is part of a larger project designed to include infants and toddlers as well as adults. As infants and toddlers require longer ISIs than adults (Richards, 2003), the longer delay between the onset of the picture and the onset of the paired word or sound was used to maintain consistency between the adult and child versions of this experiment for the larger project. The picture disappeared at the offset of the sound (300–2800 ms). A brief inter-stimulus interval grey screen was presented for 2000 ms, followed by a centrally located fixation cross for 500 ms (Fig. 1). After every 31 trials participants received a 12 s break, and a brief break was given between each run. Presentation of the Matches, Near Violations, and Far Violations was pseudo-randomized across the presentation list such that the same trial type did not appear for more than three consecutive trials.

Participants were asked to maintain their gaze toward the center of the screen and refrain from blinking during the picture presentation, but were not asked to make an overt response to the stimulus. We chose a passive listening paradigm instead of a paradigm requiring participants to respond to each stimulus for three reasons. Firstly, the infant and toddler populations we are testing for the larger project (16- and 24-months respectively) cannot execute reliable behavioral responses during the ERP experiment. Thus to maintain consistency between the adult and child samples we felt it necessary to make the ERP experiment a passive listening task. Secondly, we wanted to have the N400 as clean as possible, without obstruction by motor ERPs, which could differ significantly between conditions; there is evidence that people process environmental sounds faster than words (Cummings et al., 2006, Orgs et al., 2006), thus reaction times could have been faster for environmental sounds compared to words. Thirdly, N400 priming effects are reliably found during passive listening tasks (Ibáñez et al., 2006; Luck et al., 1996; Maki, et al., 1997; Vogel et al., 1998), such that an active task is not necessary for their elicitation, reducing task demands.

It should be noted that although we did not require a concurrent behavioral response, participants were told before the experiment that it was important for them to pay attention because questions about the pictures may be asked after the experiment. After the experiment each participant noted that some of the pictures matched the sound and when probed further, all participants were able to give word and environmental sound trial examples, indicating that they had indeed attended to the task.

2.5. EEG recording

EEG data was collected using a 64-electrode cap (Electro cap Inc.) according to the International 10–20 system. Tin electrodes were placed at the following locations (FP1, FPZ, FP2, AF3, AF4, F7, F5, F3, F1, FZ, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCZ, FC2, FC4, FC6, FT8, T7, C5, C3, C1, CZ, C2, C4, C6, T8, M1, TP7, CP5, CP3, CP1, CPZ, CP2, CP4, CP6, TP8, M2, P7, P5, P3, P1, PZ, P2, P4, P6, P8, PO7, PO5, PO3, POZ, PO4, PO6, PO8, CB1, 01, OZ, O2, CB2). EOG was recorded from electrodes positioned on the outer canthi of each eye as well as above and below the left eye. All channels were referenced to the left mastoid during data acquisition; data was re-referenced offline to the average of the left- and right-mastoid tracings. EEG was recorded at a sampling rate of 500 Hz, amplified with a Neuroscan Nuamps amplifier and low-pass filtered at 100 Hz. EEG gain was set to 20,000 and EOG gain set to 5000. Electrode impedances were kept below 5 kΩ.

2.6. EEG analysis

EEG was time locked to the auditory stimulus onset (spoken word or environmental sound) and epochs of 700 ms from auditory onset were averaged with a 100 ms pre-stimulus baseline. Trials containing eye movements, blinks, excessive muscle activity, or amplifier blocking were rejected trial-by-trial by off-line visual inspection before averaging (average rejection rate=9%). Data for three subjects in the Word run, and one subject in the Environmental Sound run were removed due to consistent, pervasive broad-spectrum noise. To analyze potential differences in distributional effects across conditions while minimizing the number of total comparisons, we coded electrodes along two dimensions: Anteriority (anterior vs posterior) and Laterality (left, central, right), effectively dividing them into six regions: Left-Anterior (F7, F5, F3, FT7, FC5, FC3, T7, C5, C3), Left-Posterior (TP7, CP5, CP3, P7, P5, P3, PO7, PO5, CB1), Central-Anterior (F1, FZ, F2, FC1, FCZ, FC2, C1, CZ, C2), Central-Posterior (CP1, CPZ, CP2, P1, PZ, P2, PO3, POZ, PO4, O1, OZ, O2), Right-Anterior (F4, F6, F8, FC4, FC6, FT8, C4, C6, T8), and Right-Posterior (CP4, CP6, TP8, P4, P6, P8, PO6, PO8, CB2); to balance the number of electrodes in each region, the following electrodes were not included in the analysis: FP1, FPZ, FP2, AF3, AF4.

Prior work indicates that N400 incongruity effects (i.e. unrelated items are more negative than related items), start earlier, and last longer in the auditory as opposed to the visual modality (Holcomb and Neville, 1990). Based on this prior work, and visual inspection of the grand average waveforms, we chose four time windows of interest: 200–300 ms, 300–400 ms, 400–500 ms, and 500–700 ms. For each sound type (Word and Environmental), mean amplitude voltage was computed separately for each condition (Match, Near Violations, and Far Violations) within the four time windows of interest. We analyzed these mean amplitude voltages using restricted maximum likelihood in a mixed-effects regression model with a random effect of subject on the intercept, fit with an unstructured covariance matrix. The model also included fixed effects of Sound Type (Word or Environmental), Featural Similarity (Match, Near Violation, Far Violation), Anteriority (Anterior/Posterior), Laterality (Left/Center/Right), and their interactions. We report Type III *F*-tests for the main effects and interactions of these factors. The three levels of Featural Similarity (Match, Near Violations, Far Violations) were contrasted within each Sound Type (Words, Environmental Sounds) for significant Sound Type x Featural Similarity interactions. For these contrasts, we report the regression coefficients (and standard error), *t*-values, *p*-values, and the 95% confidence interval. Note that the degrees of freedom are larger than in ANOVA approaches. The use of regression models offers several advantages over traditional ANOVA models, including robustness to unbalanced designs and a flexible ability to model different covariance structures, avoiding the need to correct for sphericity violations (see Newman et al., 2012, and references therein).

3. Results

The overall ERP response for Words and Environmental Sounds was similar in morphology and scalp distribution (Figs. 2 and 3; Table 1). However, condition-specific differences were present. Broadly, the two sound types show a similar pattern of ERP components across the scalp starting with a central-anterior N100 peaking near 150 ms, followed by a large centrally distributed P200 at 275 ms and an anteriorally distributed N200 peaking around 350 ms. After the N200 the ERPs are largely characterized by slower and negative-going waves that last through the end of the recording epoch. Different patterns of condition-specific effects are found for Words and Environmental Sounds on these late negative-going waves. Words showed incremental amplitude difference between conditions throughout this period, whereas Environmental Sounds did not. Additionally, differences in the amplitude of Far Violations compared to Near Violations and Matches appear earlier for Environmental Sounds (~150 ms) than for Words (~300 ms). Here we present results within each of our four time windows of interest.

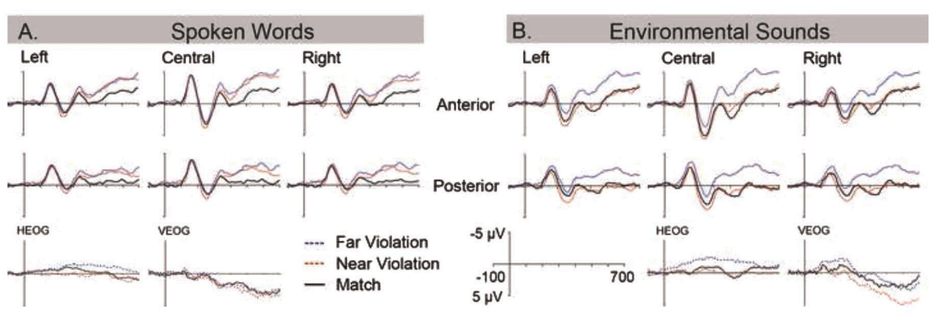

Fig. 2. Grand average ERP waveforms for the three levels of featural similarity for six regions – left-anterior, left-posterior, central-anterior, central-posterior, right-anterior, right-posterior – for Words (A) and Environment Sounds (B). ERPs from vertical electro-oculogram (VEOG) and horizontal electro-oculogram (HEOG) channels are included for illustrative purposes.

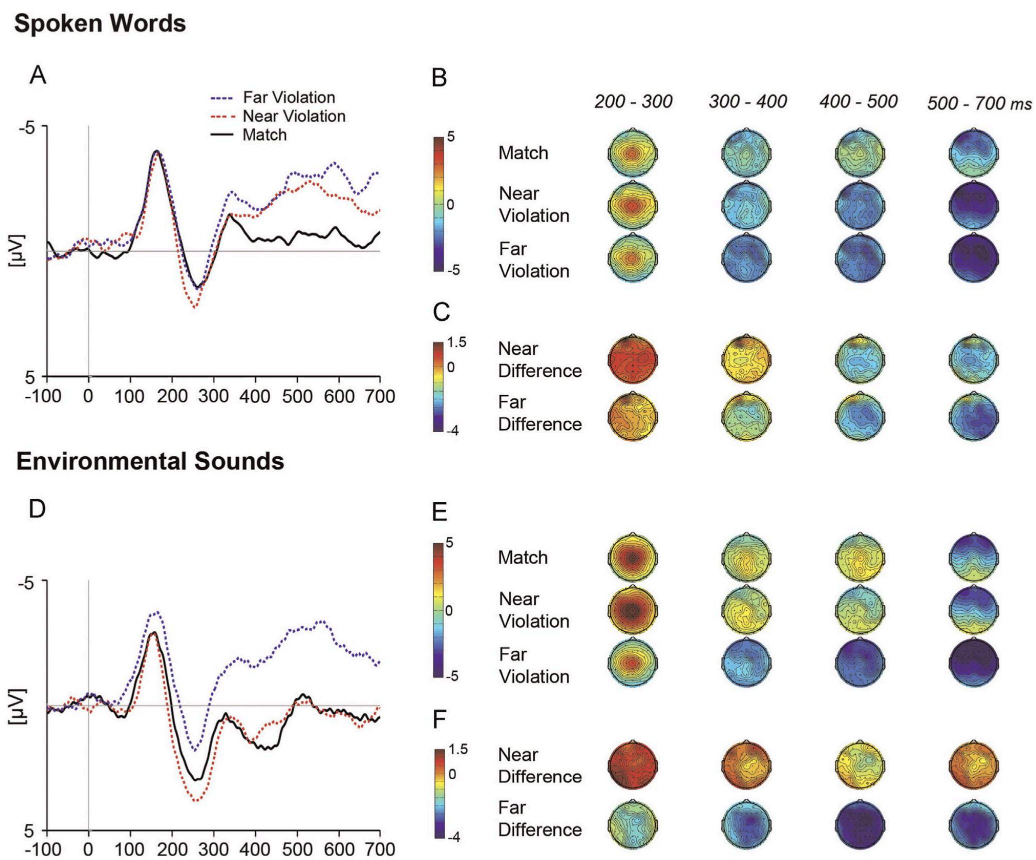

Fig. 3. ERP waveforms and voltage maps. Effect of featural similarity shown at the central-posterior region for Words (A) and Environmental Sounds (C). Voltage maps show average voltage for Words (B) and Environmental Sounds (E), and average voltage difference (measured as Violations – Matches) for Words (C) and Environmental Sounds (F) between two fixed latencies.

3.1. Time course analyses

3.1.1. 200–300 ms time window

This time window was characterized by a positive deflection in the waveform for all conditions. There were significant main effects of featural similarity [$F(2, 6747) = 132.39$, $p < 0.0001$], and sound condition [$F(1, 6751) = 298.85$, $p < 0.0001$], as well as a significant interaction [$F(2, 6747) = 35.64$, $p < 0.0001$]. Given the significant interaction, we examined contrasts between the levels of featural similarity separately for each sound type. For Words, Near Violations were significantly more positive than Matches, ($B = -0.59$ (0.13), $t(6747) = 4.45$, $p < 0.001$; 95% CI: [-0.85, -0.33]), and Far Violations ($B = -0.75$ (0.13), $t(6747) = 5.69$, $p < 0.0001$; 95% CI: [-1.01, -0.49]), and Matches and Far Violations were not significantly different ($B = -0.16$ (0.13), $t(6747) = 1.24$, $p = 0.21$; 95% CI: [-0.42, 0.10]). For Environmental Sounds, Near Violations were significantly more positive than Matches ($B = -0.91$ (0.13), $t(6747) = 7.21$, $p < 0.0001$; 95% CI: [-1.16, -0.67]) and Far Violations ($B = -2.22$ (0.13), $t(6747) = 17.61$, $p < 0.0001$; 95% CI: [-2.47, -1.97]). Matches were also significantly more positive than Far Violations ($B = -1.31$ (0.13), $t(6747) = 10.40$, $p < 0.0001$; 95% CI: [-1.56, -1.06]). There were no

Table 1
Mean Voltage (µV) for each Featural Similarity Condition for each Sound type, across all electrodes, at four time windows of interest. Standard error is shown in parentheses.

Time windows (ms)	Conditions					
	Words			Environmental Sounds		
	Match	Near	Far	Match	Near	Far
200–300	0.65 (0.10)	1.24 (0.08)	0.49 (0.09)	1.98 (1.3)	2.89 (1.2)	0.67 (0.18)
300–400	−0.75 (0.10)	-1.16 (0.07)	−1.88 (0.09)	0.53 (0.15)	0.65 (0.11)	-1.65 (0.16)
400–500	−.29 (0.10)	-1.87 (0.08)	−2.12 (0.10)	0.78 (0.17)	0.16 (0.12)	-2.46 (0.17)
500–700	−1.26 (0.10)	-2.99 (0.09)	−3.54 (0.11)	-1.08 (0.18)	-0.98 (0.14)	-3.67 (0.20)

significant interactions between featural similarity and Laterality and Anteriority.

3.1.2. 300–400 ms time window

In this time window there were significant main effects of featural similarity [$F(2, 6747)=205.46$, $p<0.0001$], sound condition [$F(1, 6752)=202.54$, $p<0.001$], and a significant interaction between them [$F(2, 6747)=39.16$, $p<0.0001$]. Given the significant interaction, we examined contrasts between the levels of featural similarity separately for each sound type. For Words, a graded effect was observed such that Far Violations were significantly more negative than both Matches ($B=-1.34$ (0.13), $t(6747)=8.66$, $p<0.0001$; 95% CI: [−1.39, −0.88]), and Near Violations ($B=-0.73$ (0.13), $t(6747)=5.53$, $p<0.0001$; 95% CI: [−0.98, −0.47]), and Near Violations were significantly more negative than Matches ($B=0.41$ (0.13), $t(6747)=3.13$, $p=0.0018$; 95% CI: [0.15, 0.67]). Environmental Sounds elicited a different pattern such that Far Violations were significantly more negative than both Matches ($B=-2.18$ (0.12), $t(6747)=17.44$, $p<0.0001$; 95% CI: [−2.42, −1.93]), and Near Violations ($B=-2.30$ (0.12), $t(6747)=18.40$, $p<0.0001$; 95% CI: [−2.54, −2.05]), however Matches and Near Violations were not significantly different ($B=-0.12$ (0.12), $t(6747)=0.97$, $p=0.33$; 95% CI: [−0.37, 0.12]). There were no significant interactions between featural similarity and Laterality or Anteriority.

3.1.3. 400–500 ms time window

In this time window there were significant main effects of featural similarity [$F(2, 6747)=351.24$, $p<0.0001$], sound condition [$F(1, 6747)=57.36$, $p<0.0001$], and a significant interaction between them [$F(2, 6747)=77.45$, $p<0.0001$]. Similar to the 300–400 ms time window, Words exhibited a graded effect such that Far Violations were marginally more negative than Near Violations ($B=-0.25$ (0.14), $t(6747)=1.79$, $p=.07$; 95% CI: [−0.52, 0.024]) and significantly more negative than Matches ($B=-1.83$ (0.14), $t(6747)=13.15$, $p<0.0001$; 95% CI: [−2.10, −1.56]), and Near Violations were significantly more negative than Matches ($B=1.58$ (0.14), $t(6747)=11.37$, $p<0.0001$; 95% CI: [1.31, 1.85]). In this time window, Environmental Sounds elicited a similar graded pattern such that Far Violations were significantly more negative than Near Violations ($B=-2.62$ (0.13), $t(6747)=19.8$, $p<0.0001$; 95% CI: [−2.88, −2.36]), and Matches ($B=-3.25$ (0.13), $t(6747)=24.5$, $p<0.0001$; 95% CI: [−3.51, −2.99]), and Near Violations were significantly more negative than Matches ($B=0.62$ (0.13), $t(6747)=4.71$, $p<0.0001$; 95% CI: [0.36, 0.88]). There were no significant interactions between Featural Similarity and Laterality or Anteriority.

3.1.4. 500–700 ms time window

In this time window there were main effects of featural similarity [$F(2, 6747)=269.81$, $p<.0001$], sound condition [$F(1, 6751)=29.55$, $p<0.0001$], and a significant interaction between them [$F(2, 6747)=58.38$, $p<0.0001$]. Consistent with the previous two time windows, Words demonstrated a graded amplitude pattern based on Featural Similarity such that Far Violations were significantly more negative than Near Violations ($B=-0.55$ (0.15), $t(6747)=3.59$, $p=0.0003$; 95% CI: [−0.86, −0.25]), and Matches ($B=-2.29$ (0.15), $t(6747)=14.77$, $p<0.0001$; 95% CI: [−2.59, −1.98]), and Near Violations were significantly more negative than Matches ($B=1.73$ (0.15) $t(6747)=11.19$, $p<.0001$; 95% CI: [1.43, 2.03]). Environmental Sounds did not exhibit a graded effect: Far Violations were significantly more negative than Matches ($B=-2.59$ (0.15), $t(6747)=17.57$, $p<.0001$; 95% CI: [−2.88, −2.30]) and Near Violations ($B=-2.69$ (0.15), $t(6747)=18.24$, $p<.0001$; 95% CI: [−2.98, −2.40]), however the Matches and Near Violations were not significantly different (B=-0.10 (0.15), $t(6747)=0.67$, $p=0.50$; 95% CI: [−0.39, 0.19]). There were no significant interactions between Featural Similarity and Laterality or Anteriority.

4. Discussion and conclusion

The primary aim of this study was to determine whether words and environmental sounds are organized similarly in semantic memory, and whether environmental sounds appear to be organized in memory according to featural similarity. We systematically varied the degree of featural similarity between an auditory stimulus and a preceding pictorial context for both sound classes. Consistent with previous work using similar methods, the overall ERP response for both sound types was similar in morphology and scalp distribution (Figs. 2 and 3). However, condition-specific differences were present. Far Violations (within-category but relatively unrelated to the preceding context), exhibited significant N400 effects for words and environmental sounds throughout all time windows of interest. This replicates and extends previous work showing that when put into similar context-dependent situations, environmental sounds elicit N400 peaks to between category violations similar to those elicited by auditory or visual words (Cummings et al., 2008; van Petten and Rheinfelder, 1995; Plante et al., 2000, Van Petten and Senkfor, 2000).

The finding that N400 effects to words and environmental sounds have similar scalp topographies is consistent with previous work using a similar cross-modal ERP priming paradigm (Cummings et al., 2006, 2008; Cummings and Čeponienė, 2010; Orgs et al., 2007). Other previous studies that report differences in scalp topography between words and environmental sounds used unimodal auditory stimulus presentation (spoken words primed environment sounds and environmental sounds primed spoken words; Van Petten and Rheinfelder, 1995). Once differences in presentation modality are taken into account, the current scalp topography results are indeed consistent with the extant literature.

Specific differences were also observed for each distinct sound type. Most relevant to the primary goal of the current paper are the differences in N400 amplitude for each sound type. For words, consistent with our prediction, we replicated previous research demonstrating gradedness in N400 amplitude modulated by featural similarity: both Near and Far Violations exhibited significant N400 effects, with Far Violations exhibiting greater N400 amplitudes than Near Violations. This graded pattern was consistent, starting in the early N400 time window (300–400 ms) and continuing throughout the epoch. In contrast to words, environmental sounds showed a pattern in which the Far Violations but not the Near Violations elicited significant N400 effects for both the early N400 time window (300–400 ms) and the late N400 time window (500–700 ms). The pattern of effects exhibited by environmental sounds is consistent with the alternative prediction based on behavioral work demonstrating that recognition (accuracy and response time) for environmental sound-object pairs is more affected by changes in feature likeness than is the recognition of word-object pairs (Saygin et al., 2005). Thus our findings lend support to the idea that the processing of environmental sounds is less sensitive to subtle changes in semantic features, and is organized in terms of a more general feature match between sounds and referents.

Environmental sounds did exhibit graded N400 amplitudes between 400–500 ms, although this effect was clearly shorter in duration than the graded effect observed for words. One possible interpretation for the relatively short-lived gradedness for environmental sounds from 400–500 ms is that the effect is an artifact of overlapping components. A slight positive deflection for the environmental sound Match condition in this time window would be consistent with a P300 component (Polich, 2012). Although this experiment did not require an explicit, task relevant response, and therefore not constructed to elicit P300 effects, the a priori probability of getting a Match was lower than the probability of getting a Violation (combined Near and Far conditions). Consistent with this interpretation, there is precedence in the literature for P300 effects being elicited in a passive listening paradigm (Bennington amd Polich, 1999). Thus, the fleeting graded effect between 400–500 ms for environmental sounds may be in part due to the Match condition demonstrating a greater positive deflection, which briefly distinguished the amplitudes of Near Violations and Matches. Indeed we observe such a putative positive deflection for words as well, however because Matches were more positive than both Violation conditions throughout the negative-going epochs, the positive deflection to Matches did not change the existing pattern of effects.

We also found differences in how feature similarity affected the processing of words and environmental sounds in the 200–300 ms time window: Matches were significantly more positive than Far Violations for environmental sounds but not words. We find two possible explanations for this early effect of featural similarity. Firstly, it may be due to low-level acoustic differences between the stimuli in the different conditions. However, we find this interpretation unlikely because the sounds used in the three conditions were highly overlapping, with virtually all stimuli present in all three conditions (e.g., *Ringing* is an environmental sound present in all three conditions, as is the corresponding word "telephone"; see Appendix). Another possible explanation is that these results reflect an early onset of the N400 component to the Far Violations for environmental sounds, causing them to be more negative in this early time window. As discussed above (see Section 1), the effects of semantic priming start earlier for auditory stimuli (Holcomb and Neville, 1990), and therefore may have affected the relative amplitudes of the featural similarity conditions in the preceding positive-going deflection. Moreover, previous research has shown that N400 priming effects appear earlier for environmental sounds than for words (Cummings et al., 2006; Orgs et al., 2006), consistent with the apparent lack of an early-onset N400 for words in this time window in our results.

Overall, our results suggest a fine-grained graded organizational structure for words based on featural similarity, with only a coarse distinction apparent for the relation between environmental sounds and concepts. However, there are several potential alternative hypotheses for the currents results: acoustic similarity between environmental sounds, precise encoding of lexical items, and differences in listener's familiarity with the sounds. Here we will discuss each alternative hypothesis in detail.

Although environmental sounds obtain their meaning through a causal production and are therefore, in principle bound to their sound source, in practice it may be quite difficult to determine the associated concept from hearing the sound alone (Ballas, 1993). It has been shown that there are more similar sounding environmental sounds that have different sources than there are similar sounding words with different meanings. For example, the environmental sound, "click–click" is ambiguous, and can be produced by a pen, a light switch, or many other possible referents (Ballas, 1993). Indeed people show more variation in their ability to recognize and identify environmental sounds than words (Gygi, 2001). Therefore it is possible that acoustic similarities between environmental sounds may be the interfering factor. That is, environmental sounds may be more coarsely organized in semantic memory than words because acoustic similarities between different environmental sounds causes an initial misread in the raw acoustic signal and results in a semantic misinterpretation. Relatedly, the acoustic stimuli for the environmental sound condition was numerically longer than the word condition (however, this difference did not reach significance). Therefore it is possible that the generation of the N400 responses was affected by the average duration of the two types of stimuli.

Another alternative interpretation is centered on lexical items being more precisely encoded. Unlike environmental sounds whose raw acoustic signal is likely directly linked to the semantic representation, word recognition undergoes multiple consecutive processing stages that translate a raw acoustic signal to a meaningful symbolic unit. Spoken words are first processed as a nonlinguistic auditory signal, then progress through a series of linguistic specific processing stages (phonetic and phonemic), and finally, higher-level brain structures compute semantic relevance (Frauenfelder and Tyler 1987; Indefrey and Levelt, 2004). The extra lexical encoding stages inherent in processing words may seek to further distinguish the raw acoustic input as referring to one possible referent. This precise level of encoding that is allowed in a relational mental lexicon (i.e., semantic and phonological relations), may not be possible for environmental sounds.

Finally, another interpretation highlights the possible differences in *a priori* levels of familiarity and exposure people have with words vs. environmental sounds. Although we pretested the materials in a group of adults to try to ensure that the environmental sounds were highly familiar, it is difficult to say with certainty whether there are differences in familiarity (and exposure) between these words and environmental sounds that could be driving the different ERP patterns. That is, participants could have more *a priori* exposure to the word stimuli than the environmental sound stimuli, and therefore, have had more time to semantically organize words in long-term memory. For the current study we included highly familiar words and environmental sounds to (a) increase the probability that participants have had long-term exposure to the auditory stimuli and, (b) ensure replication of previous findings of gradedness in N400 amplitudes for familiar words (Federmeier et al., 2002; Federmeier and Kutas, 1999, 2002; Ibáñez et al., 2006; Torkilson et al., 2006).

To our knowledge, the evidence presented here is the first to show differential semantic memory organization between linguistic and meaningful nonlinguistic auditory information. An important question for future research is whether behavioral measures would also be able to reveal the observed difference (e.g. graded vs. non-graded) in memory structure for the two types of auditory stimuli. The present study furthers our understanding of how the brain organizes meaningful auditory information and the relation between language and cognition. Additionally, this work has implications for broader theoretical accounts regarding whether the ability to procure meaning from words is subserved by semantic resources that are specific to language, or more global, cross-domain memory stores.

Acknowledgements

This research was supported by NIH R01DC009272 to the senior author, and the National Institutes of Health Training grant 5T32DC007361. We thank Zaira Flores, Kasey Palmer, and Jessica Onanian for assistance with data acquisition. Special thanks are extended to Dr. Phillip Holcomb and Dr. Jan R. Wessel for helpful comments on an earlier draft of this paper, and for their insights into data processing and analysis.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2015.01.035.

References

Amunts, K., Zilles, K., 2012. Architecture and organizational principles of Broca's area. Trends. Cogn. Sci. 16 (8), 418–426.
Aramaki, M., Marie, C., Kronland-Martinet, R., Ystad, S., Besson, M., 2010. Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. J. Cogn. Neurosci. 22 (11), 2555–2569.
Ballas, J., 1993. Common factors in the identification of an assortment of brief everyday sounds. J. Exp. Psychol.: Hum. Percep. Perform. 19 (2), 250–267.
Ballas, J., Howard, J., 1987. Interpreting the language of environmental sounds. Environ. Behav. 19 (1), 91–114.
Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. Cerebral Cortex 10 (5), 512–528.
B., Gygi, 2001. Factors in the identification of environmental sounds (Unpublished doctoral dissertation). Indiana University, Bloomington.
Chen, Y.C., Spence, C., 2011. Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. J. Exp. Psychol. Hum. Percep. Perform. 37 (5), 1554.
Cummings, A., Čeponienė, R., 2010. Verbal and nonverbal semantic processing in children with developmental language impairment. Neuropsychologia 48 (1), 77–85.
Cummings, A., Čeponienė, R., Dick, F., Saygin, A.P., Townsend, J., 2008. A developmental ERP study of verbal and non-verbal semantic processing. Brain Res. 1208, 137.
Cummings, A., Čeponienė, R., Koyama, A., Saygin, A.P., Townsend, J., Dick, F., 2006. Auditory semantic networks for words and natural sounds. Brain Res. 1115 (1), 92–107.
Cycowicz, Yael M., Friedman, David, 1998. Effect of sound familiarity on the event-related potentials elicited by novel environmental sounds. Brain Cogn. 36.1, 30–51.
Daltrozzo, J., Schön, D., 2009. Conceptual processing in music as revealed by N400 effects on words and musical targets. J. Cogn. Neurosci. 21 (10), 1882–1892.
Federmeier, K.D., Kutas, M., 1999. A rose by any other name: Long-term memory structure and sentence processing. J. Mem. Lang. 41 (4), 469–495.
Federmeier, K.D., Kutas, M., 2002. Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. Neuropsychologia. 40, 430–474.
Federmeier, K.D., McLennan, D.B., Ochoa, E., Kutas, M., 2002. The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: an ERP study. Psychophysiology 39 (2), 133–146.
Frauenfelder, U.H., Tyler, L.K., 1987. The process of spoken word recognition: an introduction. Cognition 25 (1), 1–20.
Frey, A., Aramaki, M., Besson, M., 2014. Conceptual priming for realistic auditory scenes and for auditory words. Brain Cogn. 84 (1), 141–152.
Hocking, J., Dzafic, I., Kazovsky, M., Copland, D.A., 2013. NESSTI: norms for environmental sound stimuli. PloS One 8 (9), 73382.
Holcomb, P.J., Neville, H.J., 1990. Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. Lang. Cogn. Process. 5 (4), 281–312.
Hubel, D.H., Wiesel, T.N., 1972. Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. J. Comparat. Neurol. 146 (4), 421–450.
Humphries, C., Willard, K., Buchsbaum, B., Hickok, G., 2001. Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. Neuroreport 12 (8), 1749–1752.
Ibáñez, A., López, V., Cornejo, C., 2006. ERPs and contextual semantic discrimination: degrees of congruence in wakefulness and sleep. Brain Lang. 98 (3), 264–275.
Indefrey, P., Levelt, W.J., 2004. The spatial and temporal signatures of word production components. Cognition 92 (1), 101–144.
Kay, P.,1971, Taxonomy and semantic contrast. Language, 47, 866–887.
Kimura, D., 2011. From ear to brain. Brain Cogn. 76 (2), 214–217.
Knox, C., Kimura, D., 1970. Cerebral processing of nonverbal sounds in boys and girls. Neuropsychologia 8 (2), 227–237.
Koelsch, S., 2005. Neural substrates of processing syntax and semantics in music. Curr. Opin. Neurobiol., 207–212.
Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., Friederici, A.D., 2004. Music, language and meaning: Brain signatures of semantic processing. Nat. Neurosci. 7 (3), 302–307.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science *207* (4427), 203–205.
Kutas, M., Hillyard, S.A., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cogn. *11* (5), 539–550.
Kutas, M., Van Petten, C., 1988. Event-related brain potential studies of language. . Adv. Psychophysiol. *3*, 139–187.
Kutas, M., Van_Petten, C., 1994. Psycholinguistics electrified. In: Gernsbacher (Ed.), Handbook of Psycholinguistics, pp. 83–143.
Leech, R., Saygin, A.P., 2011. Distributed processing and cortical specialization for speech and environmental sounds in human temporal cortex. Brain Lang. *116* (2), 83–90.
Lewis, J.W., Brefczynski, J.A., Phinney, R.E., Janik, J.J., DeYoe, E.A., 2005. Distinct cortical pathways for processing tool versus animal sounds. The Journal of Neuroscience *25* (21), 5148–5158.
Luck, S.J., Vogel, E.K., Shapiro, K.L., 1996. Word meanings can be accessed but not reported during the attentional blink. Nature *383* (6601), 616–618.
Maki, W.S., Frigen, K., Paulson, K., 1997. Associative priming by targets and distractors during rapid serial visual presentation: does word meaning survive the attentional blink? J. Exp. Psychol.: Hum. Percep. Perform. *23* (4), 1014.
McRae, K., Cree, G.S., Seidenberg, M.S., McNorgan, C., 2005. Semantic feature production norms for a large set of living and nonliving things. Behav. Res. Methods *37* (4), 547–559.
Murphy, G.L., Hampton, J.A., Milovanovic, G.S., 2012. Semantic memory redux: an experimental test of hierarchical category representation. J. Mem. Lang. 67 (4), 521–539.
Newman, A.J., Tremblay, A., Nichols, E.S., Neville, H.J., Ullman, M.T., 2012. The influence of language proficiency on lexical semantic processing in native and late learners of English. J. Cogn. Neurosci. *24* (5), 1205–1223.
Noppeney, U., Josephs, O., Hocking, J., Price, C.J., Friston, K.J., 2008. The effect of prior visual information on recognition of speech and sounds. Cereb. Cortex *18* (3), 598–609.
Orgs, G., Lange, K., Dombrowski, J.H., Heil, M., 2006. Conceptual priming for environmental sounds and words: an ERP study. Brain Cogn. *62* (3), 267–272.
Orgs, G., Lange, K., Dombrowski, J.H., Heil, M., 2007. Is conceptual priming for environmental sounds obligatory? Int. J. Psychophysiol. *65* (2), 162–166.
Orgs, G., Lange, K., Dombrowski, J.H., Heil, M., 2008. N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. Neurosci. Lett. *436* (2), 133–137.
Özcan, E., Egmond, R.V., 2009. The effect of visual context on the identification of ambiguous environmental sounds. Acta Psychol. *131* (2), 110–119.
Paczynski, M., Kuperberg, G.R., 2012. Multiple influences of semantic memory on sentence processing: distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. J. Mem. Lang. 67 (4), 426–448.
Plante, E., Petten, C.V., Senkfor, A.J., 2000. Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. Neuropsychologia *38* (13), 1669–1684.
Polich, J., 2012. Neuropsychology of P300. In: Luck, S.J., Kappenman, E.S. (Eds.), Oxford Handbook of Event-related Potential Components, pp. 159–188.
Polich, J., 1985. Semantic categorization and event-related potentials. Brain Lang. 2, 304–321.
Bennington, J.Y., Polich, J., 1999. Comparison of P300 from passive and active tasks for auditory and visual stimuli. Int. J. Psychophysiol. *34* (2), 171–177.
Price, C., Thierry, G., Griffiths, T., 2005. Speech-specific auditory processing: where is it? Trends Cogn. Sci. *9* (6), 271–276.
Renvall, H., Formisano, E., Parviainen, T., Bonte, M., Vihla, M., Salmelin, R., 2012. Parametric merging of MEG and fMRI reveals spatiotemporal differences in cortical processing of spoken words and environmental sounds in background noise. Cereb. Cortex *22* (1), 132–143.
Richards, J.E., 2003. Attention affects the recognition of briefly presented visual stimuli in infants: an ERP study. Dev. Sci. *6* (3), 312–328.
Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., Boyes-Braem, P., 1976. Basic objects in natural categories. Cogn. Psychol. 8 (3), 382–439.
Sajin, S.M., Connine, C.M., 2014. Semantic richness: The role of semantic features in processing spoken words. J. Mem. Lang. 70, 13–35.
Saygin, A.P., Dick, F., Bates, E., 2005. An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. Behav. Res. Methods *37* (1), 99–110.
Schirmer, A., Soh, Y.H., Penney, T.B., Wyse, L., 2011. Perceptual and conceptual priming of environmental sounds. J. Cogn. Neurosci.e *23* (11), 3241–3253.
Schneider, T.R., Engel, A.K., Debener, S., 2008. Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Exp. Psychol. 55* (2), 121–132.
Schön, D., Ystad, S., Kronland-Martinet, R., Besson, M., 2010. The evocative power of sounds: Conceptual priming between words and nonverbal sounds. J. Cogn. Neurosci. *22* (5), 1026–1035.
Shafiro, V., Gygi, B., 2004. How to select stimuli for environmental sound research and where to find them. Behav. Res. Methods, Instrum., Comput. 36 (4), 590–598.
Steinbeis, N., Koelsch, S., 2008. Comparing the processing of music and language meaning using EEG and FMRI provides evidence for similar and distinct neural representations. PLoS One *3* (5), 2226.
Stuart, G.P., Jones, D.M., 1995. Priming the identification of environmental sounds. Q. J. Exp. Psychol. *48* (3), 741–761.
Tanaka, K., 1996. Representation of visual features of objects in the inferotemporal cortex. Neural Netw. 9 (8), 1459–1475.
Thierry, G., Giraud, A.L., Price, C., 2003. Hemispheric dissociation in access to the human semantic system. Neuron *38* (3), 499–506.
Torkildsen, J.V.K., Sannerud, T., Syversen, G., Thormodsen, R., Simonsen, H.G., Moen, I., Smith, L., Lindgren, M., 2006. Semantic organization of basic-level words in 20-month-olds: An ERP study. Journal of Neurolinguistics 19 (6), 431–454.

Tranel, D., Damasio, H., Eichhorn, G.R., Grabowski, T., Ponto, L.L., Hichwa, R.D., 2003. Neural correlates of naming animals from their characteristic sounds. Neuropsychologia 41 (7), 847–854.
Van Petten, C., Rheinfelder, H., 1995. Conceptual relationships between spoken words and ES: event-related brain potential measures. Neuropsychologia *33* (4), 485–508.
Vogel, E.K., Luck, S.J., Shapiro, K.L., 1998. Electrophysiological evidence for a post-perceptual locus of suppression during the attentional blink. J. Exp. Psychol.: Hum. Percept. Perform. *24* (6), 1656.

Acknowledgments

Chapter 4, in full, is a reprint of material as it appears in Hendrickson, K., Walenski, M., Friend, M., & Love, T. (2015). The organization of words and environmental sounds in memory. *Neuropsychologia*, *69*, 67-76. The dissertation author was the primary investigator and author of this paper.

Appendix

Match		Near Violation		Far Violation	
Word	Environmental Sounds	Word	Environmental Sounds	Word	Environmental Sounds
Monkey	*Chattering*	Bear	*Roaring*	Lamb	*Baaing*
Clock	*Ticking*	Telephone	*Ringing*	Pen	*Scribbling*
Cow	*Mooing*	Bear	*Roaring*	Owl	*Hooting*
Rain	*Pitter-pattering*	Bubbles	*Gurgling*	-----	------------
Horse	*Neighing*	Cow	*Mooing*	Owl	*Hooting*
Elephant	*Trumpeting*	Horse	*Neighing*	Lamb	*Baaing*
Bus	*Idling*	Car	*Beep beep*	Bike	*Dinging*
Bee	*Buzzing*	Owl	*Hooting*	Turkey	*Gobbling*
Sheep	*Bleating*	Lamb	*Baaing*	Elephant	*Trumpeting*
Bike	*Dinging*	Truck	*Honking*	Airplane	*Propelling*
Dog	*Barking*	Cat	*Meowing*	Lion	*Roaring*
Telephone	*Ringing*	Pen	*Scribbling*	Broom	*Sweeping*
Tiger	*Growling*	Lion	*Roaring*	Sheep	*Bleating*
Bird	*Chirping*	Goose	*Honking*	Mouse	*Squeaking*
Boat	*Skimming*	Bus	*Idling*	Motorcycle	*Revving*
Duck	*Quacking*	Chicken	*Clucking*	Dog	*Barking*
Truck	*Honking*	Bus	*Idling*	Train	*Whistling*
Scissors	*Snipping*	Door	*Shutting*	Telephone	*Ringing*
Broom	*Sweeping*	Zipper	*Zipping*	Scissors	*Snipping*
Frog	*Ribbiting*	Cow	*Mooing*	Goose	*Honking*
Vacuum	*Sucking*	Bathtub	*Splashing*	Clock	*Ticking*
Cat	*Meowing*	Sheep	*Bleating*	Duck	*Quacking*
Mouse	*Squeaking*	Cat	*Meowing*	Tiger	*Growling*
Turkey	*Gobbling*	Chicken	*Clucking*	Monkey	*Chattering*
Hammer	*Hammering*	Door	*Shutting*	Pen	*Scribbling*
Pig	*Snorting*	Donkey	*Hee-hawing*	Bird	*Chirping*
Donkey	*Hee-hawing*	Horse	*Neighing*	Frog	*Ribbitting*
Owl	*Hooting*	Goose	*Honking*	Bee	*Buzzing*
Pen	*Scribbling*	Broom	*Sweeping*	Hammer	*Hammering*
Lion	*Roaring*	Elephant	*Trumpeting*	Duck	*Quacking*
Airplane	*Propelling*	Train	*Whistling*	Motorcyle	*Revving*
Door	*Shutting*	Keys	*Clinking*	Bubbles	*Gurgling*
Beach	*Crashing*	Rain	*Pitter-pattering*	Vacuum	*Sucking*
Bear	*Roaring*	Cat	*Meowing*	Pig	*Snorting*
Chicken	*Clucking*	Bird	*Chirping*	Bee	*Buzzing*
Keys	*Clinking*	Scissors	*Snipping*	Hammer	*Hammering*
Goose	*Honking*	Turkey	*Gobbling*	Dog	*Barking*
Bubbles	*Gurgling*	Beach	*Crashing*	Zipper	*Zippering*
Bathtub	*Splashing*	Rain	*Pitter-pattering*	------	------
Motorcycle	*Revving*	Bike	*Dinging*	Boat	*Skimming*
Zipper	*Zippering*	Vacuum	*Sucking*	------	------
Lamb	*Baaing*	Cow	*Moowing*	Tiger	*Growling*
Car	*Beep beep*	Truck	*Honking*	Airplane	*Propelling*
Train	*Whistling*	Car	*Beep beep*	Boat	*Skimming*

CHAPTER 5:

Organization of words and environmental sounds in toddlers

Abstract

The majority of research examining early auditory-semantic processing and organization is based on studies of meaningful relations between words and referents. However, a thorough investigation into the fundamental relation between acoustic signals and meaning requires an understanding of how meaning is associated with both lexical and non-lexical sounds. Indeed, it is unknown how meaningful auditory information that is not lexical (e.g., environmental sounds) in processed and organized in the young brain. To capture the structure of semantic organization for words and environmental sounds, we record event-related potentials (ERPs) as 20-month-olds view images of common nouns (e.g., dog) while hearing words or environmental sounds that match the picture (e.g., "dog" or barking), that are within-category violations (e.g., "cat" or meowing), or that are between-category violations (e.g., "pen" or scribbling). Results show electrophysiological markers of semantic processing are present during environmental sound processing in toddlers, however words and environmental sounds appear to be organized somewhat differently, with a more consistent fine-grained structure for words compared to environmental sounds.

1. Introduction

Auditory-semantic knowledge requires an appreciation of the relation between sounds and concepts, and an understanding of how concepts relate to one another. Indeed our ability to interpret the world depends fundamentally on how the brain organizes meaningful auditory information. In adults, lexical-semantic information exhibits a fine-grained organizational structure based on featural similarity – the perceived likeness between concepts – which aids in categorization (Kay, 1971; Murphy, Hampton, & Milovanovic, 2012; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Sajin, & Connine, 2014; Hendrickson, Walenski, Friend, & Love, 2015). There is recent evidence that the organization of the early lexical-semantic system may be mediated by featural information similar to the organization observed in adults (Ariejas-Trejo & Plunkett, 2009; 2013; Plunkett & Styles, 2009, von Koss Torkildson et al, 2006; Willits, Wojcik, Seidenberg, & Saffran, 2013).

To date, models of early auditory-semantic processing are primarily based on studies of meaningful relations between words and referents. However, auditory-semantic information can be divided into two categories: lexical (i.e., words) and non-lexical (e.g., environmental sounds such as the sound of a dog barking). Therefore, a thorough investigation into the fundamental relation between acoustic signals and meaning requires an understanding of how meaning is associated with both lexical and non-lexical sounds.

Such an investigation can further our understanding of the relation between language and cognition by examining whether an interconnected auditory-semantic network can be instantiated independent of language early in development. What's more, it has recently been suggested that the consistency with which environmental sounds are associated with their object referents may bootstrap the learning of more arbitrary word–object relations (Cummings et al., 2009). However, this claim is based on the assumption that the mechanisms of semantic integration that subserve the processing of words and environmental sounds are similar in the developing brain. Therefore, the overarching objective of this study is to use behavioral measures and event related potentials (ERPs) to compare how words and environmental sounds are organized early in language development.

1.1 Word vs. environmental sound processing in adults

Event-related potentials (ERPs) – a brain-based method that can identify well-defined stages of meaningful auditory processing – has been used to explore neural correlates of auditory-semantic integration and organization. The N400, a negative wave peaking approximately 400 ms post-stimulus onset, is an ERP component closely tied to semantic processing (Kutas & Federmeir, 2011; Kutas & Hillyard, 1980). All semantic stimuli (auditory, pictorial, orthographic) elicit an N400, whose amplitude is larger when the stimulus violates an expectancy set by a preceding semantic context. The N400 incongruity effect denotes the relative increase in N400 amplitude to a semantically unrelated stimulus.

Similar to behavioral results which show semantically congruent words or pictures can prime environmental sounds and vice versa (Ballas, 1993; Chen & Spence, 2011; Özcan & van Egmond, 2009; Schneider, Engel, & Debener, 2008), N400 incongruity effects have been found for words or pictures primed by related and unrelated environmental sounds (Schön, Ystad, Kronland- Martinet, & Besson, 2010; Daltrozzo & Schön, 2009; Frey, Aramaki, & Besson, 2014; Van Petten & Rheinfelder, 1995) and for environmental sounds primed by related and unrelated words, pictures, or other environmental sounds (Aramaki, Marie, Kronland-Martinet, Ystad, & Besson, 2010; Cummings et al.; 2006; 2008; 2010; Daltrozzo & Schön, 2009; Orgs, Lange, Dombrowski, & Heil, 2008; Orgs et al., 2006; Plante, Van Petten, & Senkfor, 2000; Schirmer, Soh, Penney, & Wyse, 2011; Schön et al., 2010; Van Petten & Rheinfelder, 1995). These effects hold from preadolescents to adulthood (Cummings, et al., 2008).

Not only is the N400 sensitive to semantic congruency (Kiefer, 2001; Kutas & Hilyard, 1980; Nigam, Hoffman, & Simons, 1992), but is also incrementally sensitive to differences in the featural similarity of concepts to which words refer (Federmeier & Kutas 1999b; 2002; Federmeir, Mclennan, Ochoa, & Kutas, 2002; Ibanez, Lopez, & Cornejo, 2006). It has been found that both within- and between-category violations exhibit significant N400 effects; however, between-category violations (e.g., 'zebra' instead of 'panda') exhibit greater N400 amplitudes than within-category violations (e.g., 'bear' instead of 'zebra') (Federmeier & Kutas, 1999b).

A recent ERP study examined whether semantic information for environmental sounds is organized in a similar fashion to words in adults (Hendrickson, et al., 2015). For this study, participants viewed a series of pictures (e.g. dog) presented with words or environmental sounds at three levels of featural similarity: those that match the picture (Match: e.g., "dog" or *barking*), those that share semantic features with the expected match (Near Violations: e.g., "cat" or *meowing*), and those that share few semantic features with the expected match (Far Violations: e.g., "lion" or *roaring*). Results showed that for words, a graded effect was replicated: starting around 300 ms N400 amplitudes were greatest for Far Violations, and smallest in the Match condition, with Near

Violations in between. For environmental sounds, significantly greater negative responses to between-category violations than matches occurred early in semantic processing (~ 200 ms), whereas differences in N400 amplitude to Near Violations and Matches occurred later (400 ms), and were short-lived (i.e., disappeared by 500 ms). These results indicate that for adults, the organization of words and environmental sounds in memory is differentially influenced by featural similarity, with a consistently fine-grained graded structure for words but not sounds.

1.2 Lexical-semantic organization in young children

Recent work examining when and how infants develop a system of words that are semantically related (featurally, functionally, associatively) comes from studies that use infant adaptations of adult lexical priming paradigms (Ariejas-Trejo & Plunkett, 2009; 2013; Hendrickson & Sundara, 2016; Plunkett & Styles, 2009, von Koss Torkildson et al, 2006; Willits, Wojcik, Seidenberg, & Saffran, 2013). Two-year-olds demonstrate a lexical priming effect such that related word primes yield longer looking times to a visual referent relative to unrelated primes. More recent evidence shows semantic priming as young as 18 months (Delle Luche, Durrant, Floccia, & Plunkett, 2014). This suggests that halfway through their 2nd year, children organize their lexical network based on the associative and featural similarities among semantic referents.

The neural architecture that underlies the N400 response develops ontogenetically early, as N400 effects have been observed in 9-month-olds viewing unanticipated action sequences (Federmeir & Kutas, 2011; Reid, Hoehl, Grigutsch, Groendahl, Parise, & Striano, 2009). In the domain of language, N400 effects in response to picture-word violations – such as those employed in the current study – have been seen as young as 12

-14 months (Friedrich & Friederici, 2008). What's more, research within the last decade has consistently shown that children in their second year exhibit an N400-like incongruity response to a variety of lexical-semantic violations. (Friedrich & Friederici, 2004, 2005; 2008; von Koss Torkildsen et al., 2006; 2009; Mills, Conboy, & Paton, 2005; Rama, Sirri, & Serres 2013). Similar to results obtained using behavioral measures (Delle Luche et al, 2014; Willits et al., 2013), there is evidence that the early lexical-semantic system is organized in a graded fashion based on featural similarity similar to the graded organization in adults (von Koss Torkildsen et al., 2006). Indeed, 20-month-olds display an N400-like incongruity effect that changes as a function of category membership; the incongruity response is earlier and larger for between category violations (e.g., dog and chair) than within-category violations (e.g., dog and cat).

1.3 Lexical vs. non-lexical auditory processing in young children

Behavioral research examining the relation between semantic processing for lexical vs. non-lexical auditory input (e.g. tones, mouth sounds, impact sounds) has demonstrated that children accept both types of auditory information as object associates well into their 2nd year (Campbell & Namy, 2003; May & Werker, 2014; Namy & Waxman, 1998). Indeed it is not until 24 months that a transition takes place such that a non-lexical auditory stimulus (e.g., gesture, whistle), formerly accepted as an object associate, is considered an unacceptable label, whereas words maintain their symbolic status (Namy & Waxman, 1998; May & Werker, 2014). Thus it has been suggested that at 24 months children undergo a form of linguistic specialization in which they begin to understand the role that language – as opposed to other types of information – plays in organizing objects according to subtle differences in featural similarity (Namy, Campbell,

& Tomasello, 2004).

Although the aforementioned studies have been fruitful for generating a broad understanding of differences and similarities between the way lexical and non-lexical auditory input is processed early in life, these studies used sounds (e.g., tones) that lack a natural semantic association with the paired visual referent, and thus are unlikely to tap into auditory-semantic systems. Therefore these studies do not directly assess infant's understanding of lexical versus inherently *meaningful* non-lexical sounds (e.g., environmental sounds) as a means of conveying semantic information.

Only one study has examined the semantic processing of environmental sounds younger than age seven. Cummings, Saygin, Bates, and Dick (2009) tested 15-, 20-, and 25-month-olds' using a looking-while-listening paradigm. Participants heard environmental sounds or spoken words when viewing pairs of images and eye movements to match versus non-match pictures were captured to determine the accuracy of object identification. Object recognition for environmental sounds and words was found to be strikingly similar across ages.

Although Cummings et al., (2009) found that recognition of sound-object associations for environmental sounds and words to be quite similar throughout the 2nd year of life, such results cannot shed light on whether there are differences in the underlying processes driving infants' overt responses to words vs. environmental sounds. Nor can the extant literature illuminate whether there are differences in how words and environmental sounds are semantically organized in the brain early in development.

1.4 Current Study

In the current study we investigate the structure of semantic organization for words and environmental sounds in the 2nd year of life by using ERPs. In order to ensure that any observed differences in organizational structure between words and environmental sounds were not due to *a priori* exposure, familiarity or comprehension, a language and environmental sounds familiarization and assessment was conducted within one and a half weeks prior to participation in the electrophysiological session. First, 20-month-olds participate in a behavioral familiarization task, to ensure that each child is familiar with both the words and environmental sounds associated with each concept tested during an ERP task. Subsequently, they participate in a behavioral task to assess word and environmental sound comprehension and speed of processing. Finally, we assess the organizational structure of words and environmental sounds by recording ERPs as participants view images of common nouns (e.g., dog) while hearing words or environmental sounds that match the picture (e.g., "dog" or barking), those that are a within-category violations (e.g., "cat" or meowing), and those that are a between-category violations (e.g., "pen" or scribbling).

This study has two aims. First, we examine whether children in their 2nd year show N400 incongruity effects for environmental sounds preceded by pictures that constitute within- or between-category semantic violations. Toddlers' ERP response to environmental sounds has not yet been examined. Therefore, the onset of the N400 – an ERP component linked to semantic integration – to this type of stimulus is unknown. We seek to establish whether N400 incongruity effects are obtained for environmental sounds given egregious semantic violations (between-category violation) since this is the contrast most likely to yield such an effect. Behavioral evidence suggests that children this age show similar performance in their recognition of familiar words and environmental sounds when presented with between category pictures (15-25-months; Cummings et al., 2009). Thus, we expect environmental sounds, like words, to exhibit significant N400 effects to between-category violations at 20 months. Further we expect children will show similar performance in their ability to comprehend referents for words and environmental sounds during the behavioral task.

The second aim of the study is to examine semantic memory organization of words and environmental sounds at 20 months. This aim primarily concerns how the N400 amplitude of within-category violations compares relative to the matches and between-category violations for each sound type. If at 24 months linguistic information undergoes specialization such that words become the predominant form of organizing semantically different but related items, we expect to observe different patterns of brain activity for each sound type at 20 months compared to adults (Hendrickson et al., 2015). Specifically, we expect that for 20-month-olds condition (between or within) will influence the ERP response for words and environmental sounds similarly: Both violations will exhibit significant N400 effects, however between-category violations will exhibit greater N400 amplitudes than within-category violations. However, like adults, environmental sounds could be organized more coarsely than words early in language development. Therefore an alternative prediction is that the processing of words and environmental sounds will be differentially influenced by the degree of semantic violation, with an earlier and more consistent fine-grained structure apparent for words than sounds (see Hendrickson, et al., 2015). From this view, we expect the ERP response to within-category violations to differ from matches and we expect this difference to be more robust (i.e., start earlier and last longer and/or evince larger effect sizes) for words compared to environmental sounds.

2. Materials and Methods

2.1 Participants

Children were obtained through a database of parent volunteers recruited through birth records, internet resources, and community events in a large metropolitan area. All infants were full-term and had no diagnosed impairments in hearing or vision. Overall, 26 children participated in this within-subjects study. All 26 participants completed the behavioral familiarization task. Of these, 18 children (8 F; 10 M) with a mean age of 20.5 months were included in the final analysis of the behavioral task. To be included in the final analysis participants had to complete at least 90% of trials. Data from 8 participants were excluded due to failure to complete the task (attrition rate = 30%). The ERP study was applied to 25 of the original 26 children. Of these 25, six children were excluded from the ERP study because of refusal to wear the cap (n = 2), and failure to obtain at least 10 artifact-free trials per condition for either words or environmental sounds (n = 4) (attrition rate = 28%). The final sample for the ERP study included 19 monolingual English–speaking children (9 F; 10 M; mean age = 20.6 months). The final within-subjects sample (including both the behavioral and ERP tasks) was 16 children (7 F; 9 M).

2.2 Stimuli

Stimuli for the behavioral and ERP tasks were colorful line drawings, auditory words, and environmental sounds of 30 highly familiar concepts (all of which were nouns). Concepts can be grouped into the following three categories: animals (dog, cat, owl, sheep, horse, cow, bird, frog, bee, elephant, duck, bear, chicken, monkey, pig), vehicles (fire truck, car, train, bicycle, motorcycle, airplane), and household objects (hammer, door, telephone, pen, clock, toothbrush, keys, zipper, broom). A female native English speaker produced the word stimuli (mean duration = 873 ms, SD = 200 ms), which were recorded in a single session in a sound-attenuating booth (sampling at 44.1 Hz, in 16-bit stereo). The average pitch of the word stimuli was 264.50 Hz (SD = 41.1 Hz). Environmental sound stimuli were obtained from several online sources (www.soundbible.com, www.soundboard.com, and www.findsounds.com) and from a freely downloadable database of normed environmental sounds (Hocking, Dzafic, Kazovsky, & Copland, 2013). Environmental sounds were standardized for sound quality (44.1 kHz, 16bit, stereo) and had a mean duration of 878 ms (SD = 251), and a mean pitch of 221.99 (SD = 119.20). The duration and pitch of the word and environmental sound stimuli did not significantly differ (duration: $t(30) = .033$, $p = .97$; pitch: $t(30) = 1.84$, $p = .07$). Visual stimuli were colorful drawings taken from Snodgrass and Vanderwart (1980).

Although comprehension norms are not available for children at 20 months of age, the words for the concepts used in this study are comprehended by an average of 54% of 16-month-olds, and produced by 62% of 24-month-olds (Fenson et al., 1994a). Therefore, these concepts should be highly familiar to children of 20 months. It must be noted that comprehension norms are not available for environmental sounds in that study. In order to ensure that the concepts were associated with easily identifiable environmental sounds, a Likert scale pretest was conducted. Ten native English-speaking college undergraduates were presented with 51 images of prototypical members of highly familiar concepts (e.g., dog) paired with an associated environmental sound (e.g., *barking*). Each image was presented twice, though in a randomized order, each time with a different exemplar of the associated environmental sound. Therefore, 102 presentations of image/sound pairs were presented one at a time with participants asked to rate, on a 1-5 scale (1= not related and 5= very/highly related), how well the picture and sound went together. Only those sounds that received a mean rating of 3.5 or higher were included as stimuli. If both sounds for the same image were above 3.5, we chose the sound with the higher score; if both sounds obtained the same score we chose the sound we thought was more stereotypical. The same 30 items were used to make match, within-category, and between-category conditions, and therefore conditions were very well controlled for word frequency, imageability, concreteness, phonology, and other properties of the stimuli.

2.3 Procedure

2.3.1 Session 1. Language and Environmental Sound Familiarization and Assessment

Within one and a half weeks prior to participation in the electrophysiological session, each subject participated in a language and environmental sound familiarization and assessment. First, the familiarization phase was used to ensure that each child was familiar with both the words and environmental sounds associated with each of the 30 concepts tested during the ERP task. Second, the assessment phase directly gauged participant's understanding of the words and environmental sounds using a two-alternative force-choice procedure. Third, we obtained parental ratings of participant's *a priori* familiarity with the words and environmental sounds associated with each concept. Together this assessment allowed us to equate exposure to the picture, word, and environmental sound stimuli, and determine whether there exist differences in participant's comprehension of the word- and environmental sound-concept associations.

2.3.1.1 *Parental Rating Scale*: Here, parents are asked to rate their child's familiarity with the list of words and environmental sounds from which the stimuli used in the ERP and behavioral tasks are drawn (30 concepts in all). They rate each on a 1 (certain their child is not familiar with it) to 7 (certain their child is familiar with it) scale. A similar rating scale has been used by Sheehan, Namy, & Mills (2007) for words and gestures.

2.3.1.2 *Behavioral Familiarization Task.* To help control for exposure and familiarity effects with the specific stimuli used in the ERP and behavioral tasks, we familiarize participants with the word-picture and environmental sound-picture combinations an equal number of times (total duration of task = 6 mins.). During this familiarization phase, participants are presented with each concept (30 in all) on a computer monitor, 6 times each, 3 with the corresponding environmental sound, and 3 with the corresponding word, in randomized order to equate *a priori* levels of exposure.

2.3.1.3 *Picture-pointing Task:* This task largely followed the protocol for the Computerized Comprehension Task (CCT; Friend & Keplinger, 2003, Friend, Schmitt, & Simpson, 2012). The CCT is two-alternative forced choice touch screen task that measures early decontextualized word knowledge. Previous studies have reported that the CCT has strong internal consistency (Form A: a = .836; Form B: a = .839), converges with parent report (partial r controlling for age = .361, $p < .01$), and predicts subsequent language production (Friend et al ., 2012). In addition, responses on the CCT are nonrandom (Friend & Keplinger, 2008) and this finding replicates across languages

(Friend & Zesiger, 2011) as well as for monolinguals and bilinguals (Poulin-Dubois, Bialystok, Blaye, Polonia & Yott, 2013).

For this procedure, infants are prompted to touch images on a touchscreen monitor by an experimenter seated to their right. Target touches (e.g. touching the image of the dog) elicit congruous auditory feedback over audio speakers (e.g. the word "dog"). Participants see each picture (30 in all) twice, once to test word comprehension (word block) and once to test environmental sound comprehension (environmental sound block). The order of blocks was counterbalanced such that half the participants received the word block first. Each block contained two training trials to ensure that participants understood the nature of the task. If the child failed to touch the screen after repeated prompts, the experimenter touched the target image for them. If a participant failed to touch during training, the two training trials were repeated once. Only participants who executed at least one correct touch during the training phase proceeded to the testing phase. For a given trial, two images appeared simultaneously on the right and left side of the touch monitor. The side on which the target image appeared was presented in pseudo-random order across trials such that the target was presented with equal frequency on both sides of the screen (Hirsh-Pasek & Golinkoff, 1996). During testing, each trial lasted until the infant touched the screen or until 7 seconds had elapsed at which point the image pair disappeared. When the participant's gaze was directed toward the touch monitor, the images appeared on the screen and the experimenter delivered a sentence prompt in infant-directed speech (*words*: Where is the___?; *environmental sounds*: Which one goes___?). The experimenter clicked the mouse so that the computer elicited the target word or environmental sound at the end of the experimenter prompt.

The criterion for ending testing was a failure to touch on two consecutive trials with two attempts by the experimenter to re-engage without success. If the attempts to re-engage were unsuccessful and the child was fussy, the task was terminated and the responses up to that point were taken as the final score. However, if the child did not touch for two or more consecutive trials but was not fussy, testing continued.

There were two measures obtained for this task: accuracy and reaction time. Accuracy was measured as the number of target touches executed during the task. Reaction time was calculated for target touch trials, and was measured at the moment the participant made contact with the touch screen upon hearing the target word or sound –i.e., the time from word or sound onset to touch response.

2.3.2 Session 2. Event-Related Potential Study

For the ERP study, sound class (word, environmental) was presented in a blocked fashion, resulting in two back-to-back runs with three conditions per run (match, within-category violation, between-category violation). Each of the runs was composed of a presentation list with 90 trials (30 trials for each condition). The presentation list was constructed so that a particular picture was not repeated on consecutive trials, and a particular sound was not repeated within 6 trials. Further, presentation of conditions was pseudo-randomized across the presentation list such that a given condition (match, within-category, between-category) did not appear for more than three consecutive trials.

2.3.2.1 ERP Testing Procedure

Participants were seated on their caregiver's lap at a distance of roughly 140 cm from a LCD computer monitor in a dimly lit, electrically shielded and sound-attenuated room. Each subject participated in two back-to-back runs, one for each sound type, each lasting approximately 8 minutes. The only difference between the runs was the type of sound presented (word or environmental). The order of the runs was counterbalanced such that half the participants received the word run first. As shown in Figure 1 (below), for each trial participants were presented with a colorful drawing of a familiar concept. The pictures were centered on screen and relatively small, so that they could be identified by central fixation (subtending a visual angle of 4.95 degrees on average). After 1750 ms participants heard a sound from one of three conditions (match, within-category violation, between-category violation). The picture disappeared at the offset of the sound (460 - 1235 ms). An inter-trial interval grey screen was then presented, its timing varied randomly between 500 and 1500 milliseconds. To maintain children's interest, an attention getter was programmed to appear on the screen every 10 trials and when the participant looked away from the screen for more than 2 seconds. Participants were video recorded during the EEG testing session to reject trials in which participants were not looking at the screen.

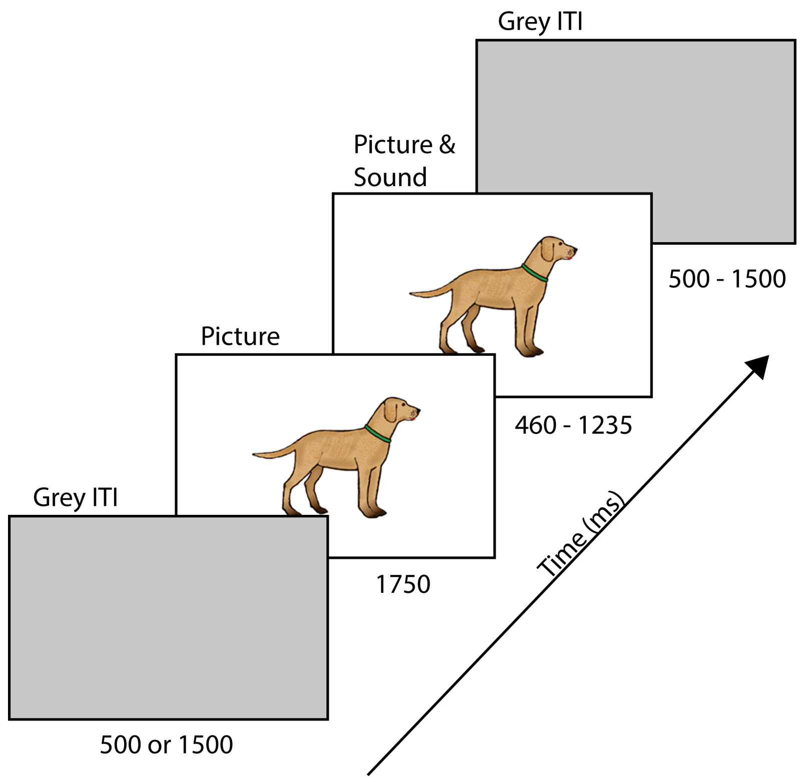

Figure 5-1. Schematic of a single trial. For each sound type (word, environmental), participants were presented with a colorful line drawing of a familiar concept for 1750 ms before hearing a sound (duration 460 – 1235 ms) from one of three conditions (match, within-category violation, between-category violation). Pictures disappeared at the offset of the sound. A variable inter-trial interval grey screen (duration 500 or 1500 ms) was presented at the offset of the picture and sound.

2.3.2.2 EEG Recording

EEG data was collected using a 21-electrode cap (Electro cap Inc.) according to the International 10-20 system. Tin electrodes were placed at the following locations (FP1, FP2, F7, F3, FZ, F4, F8, C3, CZ, C4, M1, M2, P3, PZ, P4, T3, T4, T5, T6, O1, O2) (see Figure 2). All channels were referenced to the left mastoid during data acquisition; data was re-referenced offline to the average of the left- and right-mastoid tracings. EEG was recorded at a sampling rate of 500 Hz, amplified with a Neuroscan Nuamps amplifier and low-pass filtered at 100hz. EEG gain was set to 20,000 and EOG gain set to 5,000. Electrode impedances were mostly below 5 KΩ, but at least below 20 KΩ.

2.3.2.3 EEG Analysis

EEG was time locked to the auditory stimulus onset (spoken word or environmental sound) and epochs of 1200 ms from auditory onset were averaged with a 200 ms pre-stimulus baseline. A zero-phase digital band-pass filter ranging from 0.2 to 30 Hz was applied to the EEG data. Before averaging, trials in which the child was not looking at the screen, and trials containing eye movements, blinks, excessive muscle activity, or amplifier blocking were rejected by off-line visual inspection of the EEG data and video recording. The average rejection rate was comparable between words (39.5%) and environmental sounds (36.9%). Participants were included in the final data set if they had 10 artifact-free trials per condition. Data for one subject in the Word run, and two subjects in the Environmental Sound run were removed due to insufficient data per condition (< 10 artifact-free trials). To analyze potential differences in distributional effects across conditions while minimizing the number of total comparisons, we coded frontal, central, and parietal electrodes (F3, FZ, F4, C3, CZ, C4, P3, PZ, P4) – where N400 effects have been shown to be maximal for similarly aged participant – along two dimensions: Anteriority (frontal, central, parietal) and Laterality (left, center, right). This effectively divided electrodes into nine regions: Left-Frontal (F3), Left-Central (C3), Left-Parietal (P3), Center-Frontal (FZ), Center-Central (CZ), Center-Parietal (PZ), Right-Frontal (F4), Right-Central (C4), and Right-Parietal (P4).

Prior work indicates that N400 incongruity effects (i.e. unrelated items are more negative than related items), start earlier, and last longer in the auditory as opposed to the visual modality (Holcomb & Neville, 1990). Based on this prior work, and visual inspection of the grand average waveforms, four time windows of interest were chosen: 200 – 400 ms, 400 – 600 ms, 600 – 800 ms, and 800 – 1000 ms. For each sound type (word and environmental), mean amplitude voltage was computed separately for each condition (match, within-category, and between-category) and electrode site within the four time windows of interest. For each sound type we analyzed these mean amplitude voltages using restricted maximum likelihood in a mixed-effects regression model with a random effect of subject on the intercept, fit with an unstructured covariance matrix. The model also included Condition (match, within-category, between-category), Anteriority (frontal, central, parietal), Laterality (left, center, right), and their interactions. We report Type III *F*-tests for the main effects and interactions of these factors. Condition was contrasted within each region and for significant Condition x Laterality or Condition x Anteriority interactions. For these contrasts, we report the regression coefficients (and standard error), *t*-values, *p*-values, effect size, and the 95% confidence interval. Note that the degrees of freedom are larger than in ANOVA approaches. The use of regression models offers several advantages over traditional ANOVA models, including robustness to unbalanced designs and a flexible ability to model different covariance structures, avoiding the need to correct for sphericity violations (see Newman et al., 2012 and references therein).

3. Results

3.1 Session 1: Behavioral and Parental Rating Results

To be included in analyses related to the picture-pointing task and the ERP task participants were required to complete the behavioral familiarization task. Of the 26 children who participated in the study, all were quiet and alert and maintained attention toward the screen for the duration of the behavioral familiarization task (6 mins.). Parental ratings of word and environmental sound familiarity were collected on all but one participant (N = 25).

Table 5-1. Means and standard errors of parental ratings of word and environmental sound familiarity and performance measures on picture pointing task (accuracy and reaction time).

Sound Type	Parental Rating (1 – 7 scale)	Picture Pointing Accuracy (30 items total)	Average Reaction Time
Word	5.62 (.18)	13.67 (1.31)	2426.92 (168.75)
Environmental	4.95 (.14)	12.94 (1.04)	2378.73 (226.52)

There was a significant difference in parent reported familiarity with the words and environmental sounds [t(24) = 5.67 ,p < .0001], such that parents reported that their children would be more familiar with the word stimuli compared to the environmental sound stimuli (see Table 1). In the picture pointing task, participants completed an average of 29.78 trials for words, and 29.94 trials for environmental sounds. There was no significant difference in participant's accuracy in identifying words vs. environmental sounds [t(17) = .47 , p = .64]. Further, there was no significant difference in reaction times to identify the visual referent for words vs. environmental sounds [t(17) = .22, p = .82] (see Table 1).

3.2 Session 2: ERP Results

The overall ERP response for words and environmental sounds was similar in morphology and scalp distribution (Figures 2 & 3). Broadly, the two sound types show a similar pattern of ERP components across the scalp starting with a N100 peaking near 100 ms, followed by a P200 at 175 ms, and a N200 peaking around 250 ms. After the N200 the ERPs are largely characterized by slower and negative-going waves that last through the end of the recording epoch. Further, effects were largest at centro-parietal electrode cites for both sound types (see Figure 2). However, condition-specific differences were present.

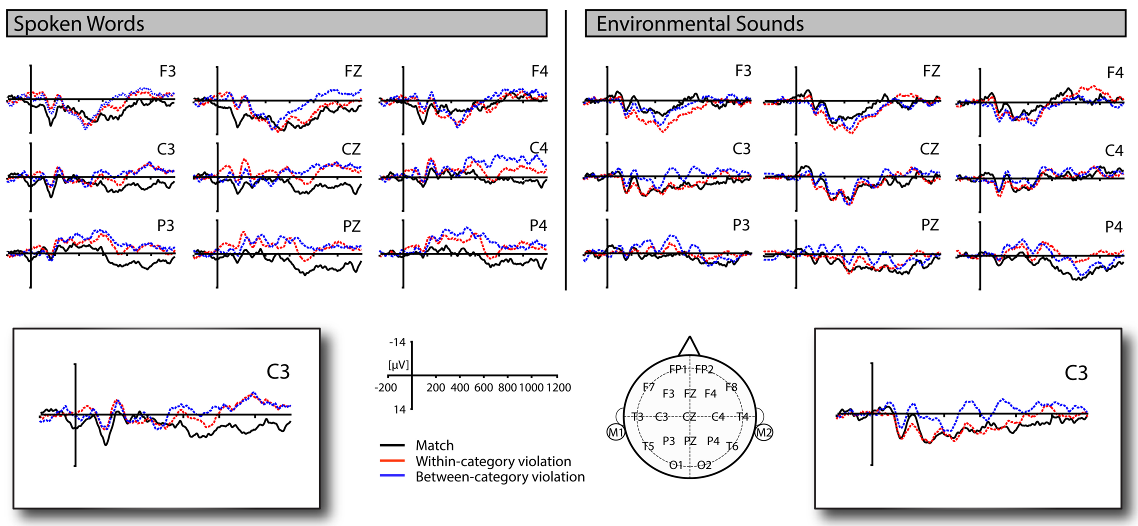

Figure 5-2. Grand average ERP waveforms for the 3 conditions (match, within-category violation, between category violation) at 9 electrodes of interest for words (left) and environmental sounds (right).

3.2.1 Time Course Analyses (see Figures 2 & 3)

3.2.1.1 200 - 400 ms Time Window

Words. There was a significant main effect of Condition [F(2, 442) = 5.33, p = 0.005], such that both within-category violations (B = -2.44 (0.91), t(442) = 2.69, p = 0.008; *d* = .30; 95% CI: [-4.24, -0.66]) and between category violations (B = -2.67 (0.91), t(442) = 2.95, p = 0.003; *d* = .31; 95% CI: [-4.48, -0.90]) were significantly more negative than matches. There was no significant difference between the amplitudes of within-and between-category violations (B = -.24 (.91), t (442) = .26, p = .79; 95% CI: [-

2.03, 1.55]). There were no significant interactions between Condition, Laterality and Anteriority.

Environmental Sounds. There was a significant main effect of Condition [F(2, 442) = 5.05, p = 0.007], in which between-category violations exhibited a significantly greater negative response than matches (B = -2.09 (0.90), t(416) = 2.34, p = 0.02; *d* = .31*;* 95% CI: [-3.85, -0.33]), but not within-category violations (B = -,62 (.90), t(416) = .7, p = .49; 95% CI: [-2.39, 1.14]). Further, between-category violations were significantly more negative than within-category violations (B = -2.71 (0.90), t(416) = 3.03, p = 0.003; *d* = .46; 95% CI: [-4.48, -0.96]). There were no significant interactions between Condition, Laterality and Anteriority.

3.2.1.2 400 - 600 ms Time Window

Both the *Words* and *Environmental Sounds* conditions revealed similar responses to the prior (200 – 400 ms) time window:

Words. There was a significant main effect of Condition [F(2,442) = 205.46, p < 0.0001], in which within-category violations were marginally more negative than matches (B = -1.79 (1.03), t(442) = 1.73, p = 0.08; *d* = .23*;* 95% CI: [-3.85, -0.33]) and between category violations were significantly more negative than matches (B = -2.69 (1.03), t(442) = 2.6, p = 0.01; *d* = .33; 95% CI: [-4.73, -0.66]). There was no significant difference between the amplitude of between- and within-category violations (B = -.90 (1.03), t(442) = .87, p = .38; 95% CI: [-2.94, 1.13]). Once again there were no significant interactions between Condition and Laterality and Anteriority.

Environmental Sounds. There was a significant main effect of Condition [F(2, 416) = 3.15, p = .04]. Again, between-category violations were significantly more negative than the matches (B = -1.84 (.96), t(416) = 1.92, p = 0.05; *d* = .28; 95% CI: [-3.72, .04]), but not within-category violations (B = -.42 (.96), t(416) = .44, p = .66; 95% CI: [-2.30, 1.46]). Further, between-category violations were significantly more negative than within-category violations (B = -2.26 (.96), t(416) = 2, p = 0.02; *d* = .34; 95% CI: [-4.14, -0.38]). There were no significant Condition, Laterality and Anteriority interactions.

3.2.1.3 600 – 800 ms Time Window

Words. There was a significant main effect of Condition [F(2, 442) = 29.35, p <.0001]. In this time window words exhibited a graded effect such that between-category violations were significantly more negative than within-category violations (B = -3.69 (1.00), t(442) = 3.67, p = .0003; *d* = .49; 95% CI: [-5.67, -1.71]) and significantly more negative than matches (B = -7.71 (1.00), t(442) = 7.66, p < 0.0001; *d* = .98; 95% CI: [-9.68, -5.73]), and within-category violations were significantly more negative than matches (B = -4.01 (1.00), t(442) = 3.99, p < 0.0001; *d* = .47; 95% CI: [-5.99, -2.04]). No significant interactions between Condition and Laterality and Anteriority were observed.

Environmental Sounds. In this time window there was no significant main effect of Condition [F(2, 416) = 1.9, p = .15], however there was a significant Condition x Anteriority interaction. Given the significant interaction, we examined contrasts between the levels of Condition separately for levels of Anteriority. The significant interaction was driven by a significantly greater negative response to between-category violations than matches at parietal electrode sites (B = -4.29 (1.72), t(416) = 2.5, p = .01; *d* = .40; 95% CI: [-7.67, -.91]), and a significantly greater negative response to matches than within-category violations at frontal electrode sites (B = -3.5 (1.72), t(416) = 2.05, p =

.04; *d* = .30; 95% CI: [-6.89, -.14]). There was no significant interaction between Condition and Laterality.

3.2.1.4 800 – 1000 ms Time Window

Words. In this final time window of interest we again found a significant main effect of Condition [F(2, 442) = 24.58, p < .0001]. Similar to results obtained in the 200 – 400 ms and 400 – 600 ms time windows, we found that both within-category violations (B = -5.42 (1.04), t(442) = 5.18, p < .0001; *d* = .61; 95% CI: [-7.46, -3.36]) and between category violations (B = -6.96 (1.04), t(442) = 6.68, p < .0001; *d* = .75; 95% CI: [-9.03 - 4.92]) exhibited a significantly greater negative response compared to matches. However there was no significant difference in the amplitude of between- and within-category violations (B = -31.57 (1.04), t(416) = 1.5, p = .13; 95% CI: [-3.62, .48]). Once again there were no significant interactions between Condition and Laterality and Anteriority.

Environmental Sounds. In this last time window there was a significant main effect of Condition [F(2, 416) = 5., p < 0.003], in which within-category violations were more negative than matches (B = -3.44 (1.02), t(416) = 3.38, p = 0.0008; *d* = .46*;* 95% CI: [-5.44, -1.44]) and between category violations were significantly more negative than matches (B = -2.00 (1.02), t(416) = 1.97, p = 0.05; *d* = .29; 95% CI: [-4.00, 0]). There was no significant difference between the amplitude of between- and within-category violations (B = 1.44 (1.02), t(416) = 1.41, p = .16; 95% CI: [-.56, 3.44]). Once again there were no significant interactions between Condition and Laterality and Anteriority.

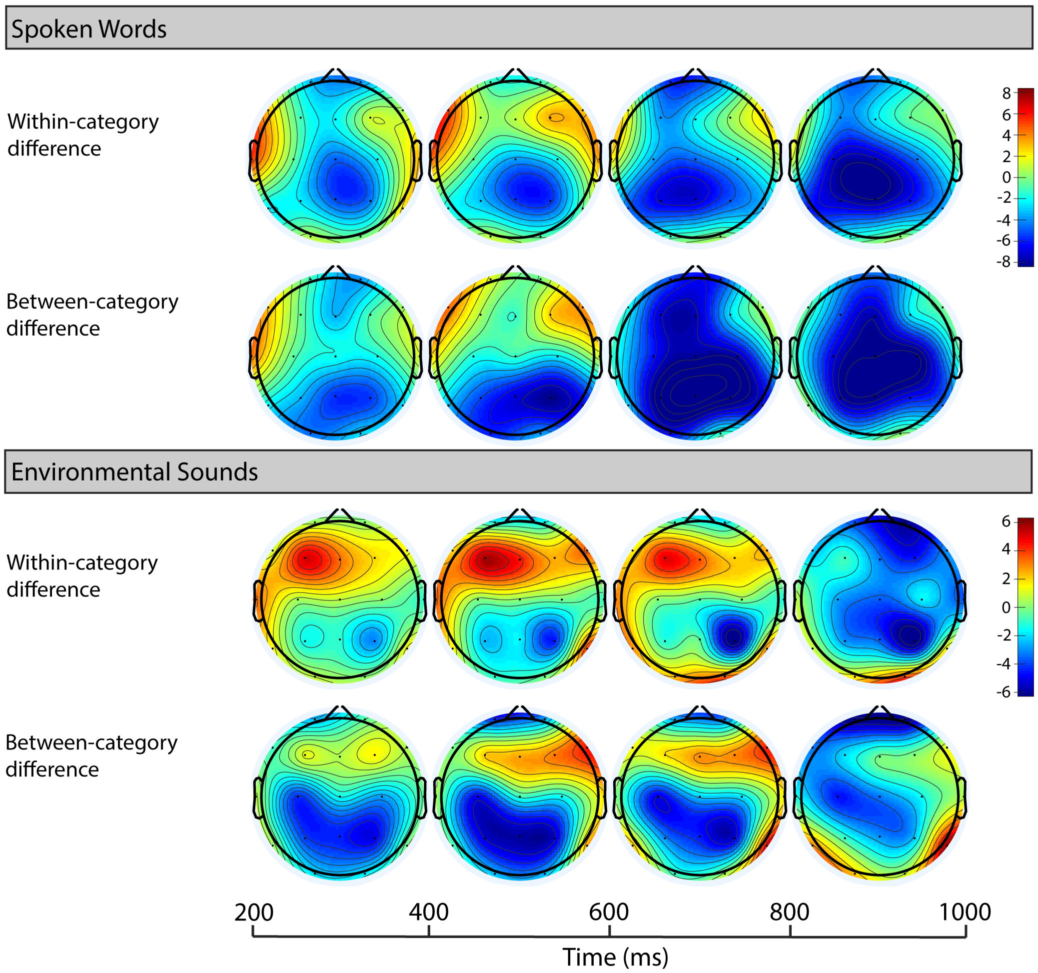

Figure 5-3. Voltage maps show average voltage difference (measured as violation – match) for within- and between category violations for words (upper) and environmental sounds (lower). Mean amplitude voltage calculated between two fixed latencies.

4. Discussion

The overarching goal of the current study was to determine whether words and environmental sounds are processed and organized similarly within the first two years of life. Prior to participating in an ERP task, participants completed a language and environmental sound familiarization and assessment. The familiarization phase was to ensure that each child had similar prior exposure to each of the 30 concepts tested during the ERP task. The assessment phase 1.) gauged children's *a prior* familiarity with the word and environmental sound stimuli through parental ratings, and 2.) measured comprehension of the words and sounds with a behavioral task. Parents rated the environmental sounds less familiar than the words, though the average rating for environmental sounds was 4.95 (out of 7), suggesting that parents were relatively certain their children would be familiar with the environmental sound stimuli. Results from the picture-pointing task revealed that recognition for sound-object associations for words and environmental sounds was quite similar: the speed and accuracy of word and environmental sound identification was statistically indistinguishable. These findings are comparable with results obtained using a visually based paradigm, which showed word and environmental sound recognition (accuracy and speed of processing) to be strikingly similar from 15 to 20 months (Cummings et al., 2009).

After the language and environmental sound assessment, children participated in an ERP task to measure how semantic relatedness effects the processing and organization of words vs. environmental sounds. We varied the degree of semantic violation between an auditory stimulus and a preceding pictorial context for both sound classes with the following two aims: 1.) To determine whether children in their 2nd year show N400 incongruity effects for environmental sounds preceded by pictures constituting between-category violations, and 2.) To examine semantic memory organization of words and environmental sounds at 20 months by evaluating between- and within-category violations. Results showed that the electrophysiological marker of semantic processing (N400) can be observed in young children's ERP response to environmental sounds (Aim 1), however the processing of words and environmental sounds is differentially influenced by the degree of the semantic violation (Aim 2).

For words, both within and between-category violations exhibited a significantly greater negative response compared to matches starting between 200 - 400 ms and lasting throughout all time windows of interest. Further we found a graded response in the mid-latency time window (600 – 800 ms), such that both within- and between-category violations exhibited significant N400 effects, with between-category violations exhibiting greater N400 amplitudes than within-category violations. This finding is consistent with previous behavioral and ERP results that show in the 2nd year of life, children organize their lexical network based on the associative and featural similarities among concepts to which words refer (Areja-Trejo et al., 2009; 2013; Hendrickson & Sundara, 2016; Plunkett & Styles, 2009, von Koss Torkildson et al, 2006; Willitz, Wojcik, Seidenberg, & Saffran, 2013).

Results showed that environmental sounds, like words, exhibit a significantly greater negative response to between-category violations compared to matches at 20 months. This effect was consistently present throughout all time windows of interest (200 – 1000 ms). These results are in line with previous behavioral evidence, and behavioral results within the current study that demonstrate children from 15-25 months display similar performance in their recognition of familiar words and environmental sounds when presented with between-category pictures (Cummings et al., 2009). Further, these results are consistent with adult research that reliably shows significant N400 effects to between-category picture-environmental sound violations.

In contrast to words however, environmental sounds showed a pattern in which matches and within-category violations were statistically indistinguishable in the early and mid-latency time windows. This pattern of results changed between 800 – 1000 ms, as both between- and within-category violations exhibited a significantly greater negative response compared to matches. Thus, unlike words that show a greater negative response to within-category violations compared to matches early and consistently, for environmental sounds such an effect occurs only in the late-latency time window.

The current results with 20-month-olds are strikingly similar to results obtained on an analogous ERP task with adults (Hendrickson et al., 2015). For adults it was found that words exhibited graded N400 amplitude starting early (300 ms) and such gradedness continued throughout all time windows of interests. Conversely, for environmental sounds, only between-category violations exhibited N400 effects in the early time window, though a graded pattern similar to that of words was exhibited at a later latency time window (400 – 500 ms). These findings are in line with the alternative hypothesis that suggests that, before children's 2^{nd} birthdays, the semantic organization of words and environmental sounds is different. Similar to adults, each sound type is differentially influenced by the degree of semantic violation, with a more consistently fine-grained structure apparent for words than sounds (see Hendrickson, et al., 2015). Indeed for words, differences between within-category violations and matches appears early in semantic processing (200 - 400 ms), continues throughout the late slow-negative going waves, and exhibits instances gradedness – i.e., both within- and between-category violations are more negative than matches, and between-category violations are more negative than within-category violations. For environmental sounds differences between within-category violations and matches do not appear until late in semantic processing (800 – 1000 ms), and no instances of gradedness were observed.

What drives these observed differences in how lexical and meaningful non-lexical sounds are temporally accessed and organized? Here we inventory several possible interpretations for the observed results. One explanation is that environmental sounds have more ambiguous relations with corresponding concepts. There are more similar sounding environmental sounds with different sources than there are similar sounding words with different meanings (Balls, 1993). Take for instance, the sound “click-click”, which can be produced by a plethora of sources (e.g., gun, door, pen). Indeed, concepts from within the same category are more likely associated with environmental sounds that are acoustically similar than concepts from different categories (Ballas, 1993). Therefore one explanation is that both 20-month-olds and adults misread the raw acoustic signal of the sound, which results in either a slowed access to or misidentification of the corresponding concept. Consistent with this possibility, behavioral evidence suggests that adults have more variation in their ability to identify environmental sound-concept associations than they do word-concept associations (Gygi, 2001, Saygin, Dick, Bates, 2005). Although previous behavioral work and the current behavioral results found object recognition for environmental sounds and words to be quite similar in toddlers, recognition was tested using between-category distractor images, and therefore is unlikely to tap more fine-grained semantic judgments.

Another interpretation for the current findings rests on the successive stages involved in encoding meaningful lexical vs. non-lexical acoustic signals. We know that word recognition progresses through a series of phonetic and phonemic processing stages that convert the raw acoustic signal to a lexical unit (Frauenfelder & Tyler 1987; Indefrey & Levelt, 2004). Unlike words, there is evidence that people do not linguistically mediate or sub-lexicalize environmental sounds (Schön, Ystad, Kronland-Martine & Besson, 2010), but instead environmental sounds themselves carry deep semantic associations with a corresponding referent (Ballas, 1993). These extra encoding stages may seek to further distinguish the raw acoustic signal of the word as referring to one possible concept. Whereas for environmental sounds, any misread in the raw acoustic signal may be linked directly with meaning, creating a more coarsely organized semantic system.

Finally, one last interpretation for the observed results involves people's experience with environmental sounds, which is largely relegated to the receptive domain. Indeed, the human vocal tract is limited in its ability to produce environmental sounds (Lewis, Brefczynski, Phinney, Janik, & DeYoe, 2005; Pizzamiglio et al., 2005). Conversely, words are both comprehended and produced, and by 20-months, children's productive vocabularies are rather large (~ 168 words; Dale & Fenson, 1996). Successful communication requires that people be able to rapidly produce words that are conceptually similar. Certainly, it is very important how the human brain organizes words for ease of retrieval during spoken word production. Although it is important that semantically different environmental sounds are comprehended differently, such communicative pressure to quickly and clearly retrieve environmental sounds with the intent to produce them is largely lacking.

A possible limitation in the current study involves potential *a priori* differences in the level of familiarity and exposure to these words and environmental sounds. That is, children could have more *a priori* experience to the words used in the study, and resultantly could have had more time to semantically organize the words. Although there was a significant difference in parent reported familiarity with the word and environmental sound stimuli, we find this interpretation unlikely for multiple reasons. First, we pretested these materials on a group of adults to ensure that the environmental sounds were highly familiar. Second, each child participated in a language and environmental sound assessment to equate levels of exposure to the stimuli used, and test comprehension of the words and environmental sounds. For this assessment, children first participated in a behavioral familiarization task, in which each concept tested during the ERP task was presented 6 times, 3 times with the associated word and 3 times with the associated environmental sound. Subsequently, each child participated in a picture pointing task to test their comprehension of the words and environmental sounds. Performance on this task revealed that object recognition and speed of processing was nearly identical for both sound types.

Although it could be argued that the sample size of the current study is small (N = 18) relative to other types of developmental work in infants, it must be noted that the N400 component is very robust, and can be observed with very few trials and participants (Luck, 2014). Indeed, prior studies using ERPs at this age suggest an N of 16 is sufficient to detect the N400 component (Friedrich & Friederici, 2008; Mani, Mills, & Plunkett, 2012; Mills, Prat, Zangl, Stager, Neville, & Werker, 2004). Importantly, effect sizes within the present study were robust, especially in the later time windows. Thus the sample size in the present study was appropriate to the research question and method with sufficient power to detect the N400 component with robust effect sizes.

5. Conclusion

The current study provides evidence that the electrophysiological marker of semantic processing (N400) can be observed in young children's ERP response to environmental sounds. However, this response does not differentiate readily between environmental sounds within the same category. Overall results from this study suggest that like adults, the young brain differentially organizes semantic information associated with words and environmental sounds, with a more consistent fine-grained structure apparent for words than sounds.

Acknowledgements

Chapter 5, in full, is being prepared for submission for publication of the material, Hendrickson, K., Love, T., Walenski, M., & Friend, M. (in preparation). The organization of words and environmental sounds in the 2nd year: Behavioral and Electrophysiological evidence. The dissertation author was the primary investigator and primary author of this paper.

We gratefully acknowledge Alyson Abel for providing facilities for this research and Leslie Carver, Philip Holcomb, and Jan R. Wessel for their seasoned review of our ERP findings. We also thank Taylor Gums and Carly Horne for assistance with data acquisition, and the families who participated in research. This research was supported by NIH Training Grant 5T32DC007361, NIH R01HD068458, and R01DC009272. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD, NIDCD, or the NIH.

References

Aramaki, M., Marie, C., Kronland-Martinet, R., Ystad, S., & Besson, M. (2010). Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. *Journal of Cognitive Neuroscience*, *22*(11), 2555-2569.

Arias-Trejo, N., & Plunkett, K. (2009). Lexical–semantic priming effects during infancy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1536), 3633-3647.

Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, *128*(2), 214-227.

Ballas, J. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2), 250–267.

Campbell, A. L., & Namy, L. L. (2003). The role of social-referential context in verbal and nonverbal symbol learning. *Child Development*, *74*(2), 549-563.

Chen, Y. C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1554.

Cummings, A., & Čeponienė, R. (2010). Verbal and nonverbal semantic processing in children with developmental language impairment. *Neuropsychologia*, 48(1), 77-85.

Cummings, A., Čeponienė, R., Dick, F., Saygin, A. P., & Townsend, J. (2008). A developmental ERP study of verbal and non-verbal semantic processing. *Brain Research*, *1208*, 137.

Cummings, A., Čeponienė, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research*, *1115*(1), 92-107.

Cummings, A., Saygin, A. P., Bates, E., & Dick, F. (2009). Infants' recognition of meaningful verbal and nonverbal sounds. *Language Learning and Development*, *5*(3), 172-190.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125-127.

Daltrozzo, J., & Schön, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience*, *21*(10), 1882-1892.

Delle Luche, C., Durrant, S., Floccia, C., & Plunkett, K. (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science*, *17*(6), 948-955.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*(4), 469-495.

Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133-146.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185.

Frauenfelder, U. H., & Tyler, L. K. (1987). The process of spoken word recognition: An introduction. *Cognition*, *25*(1-2), 1-20.

Frey, A., Aramaki, M., & Besson, M. (2014). Conceptual priming for realistic auditory scenes and for auditory words. *Brain and Cognition*, *84*(1), 141-152.

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of cognitive neuroscience*, *16*(8), 1465-1477.

Friedrich, M., & Friederici, A. D. (2005). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport*, *16*(6), 653-656.

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, *19*(18), 1757-1761.

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 302-309.

Friend, M., Schmitt, S. A., & Simpson, A. M. (2012). Evaluating the predictive validity of the Computerized Comprehension Task: Comprehension predicts production. *Developmental psychology*, *48*(1), 136.

Friend, M., & Zesiger, P. (2011). Une réplication systématique des propriétés psychométriques du Computerized Comprehension Task dans trois langues. *Enfance*, *2011*(03), 329-344.

Gygi, B. (2001). Factors in the identification of environmental sounds. Unpublished doctoral dissertation, Indiana University– Bloomington.

Hendrickson, K., & Sundara, M. (2016). Fourteen-month-olds' decontextualized understanding of words for absent objects. *Journal of child language*, 1-16

Hendrickson, K., Walenski, M., Friend, M., & Love, T. (2015). The organization of words and environmental sounds in memory. *Neuropsychologia*, *69*, 67-76.

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

Hocking, J., Dzafic, I., Kazovsky, M., & Copland, D. A. (2013). NESSTI: norms for environmental sound stimuli. *PloS One*, *8*(9), e73382.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*(4), 281-312.

Ibáñez, A., López, V., & Cornejo, C. (2006). ERPs and contextual semantic d iscrimination: degrees of congruence in wakefulness and sleep. *Brain and Language*, *98*(3), 264-275.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1), 101-144.

Kay, P. (1971). Taxonomy and semantic contrast. *Language*, 866-887.

Kiefer, M. (2001). Perceptual and semantic sources of category-specific effects: event-related potentials during picture and word categorization. *Memory & Cognition*, *29*(1), 100-116.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., & DeYoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *The Journal of Neuroscience*, *25*(21), 5148-5158.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Mani, N., Mills, D. L., & Plunkett, K. (2012). Vowels in early words: an event-related potential study. *Developmental Science*, *15*(1), 2-11.

May, L., & Werker, J. F. (2014). Can a Click be a Word?: Infants' Learning of Non-Native Words. *Infancy*, *19*(3), 281-300.

Mills, D. L., Conboy, B., & Paton, C. (2005). Do changes in brain organization reflect shifts in symbolic functioning. *Symbol use and symbolic representation*, 123-153.

Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14-and 20-month-olds. *Journal of Cognitive Neuroscience*, *16*(8), 1452-1464.

Murphy, G. L., Hampton, J. A., & Milovanovic, G. S. (2012). Semantic memory redux: An experimental test of hierarchical category representation. *Journal of Memory and Language*, 67(4), 521-539.

Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, *5*(1), 37-57.

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*,*69*(2), 295-308.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of Cognitive Neuroscience*, *24*(5), 1205-1223.

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience, 4*(1), 15-22.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2006). Conceptual priming for environmental sounds and words: an ERP study. *Brain and Cognition*, *62*(3), 267-272.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2008). N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. *Neuroscience Letters*, *436*(2), 133-137.

Özcan, E., & Egmond, R. V. (2009). The effect of visual context on the identification of ambiguous environmental sounds. *Acta Psychologica*, *131*(2), 110-119.

Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., & Di Russo, F. (2005). Separate neural systems for processing action-or non-action-related sounds. *Neuroimage*, *24*(3), 852-861.

Plante, E., Petten, C. V., & Senkfor, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia*, *38*(13), 1669-1684.

Poulin-Dubois, D., Bialystok, E., Blaye, A., Polonia, A., & Yott, J. (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, *17*(1), 57-70.

Rämä, P., Sirri, L., & Serres, J. (2013). Development of lexical–semantic language system: N400 priming effect for spoken words in 18-and 24-month old children. *Brain and language*, *125*(1), 1-10.

Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: evidence for semantic processing systems. *Developmental psychology*, *45*(3), 620.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382-439.

Sajin, S. M., & Connine, C. M. (2014). Semantic richness: The role of semantic features in processing spoken words. *Journal of Memory and Language,* 70, 13-35

Saygin, A. P., Dick, F., & Bates, E. (2005). An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. *Behavior Research Methods*, *37*(1), 99-110.

Schirmer, A., Soh, Y. H., Penney, T. B., & Wyse, L. (2011). Perceptual and conceptual priming of environmental sounds. *Journal of Cognitive Neuroscience*, *23*(11), 3241-3253.

Schneider, T. R., Engel, A. K., & Debener, S. (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)*, *55*(2), 121-132.

Schön, D., Ystad, S., Kronland-Martinet, R., & Besson, M. (2010). The evocative power of sounds: Conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience*, *22*(5), 1026-1035.

Sheehan, E. A., Namy, L. L., & Mills, D. L. (2007). Developmental changes in neural activity to familiar words and gestures. *Brain and Language*, *101*(3), 246-259.

Styles, S. J., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, *1*(1), 1-24.

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and ES: Event-related brain potential measures. *Neuropsychologia*, *33*(4), 485-508.

von Koss Torkildsen, J., Sannerud, T., Syversen, G., Thormodsen, R., Simonsen, H. G., Moen, I., Smith, L., & Lindgren, M. (2006). Semantic organization of basic-level words In 20-month-olds: An ERP study. *Journal of Neurolinguistics*, *19*(6), 431-454.

von Koss Torkildsen, J., Svangstu, J. M., Hansen, H. F., Smith, L., Simonsen, H. G., Moen, I., & Lindgren, M. (2008). Productive vocabulary size predicts event-related potential correlates of fast mapping in 20-month-olds. *Journal of Cognitive Neuroscience*, *20*(7), 1266-1282.

Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: evidence from auditory priming. *Infancy*, *18*(6), 1053-1075.

CHAPTER 6:

General Discussion and Conclusions

This dissertation investigated the nature of the early lexical-semantic system by examining the continuum of early lexical-semantic understanding and the specificity of its organizational structure. For this dissertation, I put forth a series of four studies to help answer three highly related questions: 1.) Is lexical knowledge dichotomous or continuous? 2.) How is lexical information semantically processed and organized? 3.) To what degree is such processing specific to language?

Question 1: Is lexical knowledge dichotomous or continuous?

Chapter 2 presented a study that evaluated the relation between the primary paradigms in use for measuring early word knowledge (i.e., visual reaction time, haptic response, and parent report) and evaluated a continuum of lexical-semantic knowledge in the 2nd year of life. It was found that parent report, visual RT, and haptic measures of word knowledge may give us a similar picture of general word comprehension abilities in young children. Though at the child level visual RT and haptic response appear to relate, at the trial level, participants engaged in behavioral dissociations. Convergence across modalities was observed during target touches (i.e., fast visual RT's and correct haptic response) and no touches (i.e., slow visual RT's and incorrect haptic response) revealing the most and least robust levels of comprehension, respectively. Behavioral dissociations emerged during distractor touches suggesting knowledge is strong enough to support rapid visual RTs, but too fragile to overcome a prepotent haptic response to the first image fixated (the Distractor).

Having established the presents of visual/haptic behavioral dissociations and corresponding partial knowledge at 16 months, the 2nd study of the dissertation sought to replicate these results 6 months later, and assess how the degree of word-referent understanding can influence future word recognition. Indeed results from Study 2 demonstrated that 22-month-olds, like 16-months-olds, display behavioral dissociations during distractor touches (i.e., fast visual RT's and incorrect haptic responses). According to the claims made by accumulative learning theories, learning is incremental, as past learning changes future learning (Gluck & Bower, 1988; Kruschke, 2001; Shiffrin & Schneider, 1977; Yurovsky, Fricker, Yu, & Smith, 2013). From this view, re-exposure to partially known words increases the probability of recognition compared to re-exposure to largely unknown words. Results from Study 2 generally support this notion: words that were partially known at 16 months were numerically more likely than unknown words to be known by 22 months. What's more, partially known words were significantly less likely to be unknown at 22 months than words that were unknown words to 16-month-olds.

Together the findings from studies 1 and 2 document a gradual increase in the depth of word understanding and the speed of word identification over the 2nd year of life, which is inconsistent with all-or-none view of word knowledge in early development. Instead these findings are more in line with the notion that learning is incrementally strengthened, as partial knowledge becomes more robust with experience over time. This graded structure of lexical understanding is also revealed in children's differential responses across tasks. That is, differences in task demands create scenarios in which children may show knowledge in one task but not the other. Instead of describing tasks as over- or underestimating what words children "know", it may be more useful to delineate contextual factors that contribute to these difference. Doing so will provide us with a more nuanced and dynamic picture of how lexical-semantic knowledge develops.

Consequently, these results have clinically applied implications for measuring early language delay. Approximately 10-20% of 2-year-olds are substantially behind their peers in their productive language abilities, and are at significantly elevated risk for later language impairments (Rescorla, 1989; Rescorla & Alley, 2001). Extant measures that seek to predict whether children will be identified as late-talkers, gauge early lexical-semantic understanding as all-or-none. However some of the major accounts that explain the underlying deficits of late-talkers suggest that their lexical-semantic systems contain sparse information about lexical concepts, and semantic networks have yet to be strengthened (Kail & Leonard, 1986). Therefore identifying measures that can gauge lexical-semantic knowledge across a continuum – i.e. from weak to robust – very early in development may be critical for boosting the predictive value of our current measures. Relatedly, if major theories of language delay involve weakly connected lexical-semantic networks, a better understand of the typical organizational structure of the auditory-semantic system is warranted.

Questions 2 and 3: How is lexical information semantically organized and to what degree is such processing specific to language?

The organizational structure of the lexical-semantic system has been well investigated in adult populations. A primarily factor in how words are organized in semantic memory is based on the featural similarity between concepts to which words refer, which aids in categorization. Recently, research has shown that early in development (18 - 24 months), children appreciate that certain words are more semantically related than others. In line with previous ERP research, studies 3 and 4 replicate findings of graded N400 amplitudes based on changes in featural similarity and category membership. For 20-month-olds and adults, violations that were highly related to the expected semantic stimuli (e.g., cat instead of dog) exhibited significantly greater negative responses than the expected stimuli starting early in semantic processing (200 – 400 ms), and continuing throughout the recording epoch.

Furthermore, results from studies 3 and 4 reveal language is not obligatory for creating an interconnected network of semantic information, as such a semantic network can be instantiated independent of language. However, differences exist in the timing and precision with which semantically related words vs. environmental sounds are organized in memory. For both 20-month-olds and adults, environmental sound violations that were highly related to the expected semantic stimuli resulted in a delayed N400-incongruity response that was relatively short-lived. In studies 3 and 4, several possible interpretations for the consistently fine-grained structure apparent for words but not sounds are discussed. Here I offer a more general interpretation of the findings with respect to how the functional significance of the N400 can explain the observed results.

Arguably the predominant interpretation of the functional significance of the N400 suggests that it represents the junction between feed-forward perceptual information (i.e., raw acoustic signal) and a dynamic, multi-modal semantic memory (see Kutas & Federmeir, 2011). Within 200 ms of hearing a meaningful sound, processing largely focuses on perceiving the lower-level information of the sound source, which may differ depending on the acoustic properties of the stimuli. Around the time of the N400, the low-level acoustic information begins to converge on a modality-independent semantic store. From this view, N400 amplitude is tied to the degree to which semantic information is activated in semantic memory. As a consequence, the amplitude of the N400 response to a given stimulus can be used to evaluate the degree of activation of certain semantic information: the more a semantic representation is activated by a given stimulus input, the smaller the difference between the N400 amplitude for that semantic representation will be compared to the control condition (i.e., match). Importantly, the N400 response likely exists at an early point in semantic processing, representing a crucial snapshot of accessing the semantic representation, but likely not the final state of processing in which the correct semantic information is fully and solely accessed. Consequently, the traditional time window of the N400 (300 – 500 ms) represents a point in which meaning associated with the input is still being negotiated due to competition with lexically related competitors. Therefore even after the time of the N400, the meaning associated with the stimulus is still being narrowed down. Subsequently, later-occurring processes serve to revise initial interpretations, or else update meaning representations, helping the listener to arrive at the appropriate semantic representation (Federmeier & Laszlo, 2009).

Based on this view, results from studies 3 and 4 suggest that *lexical*-semantic violations that are closely related to what is expected are pre-activated in semantic memory to a greater degree than are lexical-semantic violations that are less closely related to the match. Importantly, although highly related lexical items are pre-activated, the brain response appreciates that the highly related stimulus is indeed a violation (i.e., it is not the match). It has been argued that organizing representations of words in this way creates a processing benefit for items that are related; transitions between patterns of activation for related words are likely easier because they are already partially activated (Federmeir & Kutas, 1999). To illustrate, it is beneficial for the semantic system to pre-activate the word "mouse" after hearing the word "cat" because there is a high probability these words will co-occur.

One way in which differential activation between words and environmental sounds may occur – in terms of spreading activation – is that the more fine-grained organization for words leads to more selective activation. Conversely, the less fine-grained organization for environmental sounds means activation spreads more easily to more items, leading to greater activation throughout the network. Over-activating related items may create scenarios in which semantic competition is untenable, resulting in delayed arrival at or misidentification of the appropriate meaning.

Thus there is an initial benefit for environmental sounds because activation spreads easily, yet there is a subsequent delay in processing because there are more activated items to inhibit. As a result the fine-grained organization for words makes them more efficient to process. This interpretation is supported by the findings presented in Chapters 4 and 5. For instance, toddlers do not show significant differences between within-category violations and matches until 800 – 1000 ms post stimulus onset. Thus, it may be the case that when a 2-year-old hears an environmental sound, they over-activate semantically related items earlier in semantic processing (200 – 800 ms), after which point they refine the semantic representation and arrive at the appropriate semantic referent downstream (800 -1000 ms) (Federmeier & Laszlo, 2009). Further evidence for this interpretation comes from previous behavioral results that show that although adults are faster at processing environmental sounds, environmental sound recognition is more susceptible to interference from semantically related competitors (e.g. cow and horse) than word recognition is (Saygin et al., 2005).

Why are environmental sounds over-activated in this way? In studies 3 and 4, several possible explanations are offered: acoustic similarity between environmental sounds, precise encoding of lexical items compared to environmental sounds, and differences in listener's familiarity with sounds compared to words. To conclude this discussion, I offer one more interpretation involving the demands in processing sonic versus linguistic events.

Linguistic information is conventionally processed successively. That is, speech perception requires the listener to process fleeting sequential units (McMurray, Clayards, Tenehaus, & Aslin, 2008). Conversely, it can be argued that often times, sonic events require a level of simultaneous processing. For example, consider sitting at a café. The cacophony of a nearby couple talking, the clink of utensils, the hiss of the espresso machine, and the click of a pen may be perceived in tandem. What's more, these sounds are likely semantically related because they are all occurring in a single environmental context (the café). Therefore, it may be beneficial when processing environmental sounds to pre-activate a greater number of semantically related items to serve in the processing of highly related acoustic input in tandem – i.e., as one sonic event. However, such widespread activation is detrimental to processing when presented with environmental sounds sequentially, and therefore delays in processing may be observed (as is the case with the results for within-category violations in studies 3 and 4).

Understanding how the processing of lexical and meaningful non-lexical sounds compares throughout development has both theoretical and clinical implications. Such an investigation is vital for specifying the neural basis of language, and the nature of the auditory-semantic system. Indeed, understanding the structure and origins of semantic knowledge is the Gordian Knot in the study of cognition. By comparing how words and environmental sounds are processed early in development studies 3 and 4 further our understanding of the relation between language and cognition in two fundamental ways. First, results helped to ascertain whether conceptualization of lexical information diverges from the conceptualization of meaningful non-lexical information. Second, the current findings helped to discern whether language is obligatory for creating an interconnected auditory-semantic network, or whether such a network can be instantiated independent of language early in development.

Comparing the processing of words and environmental sounds early in development also has clinically applied implications. As previously mentioned, approximately 10-20% of 2-year-olds are significantly behind their peers in their productive language abilities (10th percentile or below on the MCDI: Words & Sentences form). Word learning is a daunting task because very early in development children must understand that the sound pattern of words bears an arbitrary relation to mental concepts and real-word referents. Environmental sounds, unlike words, have an inherent correspondence to a visual referent (Van Petten & Rheinfelder, 1995). Recently, it has been suggested that the consistency with which environmental sounds are associated with their object referents may play an important role in teaching young children about intermodal associations, which may bootstrap the learning of more arbitrary word–object relations (Cummings et al., 2009). However, this claim is based on the assumption that these sound classes are processed similarly in the developing brain. Results from study 4 demonstrate that words and environmental sounds show similarities in mechanisms of semantic integration (the presence of a robust N400-incongruity response), which is an exciting first step toward verifying the efficacy of highlighting environmental sound-object associations to serve in the learning of words.

In sum, this dissertation provided evidence for an early lexical-semantic system that is dynamic. Findings from studies 1 and 2 suggest that early word recognition is graded, and partial knowledge states play a role in future learning. Studies 3 and 4 replicate the finding from adult work that the lexical-semantic system is organized as an interconnected semantic network. Although electrophysiological markers of semantic processing are present during environmental sound processing in toddlers and adults, words and environmental sounds appear to be organized somewhat differently, with a more consistent fine-grained structure for words compared to environmental sounds.

References

Cummings, A., Saygin, A. P., Bates, E., & Dick, F. (2009). Infants' recognition of meaningful verbal and nonverbal sounds. *Language Learning and Development*, *5*(3), 172-190.

Federmeier, K. D., & Laszlo, S. (2009). Time for meaning: Electrophysiology provides insights into the dynamics of representation and processing in semantic memory. *Psychology of learning and motivation*, *51*, 1-44.

Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of memory and Language*, *41*(4), 469-495.

Gluck, M. A., & Bower, G. H. (1988). Evaluating an adaptive network model of human learning. *Journal of memory and Language*, *27*(2), 166-195.

Kail, R., & Leonard, L. B. (1986). Word-finding abilities in language-impaired children. *ASHA monographs*, (25), 1-39.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of mathematical psychology*, *45*(6), 812-863.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of psychology*, *62*, 621.

McMurray, B., Clayards, M. A., Tanenhaus, M. K., & Aslin, R. N. (2008). Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic bulletin & review*, *15*(6), 1064-1071.

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, *45*(2), 360-371.

Rescorla, L., & Alley, A. (2001). Validation of the Language Development Survey (LDS) A Parent Report Tool for Identifying Language Delay in Toddlers. *Journal of Speech, Language, and Hearing Research*, *44*(2), 434-445.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127.

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia*, *33*(4), 485-508.

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