Published in final edited form as:
Cogn Dev. 2022 ; 64: . doi:10.1016/j.cogdev.2022.101249.

Convergence and divergence in prediction from vocabulary and speed of word processing

Erin Smolak[a,b,1,*], **Jonathan Lee Helm**[c], **Tamara Patrucco-Nanchen**[d], **Pascal Zesiger**[d], **Margaret Friend**[c]

[a]University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA

[b]San Diego State University, 5500 Campanile Dr, San Diego, CA 92182, USA

[c]San Diego State University, Center for Clinical and Cognitive Neuroscience, 5500 Campanile Dr, San Diego, CA 92182, USA

[d]University of Geneva, FPSE, rue du Général-Dufour, 24, 1211 Geneva, Switzerland

Abstract

Toddler vocabulary knowledge and speed of word processing are associated with downstream language and cognition. Here, we investigate whether these associations differ across measures. At age two, 101 participants (55 monolingual French-speaking and 46 monolingual English-speaking children) completed a two-alternative forced choice task, yielding measures of decontextualized vocabulary (number of correct responses) and haptic speed of word processing (latency of correct responses). At ages three, four, and five children completed a battery of language assessments and an executive function task. Growth curve models revealed that age-two vocabulary significantly predicted age-three performance (but not growth from age three to four or four to five) across all language assessments but speed of processing did not predict language outcomes in final models. Finally, speed of processing was correlated with executive function at age three whereas vocabulary was not. Results suggest that vocabulary is associated with a range of downstream language abilities whereas haptic speed of processing may be associated with executive control.

Keywords

Vocabulary; Word processing speed; Longitudinal; Child language development

---

[*] Correspondence to: 1705 College St, Columbia, SC 29208, USA. esmolak@mailbox.sc.edu.
[1] Present Address: University of South Carolina, 920 Sumter St, Columbia, SC 29208

Declaration of interest
None.

[1] For the RT measure without outliers excluded, the final model included a significant effect of speed, language group, and a significant group by speed interaction on growth from age three to age four. Because this is the only RT measure with this result, we view these effects as spurious.

1. Introduction

Vocabulary size (Bleses et al., 2016; Friend et al., 2012; Friend et al., 2018; Friend et al., 2019; Reilly et al., 2010) and speed of word processing (Fernald & Marchman, 2012; Marchman & Fernald, 2008) in toddlerhood are associated with language and cognitive development later in childhood. Despite many papers demonstrating this link, there has been little research into the mechanisms underlying this relation and the conditions under which it emerges. The purpose of this research is to begin to address this gap in the literature.

For example, there is no consensus regarding appropriate measures of vocabulary and word processing speed, which can vary in strength of association with later outcomes (Friend et al., 2018; Friend et al., 2019; Smolak et al., 2021) and it is unknown whether different measures tap distinct aspects of the developing lexical-semantic system. The goal of the present study is to elucidate the attentional processes that haptic speed of word processing exploits by examining the types of downstream abilities it predicts (Smith et al., 2010). We investigate longitudinal models evaluating the relation between vocabulary and speed of word processing at age two and a range of language outcomes from age three to five. In addition, we assess generalizability across two linguistically and geographically distinct samples. We begin with a review of findings on the relation between vocabulary and speed of processing and their prediction of developmental outcomes, followed by a consideration of executive processes inherent in different measures of speed.

2. Vocabulary, speed, and developmental prediction

Evidence for the predictive value of vocabulary knowledge comes from larger studies using parent report and smaller studies using directly assessed vocabulary. Parent-reported vocabulary between 16 and 30 months of age predicts language, reading, and math achievement between 3 and 12 years of age (Bleses et al., 2016; Duff et al., 2015; Lee, 2011; Morgan et al., 2015). Late-talker status (below the 10th percentile in expressive vocabulary) at age two predicts language outcomes at age four beyond other risk factors (Reilly et al., 2010). Directly assessed vocabulary at age two predicts general language ability at age three and kindergarten readiness at age four (Friend et al., 2018; Friend et al., 2019).

Similar patterns are observed for speed of word processing: more efficient processing at 25 months is associated with growth in vocabulary size from 12 to 25 months (Fernald et al., 2006). Similarly, children with faster processing at 19 months have larger vocabularies at time points across 19–30 months (Peter et al., 2019). Finally, children classified as late talkers exhibit slower processing than typically developing peers (Fernald & Marchman, 2012) and this disparity persists for children who remain delayed at 30 months of age.

Despite this body of evidence, there has been relatively little investigation into the conditions under which this relation emerges. Few studies have compared prediction from vocabulary relative to speed of word processing. Moreover, the strength of these relations varies with choice of speed of word processing measure. We turn now to a discussion of speed of word processing and vocabulary, their relation, and their measurement.

3. Speed of word processing and vocabulary

Whereas vocabulary size refers to the number of word-referent relations that children know, speed of processing refers to how quickly they access those relations. Much of the work investigating the relation between vocabulary and speed has used parent-reported vocabulary and speed operationalized as latency to visually fixate a target image upon hearing a word. Both vocabulary size and speed of processing develop rapidly over the second year of life (Fenson et al., 1994; Fernald et al., 2006; Samuelson & McMurray, 2017) and larger vocabularies are generally associated with more efficient visual fixation latency (Fernald & Marchman, 2012; Fernald et al., 2006; Hendrickson et al., 2015; Hendrickson et al., 2017). Nevertheless, this relation is inconsistent across the first three years of life. For example, Fernald and colleagues (2006) found speed to be concurrently associated with parent-reported vocabulary size at two of four age points from 12 to 25 months whereas latency at 25 months was associated with vocabulary at all time points as well as growth and acceleration from 12 to 25 months. Similarly, Peter and colleagues (2019) found that visual fixation latency at 19 months was significantly associated with parent-reported vocabularies from 19 to 30 months and acceleration from 19 to 25 months. However, this study also found that latency at 25 and 31 months was not correlated with vocabulary at any time point. Donnelly and Kidd (2020) found strong evidence for an effect of 18-month vocabulary on 21-month visual fixation latency (when controlling for latency at 18 months), but weak evidence for an effect of latency at 18 months on 21-month vocabulary (when controlling for 18-month vocabulary). There were no significant effects of 21-month variables on performance at 24 months. Finally, Koenig et al. (2020) found that visual fixation latency at 36 months was significantly associated with concurrent vocabulary, although the strength of association was weaker than that found in younger children.

Two studies, to our knowledge, have investigated the relative contributions of parent-reported vocabulary and visual fixation latency to early childhood outcomes. Marchman and Fernald (2008) examined prediction from parent-reported vocabulary size and latency at 25 months of age to linguistic and cognitive outcomes at age eight. They found that vocabulary and latency each accounted for unique variance in language and working memory whereas only vocabulary size accounted for unique variance in IQ. In contrast, Koenig et al. (2020) found that visual fixation latency at 36 months was significantly associated with vocabulary and school readiness at 60 months but not when controlling for concurrent vocabulary.

Another body of work has investigated these relations using directly assessed vocabulary and speed operationalized as latency to touch a target image upon hearing a word (haptic response latency). Building on the literature on parent-reported vocabulary and visual fixation latency, haptic response latency is concurrently associated with directly assessed vocabulary in both monolingual and bilingual toddlers (DeAnda et al., 2018; Legacy et al., 2016) and haptic latency at 16 months of age is prospectively associated with vocabulary size at 22 months (DeAnda et al., 2018). However, only a single study has assessed the relative contributions of directly assessed vocabulary and haptic response latency to childhood language outcomes into the preschool period. Smolak et al. (2021) explored associations between directly assessed vocabulary, visual fixation latency, and haptic latency at age two and vocabulary at ages three and four. Vocabulary and visual fixation latency predicted vocabulary at both ages, but haptic latency did not. Moreover, only vocabulary accounted for unique variance. That haptic latency did not predict later vocabulary lead the authors to conclude that haptic response latency may reflect other cognitive processes (e.g., cognitive control), in addition to speed of word processing.

4. Characteristics of word processing speed measures

Measures of word processing speed vary in the relative ease of the response, the type of online processing demands, the association with long-term knowledge, and the extent to which they tap executive processes. Because this variation can influence conclusions regarding the relation between speed of word processing and language/cognitive development, clarity regarding the properties of these measures is required.

With regard to visual fixation latency, infants integrate auditory and visual information to search for a visual match to an auditory stimulus within milliseconds (Golinkoff et al., 1987). Although visual fixation latency is a widely used measure of speed of word processing it, like haptic response latency, has its limitations. For example, the relative ease of this response makes it poor at discriminating stable words in the lexicon from those on the cusp of comprehension. Hendrickson et al. (2015) found rapid target fixations even when word comprehension was fragile. In addition, the relation between visual fixation latency and vocabulary is strongest at around two years of age and is a better predictor of vocabulary growth over time than of vocabulary assessed at any single time point.

In contrast, the speed with which haptic responses are executed depends, in part, on decision confidence and motor control—thus, haptic responses are more computationally expensive than visual responses. In adults, response times are slower when confidence in response accuracy is low (Kiani et al., 2014). In toddlers, prospective motor control is associated with performance on inhibition and working memory tasks (Gottwald et al., 2016) and processing speed is more strongly associated with executive control in children than in adults (Cepeda et al., 2013).

One final difference between these measures concerns the timing of the trials. In the classic visual fixation latency task (aka looking while listening; Fernald et al., 2008), images appear on the screen for 2000 ms before the sentence prompt is played. In studies on haptic response latency, in contrast, images appear *as* the target word is spoken in the sentence prompt. Findings from research employing a modified looking while listening paradigm in which the images and target word are presented concurrently replicate those in the traditional paradigm suggesting that they are not influenced by this difference in the timing of stimulus presentation (Hendrickson et al., 2015; Hendrickson et al., 2017).

Smolak et al. (2021) suggested that cognitive processes underlying haptic responses may limit prediction from haptic response latency. However, a limitation to that study is that vocabulary size was the sole outcome measure. The stronger prediction from early vocabulary relative to speed of word processing may at least partly be attributed to similarities in predictor and outcome measures. This leads to an unexplored alternate conclusion: that haptic speed of word processing preferentially predicts outcomes related to working memory more so than crystallized knowledge. The goal of the current study is to further explore haptic speed of word processing by examining its prediction, relative to vocabulary, to a broader range of outcomes, ages, and to growth over time.

5. Current study

The goal of the current effort is to examine prediction from haptic latency, relative to vocabulary, to a range of developmental outcomes in order to understand the attentional processes entailed in the haptic latency measure. We chose haptic latency as our measure of interest because, although this measure entails additional cognitive processes beyond word processing, discrepant findings of concurrent and short-term longitudinal, but not long-term longitudinal, relations with vocabulary size warrants further investigation. At age two, children completed a two-alternative forced choice task from which we estimated vocabulary (number of correct responses) and haptic speed of word processing (latency to touch the correct image). This research is part of a larger longitudinal project examining predictions from early vocabulary to later cognitive development. Thus, we are somewhat constrained in our outcome measures by that earlier work. These outcomes vary in their reliance on long-term linguistic knowledge and processing/ working memory. Specifically, the four outcomes of interest are: event-related vocabulary (i.e., event knowledge, Farrar et al., 1993), vocabulary (long-term semantic memory), sentence repetition (morphosyntactic knowledge and phonological working memory), and nonword repetition (phonological short-term/working memory) at ages three, four, and five. We acknowledge that these measures entail a number of abilities beyond long-term or working memory. However, the goal was to choose a range of outcomes beyond vocabulary to which speed of processing might be more strongly associated. The current study expands on prior work in five important ways. First, we examine relative prediction from vocabulary and haptic speed of processing to a broader range of linguistic outcomes. Second, we examine associations with outcome performance at age three as well as with growth from age three to four and from four to five using structural equation growth curve models. We estimate associations with growth because speed of processing in toddlers is associated with both vocabulary size and with growth in vocabulary over time (Fernald et al., 2006; Peter et al., 2019).

Third, since haptic speed of processing is dependent on decision processes, we attempt to reduce measurement noise by correcting for responses due to guessing. Specifically, we calculate an adjusted measure of haptic speed of word processing that takes into account that children may attempt to “guess” the target response in a two-alternative task. Low confidence but correct guesses could artificially inflate estimates of processing speed. Alternatively, fast guesses could artificially deflate these estimates. We conduct analyses with this corrected measure as a within-manuscript replication. A fourth, exploratory, goal is to better understand relations between haptic speed of processing and cognition by investigating the relation between vocabulary, speed, and executive function (EF). Finally, to assess the generalizability of our models, we evaluate these hypotheses in English-speaking monolingual children in the United States and French-speaking monolingual children in Switzerland.

We hypothesize that vocabulary will predict outcomes that rely on crystallized long-term knowledge (event knowledge and vocabulary). In contrast, we hypothesize that haptic response latency will predict outcomes that entail processing and working memory (nonword and sentence repetition). Our overarching goal is to clarify the predictive capacity of vocabulary and haptic speed of processing in the second year on later language outcomes. We have four specific hypotheses:

1. Vocabulary at age two will be related to later vocabulary and event knowledge since these outcomes both involve long-term crystallized knowledge. In contrast, we hypothesize that latency will not be related to these outcomes or will no longer be related when controlling for vocabulary.
2. In contrast, latency at age two will be related to later nonword repetition whereas vocabulary will not (above and beyond speed) since this outcome is related to working memory (Archibald, 2017; Fry & Hale, 1996).
3. Both vocabulary and latency at age two will be associated with later sentence repetition, a measure of long-term morphosyntactic knowledge and working memory, because this is an intermediate outcome measure that we expect will entail both working memory and morphosyntactic knowledge (Polišenská et al., 2015).
4. Latency at age two will be associated with later EF at ages three, four, and five, given the role of cognitive control in early controlled motoric responses. In contrast, we do not expect a relation between vocabulary and EF.

We further hypothesize that results will be largely consistent across English- and French-speaking monolingual children, as we have no reason to expect the effect of early vocabulary and speed of word processing on outcomes will differ across groups.

6. Method

6.1. Participants

Seventy-nine English-speaking monolingual children (41 girls) and sixty-six Swiss French-speaking monolingual children (33 girls) participated in a large, multi-site longitudinal study investigating children's language and literacy development. English-speaking participants were recruited from Women, Infant, and Children Centers, parenting groups, swap meets, and birth records in a large city in the United States whereas French-speaking participants were recruited through birth lists in a large city in Switzerland. A $25 gift card to a major retailer and a small toy were provided as incentives at each visit. Children were included if: carried to term (at least 37 weeks), had normal hearing and vision, and had no diagnoses of language or cognitive delay at study inception. Informed consent was obtained from parents prior to study participation. Participants visited the lab on six occasions, four of which are reported here: at two (M = 22 months;16 days, range = 21;6 – 25;12), three (M = 36;27, range = 32;15 – 41;24), four (M = 49;3, range = 47;9 – 55;12), and five years of age (M = 61;13; range = 59;15 – 68;12). Thirty-three English-speaking children were excluded due to attrition (N = 25), becoming bilingual (N = 1) or missing data (N = 7). Children with full data for at least one outcome time point were included to retain as large a sample as possible but to have fully overlapping datasets across analyses. The final sample consisted of 46 children (28 girls, 18 boys) at ages two and three, 46 children at age four, and 39 children at age five. We conducted attrition analyses to determine if the group that completed data collection at age two but was not included in the final study ($N = 22$) was different from the group that was included in the final study ($N = 46$). We compared these groups on LEAT English exposure, income, maternal education, vocabulary, and haptic processing speed and found no significant differences between groups (all ps > 0.19). Eleven French-speaking children were excluded due to attrition ($N = 4$) or missing data ($N = 7$). The final sample consisted of 55 children (30 girls, 25 boys) at ages two and three, 55 children at age four, and 51 children at age five. See Table S1 in the supplementary materials for demographic data on the final sample.

6.2. Measures

6.2.1. Language Exposure Assessment Tool (LEAT)—

The LEAT, a systematic parent interview (DeAnda et al., 2016), was used to estimate relative language exposure. The LEAT gathers information on who interacts regularly with the child, the languages they speak, native speaker status, and the number of hours talking to/overheard by the child. It calculates percent exposure as a function of relative hours of exposure to each language. Internal consistency is excellent and percent exposure predicts concurrent vocabulary size above and beyond maternal education, age, and parent estimates (DeAnda et al., 2016). All children were exposed to their primary language (English or French) at least 80% of waking hours from birth.

6.2.2. Computerized Comprehension Task (CCT)—

The CCT is an experimenter-controlled, two alternative forced-choice measure of decontextualized vocabulary comprehension administered on a touch screen (Friend & Keplinger, 2003; 2008; available at https://childes.talkbank.org/). During administration, children sit on their caregiver’s lap approximately 30 cm from the monitor. The caregiver wears blackout glasses and listens to music during administration to avoid the potential for interference. The experimenter sits to the side of the child and caregiver.

On each trial, two images are presented simultaneously at left and right center screen while the experimenter delivers a prompt in a child-friendly voice naming a target image (e.g., “Where is the dog? Touch dog”). Administrators were trained to advance each trial (the appearance of the images on the screen) as they began to speak the target word in the first sentence prompt, so onset of the target word occurred just prior to the onset of the visual stimuli. Data from the English-speaking sample revealed that the average interval between target word onset and onset of visual stimuli was 238 ms (Hendrickson et al., 2015). The side on which the target appears is pseudo-randomized across trials with the restrictions that targets do not appear on the same side on more than two consecutive trials and are presented with equal frequency on each side. Each trial has a maximum duration of 7 s. Touches to the screen automatically end the trial. Target touches (e.g., dog) yield congruous auditory feedback (e.g., the sound of a dog barking) whereas distractor touches yield no feedback. Administration followed Friend et al. (2012) and allowed repetitions of trials when: 1) the child initiated a response but did not touch the screen before the end of the trial, 2) the child was distracted, not looking at the screen, and missed the trial, or 3) in the case of an accidental screen touch. If the child became fussy or disengaged at any point, the experimenter attempted to re-engage the child. If three attempts to re-engage failed, the experimenter terminated the task.

The CCT consists of 4 training trials, 41 test trials, and 13 reliability trials. Words on the CCT were obtained from the MacArthur-Bates Communicative Development Inventories (MCDI; Fenson et al., 2007). Target and distractor items are matched on color, size, saliency, word class, and word difficulty. Difficulty is defined based on 16-month normative data from the MCDI (Frank et al., 2016). There are approximately equal numbers of easy (comprehension ≥ 66 %), moderate (comprehension between 33 % and 66 %) and difficult (comprehension ≤ 33 %) words, which are randomly distributed. There are a mix of nouns ($n = 23$), adjectives ($n = 7$), and action words ($n = 11$). There are two forms of the CCT such that all word-image pairs are presented as both target and distractor. Forms were counterbalanced across participants. Paired images are high-quality, colorful digital images that are prototypical exemplars of the word.

The French adaptation of the CCT is the same in design and administration. Items on the French adaptation were selected based on comprehension norms for the French adaptation of the MCDI, Les Inventaires Français du Développement Communicatif (IFDC; Frank et al., 2016; Kern, 2007). Translation equivalents (synonyms across languages) were maintained in the adaptation to the extent that they yielded equivalent distributions of word class and difficulty as the English version. Images were selected to be prototypical for children in Switzerland.

Both adaptations of the CCT have strong test-retest reliability and internal consistency, and moderate short-term stability over a 4-month period (Friend & Keplinger, 2008; Friend & Zesiger, 2011). The CCT correlates with parent-reported vocabulary and predicts the number of unique words and the mean length of utterance in subsequent child language samples (Friend et al., 2012).

6.3. Coding vocabulary

Vocabulary size was operationalized as the number of correct target image touches (with a maximum score of 41). Coding was completed offline, and inter-rater reliability (~20 % of participants) was excellent (>.93).

6.4. Coding speed of processing

Speed of processing was operationalized as latency to touch the image after the images appeared on the screen. For the French-speaking sample, the CCT program computed reaction time automatically. For the English-speaking sample, coding of reaction time was completed offline in Eudico Linguistics Annotator (ELAN, 2018, https://archive.mpi.nl/tla/elan, Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands; Lausberg & Sloetjes, 2009). A waveform of the experimenter's prompts was extracted and synchronized to the video of children's haptic responses. We coded target word onset and offset, image onset, and the frame in which a touch was executed (touch onset). Coding of latency to touch began at image onset. Latency was coded in the frame during which the child touched the screen (i.e., when forward momentum ceased and the child's finger was slightly bent, indicating contact). Inter-rater reliability was established for ~20 % of participants. Reliability was calculated as the number of codes within three frames for both coders divided by the total number of codes and exceeded .90 for all code types.

6.4.1. Raw reaction time (RT)—Following DeAnda et al. (2018), only trials during which the child touched the correct image were included in this calculation. Because trials with short latencies may reflect a decision to touch prior to accessing word meaning, trials with latencies less than 400 ms were excluded, corresponding to longest cutoffs typically used in eye-tracking studies (DeAnda et al., 2018; Fernald et al., 2008). Additionally, each child's distribution of RT latencies was examined for outliers. Removal of outliers is common in RT studies to reduce the effect of uncharacteristically quick or slow responses (e.g., Montgomery, 2008). For each child, trials with a latency + /− 3 SD from that child's mean latency were excluded. This resulted in the removal of 17 trials in the English-speaking sample (1.4 % of trials) and 5 trials in the French-speaking sample (0.3 %). We then took the mean RT as our measure of processing speed. Final model results were largely equivalent whether we used the original latency variable or the variable with outliers excluded; we report results from the variable with outliers excluded. Discrepancies between results are marked with footnotes.

6.4.2. Adjusted reaction time (RT)—If children guess in a two-alternative task they have a 50% chance of touching the correct image. Thus, mean reaction time includes response times for guesses and may not be an accurate estimate of processing speed. To control for the effect of guessing, we calculated a measure of adjusted RT as seen in (1) that corrects for guessing by considering reaction times on incorrect trials. This calculation is based on the fast guess model and is a method for estimating the latency of stimulus-controlled responses (Yellott, 1971; see discussions in Thomas, 1974 and Vandierendonck, 2021).

$$t1 = (k-1)\left(pc \times tc - \frac{pe \times te}{k-1}\right) \div (k - 1 - [k \times pe]) \tag{1}$$

Here, k is the number of alternatives, pc is the proportion of responses that were correct, tc is reaction time on correct trials, pe is the proportion of responses that were errors, and te is reaction time on error trials. In a two-alternative task, this simplifies to (2):

$$t1 = \frac{(pc \times tc - pe \times te)}{(1 - 2pe)} \tag{2}$$

Similar to raw RT, we removed trials with latencies less than 400 ms and latencies + /− 3 SD from individual mean error latency. Final model results from the adjusted RT measure were consistent whether outliers were included or excluded. One child performed at chance and was faster when choosing the target, resulting in a negative adjusted RT when applying the formula. In the interest of accurate estimations of RT, which should always be positive, this child's data were removed from analyses of adjusted RT. Another child had no incorrect trial touches with a latency greater than 400 ms. For this child, we used their raw reaction time for adjusted RT. Adjusted RT was significantly correlated with raw RT, $r(95) = 0.80$, $p < 0.001$.

6.5. Event knowledge

Event knowledge was coded from a free play language sample. Caregivers were instructed to play with their children as they would at home using a Fisher-Price farm play set including a rich assortment of farm structures (barn, silo, etc.), animals, people, and vehicles. The duration of the language sample was 15 min for the 3-year visit and 10 min for the 4- and 5-year visits. The session was recorded using a Zoom H2n Handy Recorder microphone. The audio files were transcribed and coded for grammatical morphemes (e.g., tense markers) using the Systematic Analysis of Language Transcripts software (SALT; Miller & Iglesias, 2012). All transcribers completed the online SALT training platform and trained to a reliability of > = 90 % with a sample transcript. Assistants were fluent in the language of the sample. Finally, ~20% of the transcripts were transcribed for reliability by a second assistant. Reliability was calculated as the number of agreements divided by the number of agreements plus the number of disagreements. For the English-speaking sample, word-level reliability was .91 at the 3-year visit, .94 at the 4-year visit, and .91 at the 5-year visit. For the French-speaking sample, reliability was .90 at the 3-year visit, .91 at the 4-year visit, and .96 at the 5-year visit.

Event knowledge was assessed by coding context-related words. First, we generated a list of farm-related words by searching the MCDI and its French adaptation (IFDC), online farm word lists in English and French (https://relatedwords.org/relatedto/farm; https://www.frenchlearner.com/vocabulary/farm/), and commonly used words in the English and French transcripts. Once this list was generated, the first author and a coder fluent in French confirmed the appropriateness of words included and added synonyms (e.g., tractor/truck, étable/écurie). The final list contained 303 words in English and 264 words in French, a combination of farm-related nouns (e.g., horse, barn, fence) and verbs (e.g., feed, plow, grow).

Two coders (the first author, who coded the English transcripts, and an assistant fluent in French and English, who coded the French transcripts) coded each utterance for words that were farm-related, coding only content words (i.e., nouns and verbs) but excluding function words (articles, prepositions, etc.). Approximately 20 % of the transcripts were coded for reliability. For the English sample, utterance- and word-level agreement was .93 and .93 for the 3-year visit, .95 and .96 for the 4-year visit, and .98 and .98 for the 5-year visit. Agreement for the French sample was established on transcripts that were translated to English and was .98 and .97 for the 3-year visit, .96 and .97 for the 4-year visit, and .98 and .98 for the 5-year visit.

Our primary variable of interest was the average number of unique farm-related words produced. Therefore, we divided the number of different farm words by the total number of utterances to develop a standard measure of event knowledge: the average number of different farm words produced per utterance. We expect farm-related vocabulary to increase over time reflecting long-term knowledge of the farm context or event (e.g., Farrar et al., 1993). Specifically, Farrar and colleagues found that children produced an increased number of different words over time during free play with a familiar event (the same toy set repeated across sessions) compared to an unfamiliar event (a novel toy set at each session). They concluded that gains in lexical diversity were due to increased experience with an event (and thus increased knowledge of the event). This measure differs conceptually from overall vocabulary because of the effect of the repeated context/event. Whereas Farrar et al. (1993) showed that increased experience with a specific event facilitated language use about that event, we expected that general facility with word meanings would support learning event-specific words. Thus, we expected early vocabulary to be associated with the subsequent acquisition of event knowledge. In the present case, we sampled only a single event context but would expect this to be true across contexts.

6.6. Peabody Picture Vocabulary Test (PPVT) and Échelle de Vocabulaire en Images Peabody (EVIP)

The PPVT is a measure of vocabulary comprehension for ages 30 months through adulthood (Dunn & Dunn, 1997) and the EVIP is its French adaptation (Dunn et al., 1993). During administration of the PPVT/EVIP an experimenter displays four pictures and asks the child to point to the one that corresponds to a target word. The test is adaptive such that difficulty increases with age and the test ends once the ceiling criterion is met. The final score is the number of items to reach ceiling minus the number of errors. The PPVT and EVIP were standardized on representative samples and have strong reliability for all age ranges tested, as well as strong internal consistency (Dunn & Dunn, 1997; Dunn et al., 1993). Like the CCT, the PPVT yields a direct estimate of receptive vocabulary. The measure of interest is number of words correct, expected to capture long-term semantic knowledge.

6.7. Sentence repetition (SR)

This task is a modification of the Italian task by Devescovi and Caselli (2007). In the current study, sentences in each category (e.g., simple sentences with copula, simple sentences with one argument) were created in English and French in order to equate, as much as possible, the number of words and morphemes in each sentence across languages. At the year-3 visit, the test included 27 sentences of varying complexity and length. An additional 11 sentences were added at the year-4 and year-5 visits to increase complexity and reduce the chances of a ceiling effect. During test, the experimenter told the child to repeat after her, modeled the sentence, and revealed an image depicting sentence-level meaning. The sentence was modeled up to three times however only the child’s first attempt at repetition was scored. The primary variable of interest is proportion of morphemes repeated correctly. Approximately 20 % of the SR data were coded for reliability. For the English-speaking sample, sentence-level agreement was .92 at the year-3 visit, .91 at the year-4 visit, and .92 at the year-5 visit. For the French-speaking sample, sentence-level agreement was .92 at the year-3 visit, .99 at the year-4 visit, and .99 at the year-5 visit. Research suggests that SR performance requires diverse skills in language processing, comprehension, and production (Klem et al., 2015). Therefore, this task entails long-term morphosyntactic knowledge and well as processing efficiency and working memory (Polišenská et al., 2015).

6.8. Nonword repetition (NW)

Pseudowords were created using the MCDI adaptation in English/French and syllables were shuffled to create phonotactically plausible words for each language separately. Nonword repetition measures phonological working/short term memory (i.e., memory for legal sound sequences; Gathercole et al., 1994). Repetition of shorter nonwords involves short-term memory (short-term storage of information) whereas repetition of longer nonwords involves working memory (short term storage and manipulation of information; Unsworth & Engle, 2007). Nonwords increased in length and complexity (i.e., inclusion of consonant clusters) with age. At 3 years, there were four one-syllable words, four two-syllable words, four three-syllable words, and four four-syllable words. At 4 years, four five-syllable words were added. At 5 years, one-syllable words were removed (because 5-year-olds would likely be at ceiling on the one syllable items), and four additional five-syllable words were added. However, scores on the one syllable items were credited to children at age 5 and were therefore included in the scoring.

During test, the experimenter asked the child to repeat after her. She presented a toy figure and said, “This one’s name is ___. Can you say ___?” The child’s first attempt at repetition was scored. The measure of interest was proportion of consonants repeated correctly (PCC). For the English-speaking sample, phoneme-level inter-rater agreement was .88 at the 3-year visit, .91 at the year-4 visit, and .94 at the year-5 visit. For the French-speaking sample, phoneme-level inter-rater agreement was .97 at the year-3 visit, .99 at the year-4 visit, and .96 at the year-5 visit. Although there is evidence that long-term semantic knowledge is associated with nonword repetition performance (Jones et al., 2007), recent clinical evidence indicates that the task is primarily a measure of phonological working/short term memory (Archibald, 2017).

6.9. Executive function (EF)

The Dimensional Change Card Sort Task (DCCS) is a measure of the cognitive flexibility (or task switching) component of EF. We implemented DCCS in the NIH Toolbox, a measure which was designed for individuals aged 3 and up and validated for children 3–15 years of age (Gershon et al., 2013; Zelazo & Bauer, 2013; Zelazo et al., 2014). Children sort a series of bivalent cards (e.g., a yellow ball, a blue truck) first by one dimension (e.g., color), and then by the other (e.g., shape). The task consists of four blocks (practice, pre-switch, post-switch, and mixed). On each practice trial, a test card (a green rabbit or a white boat) appeared in the top center of the screen along with two target cards (a white rabbit and a green boat) at the bottom of the screen. The child was instructed to sort the test card first by color (4 trials) and then by shape (4 trials) and to touch the target that matched the color/shape of the test card. Feedback was provided on practice trials. Children who scored at least 3 trials correct out of four proceeded and were given three chances to reach criterion. Test trials were similar to practice trials except there were five trials for each dimension, the cards were blue or yellow balls or trucks, and no feedback was provided. Children who scored at least 3 trials correct out of 5 proceeded to complete 30 mixed trials, 22 of which were shape and 8 of which were color.

On the first trial of a given sorting type, the examiner said “Let’s play the color/shape game! Where does this one go in the color/shape game?” and on subsequent trials of the same type, “Where does this one go in the color/shape game?” The measure of interest was performance on the post-switch trials at ages three, four, and five (hereafter referred to as post-switch). We assessed EF in the English-speaking sample only due to technical difficulties at the French-speaking study site.

7. Results

7.1. Analytic approach

The data and model code for this project are available at: https://osf.io/rabez/. We began by identifying plausible models for change in each of the four dependent variables (i.e., event knowledge, vocabulary, sentence repetition, and nonword repetition). Plausible models included intercept only, linear, and two-part linear spline models (see Ram & Grimm, 2007; 2015). As an aside, a quadratic could also be used to analyze the data, but this model was discarded because (1) it will show identical fit as the two-part linear spline, and (2) the parameters of the two-part linear spline are more interpretable than the quadratic model in this empirical example (again, see Ram & Grimm, 2007; 2015).

Indicators of model fit rejected the intercept only model for all dependent variables (all p-values for χ^2 were less than .001; all CFI and TLI values were less than .50); suggesting that all dependent variables changed over time. Except for PPVT, indicators of model fit also rejected the linear model for all dependent variables (all p-values for χ^2 were less than .001; all CFI and TLI values were less than .51). For PPVT, the linear model was rejectable according to the χ^2 fit statistic ($p = 0.02$), but CFI and TLI showed reasonable fit (both equaled .95). However, the model showed estimation challenges (i.e., the model showed a non-positive definite latent variable covariance matrix; see Wothke, 1993), and was therefore discarded.

Therefore, for all dependent variables, the two-part linear spline was chosen as the best model for change. Here, the model quantifies change as three latent variables: (1) an intercept reflecting the level of the dependent variable at age 3, (2) a shift in the dependent variable from age 3–4, and (3) a second shift from age 4–5. Although the two-part linear spline model is saturated in this empirical example (and therefore cannot be rejected: for all dependent variables, $p = 1.0$ for χ^2 fit statistic; CFI = TLI = 1.0), all three dependent variables showed significant between individual variation for all three components (all ps < 0.01 for variances of latent intercepts/slopes), suggesting that each may be used for examining between-person differences (i.e., predictors of these differences may be examined), and also showing significantly better fit than either the intercept-only or linear models. See Fig. 1 for a diagram of the SEM model we employed.

Finally, for both slopes of each outcome measure, we calculated two metrics to determine power to detect individual differences in growth: Growth Rate Reliability (GRR; Rast & Hofer, 2014) and d (Fan & Fan, 2005). Growth Rate Reliability is calculated using the intraclass correlation coefficient (ICC) but takes study design into account by dividing the residual variance by the sum of squared deviations of time points about the mean time point.

This metric can be considered a measure of the ability to identify individual differences in slopes. Similarly, d is the standardized mean difference between the last measurement and the first measurement and identifies the magnitude of growth. For event knowledge, GRR = 0.91 and − 0 .12, d=0 .45 and 0.02. For PPVT vocabulary, GRR= 0.96 and 0.93, d= 0.70 and 0.54. For SR, GRR= 0.83 and 0.79, d=0 .34 and 0.32. For NW repetition, GRR= 0.85 and 0.14, d=0 .35 and 0.08. For our sample size, Monte Carlo simulations reveal that d=> 0.30 achieves power greater than .8 to detect linear growth (Fan & Fan, 2005). Similarly, our sample size is sufficient to achieve power = 0.8 to detect linear slope variances at GRR = 0.6 to 0.8 (Rast & Hofer, 2014). These metrics indicate sufficient power to detect statistical differences in slope in our combined sample with the exception of the second slope of NW repetition and event knowledge. Thus, we do not report results from these slopes.

The two-part linear spline was further examined with three steps in a model-building approach to the data. First, we added predictor and control variables to the model one by one. Predictor variables were vocabulary size and speed of processing at age two. Control variables were age in months at age two, maternal education, and biological sex as each has been shown to predict language measures in larger studies (e.g., Reilly et al., 2010). This step was taken in order to determine independent effects on the three latent variables as a first look. Second, we constructed two models with the predictors (vocabulary and speed of processing), language group (English or French), and a language group by predictor interaction. These models evaluate the effect of language group on latent variables and whether effects of predictors differed by language group. Third, we constructed a final model including both predictors and all significant variables from prior models. This step was taken in order to clarify the effects of each of our predictors of interest (vocabulary and speed) while controlling for any significant effects of group or control variables.

We report all significant effects in the models constructed in steps 1 and 2 but interpret the nature of effects from the final models only, which we consider to be the central findings. All analyses were run in R 4.0.0 (R Core Team, 2020) using RStudio 1.3.959 (RStudio Team, 2020). We used the lavaan package to construct and fit all SEM models (Rosseel, 2012) using the default maximum likelihood estimator and full information maximum likelihood for missing data. However, given the small sample size, we also examined final models using maximum likelihood estimation of standard errors, which provides robust standard errors and a scaled test statistic. The substantive interpretation of the results was consistent across these analyses.

All continuous independent variables were mean centered. We rescaled the speed of processing measure from milliseconds to seconds to maintain comparable numeric ranges across predictors and facilitate accurate computation of standard errors. Language group and sex were dummy-coded numeric variables (language group 0 = English; sex 0 = female).

7.2. Descriptive data

See Table 1 for descriptive data for the English and French samples. We performed t-tests across groups to determine if group differences in variables of interest existed. A family-wise correction for multiple tests was applied using a sequential Bonferroni procedure (Benjamini & Hochberg, 1995) for 20 comparisons. The false discovery rate was set at

.05. The predictor and outcome variables were approximately normally distributed with a few exceptions (see Table S2 in the supplementary materials for skewness and kurtosis information). For adjusted RT, there were two outliers (English n = 1, French n = 1) with scores ~4 SD above the group mean. These scores were excluded. For nonword repetition, there were five outliers (English n = 1, French n = 4). These children scored > 4 SD below the group mean and were excluded as outliers.

In sentence repetition (both samples) and nonword repetition (French sample only), we observed negative skewness indicating ceiling effects at some ages. Transformation would render growth over time uninterpretable. Fortunately, SEM methodology is generally robust to non-normality. Therefore, we present results from raw variables.

7.3. Prediction to outcomes

See Table 2 for zero-order correlation results between predictors of interest (vocabulary and speed of processing using the raw RT measure), control variables (age in months at initial visit, maternal education, and biological sex), and the outcomes (performance at age three, growth from age three to four, and growth from age four to five). Note these are results from the rescaled and mean-centered variables. We applied the sequential Bonferroni correction (Benjamini & Hochberg, 1995) for 28 correlations for the inter-relations between the predictors and 24 correlations per family for correlations between predictor and outcome variables. The false discovery rate was set at .05.

7.4. Two-part linear spline models

7.4.1. Event knowledge—Models investigating the independent effects of the predictor and control variables revealed that age two vocabulary and maternal education significantly predicted the event knowledge intercept, $\beta = 0.347$, $p < 0.001$ and $\beta = 0.235$, $p = 0.02$, respectively. Maternal education predicted growth from age three to four, $\beta = 0.200$, $p = 0.04$. There were no other significant effects and no effects of group or group by predictor interactions. Thus, the final model included vocabulary, speed of processing, and maternal education. Results from this model can be found in Table 3. Vocabulary significantly predicted the intercept: larger vocabularies at age two were associated with larger event knowledge scores at age three. The effect of maternal education on the intercept was marginally significant such that higher levels of maternal education were associated with larger event knowledge scores at age three. Additionally, maternal education predicted growth from age three to four such that higher levels of maternal education were associated with more rapid growth. There were no other significant effects[1]. The final model was consistent across the raw and adjusted RT measures. That vocabulary, but not speed of processing, predicted event knowledge was consistent with our first hypothesis.

7.4.2. PPVT vocabulary—In PPVT models with a binary variable (sex, language group), PPVT variance was at least 1000 times larger than others. Therefore, in these models, we divided PPVT vocabulary by 10 to put it on a similar scale with other variables.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cogdev.2022.101249.

Models investigating independent effects revealed that vocabulary, $\beta = 0.340$, $p < 0.001$, speed of processing, $\beta = -0.226$, $p = 0.02$, and age, $\beta = 0.357$, $p < 0.001$, significantly predicted the PPVT vocabulary intercept. There were no other significant effects. In the vocabulary, language group, and language group by vocabulary interaction model, vocabulary and language group significantly predicted the PPVT vocabulary intercept, $\beta = 0.520$, $p < 0.001$ and $\beta = -0.655$, $p < 0.001$, respectively. Further, there was a significant language group by vocabulary interaction for growth from age three to four, $\beta = 0.264$, $p = 0.037$. In the speed, language group, and language group by speed interaction model, language group significantly predicted the PPVT vocabulary intercept, $\beta = -0.588$, $p < 0.001$. The final model included vocabulary, speed, language group, language group by vocabulary, and age. Results from this model can be found in Table 4. Age two vocabulary and language group significantly predicted the PPVT vocabulary intercept: larger vocabularies at age two were associated with larger vocabularies at age three. Additionally, children in the English group had larger vocabulary scores at age three than children in the French group. The final model was consistent across the raw and adjusted RT measures. That vocabulary, but not speed of processing, predicted PPVT vocabulary was consistent with our first hypothesis.

7.4.3. Sentence repetition—Models investigating independent effects revealed that vocabulary, $\beta = 0.517$, $p < 0.001$, and sex, $\beta = -0.219$, $p = 0.030$, significantly predicted the SR intercept. Additionally, vocabulary significantly predicted growth from age three to four, $\beta = -0.523$, $p < 0.001$, and marginally predicted growth from age four to five, $\beta = -0.203$, $p < 0.05$. Finally, sex significantly predicted growth from age four to five, $\beta = -0.208$, $p = 0.040$. There were no other significant effects and no significant effects of language group or language group by predictor interactions. The final model included vocabulary, speed, and sex. Results from this model can be found in Table 5. Age two vocabulary significantly predicted the SR intercept and growth from age three to four: larger vocabularies at age two were associated with higher SR scores at age three but less growth from age three to four. There were no other significant effects. The final model was consistent across the raw and adjusted RT measures. In contrast to our hypothesis that both vocabulary and speed of processing would predict later sentence repetition, vocabulary and sex were the only significant predictors.

7.4.4. Nonword repetition—Models of independent effects revealed that vocabulary significantly predicted the NWR intercept, $\beta = 0.356$, $p = 0.001$. In the vocabulary, language group, and language group by vocabulary interaction model, vocabulary, $\beta = 0.295$, $p = 0.022$, and language group, $\beta = 0.383$, $p < 0.001$, both significantly predicted the intercept. In the speed, language group, and language group by speed interaction model, speed of processing, $\beta = -0.333$, $p = 0.036$, and language group, $\beta = 0.473$, $p < 0.001$, significantly predicted the intercept. There were no other significant effects. The final model included vocabulary, speed, and language group. Results from this model can be found in Table 6. Age two vocabulary and language group significantly predicted the intercept: larger vocabularies at age two were associated with higher NWR scores at age three. The effect of language group indicated that children in the French group had higher NWR scores than their English peers. There were no other significant effects. The final model was consistent across the raw and adjusted RT measures. In contrast to our hypothesis that speed of processing would predict later nonword repetition, vocabulary and language group were the only significant predictors.

7.4.5. Executive function (EF)—

Because there were no EF data for the French sample, we calculated zero-order correlations to examine the relative association of processing speed and vocabulary with EF. At age three, 26 children did not complete the post-switch trials on the DCCS due to fuss-out or failing to meet criterion on training or pre-switch trials, leaving a final sample of 20. At age four, 6 children did not complete the DCCS, leaving a final sample of 40. At age five, 8 children did not complete the DCCS, leaving a final sample of 38. Because of these small samples, we present these results as purely exploratory.

Vocabulary was uncorrelated with DCCS post-switch performance at age three, $r(18) = -0.13$, $p = 0.58$, four, $r(38) = 0.17$, $p = 0.28$, or five, $r(36) = 0.13$, $p = 0.43$. In contrast, speed of processing evinced a significant correlation at age three, $r(18) = -0.46$, $p = 0.04$, but not at age four, $r(38) = -0.06$, $p = 0.70$, or five, $r(36) = -0.18$, $p = 0.28$. Faster reaction times at age two were associated with more correct post-switch trials on the DCCS at age three. This finding did not hold for the adjusted RT measure. Adjusted RT was uncorrelated with DCCS performance at any age, $r(18) = -0.17$, $p = 0.47$, $r(35) = 0.06$, $p = 0.74$, and $r(33) = -0.04$, $p = 0.83$, respectively. Thus, our hypothesis that speed of processing, but not vocabulary, would be associated with later EF was partially supported: raw RT was correlated with EF at age three. Further, removing the effect of guessing in the adjusted measure also removed the correlation between RT and EF, suggesting that correcting RT for guessing removed variance associated with EF.

In summary, vocabulary at age two was associated with age-three performance on all linguistic outcome variables. However, it was not associated with growth over time with the exception of growth in sentence repetition performance from age three to four. In contrast, haptic speed of processing was not associated with any linguistic outcome. Instead, haptic speed of processing was associated with later executive function at age three whereas vocabulary was not. See Table 7 for a summary of the main results.

8. Discussion

The primary goal of this research was to clarify the processes that common toddler language assessments entail by investigating longitudinal prediction from haptic speed of processing relative to vocabulary size at age two to an array of outcomes at ages three, four, and five. In doing so we examined the speed of processing measure in more detail by adjusting for the potential confound of guessing and exploring the relation of haptic processing speed to later executive function. The impetus for this work was threefold. First, there is a body of evidence indicating that both vocabulary and speed of processing significantly predict downstream language and cognition and are therefore foundational skills for later learning (Bleses et al., 2016; Fernald & Marchman, 2012; Friend et al., 2012; Friend et al., 2018; Friend et al., 2019; Marchman & Fernald, 2008; Morgan et al., 2015). However, there has been less work investigating the conditions under which this relation emerges.

Second, although vocabulary size and speed of processing are correlated their shared variance is only about 20% in prior research (DeAnda et al., 2018) and 5% in the present study. Therefore, they appear to be dissociable components of the lexical-semantic system (see also McMurray et al., 2012). Finally, relations between haptic processing speed and vocabulary do not hold longitudinally for intervals greater than six months in the preschool period (Smolak et al., 2021). The current study adds to extant work on the stability of early language abilities by evaluating prediction to a range of linguistic outcomes beyond vocabulary size at age three and growth from age three to five in two linguistically and geographically distinct groups of children.

8.1. Prediction to language outcomes

The primary language outcomes in the present study were event knowledge, vocabulary size, sentence repetition, and nonword repetition. These outcome measures were chosen partially because we expected that event knowledge and vocabulary tap long-term crystallized knowledge, sentence repetition taps morphosyntactic long-term knowledge as well as language processing and working memory, and nonword repetition primarily taps phonological working memory. One way in which early vocabulary and speed of processing might be dissociable is in their association with long-term word knowledge versus active language processing. Consistent with this idea, we expected vocabulary to predict event knowledge and later vocabulary, both vocabulary and speed of processing to predict sentence repetition, and speed of processing to predict nonword repetition.

However, only the first hypothesis was supported. Vocabulary at age two predicted the age-three intercept of event knowledge and PPVT vocabulary. In contrast, although speed of processing was associated with the age three intercept of PPVT vocabulary in a preliminary model, it had no significant effects on these outcomes in final models. Contrary to our second hypothesis, although vocabulary at age two predicted the age-three intercept of sentence repetition, speed of processing did not. The same was true of nonword repetition. As was the case for PPVT vocabulary, speed of processing did predict the age-three intercept of nonword repetition in an intermediate model but was not significant in the final model.

These results are consistent with prior work showing that vocabulary predicts a range of downstream language abilities. Indeed, the quantity and quality of word representations are associated with more complex language (Friend et al., 2019; Lee, 2011; Reilly et al., 2010), kindergarten readiness (Friend et al., 2018) reading accuracy and comprehension (Duff et al., 2015; Perfetti & Hart, 2002), and academic and mathematics achievement (Bleses et al., 2016; Morgan et al., 2015). Moreover, early decontextualized vocabulary comprehension (as measured with the CCT) has been shown to be a more robust predictor of downstream outcomes than other measures of vocabulary or processing efficiency (Friend et al., 2018; Friend et al., 2019; Smolak et al., 2021). Although we hypothesized that speed of processing would predict later nonword repetition, the effect of vocabulary on nonword repetition is perhaps not surprising given that long-term semantic and phonotactic knowledge also contribute to performance on this task (Jones et al., 2007).

The lack of prediction from speed of processing is inconsistent with our hypotheses and prior work showing that haptic response times predict vocabulary over a six-month interval

(DeAnda et al., 2018). However, it is consistent with the finding that haptic processing speed does not predict vocabulary knowledge over longer intervals in the preschool period (Smolak et al., 2021). In the current study it was a poorer predictor than early vocabulary even for outcomes (like phonological working memory) that we expected to be particularly dependent on processing speed. In contrast to prior work finding a longitudinal relation between visual speed of processing and later vocabulary/language (Koenig et al., 2020; Marchman & Fernald, 2008; Smolak et al., 2021), the current study used a haptic response latency measure. As we have argued, whereas both visual and haptic latency measures of speed of word processing have both strengths and weaknesses, of interest here is the potential contribution of executive control to haptic latency.

Previous work suggests that accurate haptic responses reflect relatively stable word knowledge (Hendrickson et al., 2015; Hendrickson et al., 2017). Additionally, prospective motor control in toddlerhood is related to inhibition and working memory components of executive function (Gottwald et al., 2016). Accordingly, haptic responses may reflect deliberative processes and/or recruitment of executive control in addition to processing speed. Thus, the haptic measure may be partially tapping deliberative processes and decision-making or motoric and executive control. The present findings are consistent with this interpretation.

Finally, we expected haptic processing speed at age two to be associated with later executive function (EF). Exploratory correlation analyses supported an association with the cognitive flexibility or task switching component of EF at age three but not at age four or five whereas vocabulary size at age two was not associated with cognitive flexibility at either age. These results offer preliminary support for the idea that haptic processing speed involves executive control, although further research is required to confirm this result.

Finally, the raw haptic processing speed measure correlated with a measure that corrected for guessing. Final models were consistent across the raw and corrected measures with only one exception: adjusted RT was uncorrelated with executive function. This suggests that removing the effect of guessing from the adjusted measure also removed the variation in haptic speed associated with executive control by estimating speed on stimulus-driven responses only. In other words, children with lower EF might be more prone to guessing (i.e., less likely to resist touching the screen when they do not know the answer) and reducing the effect of guessing in the adjusted measure also reduced the effect of cognitive control in this measure.

8.1.1. Prediction of growth—We were interested in assessing prediction to performance at a single time point as well as growth on the basis of prior research finding that, in toddlers, processing speed is associated both with vocabulary at a single time point and growth and acceleration of vocabulary size (Fernald et al., 2006; Peter et al., 2019). We sought to extend this research later in developmental time to additional language measures. This approach, using one metric at age two (vocabulary or speed) to assess outcomes on different but related metrics, attempts to assess the stability (in the relative rank order of individuals in a group) of an underlying language construct (Bornstein et al., 2017). Although latent change models do not directly test stability, we can infer stability from the effect of the predictors on growth. For example, a positive association between predictors and growth suggests stability in the preservation of relative rank order between children over time. In contrast, a negative effect would suggest instability where children with higher scores on the predictor are exhibiting more protracted growth relative to those with lower scores. Whereas stability can be modeled using the same metric over time, assessing stability using different methods yields a relatively conservative estimate of stability since different measures introduce variance to the growth estimate. In other words, our methods assess stability in language over time using a conservative approach that may underestimate true stability in language development, resulting in a null effect. In our final models, vocabulary predicted growth only for sentence repetition. Haptic speed of processing did not predict growth in any model. We explore potential explanations for these findings below.

Interestingly, the prediction from vocabulary to growth in sentence repetition was in the opposite direction of expectations such that larger vocabulary at age two was associated with less growth. This may be due to instability or to ceiling effects such that children with larger vocabularies have better sentence repetition scores at age three and therefore less room to grow. Indeed, sentence repetition performance at age three was also negatively associated difference scores between age three and age four, $r(80) = -0.88$, $p < 0.001$, and between age four and age five, $r(64) = -0.31$, $p = 0.01$. Although sentence repetition performance reached ceiling by age four or five, this cannot fully explain the direction of prediction. One possibility is that children with large vocabularies at age two reach ceiling earlier than their lower vocabulary peers leading to the observed positive relation between vocabulary at age two and sentence repetition at age three and the negative relation between vocabulary and growth from three to four and four to five. Consistent with this interpretation, children who were at or above the median on CCT vocabulary at age two scored 90% correct on the sentence repetition task at age three, whereas children who were below the median on CCT vocabulary at age two scored 77%. Thus, children with relatively large vocabularies at age two were close to ceiling on sentence repetition by age three and had less room to grow from age three to five.

How can we account for the inability to predict growth on other measures? Bornstein et al. (2017) suggest that such estimates of stability present a lower bound. Prediction from vocabulary and speed to related, but distinct, measures necessarily limits estimates of stability. Further, as the time between assessments increases, the likelihood of stability decreases. On the other hand, prediction across similar measures leads to inflated estimates of stability and fails to capture the dynamic relation between early vocabulary and language outcomes. One solution to this conundrum would be to utilize composite factors in prediction to reduce method variance. For example, one could construct composite variables from visual (fast, non-deliberative) and haptic (slow, deliberative) latencies to obtain more stable estimates of speed of word processing over time. A similar approach could be taken to conceptual outcome measures. This approach would yield a clearer picture of the longitudinal relations between speed of word processing and later crystallized word knowledge in the preschool period (Bornstein et al., 2017).

8.2. Language group differences

A secondary goal was to evaluate the generalization of our models across English- and French-speaking monolingual children. On the whole, and consistent with our hypothesis, results were similar across groups. However, the descriptive data revealed that French-speaking children exhibited slower response times at age two and smaller vocabularies at ages three, four, and five. The latter result may reflect differences across English and French versions of the PPVT vocabulary measure. However, this finding is consistent with evidence that French-speaking children have smaller vocabularies than English-speaking children between two and four years of age (Thordardottir, 2005). On the other hand, French-speaking children performed better than the English sample on nonword repetition possibly reflecting differences in English and French phonology. For example, the English version of the nonword repetition task contained more complex consonants (e.g., / dʒ/) and consonant clusters that do not exist in French. There were no other significant effects of language group or language group by predictor interactions, suggesting that the effects of vocabulary and speed of processing on outcomes largely generalized across groups.

8.3. Limitations and directions for future research

There are a few limitations of the current study. First, the tasks became more challenging with age to account for developmental changes in ability and to avoid ceiling effects. This design may have influenced growth estimates such that actual growth was underestimated. Second, due to between-subjects variability and limited growth in event knowledge and nonword repetition, we did not have the power to predict individual differences from age four to five for these outcomes. Designs that incorporate data from children approaching school age may need to be more highly powered or utilize composite measures to detect individual difference in growth curve trajectories.

Several lines of evidence (i.e., prediction from vocabulary to outcomes at age three, the association between raw and adjusted haptic measures of speed, and sufficient power across all outcomes from age three to four) suggest that haptic speed of processing in the toddler period is a relatively weak predictor of developmental outcomes. However, another limitation of the current study is that all outcome measures were dependent on long term language knowledge to at least some degree (e.g., our phonological working memory measure taps phonotactic knowledge and can be influenced by semantic knowledge). A different test of working memory less dependent on long term semantic knowledge may have yielded a different result. Future research would benefit from investigating associations between haptic speed of processing and other outcomes (e.g., visuo-spatial working memory). Moreover, given that prior research has found an association between visual speed of processing and later language outcomes, one explanation for the lack of relations in the present study is that the relation was obscured due to noise introduced by motor planning and movement. It is possible that if we were to control for motor skill, or developed more stable, composite estimates of speed of word processing, this relation would re-emerge. This is a direction for future research.

Finally, lack of data on EF for the French-speaking sample does not permit us to generalize beyond the English-speaking sample. Moreover, high rates of fussiness and failure to reach criterion led to a small sample on which to calculate correlation coefficients. We view our findings on the relation between haptic processing and EF as exploratory. Future research should investigate this question in larger samples to elucidate deliberative processes in word processing.

9. Conclusion

Vocabulary at age two significantly predicted a range of downstream language measures, although it rarely predicted growth. In contrast, haptic speed of processing did not predict any downstream abilities in our final models. Haptic processing speed was however, correlated with cognitive flexibility at age three (where vocabulary was not). These results provide evidence on the relative predictive utility of vocabulary and haptic response times for preschool outcomes. Further, they offer a tentative clarification of the contribution of executive processes to offline measures of speed of processing, informing our understanding of the early lexical-semantic system as well as the cognitive processes that toddler language measures entail.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Stephanie DeAnda, Brooke Tompkins, Tara Rutter, Kasey Palmer, Kelly Kortright, Hanna Moon, and Emma Healy for assistance in data collection and coding, Adriana Weisleder for her contribution to an earlier version of this manuscript, and to all our participant families without whom this research would not be possible.

Funding

This research was supported by the National Institute of Child Health and Human Development [NICHD HD468058] to the senior authors and the National Institute on Deafness and Other Communication Disorders [NIDCD DC007361] to the first author and does not necessarily represent the views of the National Institutes of Health.

Data Availability

I have shared the link to the data on OSF in the manuscript.

References

Archibald LM (2017). Working memory and language learning: A review. Child Language Teaching and Therapy, 33(1), 5–17. 10.1177/0265659016654206

Benjamini Y, & Hochberg Y.(1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300. 10.1111/j.2517-6161.1995.tb02031.x

Bleses D, Makransky G, Dale PS, Højen A, & Ari BA (2016). Early productive vocabulary predicts academic achievement 10 years later. Applied Psycholinguistics, 37(6), 1461–1476. 10.1017/S0142716416000060

Bornstein MH, Putnick DL, & Esposito G.(2017). Continuity and stability in development. Child Development Perspectives, 11(2), 113–119. 10.1111/cdep.12221 [PubMed: 29503666]

Cepeda NJ, Blackwell KA, & Munakata Y.(2013). Speed isn't everything: Complex processing speed measures mask individual differences and developmental changes in executive control. Developmental Science, 16(2), 269–286. 10.1111/desc.12024 [PubMed: 23432836]

DeAnda S, Bosch L, Poulin-Dubois D, Zesiger P, & Friend M.(2016). The language exposure assessment tool: Quantifying language exposure in infants and children. Journal of Speech, Language, and Hearing Research, 59(6), 1346–1356. 10.1044/2016_JSLHR-L-15-0234

DeAnda S, Hendrickson K, Zesiger P, Poulin-Dubois D, & Friend M.(2018). Speed of processing in the second year: A study of monolingual and bilingual vocabulary development. Bilingualism: Language and Cognition, 21(2), 314–327. 10.1017/S1366728917000220 [PubMed: 29731683]

Devescovi A, & Caselli MC (2007). Sentence repetition as a measure of early grammatical development in Italian. International Journal of Language & Communication Disorders, 42(2), 187–208. 10.1080/13682820601030686 [PubMed: 17365093]

Donnelly S, & Kidd E.(2020). Individual differences in lexical processing efficiency and vocabulary in toddlers: A longitudinal investigation. Journal of Experimental Child Psychology, 192, Article 104781. 10.1016/j.jecp.2019.104781

Duff FJ, Reen G, Plunkett K, & Nation K.(2015). Do infant vocabulary skills predict school-age language and literacy outcomes? Journal of Child Psychology and Psychiatry, 56(8), 848–856. 10.1111/jcpp.12378 [PubMed: 25557322]

Dunn LM, & Dunn LM (1997). PPVT-III: Peabody Picture Vocabulary Test. Circle Pines. MN: American Guidance Service.

Dunn LM, Thériault-Whalen CM, & Dunn LM (1993). Échelle de vocabulaire en images Peabody: série de planches. Toronto: Psycan,.

ELAN (Version 5.2) [Computer software]. (2018). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan.

Fan X, & Fan X.(2005). Power of latent growth modeling for detecting linear growth: Number of measurements and comparison with other analytic approaches. The Journal of Experimental Education, 73(2), 121–139. 10.3200/JEXE.73.2.121-139

Farrar MJ, Friend MJ, & Forbes JN (1993). Event knowledge and early language acquisition. Journal of Child Language, 20(3), 591–606. 10.1017/S0305000900008497 [PubMed: 8300777]

Fenson L, Dale PS, Reznick SJ, Bates E, Thal DJ, Pethick SJ, & Stiles J.(1994). Variability in early communicative development. i-185 Monographs of the Society for Research in Child Development, 59(5). 10.2307/1166093.

Fenson L, Marchman VA, Thal D, Dale P, Reznick JS, & Bates E.(2007). MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual (2nd ed.). Baltimore, MD: Brookes Publishing Co.

Fernald A, & Marchman VA (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. Child Development, 83(1), 203–222. 10.1111/j.1467-8624.2011.01692.x [PubMed: 22172209]

Fernald A, Perfors A, & Marchman VA (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. Developmental Psychology, 42(1), 98–116. 10.1037/0012-1649.42.1.98 [PubMed: 16420121]

Fernald A, Zangl R, Portillo AL, & Marchman VA (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young chidlren. In Sekerina I, Fernandez EM, & Clahsen H.(Eds.),´ Developmental psycholinguistics: On-line methods in children's language processing (pp. 97–135). Amsterdam: J. Benjamins.

Frank MC, Braginsky M, Yurovsky D, & Marchman VA (2016). Wordbank: An open repository for developmental vocabulary data. Journal of Child Language, 44 (3), 677–694. 10.1017/S0305000916000209 [PubMed: 27189114]

Friend M, & Keplinger M.(2003). An infant-based assessment of early lexicon acquisition. Behavior Research Methods, Instruments, & Computers, 35(2), 302–309. 10.3758/BF03202556

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(1), 77–98. 10.1017/S0305000907008264 [PubMed: 18300430]

Friend M, Schmitt SA, & Simpson AM (2012). Evaluating the predictive validity of the computerized comprehension task: Comprehension predicts production. Developmental Psychology, 48(1), 136–148. 10.1037/a0025511 [PubMed: 21928878]

Friend M, Smolak E, Liu Y, Poulin-Dubois D, & Zesiger P.(2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. Developmental Psychology, 54(7), 1317–1333. 10.1037/dev0000514 [PubMed: 29620386]

Friend M, Smolak E, Patrucco-Nanchen T, Poulin-Dubois D, & Zesiger P.(2019). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year. Developmental Psychology, 55(1), 9–22. 10.1037/dev0000617 [PubMed: 30359059]

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. 10.4074/S0013754511003041

Fry AF, & Hale S.(1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. Psychological Science, 7(4), 237–241. 10.1111/j.1467-9280.1996.tb00366.x

Gathercole SE, Willis CS, Baddeley AD, & Emslie H.(1994). The children's test of nonword repetition: A test of phonological working memory. Memory, 2(2), 103–127. 10.1080/09658219408258940 [PubMed: 7584287]

Gershon RC, Wagster MV, Hendrie HC, Fox NA, Cook KF, & Nowinski CJ (2013). NIH toolbox for assessment of neurological and behavioral function. Neurology, 80(11), S2–S6. 10.1212/WNL.0b013e3182872e5f [PubMed: 23479538]

Golinkoff RM, Hirsh-Pasek K, Cauley KM, & Gordon L.(1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. Journal of Child Language, 14(1), 23–45. 10.1017/S030500090001271X [PubMed: 3558524]

Gottwald JM, Achermann S, Marciszko C, Lindskog M, & Gredeback G.(2016). An embodied account of early executive-function development: Prospective ¨ motor control in infancy is related to inhibition and working memory. Psychological Science, 27(12), 1600–1610. 10.1177/0956797616667447 [PubMed: 27765900]

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. 10.1111/desc.12250 [PubMed: 25444711]

Hendrickson K, Poulin-Dubois D, Zesiger P, & Friend M.(2017). Assessing a continuum of lexical-semantic knowledge in the second year of life: A multimodal approach. Journal of Experimental Child Psychology, 158, 95–111. 10.1016/j.jecp.2017.01.003 [PubMed: 28242363]

Jones G, Gobet F, & Pine JM (2007). Linking working memory and long-term memory: A computational model of the learning of new words. Developmental Science, 10(6), 853–873. 10.1111/j.1467-7687.2007.00638.x [PubMed: 17973801]

Kern S.(2007). Lexicon development in French-speaking infants. First Language, 27(3), 227–250. 10.1177/0142723706075789

Kiani R, Corthell L, & Shadlen MN (2014). Choice certainty is informed by both evidence and decision time. Neuron, 84(6), 1329–1342. 10.1016/j.neuron.2014.12.015 [PubMed: 25521381]

Klem M, Melby-Lervåg M, Hagtvet B, Lyster SAH, Gustafsson JE, & Hulme C.(2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. Developmental Science, 18(1), 146–154. 10.1111/desc.12202 [PubMed: 24986395]

Koenig A, Arunachalam S, & Saudino KJ (2020). Lexical processing of nouns and verbs at 36 months of age predicts concurrent and later vocabulary and school readiness. Journal of Cognition and Development, 21(5), 670–689. 10.1080/15248372.2020.1802277 [PubMed: 33727892]

Lausberg H, & Sloetjes H.(2009). Coding gestural behavior with the NEUROGES-ELAN system. Behavior Research Methods, 41(3), 841–849. 10.3758/BRM.41.3.841

Lee J.(2011). Size matters: Early vocabulary as a predictor of language and literacy competence. Applied Psycholinguistics, 32(1), 69–92. 10.1017/S0142716410000299

Legacy J, Zesiger P, Friend M, & Poulin-Dubois D.(2016). Vocabulary size, translation equivalents, and efficiency in word recognition in very young bilinguals. Journal of Child Language, 43(4), 760–783. 10.1017/S0305000915000252 [PubMed: 26044885]

Marchman VA, & 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. 10.1111/j.1467-7687.2008.00671.x [PubMed: 18466367]

McMurray B, Horst JS, & Samuelson LK (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. Psychological Review, 119(4), 831–877. 10.1037/a0029872 [PubMed: 23088341]

Miller J, & Iglesias A.(2012). Systematic Analysis of Language Transcripts (SALT), Research Version 2012 [Computer Software]. Middleton, WI: SALT Software, LLC.

Montgomery JW (2008). Role of auditory attention in the real-time processing of simple grammar by children with specific language impairment: A preliminary investigation. International Journal of Language & Communication Disorders, 43(5), 499–527. 10.1080/13682820701736638 [PubMed: 22612629]

Morgan PL, Farkas G, Hillemeier MM, Hammer CS, & Maczuga S.(2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. Child Development, 86(5), 1351–1370. 10.1111/cdev.12398 [PubMed: 26283023]

Perfetti CA, & Hart L.(2002). The lexical quality hypothesis. In Vehoeven L, Elbro C, & Reitsma P.(Eds.), Precursors of functional literacy (pp. 189–213). Amsterdam: John Benjamins.

Peter MS, Durrant S, Jessop A, Bidgood A, Pine JM, & Rowland CF (2019). Does speed of processing or vocabulary size predict later language growth in toddlers. Cognitive Psychology, 115, Article 101238. 10.1016/j.cogpsych.2019.101238

Polišenská K, Chiat S, & Roy P.(2015). Sentence repetition: What does the task measure? International Journal of Language & Communication Disorders, 50(1), 106–118. 10.1111/1460-6984.12126 [PubMed: 25208477]

Reilly S, Wake M, Ukoumunne OC, Bavin E, Prior M, Cini E, & Bretherton L.(2010). Predicting language outcomes at 4 years of age: Findings from early language in victoria study. Pediatrics, 126(6), E1530–E1537. 10.1542/peds.2010-0254 [PubMed: 21059719]

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL 〈https://www.R-project.org/〉 .

Ram N, & Grimm K.(2007). Using simple and complex growth models to articulate developmental change: Matching theory to method. International Journal of Behavioral Development, 31(4), 303–316. 10.1177/0165025407077751

Ram N, & Grimm KJ (2015). Growth curve modeling and longitudinal factor analysis. In Overton WF, Molenaar PCM, & Lerner RM (Eds.), Handbook of child psychology and developmental science: Theory and method (pp. 758–788). John Wiley & Sons, Inc. 10.1002/9781118963418.childpsy120.

Rast P, & Hofer SM (2014). Longitudinal design considerations to optimize power to detect variances and covariances among rates of change: Simulation results based on actual longitudinal studies. Psychological Methods, 19(1), 133. 10.1037/a0034524 [PubMed: 24219544]

Rosseel Y.(2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). Journal of Statistical Software, 48(2), 1–36 (URL) L) 〈http://www.jstatsoft.org/v48/i02/〉 .

Team RStudio (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. URL 〈http://www.rstudio.com/〉 .

Samuelson LK, & McMurray B.(2017). What does it take to learn a word? WIREs Cognitive Science, 8(1–2), Article e1421. 10.1002/wcs.1421

Smith LB, Colunga E, & Yoshida H.(2010). Knowledge as process: Contextually cued attention and early word learning. Cognitive Science, 34(7), 1287–1314. 10.1111/j.1551-6709.2010.01130.x [PubMed: 21116438]

Smolak E, Hendrickson K, Zesiger P, Poulin-Dubois D, & Friend M.(2021). Visual and haptic responses as measures of word comprehension and speed of processing in toddlers: Relative predictive utility. Journal of Experimental Child Psychology, 203, 105032. 10.1016/j.jecp.2020.105032

Thomas EA (1974). The selectivity of preparation. Psychological Review, 81(5), 442. 10.1037/h0036945

Thordardottir ET (2005). Early lexical and syntactic development in Quebec French and English: Implications for cross-linguistic and bilingual assessment. International Journal of Language & Communication Disorders, 40(3), 243–278. 10.1080/13682820410001729655 [PubMed: 16195189]

Unsworth N, & Engle RW (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. Psychological Review, 114(1), 104–132. 10.1037/0033-295X.114.1.104 [PubMed: 17227183]

Vandierendonck A.(2021). On the utility of integrated speed-accuracy measures when speed-accuracy trade-off is present. Journal of cognition, 4(1), 22. 10.5334/joc.154 [PubMed: 33778368]

Wothke W.(1993). Nonpositive definite matrices in structural modeling. In Bollen KA, & Long JS (Eds.), Testing structural equation models (pp. 256–293). SAGE Publications, Inc.

Yellott JI Jr, (1971). Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. Journal of Mathematical Psychology, 8(2), 159–199. 10.1016/0022-2496(71)90011-3

Zelazo PD, Anderson JE, Richler J, Wallner-Allen K, Beaumont JL, Conway KP, & Weintraub S.(2014). NIH Toolbox Cognition Battery (CB): validation of executive function measures in adults. Journal of the International Neuropsychological Society: JINS,20(6), 620–629. 10.1017/S1355617714000472 [PubMed: 24960301]

Zelazo PD, & Bauer PJ (Eds.). (2013). National Institutes of Health Toolbox cognition battery (NIH Toolbox CB): Validation for children between 3 and 15 years. Hoboken, NJ: Wiley.

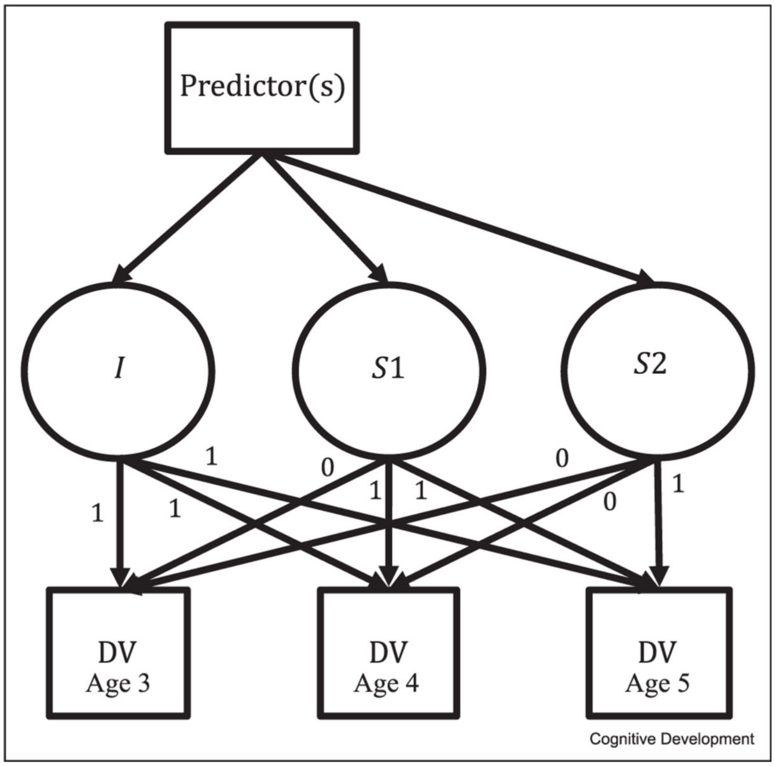

Fig. 1.

Depiction of parameter estimates from SEM models. Please note that intercepts, variances, and covariances are not depicted (see Tables 3–6 for more parameter estimates). Labels ‘I′, ‘S1′, and ‘S2′ refer to ‘intercept’, ‘slope 1′, and ‘slope 2′, respectively. The label ‘Predictor(s)’ refers to any predictors retained in the final models. The label ‘DV’ refers to measured dependent variables (Event Knowledge, PPVT Vocabulary, Sentence Repetition, and Nonword Repetition). Factor loadings are in raw metric.

Table 1

Descriptive information of all measures of interest by sample. T-tests conducted across samples.

	English Sample				French Sample					
Variables	**N**	**Mean**	**SD**	**Range**	**N**	**Mean**	**SD**	**Range**	***t*-test**	**Sig.**
CCT Vocabulary	46	27.74	6.89	10.00–39.00	55	29.8	5.20	19–40	−1.67	0.10
Speed of Processing (RT)	46	2600.14	589.83	1264.04– 3812.00	55	3094.57	688.68	1980.06–4739.50	−3.83	< 0.001 *
True RT	43	2522.78	839.13	478.42–4519.00	54	3134.22	1006.00	1530.50–5616.25	−3.20	0.002 *
Event Knowledge										
Age Three	44	0.21	0.10	0.02–0.42	54	0.22	0.08	0.06–0.41	−0.43	0.67
Age Four	46	0.29	0.11	0.10–0.52	54	0.33	0.13	0.10–0.77	−1.70	0.09
Age Five	39	0.30	0.11	0.14–0.60	50	0.32	0.10	0.14–0.60	−1.24	0.22
PPVT Vocabulary										
Age Three	46	49.76	18.30	11–93	55	28.71	11.28	10–61	6.80	< 0.001 *
Age Four	46	71.80	16.09	31–106	54	55.65	14.69	27–92	5.25	< 0.001 *
Age Five	39	89.44	13.01	54–124	51	73.76	13.14	39–101	5.63	< 0.001 *
Sentence Repetition										
Age Three	42	0.84	0.14	0.36–0.99	42	0.86	0.15	0.42–0.99	−0.39	0.70
Age Four	45	0.93	0.08	0.54–1.00	51	0.92	0.10	0.55–1.00	0.53	0.60
Age Five	39	0.96	0.05	0.73–1.00	46	0.97	0.03	0.87–1.00	−1.12	0.27
Nonword Repetition										
Age Three	40	0.81	0.08	0.61–0.93	50	0.89	0.10	0.51–1.00	−4.14	< 0.001 *
Age Four	45	0.89	0.08	0.62–0.97	54	0.94	0.08	0.68–1.00	−3.09	0.003 *
Age Five	38	0.91	0.05	0.81–1.00	50	0.96	0.04	0.85–1.00	−4.33	< 0.001 *
DCCS Post-Switch										
Age Three	20	3.45	1.99	0.00–5.00						
Age Four	40	3.58	1.68	0.00–5.00						
Age Five	38	3.55	1.25	0.00–5.00						

Note.

[*] significant after family-wise FDR correction

Table 2

Correlation matrix of predictive and longitudinal relations in the full sample.

	1.	2.	3.	4.	5.	6.
1. Vocabulary	−0.22					
2. Speed of Processing	−0.22					
3. Language Group	0.17	0.36 *				
4. Age (at 2 yrs)	0.02	−0.26 *	−0.55 *			
5. Maternal Education	0.17	−0.11	−0.04	0.04		
6. Biological Sex	−0.16	0.25 *	0.06	−0.06	−0.03	
Event Knowledge Intercept	0.35 *	−0.04	0.04	−0.03	0.24	0.04
Event Knowledge Slope 3–4	−0.04	−0.02	0.12	−0.03	0.22	0.04
Event Knowledge Slope 4–5	−0.08	−0.01	−0.04	0.02	−0.27	−0.02
PPVT Vocab Intercept	0.34 *	−0.23	−0.58 *	0.36 *	0.14	−0.04
PPVT Vocab Slope 3–4	−0.02	−0.0001	0.17	−0.14	−0.02	0.07
PPVT Vocab Slope 4–5	−0.13	−0.06	0.02	−0.07	−0.06	−0.09
SR Intercept	0.54 *	−0.07	0.04	0.02	−0.007	−0.18
SR Slope 3–4	−0.52 *	0.07	−0.15	0.0009	−0.08	0.09
SR Slope 4–5	−0.18	0.02	0.11	−0.11	−0.09	0.16
NWR Intercept	0.31 *	0.03	0.40 *	−0.18	0.08	−0.06
NWR Slope 3–4	−0.14	0.11	−0.11	0.09	0.03	−0.04
NWR Slope 4–5	−0.003	−0.22	−0.10	0.02	−0.20	0.04

Notes.

[*] significant after family-wise FDR correction. Language Group = dummy coded as 0 = English, French = 1. Group X Vocab = variable created by multiplying Language Group by Vocabulary. Group X Speed = variable created by multiplying Language Group by Speed of Processing. Slope = variables created by subtracting scores at time point 1 from scores at time point 2.

Table 3

Standardized regression coefficient estimates and p-values for the final models predicting event knowledge over time. Variance and covariance estimates are included in each model for latent variables (i.e., Intercept, Slope 1, and Slope 2).

	Intercept		**Slope 1**	
	Std. Est.	***p***	***Std. Est.***	***p***
Regressions				
Maternal Education	**0.187**	**0.047**	**0.209**	**0.036**
Vocabulary	**0.328**	**0.001**	−0.080	0.427
Speed of Processing	0.052	0.586	−0.032	0.8750
Variance	**0.844**	**< 0.001**	**0.954**	**< 0.001**
Covariance				
Intercept			**−0.417**	**< 0.001**
Slope 1				

Table 4

Standardized regression coefficient estimates and p-values for the final models predicting PPVT vocabulary over time. Variance and covariance estimates are included in each model for latent variables (i.e., Intercept, Slope 1, and Slope 2).

	Intercept		**Slope 1**		**Slope 2**	
	Std. Est.	***p***	***Std. Est.***	***p***	***Std. Est.***	***p***
Regressions						
Age	−0.024	0.764	−0.043	0.712	−0.074	0.549
Vocabulary	**0.547**	**< 0.001**	−0.211	0.103	−0.114	0.387
Speed of Processing	0.129	0.089	−0.075	0.494	−0.196	0.093
Language Group	**−.722**	**< 0.001**	0.170	0.2174	0.068	0.602
Language Group X Vocabulary	−0.087	0.327	0.244	0.057	−0.110	0.408
Variance	**0.459**	**< 0.001**	**0.932**	**< 0.001**	**0.976**	**< 0.001**
Covariance						
Intercept			**−0.470**	**< 0.001**	−0.088	0.397
Slope 1					**−0.518**	**< 0.001**

Table 5

Standardized regression coefficient estimates and p-values for the final models predicting sentence repetition over time. Variance and covariance estimates are included in each model for latent variables (i.e., Intercept, Slope 1, and Slope 2).

	Intercept		**Slope 1**		**Slope 2**	
	Std. Est.	***p***	***Std. Est.***	***p***	***Std. Est.***	***p***
Regressions						
Sex	−0.152	0.098	0.043	0.665	0.196	0.059
Vocabulary	**0.511**	**< 0.001**	**−0.523**	**< 0.001**	−0.190	0.063
Speed of Processing	0.072	0.453	−0.025	0.813	−0.063	0.541
Variance	**0.708**	**< 0.001**	**0.723**	**< 0.001**	**0.921**	**< 0.001**
Covariance						
Intercept			**−0.821**	**< 0.001**	**−0.505**	**< 0.001**
Slope 1					−0.031	0.807

Table 6

Standardized regression coefficient estimates and p-values for the final models predicting nonword repetition over time. Variance and covariance estimates are included in each model for latent variables (i.e., Intercept, Slope 1, and Slope 2).

	Intercept		**Slope 1**	
	Std. Est.	***p***	***Std. Est.***	***p***
Regressions				
Vocabulary	**0.307**	**0.003**	−0.095	0.393
Speed of Processing	−0.015	0.880	0.124	0.272
Language Group	**.390**	**< 0.001**	−0.130	0.243
Variance	**0.715**	**< 0.001**	**0.961**	**< 0.001**
Covariance				
Intercept			**−0.594**	**< 0.001**
Slope 1				

Table 7

Summary of the main results in final models.

	Predictor Variable		
	Vocabulary	**Haptic Speed of Processing**	**Language Group**
Outcome			
Event Knowledge Intercept	**sig.**	n.s.	
Event Knowledge Slope 1	n.s.	n.s.	
Event Knowledge Slope 2	n.s.	n.s.	
Vocabulary Intercept	**sig.**	n.s.	**sig.**
Vocabulary Slope 1	n.s.	n.s.	n.s.
Vocabulary Slope 2	n.s.	n.s.	n.s.
Sentence Repetition Intercept	**sig.**	n.s.	
Sentence Repetition Slope 1	**sig.**	n.s.	
Sentence Repetition Slope 2	n.s.	n.s.	
Nonword Repetition Intercept	**sig.**	n.s.	**sig.**
Nonword Repetition Slope 1	n.s.	n.s.	n.s.
Nonword Repetition Slope 2	n.s.	n.s.	n.s.
EF Age 3	n.s.	**sig.**	
EF Age 4	n.s.	n.s.	
EF Age 5	n.s.	n.s.	

Notes. sig. = significant. n.s. = not significant. Cells in the Language Group column are blank for Event Knowledge and Nonword Repetition because this variable was not included in the final SEM model. Cells in the Language Group column are blank for EF because these analyses were conducted in the English-speaking monolingual group only.