Assessing a continuum of lexical–semantic knowledge in the second year of life: A multimodal approach

Kristi Hendrickson a,⇑,1, Diane Poulin-Dubois b, Pascal Zesiger c, Margaret Friend d

a Joint Doctoral Program in Language and Communicative Disorders, San Diego State University and University of California, San Diego, San Diego, CA 92120, USA
b Department of Psychology, Concordia University, Montreal, Quebec H3G 1MB, Canada
c Faculty of Psychology and Educational Sciences, Université de Genève, 1211 Genève 4, Switzerland
d Department of Psychology, San Diego State University, San Diego, CA 92182, USA

Article info
Article history:
Received 24 May 2016
Revised 4 January 2017
Available online 24 February 2017

Keywords:
Words
Partial knowledge
Behavioral dissociations
Incremental learning
Lexical processing

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 instead exists on a continuum from absence of knowledge, to partial knowledge, to robust knowledge. This longitudinal study tested a group of 16- to 18-month-olds, 6 months after their initial visit, to replicate results of partial understanding as shown by visual–haptic dissociations and to determine whether 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. Furthermore, results suggest that partial word knowledge at one time predicts the degree to which that word will be understood in the future.

Published by Elsevier Inc.
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 (Carey & Bartlett, 1978; Heibeck & Markman, 1987; Houston-Price, Plunkett, & Harris, 2005; Markson & Bloom, 1997; Trueswell, Medina, Hafri, & Gleitman, 2013; Woodward, Markman, & Fitzsimmons, 1994). Evidence for single-shot word learning comes from Trueswell and colleagues (2013), who proposed a model in which each time a word–object pair is encountered, a single conjecture regarding word meaning is made in a binary fashion. If the conjecture is correct, the association between word and object is strengthened; if the conjecture is incorrect, the word–object association is abandoned. From this view, word meaning is not gradually accrued by accumulated partial knowledge but instead is confirmed or altogether disconfirmed with multiple exposures.

An alternate view suggests that word knowledge is not dichotomous but instead exists on a continuum from absence of knowledge, to partial knowledge, to robust knowledge (Frishkoff, Perfetti, & Westbury, 2009; Henderson, Weighall, & Gaskell, 2013; Hendrickson, Mitsven, Poulin-Dubois, Zesiger, & Friend, 2015; Ince & Christman, 2002; McClelland & Elman, 1986; McMurray, 2007; Schwanevflugel, Stahl, & McFalls, 1997; Shore & Durso, 1991; Steele, 2012; Stein & Shore, 2012; Suanda, Mugwanya, & Namy, 2014; Whitmore, Shore, & Smith, 2004; Zareva, 2012). In contrast to theories that suggest that word learning is all-or-none (Gallistel, Fairhurst, & Balsam, 2004; Trueswell et al., 2013), incremental learning theories of word comprehension rely on the assumption that word knowledge is incremental and unfolds over time (McMurray, Horst, & Samuelslon, 2012; Rogers & McClelland, 2004; Siskind, 1996; Spivey et al., 2010; Yu, 2008; Yu & Smith, 2007; Yurovsky, Fricker, Yu, & Smith, 2014). For instance, studies that examine arm movements during mouse-click responses reveal a graded competitive effect in categorization tasks (Dale et al., 2007). Recent connectionist models have corroborated the view of partial knowledge as a central component in characterizing lexical development as well (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 over time, which has been shown to result in unforeseen outcomes; for example, 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 (Gurteen, Horne, & Erjavec, 2011; 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 touch (touched the picture of the word referent), distractor touch (touched the picture of the unprompted word), or no touch (failed to touch either image). Importantly, all of the target and distractor pairs were selected to be of equal difficulty, from the same word class (noun, verb, or adjective), and from the same category (animal, artifact, activity, color, or size). This design maximizes the similarities between target and distractor, such that only established word–referent relations are likely to yield a target touch. In the study, 16-month-olds’ visual reaction times to fixate a prompted image were significantly slower during no touches compared with distractor and target touches, which were statistically indistinguishable. Therefore, in the case of distractor touches, the visual and haptic response modalities conflict; that is, children are quick to fixate the target image but touch the distractor image. So, although evidence within the same study demonstrates a significant relation among different measures of word comprehension (visual reaction time, haptic response, and parent report), at the item level word knowledge is highly task dependent; children can demonstrate knowledge in one modality (e.g., visual) but not in 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; Munakata, 2001; Munakata & McClelland, 2003). The model demonstrates that when active memory for currently relevant knowledge is fragile, the system is not strong enough to surmount a prepotent response established by previous experience, resulting in a behavioral dissociation. For example, a highly interconnected semantic system allows for the activation of related representations from just partial information (Munakata, 2001). This could result in an incomplete representation (e.g., an overgeneralized mapping of the word dog) becoming activated by a related exemplar (e.g., a cat). Such dissociations may reflect the amount of coactivation between active (current relevant information) and latent (semantically related competitor).

According to this view, incorrect (distractor touch) and absent (no touch) haptic responses 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, using 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 and colleagues (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 with 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. Whereas 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; Yurovsky & Frank, 2015; Yurovsky et al., 2014), 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.

The 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 where investigators selectively use one measure or possibly two measures (DeAnda, Arias-Trejo, Poulin-Dubois, Zesiger, & Friend, 2016; Fernald, Perfors, & Marchman, 2006; Houston-Price, Mather, & Sakkalou, 2007; Hurtado, Marchman, & Fernald, 2008; Legacy, Zesiger, Friend, & Poulin-Dubois, 2016; Marchman & Fernald, 2008; Poulin-Dubois, Bialystok, Blaye, Polonia, & Yott, 2013). Indeed, no study to date has examined the relation among haptic, visual, and parent report measures of word knowledge within the same cohort of children over time. Therefore, the current study had three primary aims. The first aim (Aim 1) was to examine developmental changes in the speed and accuracy of word recognition across these three measures in the same cohort of children throughout the second year of life. The second aim (Aim 2) was to replicate and extend the finding of visual and haptic behavioral dissociations and corresponding partial knowledge states across the second year (Hendrickson et al., 2015). Finally, the third aim (Aim 3) was to investigate the role that partial knowledge plays in early word comprehension over this same time period.

This study is part of a larger longitudinal project examining early language and literacy development (N = 62). The analyses related to Aims 1 and 2 included a group of children for whom we had visual reaction time (visual RT), haptic response, and parent report measures of vocabulary knowledge at 16 months and 22 months (N = 39). For Aim 1, we examined the correlations between children’s vocabulary knowledge indexed by the haptic modality, speed of lexical access indexed by visual RT, and the well-documented MacArthur–Bates Communicative Development Inventories (MCDI; Fenson et al., 1993) at 16 and 22 months. We anticipated these correlations to reveal stability from 16 to 22 months for all measures. For Aim 2, 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 is
processed (visual RT) as a function of the type of haptic response: target touch (touched the picture of the prompted image), distractor touch (touched the picture of the unprompted image), or no touch (failed to touch either image). In line with previous results, we predicted that visual RT would vary as a function of haptic response (Hendrickson et al., 2015). Specifically, we predicted that visual RTs would be fastest for target touches and slowest for no touches, with an intermediate speed of processing for visual RTs associated with distractor touches.

For analyses related to Aim 3, we included a larger sample ($N=62$) for whom we had haptic (but not visual RT) performance data at 16 and 22 months. We assessed 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 haptic responses (target touches), incorrect haptic responses (distractor touches), and absent haptic responses (no touches) have been hypothesized 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 modality but not in the 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 participants’ comprehension of 41 words. In contrast to previous studies that focused on correctly selected referents to gauge vocabulary knowledge and size, for this analysis we instead focused on the words for which participants give incorrect answers: distractor touches (partially known) and no touches (unknown). Critically, participants were 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 with unknown words (no touches) (see Fig. 1).

**Method**

**Participants**

In this study, we brought back toddlers who participated in a larger, multi-institutional longitudinal project assessing language comprehension in the second year of life ($N=62$; mean age = 23.0 months, range = 21.3–24.9; 30 female and 32 male). Touch response data at 16 and 22 months were collected on all 62 participants. Of these 62 participants, for Aims 1 and 2 we also obtained visual response data on the first 50 of these participants. Of these 50 participants, 46 (mean age = 22.98 months, range = 21.2–24.6; 20 female and 26 male) contributed visual and haptic data at 16 months ($n=4$ participants were excluded due to fussiness), and 39 contributed visual and haptic data at both 16 and 22 months ($n=3$ were lost to attrition and $n=4$ were excluded due to fussiness at 22 months). 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 infants. Only those infants with at least 80% language exposure to English were included in the study (DeAnda, Bosch, Poulin-Dubois, Zesiger, & Friend, 2016; DeAnda, Hendrickson, Zesiger, Poulin-Dubois, & Friend, 2016).

**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 high-definition 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 participants’ haptic response and the stimulus pair presented on the touch monitor. Two audio speak-
ers 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**

On 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-participants forms of the procedure, such that distractors on one form serve as targets on the other. All image pairs presented during training, testing, and reliability were matched for word difficulty (easy, medium, or hard) based on MCDI: Words and Gestures (WG) norms (Dale & Fenson, 1996; Frank, Braginsky, Yurovsky, & Marchman, in press), part of speech (noun, adjective, or verb), category (animal, human, artifact, activity, color, or size), and visual salience (color, size, or luminance). A previous attempt had been made to automate the procedure, such that verbal prompts come from the audio speakers positioned behind the touch monitor instead of from the experimenter seated to the right of children. 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, 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 & Keplinger, 2003; Friend et al., 2012). Previous studies have reported that the CCT has strong internal consistency (Form A: \( \alpha = .836 \); Form B: \( \alpha = .839 \)), 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 et al., 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 4 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 sides of the touch monitor. The side on which the target image appeared was presented in pseudorandom order across trials, such that target images could not appear on the same side on more than 2 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 needed to be hard enough, such that children at 22 months did not perform at ceiling. To ensure this outcome, there were equal numbers of easy words (comprehension > 66%), moderately difficult words (comprehension = 33–66%), and difficult words (comprehension < 33%) based on normative data at 16 months (Dale & Fenson, 1996; Jorgensen, Dale, Bleses, & Fenson, 2010).

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; Jorgensen et al., 2010) and prompted by the experimenter to touch the target. If infants failed to touch the screen after repeated prompts, the experimenter touched the target image for them. If participants failed to touch during training, the 4 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 s 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 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 2 consecutive trials with two attempts by the experimenter to reengage without success. If the attempts to reengage 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 2 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 where 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 MCDI: WG, a parent report checklist of language comprehension and production, and at 22 months using the MCDI: Words and Sentences (WS), a parent report checklist of language production developed by Fenson and colleagues (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, the 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 multimedia 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 on 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 _____?”, “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 when 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 approximately 50% of the data.

Trials with short latencies (200–400 ms) likely reflect eye movements that were planned prior to hearing the target word (Bailey & Plunkett, 2002; Ballem & Plunkett, 2005; Fernald et al., 2008). For this reason, trials were included in subsequent analyses if participants looked at the screen for at least 400 ms. In addition, looking responses were coded during the first 2000 ms of each trial. As mentioned previously, looking responses that take place 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 were 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 that the 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. 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 with gaze shifts (Fernald, Pinto, Swingley, Weinbergy, & McRoberts, 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 with 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: \( \alpha = .931 \); Form B: \( \alpha = .940 \)).

At the age of 16 months, children executed target touches on 11.78 trials, executed 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 mentioned previously, 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 while being difficult enough so that 22-month-olds did not perform at ceiling. Side bias effects were assessed by analyzing the number of looks and touches to the right versus left side of the screen at 16 and 22 months. For the 16-month-olds, there was no significant effect of side for touch, \( t(44) = 1.80, p = .08 \), or number of first looks, \( t(44) = 1.50, p = .13 \), for images presented on the right relative to the left. Again, for children at 22 months, there was no significant effect of side for touch location, \( t(39) = 1.70, p = .10 \). However, children at 22 months tended to have more first looks to the image on the right side, \( t(37) = 6.89, p < .0001 \). However, we expected the effect of this tendency to be mitigated by task design because visual RT was concerned only with looks that initiated on distractors that appeared on the right and left sides of the screen with equal frequency. Indeed, a paired \( t \) test revealed that there was no significant difference in the number of first fixations to the target compared with the distractor (\( M_{\text{Target}} = 20.31, SD = 2.76; M_{\text{Distractor}} = 20.00, SD = 3.12 \)), \( t(37) = 0.38, p = .70 \). Finally, to ensure that our findings were uncontaminated by the tendency of 22-month-olds to look more to the right, we examined visual RTs as a function of haptic response type separately by side of initial fixation (left vs. right). The pattern of results was identical regardless of whether a first fixation was directed to the left or right; side of initial fixation did not influence our findings.
Aim 1: Relation of visual, haptic, and parent report measures in second year of life

Speed of processing and word recognition from 16 to 22 months

One goal of this research was to directly test interpretations reached in our earlier research (Hendrickson et al., 2015). To compare speed of word processing and word recognition in the same group of children at different ages, we performed correlations among visual RT, haptic response, and parent report at 16 and 22 months (see Fig. 2 for summary of results).

The average visual RT was calculated for each child at 16 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 and 22 months. Following our findings for visual RT, the correlation between the number of target touches executed at 16 and 22 months ($M = 12.8, SE = 0.98$ 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 (Friend & Keplinger, 2008; Friend & Zesiger, 2011; Legacy et al., 2016).

Finally, parent-reported vocabulary comprehension and production was obtained by using the MCDI: WG at 16 months and the MCDI: WS 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 ($M = 247, SE = 24.41$).

Intercorrelation among 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 children processed words, the more words they were reported to produce. In addition, 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 parents reported they produced. Finally, although the correlation between visual RT and the haptic measure was in the expected direction, it was not significant ($r = -.19, p = .25$). It is possible that each measure may be differentially sensitive to knowledge across a hypothetical continuum due to differences in task demands (Hendrickson et al., 2015; Munakata, 2001). See the Discussion for further consideration of this issue.

Aim 2: Concurrent analyses of visual and haptic responses at 22 months

We compared speed of processing across the haptic types (target, distractor, or no touch) using visual RT. Calculating visual RT 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. Due to the different number of observations per participant, we analyzed visual RT using restricted maximum likelihood in a mixed-effects regression model with a random effect of participant on the intercept fit with an unstructured covariance matrix. The use of regression models offers several advantages over traditional analysis of variance (ANOVA) models, including robustness to unbalanced designs (see Newman et al., 2012, and references therein). For the contrasts, we report the regression coefficients (and standard errors), $z$ values, $p$ values, and 95% confidence intervals (using a Tukey correction for multiple comparisons).

As mentioned previously, average visual RTs were calculated for distractor-initial trials in which a shift in gaze occurred between 400 and 2000 ms post-visual onset. A test of the full model against the null model revealed that touch type varied as a function of visual RT, $\chi^2(2) = 6.99, p = .03$. Specifically, visual RTs preceding target touches were significantly faster compared with no touches ($B = -189.71, SE = 72.5, z = 2.68, p = .02$) but not different from distractor touches ($B = -65.16, SE = 42.53, z = 1.53, p = .26$). However, there was no significant difference between the visual RTs for distractor touches and no touches ($B = -124.54, SE = 78.49, z = 1.59, p = .24$). Therefore, participants demonstrated a similar pattern...
of looking times across the three response types, as reported previously at 16 months, such that visual RTs were fastest before target touches and slowest during no touches, with an intermediate speed of processing during distractor touches (Hendrickson et al., 2015) (see Fig. 3).

Aim 3: Influence of partial knowledge on word recognition

To address Aim 3, we evaluated the role that partial knowledge plays in early word comprehension in the second 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 the child executed target touches that we refer to as known words (e.g., dog, shoe, running), a list of words for which the child executed distractor touches that we refer to as partially known words (e.g., cat, red, hand), and a list of words for which the child made no haptic response that we refer to as unknown words (no touch; e.g., truck, bubbles, jumping).

A logistic regression analysis was conducted to predict whether a word was known at 22 months using knowledge level at 16 months (known, partially known, or unknown) as a predictor. The model consisted of a binary outcome variable (known or unknown at 22 months), a random effect of participant, and a fixed effect of knowledge level at 16 months (known, partially known, or unknown). Therefore, knowledge level at 16 months was a categorical predictor variable; known words were entered as the first predictor and therefore acted as the reference level. A test of the full model against the null model was statistically significant, indicating that word knowledge at 16 months as a set reliably distinguished between whether a word was known and whether it was unknown at 22 months, $\chi^2(2) = 33.65, p < .0001$ (see Fig. 4 for uncorrected coefficient estimates for the individual levels of the predictor). Furthermore, we used the logistic regression to transform the estimate coefficient comparing the predictors (partially known and unknown) and reference level (known) with odds ratios [ORs; $\text{Exp}(\hat{B}_x)$]. Results show that the odds of knowing a word at 22 months were nearly identical regardless of whether the word was known or partially known at 16 months ($\text{OR} = 1.08$). If words were unknown at 16 months, they were only 33% as likely as known words, and 31% as likely as partially known words, to be known at 22 months ($\text{OR} = 0.33$ and $\text{OR} = .31$, respectively).

In total, these findings suggest that partial knowledge provides a basis for developing more robust word representations. Known and partially known words at 16 months were equally likely to be known by 22 months. Conversely, unknown words at 16 months were much less likely than known and partially known words at 16 months to be known by 22 months.

**Discussion**

The overarching goal of the current study was threefold: (a) to evaluate the relation among visual RT, haptic response, and parent report as measures of early word knowledge throughout the second year of life, (b) to assess whether there is a continuum of word knowledge over this same period, and (c) to examine the role of partial knowledge in future word recognition. This longitudinal study provides the first data on convergent and predictive associations among the three primary paradigms in use for measuring word processing and word comprehension across the second year. Furthermore, this research provides new evidence of a continuum of word knowledge and the role that knowledge states play in future word comprehension in early development.

**Relation of visual, haptic, and parent report measures in the second year of life**

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, and parent
report) in the same cohort of children throughout the second year of life. 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 and parent report) increased significantly over this period, consistent with earlier research, which examined this trend for each measure separately (Fernald et al., 2006; Houston-Price et al., 2007; Hurtado et al., 2008; Legacy et al., 2016; Marchman & Fernald, 2008).

We also conducted a series of comparisons among 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, the current results revealed a significant relation between visual RT and parent report, such that children who are faster at processing words also exhibit more parent-reported vocabulary knowledge at 22 months (Hendrickson et al., 2015). 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, as well as in the current study, that visual RT as a measure of processing speed may be less stable early in the second year of life because 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 may 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 with haptic responses. Due to the design of the task, 22-month-olds' understanding of the words tested was rather robust; that is, 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 measures was due to the fact that certain tasks tap weaker representations (visual RT), whereas 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 findings of visual and haptic behavioral dissociations and corresponding partial knowledge states across the second year of life (Hendrickson et al., 2015). 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 touches), followed by incorrect haptic responses (distractor touches), 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 behaviors 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; Munakata, 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 is too weak to surmount a predominant response to touch the first image fixated (the distractor). This replicates and extends findings from 16-month-olds that word knowledge is not all-or-none but instead exists on a continuum from absence of knowledge, to partial knowledge, to robust knowledge (Hendrickson et al., 2015). Specifically, an all-or-none approach would not consider the observed behavioral dissociations at 16 months as indicative of word knowledge. A distractor response would be considered an error. With an all-or-none approach, we would not necessarily expect (or look for) any learning advantage to derive from behavioral dissociations as contrasted with haptic nonresponses (no touches). However, the inclusion of looking data in the current research allowed us to take a more nuanced approach. Using this approach, we found that the odds of knowing a word at 22 months were nearly identical regardless of whether the word was known or partially known at 16 months. Words that were unknown (haptic nonresponse and slow visual RT) at 16 months were much less likely to be known by 22 months. These results also provide evidence that children's behavioral responses in vocabulary assessments are more appropriately viewed on a continuum from absence of knowledge, to partial knowledge, to full decontextualized knowledge. Incorrect and absent responses represent different abilities in lexical access that meaningfully measure knowledge across this hypothetical continuum.

Influence of partial knowledge on word recognition

The third aim investigated the role that partial knowledge plays in early word comprehension throughout the second year of life. As opposed to one-shot learning theories, 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 a word when they executed a distractor touch at 16 months, it could contribute to future learning by helping them to demonstrate knowledge of the same word at 22 months. That is, even when participants fail to demonstrate knowledge of a 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 not responding). However, if participants have very little or no knowledge of the word–object relation—as we argue is the case in no touch 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 known at 22 months if they were partially known, as opposed to unknown, 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 that word at a later point in development. This suggests that even when participants fail to correctly identify a word's referent and touch the distractor, they may nevertheless have encoded, and continue to represent, partial information about the word–object relation.

How then is partial information encoded, stored, and used in online processing? It has been shown in a recent computational model that the gradual encoding of partial knowledge states, which appears as gains in the efficiency of processing familiar words, is a result of changes in levels of activation (McMurray et al., 2012). There is evidence that when hearing a word–referent pairing, young children encode not only the correct referent but also nontarget items that co-occur with the word (Kucker, 2013).
McMurray, & Samuelson, 2015). For example, on hearing the label “dog”, a child may be fixating a dog and a cat or a dog and a bone; therefore, an association is established not only between the word and target but also between related nontarget items. With repeated exposures, the correct referent for dog will co-occur with the label more often, resulting in a stronger association between the label and the correct referent.

Of particular importance for the current results, the model showed that the majority of word learning and recognition of familiar words is not so much in strengthening the activation between the word and referent but rather in inhibiting incorrect activations—that is, identifying those words and referents that are not associated. Word recognition in this task (and in general) necessitates a change from a point at which multiple words are being considered (e.g., on hearing the word dog, both dog and cat are activated) to a point at which only one word (dog) is considered and all spurious activations (cat) must be inhibited. Thus, for known words, understanding is sufficiently robust to inhibit the previously activated incorrect response. However, fragile understanding of word meaning (as in the case of words that are partially known) results in untenable semantic competition; that is, within-category competitors are left activated and are still considered, resulting in a misidentification of the appropriate referent.

This interpretation is in line with research on word learning in preadolescents (McGregor et al., 2007). It was found that interference with word retrieval was strongest after the initial exposure to the new words. At this stage of learning when knowledge is fragile, children demonstrated increased rates of semantically related target substitution errors (e.g., recalling the definition for pharaoh instead of sphinx) (McGregor et al., 2007). As we argue is the case for the current study, fragile meanings had been mapped, but associations between words and meanings were not robust enough to compete with semantically related word forms.

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 by helping the participant to recognize what referents are not related to the word. However, the exact mechanisms underlying how young children leverage partial word knowledge for future word learning, and more specifically what information they use during online processing, are still being debated. From adult work, we know that words that are semantically related can facilitate the processing of one another (McRae & Boisvert, 1998). To illustrate, when a word cannot be remembered, adults often show retrieval of partial word meaning (e.g., has four legs and fur), which when followed by a related item (e.g., cat) can result in complete retrieval of the unrecalled word (Brown & Kulik, 1977; Durso & Shore, 1991; Hicks & Marsh, 2002; Koriat, 1997; Meyer & Bock, 1992). This suggests that once partial information about the structure of the concept to be learned has been accumulated, the acquisition of threshold knowledge can be acquired rapidly (Yurovsky et al., 2014). Importantly, these findings are at odds with approaches to word learning that suggest a form of single-shot hypothesis testing in which misidentifying a word’s meaning can have detrimental effects for future learning (Trueswell et al., 2013). Instead, these results are more consistent with research showing that when adults reencounter a word they misidentified previously, word–object identification improves dramatically (Yurovsky et al., 2014).

One potential limitation of this interpretation concerns the role that frequency of exposure plays in word learning. That is, input frequency may ultimately underlie knowledge level at both ages in the current research because frequency of exposure may be somewhat stable over time. Indeed, increasing exposure to any stimulus improves learning, and the current study is unlikely to be an exception. However, the relation between frequency and word learning is not straightforward. Frequency has been shown to differentially influence enhancements in lexical versus semantic memory and for fragile versus more established levels of understanding (McGregor et al., 2007). In the current task, understanding rather fine semantic details related to the word form is essential for success due to the within-category design of the stimulus pairs. Importantly, McGregor and colleagues (2007) found that frequency of exposure had a greater effect on later phases, relative to early phases, of lexical learning. This suggests that frequency is less influential for words that are largely unknown (i.e., the learner has no a priori knowledge of the word–referent association). Once some amount of semantic understanding has been established, frequency of exposure exerts greater effects. This is not to say that input frequency is not a contributing factor to the observed results; instead, frequency might not drive
knowledge level at both age groups in a similar fashion. Indeed, future research is needed to address the influence of level of knowledge independently of input frequency to determine the unique contribution that each makes in word learning.

Practical implications

Of interest are the applied implications of these results for the structure of early lexical–semantic knowledge, particularly with respect to whether both incorrect and absent volitional responses should be treated as reflecting lack of knowledge. Correct responses in picture pointing tasks have been shown to predict subsequent language outcomes and readiness for school (Friend, (in review), Smolak, Liu, Poulin-Dubois, & Zesiger, 2016). Therefore, children’s most robust word knowledge may be a stronger predictor of developmental outcomes than is partial knowledge. However, it may be the case that including partially known words along with known words can boost the predictive validity of extant behavioral measures. It may be that two children who look similar when measuring robust word knowledge will have very different developmental gains depending on the number of partially known words in their vocabulary. Evaluating the utility of including partial knowledge states to predict future vocabulary knowledge is a direction for future research.

Finally, visual RTs were slowest when children failed to make a haptic response. This suggests that, in these cases, children truly do not know the word or cannot disambiguate the target from the referent. However, this does not exclude the possibility that children failed to make a haptic response for reasons unknown (e.g., lack of cooperation, disengagement). We limited the influence of disengagement by including only those children who completed the training phase and by using criteria for ending the task when necessary so that we could be confident of the responses that contributed to the final dataset.

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. Furthermore, this work demonstrates that partial knowledge influences future word understanding and offers evidence through accumulative learning theories concerning the role that 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

We are grateful to the participant families and also to Kelly Kortright for assistance in data collection and coding. This work was supported by National Institutes of Health (NIH) Grant No. HD068458 to the senior author (Margaret Friend) and by NIH Training Grant No. 5T32DC007361 to the first author. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or National Institute of Child Health and Human Development.

References


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, 144–166.


