Journal of Child Language (2024), **51,** 721–740
doi:10.1017/S0305000924000060

ARTICLE

Internal state language factor structure and development in toddlerhood: Insights from WordBank

Tyler C. MCFAYDEN[1] and Madeleine BRUCE[2]

[1]University of North Carolina at Chapel Hill, Carolina Institute for Developmental Disabilities, Chapel Hill, NC, USA
[2]University of Utah, Department of Psychology, Salt Lake City, UT, USA
Corresponding author: Tyler C. McFayden; Email: tyler_mcfayden@med.unc.edu

(Received 12 April 2023; revised 05 February 2024; accepted 14 February 2024)

Abstract

Internal state language (ISL) research contains knowledge gaps, including dimensionality and predictors of growth, addressed here in a two-aim study. Parent-reported expressive language from N = 6,373 monolingual, English-speaking toddlers (M_{age} = 23.5mos, 46% male, 57% white) was collected using cross-sectional and longitudinal data in WordBank. Exploratory and confirmatory factor analyses suggested a best-fitting one-factor model of ISL. The single-factor model of ISL was then submitted to hierarchical linear modeling to evaluate predictors of ISL development. Age 2 ISL production was predicted by child sex, wherein females outperform males, and maternal education, wherein higher education contributes to higher ISL. Only maternal education emerged as a significant predictor of ISL growth. These results provide support to theory suggesting a unitary construct of ISL, as opposed to considering ISL as categorical, and further illustrate linear growth through the second postnatal year that varies as a function of child sex and maternal education.

Keywords: internal state language; factor structure; language development; toddlerhood

Introduction

As children age, they begin to label their own and others' emotional, cognitive, and physical states, otherwise known as internal state language (ISL; Bretherton & Beeghly, 1982). Children's communication about the mind signals an ability to understand and reason about others' experiences, making one's theory of mind observable and explicit. As such, ISL has emerged as a marker of cognitive and social development in early toddlerhood, with relations to expressive and receptive language (Bretherton & Beeghly, 1982), executive function (Bellagamba et al., 2014), theory of mind (Carlson et al., 2004; Chiarella et al., 2013; Olineck & Poulin-Dubois, 2007), social understanding (e.g., Ensor & Hughes, 2008), and even reading comprehension at the high school and college levels (Booth & Hall, 1994). The developmental phenomenon of ISL has been evidenced in western (Kristen et al., 2014) and eastern (Tardif & Wellman, 2000) cultures, suggesting a universally shared developmental experience.

Despite its developmental relevance, there exists a core knowledge gap in the ISL literature: empirical and theoretical consensus on what comprises ISL. Measured via parent-report (e.g., Kristen et al., 2012), naturalistic observation (e.g., Pascual et al., 2008), or elicitation-specific tasks (e.g., LaBounty et al., 2008; Ruffman et al., 2003), ISL has been argued to have anywhere from two (Roger et al., 2012) to eleven (Lemche et al., 2007) categories, with no statistical evidence to date of an established factor structure. This knowledge gap has implications for how we understand ISL development in early toddlerhood, including how it develops, how it is measured, and its relations with other demographic factors. The current study seeks to address this knowledge gap by systematically evaluating ISL categories and their development in toddlerhood via two aims. Aim 1 uses a data-driven approach to evaluate the factor structure of ISL in monolingual, American English speaking toddlers from ages 16 to 30 months using a large repository of parent-report data from WordBank (Frank et al., 2016). Aim 2 uses a hierarchical linear modeling approach to evaluate predictors of the factor structure of ISL obtained from Aim 1. These interrelated aims serve important theoretical and empirical foundations for understanding ISL development in typical development as precursors to other developmental phenomena including theory of mind and atypical social development.

Aim 1: Internal state language structure

Despite layman's conceptualizations, "vocabulary" is not a unitary structure comprising all words you know or use (Bowles et al., 2005). Several investigations into adult vocabulary have yielded two-factor structures of vocabulary, and these underlying latent factors are thought to capture vocabulary "breadth and depth", "basic and advanced" vocabulary (Bowles et al., 2005), or groupings by origin (e.g., Latin or Swedish; Gustafsson & Holmberg, 1992). Although no empirical work to date has evaluated the factor structure of vocabulary in youth, it stands to reason that during the temporal development of expressive language, different "factors" or dissociable aspects of child language may develop before others. If, indeed, the "general" and "advanced" vocabulary structures observed in adults extend to youth, understanding the development of these structures would be of great theoretical and applied interest.

One such domain of language, ISL, emerges in the second postnatal year (Bretherton et al., 1981), with 30% of 20-month-olds using linguistic labels for fear, pain, disgust, morality, and affection, and rapid growth of these terms occurring during the child's third year (Bretherton & Beeghly, 1982). By 30-months, 81% of toddlers use ISL; by 36 months, 97% of toddlers are reliably producing internal state terms (Kristen et al., 2012). Language about internal states may range from labeling physiology cues (e.g., "I'm hungry"), to commenting on another person's knowledge (e.g., "she knows"). Variation in ISL production may be best accounted for by a common latent factor, or ISL itself may comprise multiple interrelated yet distinct factors. Given this wide range of ways in which one can describe internal states of selves and others, researchers frequently parse ISL into different domains to capture this variance. Interestingly, ISL development often differs as a function of these linguistic categories.

In a longitudinal study of ISL production at 24, 30, and 36 months, toddlers started predominantly producing physiological (e.g., "tired", "hungry"), perceptual (e.g., "taste", "see", "touch"), and desire (e.g., "want", "need") language, and then expanded to produce a greater variety of mental terms during the third year (Kristen et al., 2012). These results have been replicated in other research, underscoring a higher prevalence of volition/ability (e.g., "want", "can") and physiological terms during the second postnatal year followed by emotional/affect (e.g., "sad", "happy", "mad") and moral/obligation words (e.g., "bad", "nice"), with cognition words (e.g., "know", "understand") lagging significantly behind (Kristen et al., 2012). The relatively late use of cognition words has been replicated and extended in English and other languages, including Spanish and Chinese dialects (Ferres, 2003; Pascual et al., 2008; Shatz et al., 1983; Tardif & Wellman, 2000). Taken together, evidence suggests a developmental pathway starting with labeling physiological states, then moving into ability and emotional states, and culminating in cognitive states, with this last category developing in the third year. Exceptions to the traditional ISL developmental trajectory have been established in clinical groups, including autistic toddlers and children with attention-deficit/hyperactivity disorder, especially as it relates to cognitive (Rumpf et al., 2012) and emotion labels (Siller et al., 2014), which are produced at a significantly lower frequency compared to their typically-developing, same-aged peers. Indeed, categorical development of ISL takes a predictable pathway through early childhood that, when altered or presented differently, may reflect atypical social communication (Rumpf et al., 2012).

These categorical developmental trajectories suggest that ISL may not be a unitary construct. In other words, patterns in ISL production may be better represented using a more complex factor structure. From its earliest conception, Bretherton et al. (1981) proposed a six-factor model of ISL including perception (e.g., sight, hearing, taste), physiology (hunger, thirst), positive and negative affect (joy, anger, fear), volition and ability (desire, need), cognition (know, dream), and morality (permission, obligation). Admittedly, Bretherton and Beeghly (1982) suggested that many of these terms could reside in alternate categories, such as "know", which could be considered either a cognitive term or volition term. Thus, these categories were suggested as recommendations as opposed to definite groupings. Since the 1980s, ISL has been defined in a multitude of ways, ranging from several categories (i.e., 7 categories: cognitions, desires, emotions, intentions, preferences, perceptions, and physiology [Hashmi et al., 2021]; 11 categories; positive emotion, negative emotion, valence reversal, physiology, ability, volition, obligation, moral, cognition, emotion-modulatory particles and cognitive-contrast particles [Lemche et al., 2007]), to smaller groupings of two or three constructs (Roger et al., 2012; Siller et al., 2014; Tager-Flusberg & Sullivan, 1995). Some ISL theorists have proposed that physiological terms may not be considered ISL, as physiological states (e.g., smiling, crying) are observable and thus do not imply any mentalistic inference (Meins et al., 2001, 2006). This exclusion of physiological labels from ISL categorization is argued to be supported in the literature, as evidenced by temporally earlier emergence developmentally (Kristen et al., 2012). Therefore, some works do not consider physiological terms as part of ISL (e.g., Meins et al., 2006), whereas other studies simply group them as a distinct ISL category (e.g., Tager-Flusberg & Sullivan, 1995).

Despite empirical evidence suggesting differential categorical development of ISL in toddlerhood, theoretical evidence suggesting some terms should be included/excluded, and lack of consensus in measurement across studies, no work to date has statistically evaluated the factor structure of ISL. Thus, the first aim of the current study was to evaluate dimensions of ISL produced by English-speaking toddlers from ages 16 to 30 months. As studies have suggested the second and third postnatal years are ripe with ISL development with 81% of 30-month-olds producing ISL from all measured

https://doi.org/10.1017/S0305000924000060 Published online by Cambridge University Press

categories (Kristen et al., 2012), this developmental period represents a dynamic time to evaluate ISL dimensionality.

Aim 2: Longitudinal development of ISL

Much of the previous longitudinal research on ISL development has focused on the differential development of ISL CATEGORIES from ages two through four years of age, highlighting the predictable categorical development and late emergence of cognitive terms (Ferres, 2003; Kristen et al., 2012; Pascual et al., 2008; Shatz et al., 1983; Tardif & Wellman, 2000). However, the previous research into ISL categorical development has used theory-driven categories as opposed to data-driven categories. Thus, the second aim, building from the first, seeks to evaluate the longitudinal growth of ISL factors derived from aim 1, while accounting for differences in developmental trajectories due to maternal (i.e., maternal education) or child (i.e., sex) factors.

Maternal education has been considered as a proxy for socioeconomic status and a child's language environment (Hoff, 2006), with the argument that mothers with higher degrees of education may gain employment at a higher rate, thus affording greater resources, and may produce speech at a higher quantity and quality. Higher maternal education has been consistently associated with stronger child language skills across numerous forms of measurement (e.g., mean length of utterance, total vocabulary size, vocabulary diversity; Dollaghan et al., 1999). With regard to ISL, studies consistently report positive associations between maternal education and mothers' use of mental state terms (e.g., Adrian et al., 2005), which are then related to the child's use of ISL (Dunn et al., 1991; LaBounty et al., 2008; Longobardi et al., 2018; Roger et al., 2012). Indeed, the parent-child context is crucial for the examination of ISL development, in which maternal education may play a meaningful role as a proxy for language and resources.

The dyadic context of mother-child interactions also calls upon child characteristics that may influence parental ISL use, such as child sex. Much research has noted the gendered difference in how males and females discuss emotions (Roger et al., 2012). Research indicates that females use more emotion labels with parents than males at 2 years of age (Cervantes & Callanan, 1998; Dunn et al., 1991), and that preschool females use more mental states when talking with friends than males (Hughes & Dunn, 1998, 2002; Hughes et al., 2007). Kristen and colleagues (2014) evaluated toddlers' ISL using parent-report across four languages and also found a significant sex difference, wherein females produced more ISL terms than males. However, the authors attribute this difference to females having a larger vocabulary than males at this age as there was no significant difference in proportion of subcategories of ISL. These early sex differences have been largely attributed to gendered emotion socialization, as many findings have been related to differential use of ISL by mothers and fathers with their children (Kuebli & Fivush, 1992). In fact, research has largely failed to support sex differences in ISL production from 28 months (Bretherton & Beeghly, 1982) through preschool age (e.g., 3- to 4-year-olds; Jenkins et al., 2003; Recchia & Howe, 2008). Although these cross-sectional studies suggest no significant sex differences in ISL production in early toddlerhood, no works have evaluated sex as a predictor of ISL growth over time, which may further elucidate developmental phenomena related to gendered socialization or differences in language acquisition.

Aim 2 hinges on the results obtained from Aim 1 to investigate how factor(s) of ISL develop from ages 16 to 30 months in monolingual English toddlers. Importantly, relevant child and parent characteristics are evaluated as predictors of ISL individual differences (i.e., intercept), and growth (i.e., slope).

Method

Participants

Participants included N = 6,373 toddlers from ages 16 to 30 months (45.5% male) and their primary caregivers who contributed cross-sectional or longitudinal data to WordBank. Children were predominantly white (57%) with educated mothers (52% received a college degree or higher). Toddlers included in the study were born full-term and were reported to be monolingual speakers of American English without any known or suspected developmental differences. Caregivers provided their informed consent for data to be shared in WordBank and published.

Given the nature of the WordBank dataset, both the age at which children's language data were first recorded and the number of data collection waves children participated in varied across children. At the first data collection wave, children in the full sample on average were 23.5-months of age (SD = 4.37). Among the longitudinal subsample (n = 817), children ranged from 16- to 29-months of age when their caregiver first reported their child's language ability (M = 19.24 months, SD= 2.79), and the number of data collection waves each child contributed usable language data for is as follows: two (n = 577, 70.6%), three (n = 160, 19.6%), four (n = 38, 4.7%), five (n = 25, 3.1%), six or more (n = 17, 2.1%). Demographics are available in Table 1; the breakdown of age range by sex is available in Figure 1.

Data source

Data were downloaded from WordBank (Frank et al., 2016) on 9/1/2022. For the following analyses, full child-level and item-level responses were extracted from the MacArthur-Bates Communicative Development Inventory Words and Sentences Forms in American English for typically-developing, monolingual samples.

Measures

Demographics

Child and parent demographics available in WordBank were used for the current analyses. Relevant demographic factors identified a priori with established theoretical connections to language were included in the subsequent analyses including child sex (0 = *male*, 1 = *female*) and maternal education. Maternal education was coded categorically (1 = *no high school diploma*, 2 = *high school diploma*, 3 = *some college*, 4 = *college degree*, 5 = *some graduate school*, 6 = *graduate school degree*, 7 = *missing*) and then dichotomized (0 = *no college education (codes 1-3)*, 1= *college education (codes 4-6)*); an approach established in previous works as sensitive enough to detect group differences in child language (e.g., Gilkerson et al., 2017; Hoff-Ginsberg, 1998).

https://doi.org/10.1017/S0305000924000060 Published online by Cambridge University Press

Table 1. Demographic characteristics of study participants

	Total Sample (*N* = 6,373)
Child Sex (*n*, %)	
Male	2901 (45.5%)
Female	2689 (42.2%)
Did Not Report	783 (12.3%)
Maternal Education (*n*, %)	
No High School Diploma	175 (2.7%)
High School Diploma	573 (9.0%)
Some College	816 (12.8%)
College Degree	1575 (24.7%)
Some Graduate School	227 (3.6%)
Graduate Degree	1560 (24.5%)
Did Not Report	1447 (22.7%)
Child Race (*n*, %)	
White	3628 (56.9%)
Black	252 (4.0%)
Asian	170 (2.7%)
Other	504 (7.9%)
Did Not Report	1819 (28.5%)
Child Ethnicity (*n*, %)	
Hispanic	469 (7.4%)
Non-Hispanic	2076 (32.6%)
Did Not Report	3828 (60.1%)

MacArthur-Bates Communicative Development Inventory (CDI)

Primary caregivers completed the CDI Words and Sentences (CDI-WS; Fenson et al., 2007), which is normed for children from ages 16 to 30 months to evaluate expressive language skills. Although toddlers from ages 16 to 18 months can also be administered the CDI Words and Gestures form, the CDI-WS is normed for the younger range of this demographic. The CDI-WS reports excellent metrics of reliability, internal consistency, and validity (Fenson et al., 2007). Previous research supports the use of parent-report questionnaires for the measurement of ISL in this age range (Kristen et al., 2012) and suggests parent-report measures better capture ISL production compared to naturalistic observations or semi-structured interaction tasks (Pascual et al., 2008), as naturalistic or elicitation tasks may represent a skewed perspective of the child's ISL knowledge (Drummond et al., 2014).

The CDI-WS comprises three sections; the first section evaluates the words children use with a 680-word vocabulary checklist. Parents select 'yes' or 'no' for whether their child produces words on the checklist. From this section, 43 ISL terms were extracted

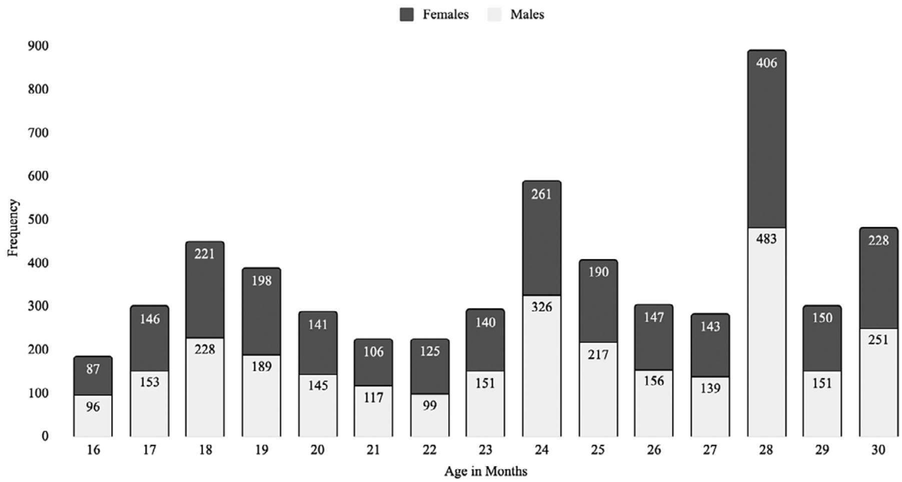

Figure 1. Histogram of participants by age and sex.

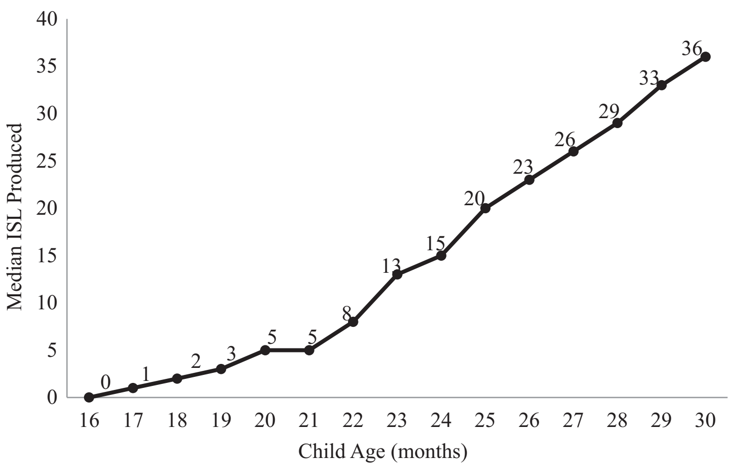

Figure 2. Median number of ISL terms produced over time.

(Figure 2). The 43 terms were selected due to their overlap with other measures of ISL (i.e., the Internal State Language Questionnaire; Olineck & Poulin-Dubois, 2005) and due to their use in previous ISL research using the CDI-WS (e.g., Lamb, 1991; Longobardi et al., 2018; Raines, 2014). Children received a score of "1" for every word the caregiver endorsed, otherwise they received a "0" for that item.

Analytic plan

The analytic plan for aims one and two were pre-registered on Open Science Framework: https://doi.org/10.17605/OSF.IO/TGB2Y.

Aim 1: the factor structure of ISL

To evaluate the latent factor structure of ISL, exploratory (EFA) and confirmatory factor analyses (CFA) were conducted in Mplus (Version 8; Muthén & Muthén, 2017). EFA and CFA can be used in succession to provide a more rigorous and comprehensive examination of the factor structure in an observed dataset. When there is limited knowledge of the underlying factor structure of a construct, as is the case for ISL, EFA first identifies the number of latent factors that best explain the pattern of correlations in the data. Once a tentative factor structure has been identified, CFA is then used to confirm, or check, that the exploratory structure still holds. Splitting the larger study sample in half prior to conducting EFA and CFA enhances the robustness of the factor structure exploration and model confirmation by independently validating the findings in one subset and testing their replicability in the second subset. Thus, the WordBank sample of 6,373 children was split into two subsamples using simple random sampling in SPSS (Version 27). One sample was used to conduct an exploratory factor analysis (EFA) to identify the best fitting factor model (Sample A: $N = 3{,}121$; $M_{age} = 23.49$ months, $SD_{age} = 4.39$; 47.2% female). The second sample was then used to perform confirmatory factor analyses (CFA) to confirm the factorial structure obtained from the EFA (Sample B: $N = 3{,}252$; $M_{age} = 23.52$ months, $SD_{age} = 4.35$; 49% female). As anticipated, the subsamples did not differ in terms of child sex, $\chi^2(1) = 1.77$, $p =. 19$, child race, $\chi^2(4) = 4.72$, $p =. 32$, or maternal education, $\chi^2(1) = 1.23$, $p =. 27$.

For the EFA, given that our data are dichotomous (yes/no), a tetrachoric correlation matrix was generated and we utilized the robust weighted least squares (WLSMV) estimator with an oblique geomin factor rotation (Yates, 1987). The scree plot, eigenvalues, and factor loadings were examined to identify the optimal number of factors. Factors with an eigenvalue ≥ 1 were retained (Kaiser, 1960) and the factor loadings were inspected and considered acceptable if ≥. 45 (Hair et al., 2010). CFA using the WLSMV estimation method was then performed to validate the identified one-factor model, where all 43 items were loaded onto a single latent ISL factor. The factor loadings were once again inspected and model fit was assessed using the following descriptive and inferential indices of model fit: Chi-square (χ^2) test, root mean square of approximation (RMSEA, values ≤. 06 reflect acceptable fit; MacCallum et al., 1996), standardized root mean squared residual (SRMR, values ≤. 08 reflect acceptable fit; Hu & Bentler, 1999), and confirmatory fit index (CFI, values ≥. 90 reflect acceptable fit; Hu & Bentler, 1999).

Aim 2: The early development of ISL

To evaluate ISL growth across early childhood, multilevel modeling was used to account for the nested structure of the data (i.e., repeated observations within children over time). Multilevel modeling can be used to examine whether and how children's ISL production grows over time (i.e., within-person changes). By introducing potential predictors of ISL development into the multilevel model, we can further evaluate whether ISL growth patterns or children's starting points are influenced by child- or maternal-factors such as sex and maternal education (i.e., between-person differences). A balanced dataset is not a requirement of multilevel modeling, and it is therefore well-suited to analyzing data with varying numbers of measurement waves that are unevenly spaced across individual participants (Singer & Willett, 2003). Furthermore, partially missing data (e.g., when some children have just one or two observations) can be handled using full information maximum likelihood, which estimates population parameters that would most likely produce the estimates from the sample data (Curran et al., 2010). This advantage of multilevel modeling is particularly important as both cross-sectional and longitudinal data are available within the WordBank dataset.

Analyses were conducted in HLM Version 7.0 (Raudenbush et al., 2011) using a taxonomy of two-level statistical models, which provides a baseline from which to compare model fit between nested models using model deviance (Bryk & Raudenbush, 1992), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC). In other words, increasingly complex models can be generated by introducing new variables into the original multilevel model. This approach allows us to evaluate whether the additional variance accounted for by a new variable is significant, above and beyond the previous model. In the two-level hierarchical linear model, the Level 1 (L1) equation represents the within-person model, the Level 2 (L2) equation represents the between-person model, *t* represents repeated observations, and *i* represents individual children. To quantify how much of the total variation in ISL is within-person vs. between-person variance using the intraclass correlation coefficient (ICC; Bryk & Raudenbush, 1992), we first estimated an unconditional means model (Model A) as follows:

$$\text{Level 1}: \quad \mathrm{ISL_{ti}} = \pi_{0\mathrm{i}} + e_{\mathrm{ti}}$$

$$\text{Level 2}: \quad \pi_{0i} = \beta_{00} + r_{0\mathrm{i}}$$

In the unconditional means model, the total variance in ISL aggregated across time-points is partitioned into its within, σ^2, and between, τ_{00}, components, which can then be used to calculate ICC. Next, to test our research question regarding the rate of ISL change across early childhood, we estimated an unconditional linear growth model (Model B) as follows:

$$\text{Level 1}: \quad \mathrm{ISL_{ti}} = \pi_{0i} + \pi_{1i}*(\mathbf{AGE}_{\mathrm{ti}}) + e_{\mathrm{ti}}$$

$$\text{Level 2}: \quad \pi_{0i} = \beta_{00} + r_{0\mathrm{i}}$$
$$\pi_{1i} = \beta_{10} + r_{1\mathrm{i}}$$

In the unconditional linear growth model, AGE (in months) was the only variable entered at L1 in order to generate a baseline model depicting children's rate of ISL growth (slope) without any L2 predictors entered. To set a meaningful and variable intercept (initial status), AGE was centered at the grand mean so that the initial status value corresponds to children's average ISL at the end of the second postnatal year (i.e., at 23.5-months). Finally, to address whether variation in ISL initial status or slope is accounted for by either child sex or maternal education status, these time-invariant covariates were dummy-coded and entered into the model. That is, we parameterized a conditional linear growth model (Model C) as follows:

$$\text{Level 1}: \quad \mathrm{ISL_{ti}} = \pi_{0i} + \pi_{1i}*(\mathbf{AGE}_{\mathrm{ti}}) + e_{\mathrm{ti}}$$

$$\text{Level 2}: \quad \pi_{0i} = \beta_{00} + \beta_{01}*(\mathrm{SEX_i}) + \beta_{02}*(\mathrm{EDUCATION_i}) + r_{0\mathrm{i}}$$
$$\pi_{1i} = \beta_{10} + \beta_{11}*(\mathrm{SEX_i}) + \beta_{12}*(\mathrm{EDUCATION_i}) + r_{1\mathrm{i}}$$

In the conditional linear growth model, ISL for child *i* at age *t* is modeled as a function of the intercept parameter (β_{00}; i.e., ISL at age 2), the slope parameter (β_{10}; i.e., ISL rate of

https://doi.org/10.1017/S0305000924000060 Published online by Cambridge University Press

Table 2. Multilevel models: fixed effects, variance components, and goodness-of-fit

		Model A	Model B	Model C
Fixed Effects				
Initial Status (π_0)	Intercept (β_{00})	17.31*** (0.18)	17.58*** (0.19)	13.77*** (0.41)
	SEX (β_{01})			3.08 *** (0.41)
	EDUCATION (β_{02})			1.53*** (0.44)
Rate of Change (π_1)	Intercept (β_{10})		2.58*** (0.06)	1.85*** (0.16)
	SEX (β_{11})			0.09 (0.15)
	EDUCATION (β_{12})			0.48** (0.18)
Variance Components				
Level 1	Within-person (σ^2)	157.31*** [12.54]	26.01*** [5.10]	65.59*** [8.10]
Level 2	Initial status (τ_{00})	74.90*** [8.65]	196.52*** [14.02]	141.57*** [11.90]
	Rate of change (τ_{11})		2.92*** [1.71]	1.02*** [1.01]
	Covariance (τ_{01})		16.68 (0.96)	11.93 (1.03)
Goodness-of-Fit Statistics				
	n	6,373	6,373	4,926
	p	3	6	10
	Deviance	62805.62	61310.72	44508.98
	AIC	62811.62	61322.72	44528.98
	BIC	62831.90	61363.28	44594.00

change), the individual-specific, between-person predictors ($\beta_{\_1}$ and $\beta_{\_2}$; i.e., child sex and maternal education), and the residual error (e_{ti}). As mentioned above, missing observations at L1 (within-person language data) were handled using full information maximum likelihood as this estimation method enables all children with at least one wave of ISL data to be included in the models. Children with missing observations at L2 (n = 1,447; between-person predictor data) were excluded only from the conditional linear growth model, hence the change in sample size between Models A and B vs. Model C in Table 2.

Results: Aim 1

Exploratory factor analysis

To evaluate the factor structure of ISL in early childhood, an EFA was performed on sample A with one to four factors extracted. A one-factor model, RMSEA =. 028 (90% CI [0.027, 0.029]); CFI =. 996, was identified by the analysis based on our examination of the scree plot and eigenvalues. The eigenvalues for the first three factors were 34.8, 0.7, 0.6, which suggests a one-factor solution based on the substantial separation between factors one and two, having applied the eigenvalue ≥ 1 standard for factor retention (Kaiser, 1960). For the one-factor model, the factor loadings for all 43 items ranged from 0.74 to 0.95.

Confirmatory factor analysis

To validate the identified ISL factor structure, the one-factor model was then tested by performing a CFA using sample B. Results supported a unidimensional factor structure of ISL as the model fit the data well, $\chi^2(860) = 2982.08$, $p < .001$; RMSEA =. 028 (90% CI [0.027, 0.030]); SRMR =. 021; CFI =. 997. All the ISL items loaded significantly on the respective latent factor (*p*s <. 05) and had acceptable factor loadings (i.e., ≥. 45; see Supplementary Figure 1).

Results: Aim 2

Preliminary analysis

Due to the results of Aim 1 indicating a best-fitting one-factor structure of ISL, children's ISL scores (range = 0-43) were summed to create a composite score of ISL. We then examined the descriptive statistics for the language variables as well as the participant demographics (i.e., the between-person predictors). Most children were not producing any ISL items at 16-months of age ($\text{Median}_{16\text{-mo}} = 0$). However, having graphically inspected the data, children's rate of ISL growth appeared to proceed in a relatively linear fashion (Figure 2). Because the majority of children were producing few if any ISL terms at the earliest ages in this study, the data are moderately right skewed at 16-, 17-, and 18-months of age. Although useful in approximating a normal distribution, data transformations can reduce the interpretability of the coefficients. Thus, to protect against violations of assumptions (i.e., normality and homogeneity of variance), we calculated robust standard errors in place of regular standard errors in our analyses (Bryk & Raudenbush, 1992). Additionally, several outliers (i.e., values that are ± 3*SD* of the mean) were detected in our dataset. Given evidence of extensive variability in young children's productive vocabulary and ISL development (e.g., Kristen et al., 2012), we opted to retain the outliers in the dataset to reflect the natural variation in early childhood ISL. It should be noted, however, that we conducted our analyses with and without the outliers included in the dataset and no differences emerged as a function of outlier management strategy.

Multilevel models

Unconditional means model

As a first step toward building the final HLM growth model, we estimated an unconditional means model (Model A) to confirm that the variability in ISL at the individual (L1) and between-person (L2) levels were significantly different from zero. Table 2 shows the variance components, which resulted in an ICC of 0.323. Thus, 32% of the variance in ISL occurs at L2 between children, which is visually represented in the individual growth trajectories depicted in Figure 3.

Unconditional linear growth model

Next, we estimated an unconditional linear growth model (Model B) to evaluate the rate of change in children's ISL over time, and the results are presented in Table 2. Our results suggest that on average by age 2, children produced approximately 17 ISL terms. Additionally, the average rate of growth per month was 2.58, which suggests that children are producing approximately 5 new ISL terms every two months across this age range. We

https://doi.org/10.1017/S0305000924000060 Published online by Cambridge University Press

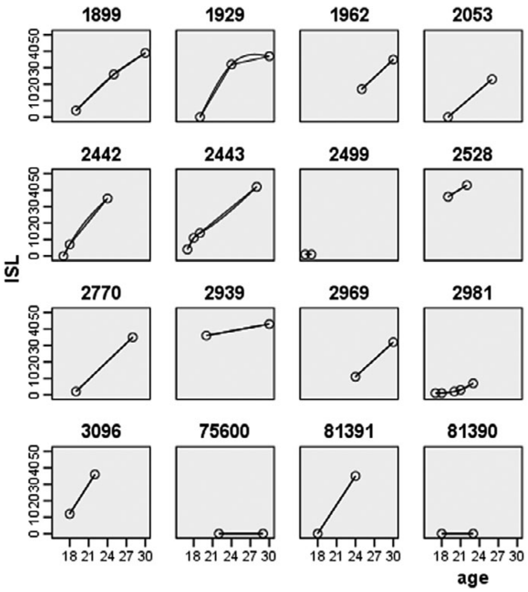

Figure 3. ISL growth trajectories for a randomly selected subset of children with longitudinal data.

also examined the association between initial status (π_0) and slope (π_1), $r = 0.69$, $p <. 01$. This value indicates that children with a larger ISL vocabulary by the end of their second year exhibited a faster rate of ISL growth over time across toddlerhood. As expected, the deviance test verified that the linear model better fit the data in comparison to the means model, $\chi^2(3) = 1494.28$, $p <. 001$.

Conditional linear growth model

Finally, we ran a conditional linear growth model (Model C) that included two L2 predictors to examine whether children's initial status or rate of change differed as a function of child sex and/or maternal education status. In addition to comparing model AIC and BIC, the deviance test was once again used to verify that model fit improved in moving from Model B to Model C, $\chi^2(4) = 16801.74$, $p <. 001$. As can be seen from Table 2, sex, $\beta_{01} = 3.08$, $p <. 001$, and maternal education, $\beta_{02} = 1.53$, $p <. 001$, predicted the number of ISL terms children produced on average at 23.5-months of age. This finding implies that both girls as well as children of mothers with a college degree produced more ISL terms on average near age 2. Together, these variables accounted for 27.96% of the variance in ISL initial status as determined by the pseudo R^2 value. By contrast, although maternal education predicted ISL slope, $\beta_{12} = 0.48$, $p <. 001$, child sex did not, $\beta_{01} = 0.09$, $p =. 54$. Thus, children of more highly educated mothers exhibited a faster rate of ISL growth in comparison to children of mothers who do not hold a college degree. In total, our results suggest that 65.07% of the variance in ISL growth was accounted for by the variables in Model C.

Using the fixed effects generated by Model C, we constructed a model-based growth trajectory starting at the intercept to visualize ISL development between the subgroups in the current study (Figure 4). Four groups were created by pairing child sex with maternal

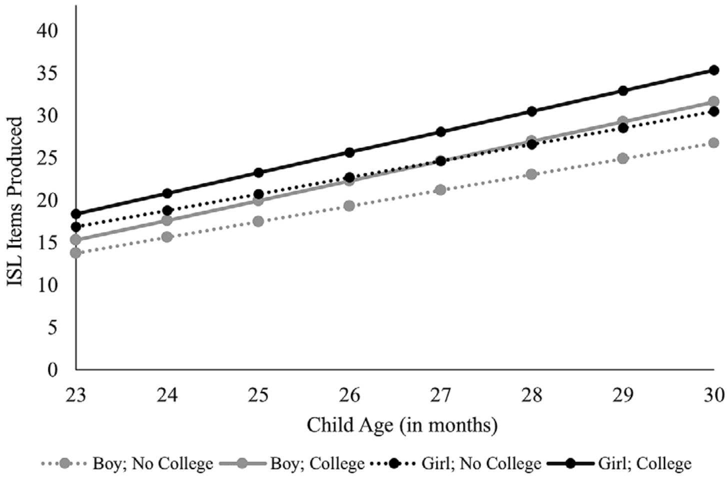

Figure 4. Model-based ISL growth trajectories as a function of child sex and maternal education.

education status. Therefore, Figure 4 depicts how the ISL starting point and trajectories differ between boys and girls as a function of whether their mother holds a college degree.

Discussion

The current two-aim study sought to address theoretical and empirical gaps in the ISL literature by evaluating (1) ISL factor structure and (2) predictors of ISL development between 16- and 30-months of age in typically-developing, monolingual English-speaking toddlers. Results will be discussed in accordance with aim, and then summarized together.

Aim 1. Factor structure of ISL

Aim 1 evaluated the factor structure of ISL between the ages of 16- and 30-months to investigate the validity of the wide range of ISL categorization practices (Hashmi et al., 2021; Lemche et al., 2007; Meins et al., 2001, 2006; Roger et al., 2012; Siller et al., 2014; Tager-Flusberg & Sullivan, 1995). Results from EFA and CFA suggested a one-factor structure best fit the ISL data with all statistical indices of model fit meeting cutoff criteria, aside from the chi-square statistic which was expected given this statistic's sensitivity to sample size (Babyak & Green, 2010). In other words, collapsing across this narrow age range, results suggest that ISL can best be considered a unitary factor.

Our results can lend a response to an ongoing debate in the ISL literature about whether physiological terms should be considered as internal state language (Kristen et al., 2012; Meins et al., 2001, 2006). Although they may emerge in a distinct temporal window with an onset prior to other ISL terms (Kristen et al., 2012), our results suggest physiological terms warrant consideration as ISL. Further, our results suggest a categorical consideration of ISL in toddlerhood may not be rooted in the underlying factor structure of ISL at this age. Continuing to parse by linguistic category may serve some theoretical relevance for studies interested in certain terms or groups of terms based on a priori research questions but is not empirically supported in this demographic for these ISL terms in this age range.

Our results could be seen as presenting findings that are contradictory to previous works suggesting a developmental order of acquisition of internal state words (and maternal/paternal use of ISL words differentially by category and by age). However, it is important to note that our findings do not discount previous research showing that children begin producing some ISL terms before others. Rather, our results suggest that there appears to be a common, underlying mechanism for ISL production in toddlerhood. That is, children's production of cognitive ISL terms may not be entirely dissociable from perception ISL terms, for example. These findings may have meaningful implications for targeting the core, underlying mechanism of ISL development in clinical groups such as autism and ADHD. An additional explanation to the developmental timing of certain words could be the general growth of vocabulary. In other words, perhaps as children gain greater vocabulary knowledge with age, this provides access to new ISL terms. Future research may wish to empirically evaluate the onset of ISL terms as a function of vocabulary size (e.g., 50 words, 200 words, 500 words) rather than simply as a function of age in toddlerhood.

Furthermore, although a single factor of ISL was observed in the current study, it is possible that the factor structure of ISL changes over time, which in turn could yield different trajectories during later toddlerhood or the preschool years. Indeed, this pattern would be similar to that which has been uncovered by early childhood executive function researchers. For example, despite evidence that inhibitory control and working memory emerge before cognitive flexibility, a unitary factor model has been identified in children ages 3 years and under, whereas a two- or three-factor model has been identified in later stages of development (Lehto et al., 2003; Wiebe et al., 2008, 2011).

One limitation of the current approach was the use of a parent-report form to capture ISL knowledge. Although the MCDI has been used in previous ISL research with this specific word list (e.g., Lamb, 1991; Longobardi et al., 2018; Raines, 2014), it does have noted limitations. First, a few words have multiple meanings, such as "can" ("*Can* I have this?" is an ISL term, but a "tinned *can*" is not), and these varied meanings are not captured on the parent-report form. However, due to the observed results of a unitary factor structure, there is greater confidence that these few, polysemous words did not change or impact our results. Second, the MCDI comprises a varied number of words from different proposed ISL categories. The large majority of words are dispositional or perceptual; there are only a few volition and cognitive words such as "need" and "want". Interestingly, these items had the weakest factor loading (although these values were still quite high and well-above threshold,. 86-.87). Future research may wish to use other vocabulary measures with a greater variety of volition and cognitive terms to replicate the current results.

Aim 2. Longitudinal development of ISL

The second aim of the current study sought to evaluate the individual predictors of ISL growth given the unitary factor structure of ISL observed via Aim 1. The results from our best-fitting model (Model C) suggested that right around the second postnatal year (23.5-months, intercept), children are producing an average of 13 ISL terms of the available 43 coded from the MCDI. However, the number of ISL terms children produced at this age varied considerably as a function of both sex and maternal education. Female toddlers

https://doi.org/10.1017/S0305000924000060 Published online by Cambridge University Press

had higher ISL production than males at age 2 by approximately three ISL terms, and toddlers with college-educated mothers had higher ISL production than toddlers of a mother without a college education by approximately two ISL terms. Consequently, the greatest difference emerged at age 2 between male toddlers of mothers without a college-education (~13 ISL terms) and female toddlers with college-educated mothers (~17 ISL terms). Furthermore, we illustrated that children are producing approximately 5 new ISL terms every two months across early childhood, and that this growth occurs at a linear rate in early childhood. It is worth noting that we inspected the data for non-linear patterns of growth; however, the linear model emerged as the best-fitting model as evidenced by non-significant quadratic terms and decreased model fit. Future research may wish to evaluate more complex patterns of growth via piecewise linear modeling or the use of linear splines.

Maternal education also emerged as a significant predictor of ISL growth (slope). These results suggested that college-educated mothers reported their toddlers as having faster rates of growth in ISL terms compared to less than college-educated mothers. That is, by 30-months of age, female toddlers of mothers without a college degree produced an average of 30 ISL terms whereas female toddlers of mothers with a college degree produced an average of 35 ISL terms. Interestingly, sex did not significantly predict the rate of growth in ISL, meaning males and females experienced similar growth trajectories of ISL over time. Given the significant difference in intercept, however, females in this sample maintained their advantage in ISL production across this developmental period compared to males. Of note, when comparing the model fit indices, adding these two individual-level predictors significantly improved the overall fit of the model, suggesting that child sex and maternal education are significant factors to consider when evaluating between-person differences in ISL development.

Related to the significant sex findings, our results replicate previous literature suggesting a female advantage in ISL and overall language during this developmental period (Hughes & Dunn, 2002; Kristen et al., 2014). However, it is important to keep in mind that this is simply a snapshot of a 14-month period, wherein females show this ISL advantage. Indeed, the use of a narrow age range may be artificially highlighting a sex difference that ebbs overtime. In a systematic review and meta-analysis (Etchell et al., 2018), sex differences were most apparent in language when using tight age ranges, suggesting sex differences may be most evident at certain developmental stages and negligible in others, likely due to differential rates of maturation. Certainly, these sex differences may be transient or specific to this developmental period, as previous research suggests no significant sex differences in ISL production from 28-months (Bretherton & Beeghly, 1982) through age 5 (Jenkins et al., 2003).

Taking a developmental approach, one should consider the role of equifinality, wherein with different starting points, groups may emerge at the same endpoint. Previous research has underscored the equifinality of ISL development across males and females when using extended age ranges. Cervantes and Callanan (1998) found that 2-year-old girls talked more frequently about emotions than boys with their mothers, but by age 4, boys had increased their frequency of emotion talk and no sex differences were evident. More recent research has supported this finding, suggesting that sex differences in language competence decreased from ages 3-6 years, with males showing greater variance in language outcomes than females (Lange et al., 2016). Despite previous research arguing for equifinality in ISL development as age increases, we did not see any trend to suggest this, as males and females did not have significantly different rates of growth. The lack of different growth trajectories suggests that during this developmental period, females are

https://doi.org/10.1017/S0305000924000060 Published online by Cambridge University Press

showing a clear advantage at each age point. Future research could continue to extend these growth trajectories beyond 30-months, as children were not yet at ceiling on ISL, to evaluate if males and females reach equifinality in their development. Extending these growth trajectories may reveal greater insights into the sex differences of longitudinal ISL development.

Although sex emerged as a significant predictor of intercept, maternal education was the only factor that significantly predicted intercept AND slope, or growth. These findings replicate previous literature highlighting the connection between maternal education and child ISL production across developmental periods (Dunn et al., 1991; LaBounty et al., 2008; Longobardi et al., 2018; Roger et al., 2012), which has been postulated to be mediated by maternal ISL production (Adrian et al., 2005). Maternal education, coded here dichotomously around a college education, is a coarse proxy for access to resources or maternal language production, with ample room for future research to investigate the mechanism behind this relationship, including investigating maternal ISL production, book-reading practices, or the gendered relationship between parent and child ISL discourse. Finally, in our longitudinal analyses, maternal education was entered as a fixed characteristic (L2 predictor) based on the data that were available. As many mothers will pursue continued education after the transition to motherhood (Augustine, 2016), future research should take a more dynamic approach by treating maternal education as a time-varying covariate to better capture changes in mothers' education status as they may relate to early ISL development.

Integrative summary

Combining the results of Study 1 and Study 2, our findings suggest a unitary structure of ISL between the ages of 16- to 30-months in terms of expressive vocabulary. This unitary structure develops linearly during this developmental period, with variations in start point and growth due to infant factors (sex) and maternal factors (education). Results of our hierarchical linear modeling suggests that females produce more ISL terms than males, with linear growth that continues to be significantly higher in females through the second postnatal year. Toddlers of college-educated mothers also present with a higher ISL vocabulary at the end of the second postnatal year (23mos), with faster rates of growth over time compared to mothers without a college education. These results are specific to the demographics of the current sample, including monolingual, English-speaking toddlers who are typically-developing between the ages of 16- and 30-months. Despite this specificity, these results have practical and theoretical implications. Practically, these results provide age-normed insights for expected ISL production that can be assessed using a common parent-reported measure, the MCDI, with anticipated averages around 13 ISL terms at 23-months of age. These results also provide data to suggest that certain demographics may show differential rates of growth and development in this domain. Theoretically, these results are the first to lend empirical support to the idea that ISL should be considered a unitary construct at this age and does not require further binning into different term types, such as physiological, cognitive, or emotional. Future research avenues are numerous and include multiple ideas for generalizing these findings further, including to novel groups of non-English speakers, multilingual households, and neuro-diverse toddlers. Future work could also extend these findings into later toddlerhood and early preschool years to evaluate the multi- or equi-finality of ISL development by sex, evaluate other predictors of ISL development such as birth order, or relate these findings to other domains of interest such as theory of mind and social skills in childhood.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0305000924000060.

Acknowledgements. We thank the contributors to the WordBank open repository project: we are grateful for your efforts in creating WordBank and disseminating scholarship via open science. The authors have no conflicts of interest to disclose. Author T.C.M. was supported by an NICHD training fellowship awarded to UNC-CH (T32 HD040127-21). Author M.B. was supported by the National Institute of Mental Health R01 MH119070 awarded to Dr. Sheila E. Crowell and Dr. Elisabeth Conradt. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD or NIMH.
This project was pre-registered on OSF: https://doi.org/10.17605/OSF.IO/TGB2Y

References

Adrian, J. E., **Clemente, R. A.**, **Villanueva, L.**, & **Rieffe, C.** (2005). Parent–child picture-book reading, mothers' mental state language and children's theory of mind. *Journal of Child Language*, **32**(3), 673–86. https://doi.org/10.1017/S0305000905006963

Augustine, J. M. (2016). Exploring new life course patterns of mother's continuing secondary and college education. *Population Research and Policy Review*, **35**, 727–755. https://doi.org/10.1007/s11113-016-9401-5

Babyak, M. A., & **Green, S. B.** (2010). Confirmatory factor analysis: an introduction for psychosomatic medicine researchers. *Psychosomatic Medicine*, **72**(6), 587–597. https://doi.org/10.1097/PSY.0b013e3181-de3f8a

Bellagamba, F., **Laghi, F.**, **Lonigro, A.**, **Pace, C. S.**, & **Longobardi, E.** (2014). Concurrent relations between inhibitory control, vocabulary and internal state language in 18-and 24- month-old Italian-speaking infants. *European Journal of Developmental Psychology*, **11**(4), 420–32. https://doi.org/10.1080/17405629.2013.848164

Booth, J. R., & **Hall, W. S.** (1994). Role of the cognitive internal state lexicon in reading comprehension. *Journal of Educational Psychology*, **86**, 413–22. https://doi.org/cm6vbt

Bowles, R. P., **Grimm, K. J.**, & **McArdle, J. J.** (2005). A structural factor analysis of vocabulary knowledge and relations to age. *The Journals of Gerontology: Series B*, **60**(5), 234–41. https://doi.org/10.1093/geronb/60.5.P234

Bretherton, I., & **Beeghly, M.** (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, **18**(6), 906–21. https://doi.org/brkgg5

Bretherton, I., **McNew, S.**, & **Beeghly-Smith, M.** (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a "theory of mind?" In M. E. Lamb & L. R. Sherrod (Eds.), *Infant Social Cognition*. Hillsdale, N.J.

Bryk, A. S., & **Raudenbush, S. W.** (1992). *Hierarchical linear models: Applications and data analysis methods*. Sage Publications, New York.

Carlson, S. M., **Mandell, D. J.**, & **Williams, L.** (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology*, **40**, 1105–22.

Cervantes, C., & **Callanan, M.** (1998). Labels and explanations in mother-child emotion talk: Age and gender differentiation. *Developmental Psychology*, **34**, 88–98. https://doi.org/crx93j

Chiarella, S. S., **Kristen, S.**, **Poulin-Dubois, D.**, & **Sodian, B.** (2013). Concurrent relations between perspective-taking skills, desire understanding, and internal-state vocabulary. *Journal of Cognition and Development*, **14**(3), 480–98. https://doi.org/ggmtbn

Curran, P. J., **Obeidat, K.**, & **Losardo, D.** (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, **11**(2), 121–136. https://doi.org/10.1080/15248371003699969

Dollaghan, C. A., **Campbell, T. F.**, **Paradise, J. L.**, **Feldman, H. M.**, **Janosky, J. E.**, **Pitcairn, D. N.**, & **Kurs-Lasky, M.** (1999). Maternal education and measures of early speech and language. *JSLHR*, **42**(6), 1432–43. https://doi.org/10.1044/jslhr.4206.1432

Drummond, J., **Paul, E. F.**, **Waugh, W. E.**, **Hammond, S. I.**, & **Brownell, C. A.** (2014). Here, there and everywhere: Emotion and mental state talk in different social contexts predicts empathic helping in toddlers. *Frontiers in Psychology, Developmental Psychology*, **5**. https://doi.org/10.3389/fpsyg.2014.00361

Dunn, J., **Brown, J.**, **Slomkowski, C.**, **Tesla, C.**, & **Youngblade, L. M.** (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, **62**, 1352–66. https://doi.org/10.2307/1130811

Ensor, R., & **Hughes, C.** (2008). Content or connectedness? Mother-child talk and early social understanding. *Child Development*, **79**(1), 201–16. https://doi.org/10.1111/j.1467-8624.2007.01120.x

Etchell, A., **Adhikari, A.**, **Weinberg, L. S.**, **Choo, A. L.**, **Garnett, E. O.**, **Chow, H. M.**, & **Chang, S.** (2018). A systematic literature review of sex differences in childhood language and brain development. *Neuropsychologia*, **114**, 19–31. https://doi.org/gd9nrx

Fenson, L., **Marchman, V. A.**, **Thal, D. J.**, **Dale, P. S.**, **Reznick, J. S.**, & **Bates, E.** (2007). *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual* (2nd ed.) Baltimore: Paul H. Brookes.

Ferres, L. A. (2003). Children's early theory of mind: Exploring the development of the concept of desire in monolingual Spanish children. *Developmental Science*, **6**, 159–165.

Frank, M. C., **Braginsky, M.**, **Yurovsky, D.**, & **Marchman, V. A.** (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, **44**(3), 677–94. https://doi.org/10.1017/S0305000916000209

Gilkerson, J., **Richards, J. A.**, **Warren, S. F.**, **Montgomery, J. K.**, **Greenwood, C. R.**, **Kimbrough Oller, D.**, **Hansen, J.**, & **Paul, T. D.** (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, **26**(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Gustafsson, J.-E., & **Holmberg, L. M.** (1992). Psychometric properties of vocabulary test items as a function of word characteristics. *Scandinavian Journal of Educational Research*, **36**, 191–210. https://doi.org/10.1080/0031383920360303

Hair, J. F., **Black, W. C.**, **Babin, B. J.**, & **Anderson, R. E.** (2010). *Multivariate Data Analysis: A Global Perspective*. 7th Edition, Pearson Education, Upper Saddle River.

Hashmi, S., **Vanderwert, R. E.**, **Paine, A. L.**, & **Gerson, S. A.** (2021). Doll play prompts social thinking and social talking: Representations of internal state language in the brain. *Developmental Science*, **25**(2), e13163. https://doi.org/10.1111/desc.13163

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, **26**, 55–88. https://doi.org/10.1016/j.dr.2005.11.002

Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics*, **19**, 603–29. https://doi.org/10.1017/S0142716400010389

Hu, L.-t., & **Bentler, P. M.** (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, **6**(1), 1–55. https://doi.org/10.1080/10705519909540118

Hughes, C., & **Dunn, J.** (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology*, **34**, 1026–37. https://doi.org/10.1037/0012-1649.34.5.1026

Hughes, C., & **Dunn, J.** (2002). 'When I say a naughty word': A longitudinal study of young children's accounts of anger and sadness in themselves and close others. *British Journal of Developmental Psychology*, **20**, 515–35. https://doi.org/10.1348/026151002760390837

Hughes, C., **Lecce, S.**, & **Wilson, C.** (2007). "Do you know what I want?" Preschoolers' talk about desires, thoughts and feelings in their conversations with sibs and friends. *Cognition & Emotion*, **21**, 330–350. https://doi.org/10.1080/02699930600551691

Jenkins, J. M., **Turrell, S. L.**, **Kogushi, Y.**, **Lollis, S.**, & **Ross, H. S.** (2003). A longitudinal investigation of the dynamics of mental state talk in families. *Child Development*, **74**, 905–920. https://doi.org/10.1037/0893-3200.19.4.533

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, **20**, 141–151. https://doi.org/10.1177/001316446002000116

Kristen, S., **Chiarella, S.**, **Sodian, B.**, **Aureli, T.**, **Genco, M.**, & **Poulin-Dubois, D.** (2014). Crosslinguistic developmental consistency in the composition of toddlers' internal state vocabulary: Evidence from four languages. *Child Development Research*, Article ID 575142. https://doi.org/10.1155/2014/575142

Kristen, S., **Sodian, B.**, **Licata, M.**, **Thoermer, C.**, & **Poulin-Dubois, D.** (2012). The development of internal state language during the third year of life: A longitudinal parent report study. *Infant and Child Development*, **21**(6), 634–645. https://doi.org/jh9q

Kuebli, J., & **Fivush, R.** (1992). Gender differences in parent-child conversations about past emotions. *Sex Roles*, **27**, 683–698. https://doi.org/10.1007/BF02651097

LaBounty, J., **Wellman, H. M.**, **Olson, S.**, **Lagattuta, K.**, & **Liu, D.** (2008). Mothers' *and* fathers' use of internstal state talk with their young children. *Social Development*, **17**(4), 757–75. https://doi.org/10.1111/j.1467-9507.2007.00450.x

Lamb, S. (1991). Internal state words: Their relation to moral development and to maternal communications about moral development in the second year of life. *First Language*, **11**, 391–406. https://doi.org/10.1177/014272379101103306

Lange, B., **Euler, H.**, & **Zaretsky, E.** (2016). Sex differences in language competence of 3- to 6-year-old children. *Applied Psycholinguistics*, **37**(6), 1417–1438. https://doi.org/f9c2zz

Lehto, J. E., **Juujärvi, P.**, **Kooistra, L.**, & **Pulkkinen, L.** (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, **21**(1), 59–80. https://doi.org/10.1348/026151003321164627

Lemche, E., **Kreppner, J. M.**, **Joraschky, P.**, & **Klann-Delius, G.** (2007). Attachment organization and the early development of internal state language: A longitudinal perspective. *International Journal of Behavioral Development*, **31**(3), 252–62. https://doi.org/bmj6dk

Longobardi, E., **Spataro, P.**, & **Colonnesi, C.** (2018). Maternal communicative functions and mind-mindedness at 16 months as predictors of children's internal and non-internal language at 20 months. *Inf. Behavior and Development*, **50**, 52–63. https://doi.org/gddvhc

MacCallum, R. C., **Browne, M. W.**, & **Sugawara, H. M.** (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, **1**(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Meins, E., **Fernyhough, C.**, **Fradley, E.**, & **Tuckey, M.** (2001). Rethinking maternal sensitivity: Mothers' comments on infants' mental processes predict security of attachment at 12 months. *Journal of Child Psychology and Psychiatry*, **42**, 637–48. https://doi.org/dr3nzq

Meins, E., **Fernyhough, C.**, **Johnson, F.**, & **Lidstone, J.** (2006). Mind-mindedness in children: Individual differences in internal-state talk in middle childhood. *British Journal of Developmental Psychology*, **24**, 181–96. https://doi.org/10.1348/026151005X80174

Muthén, L. K., & **Muthén, B.** (2017). *Mplus User's Guide* [Computer Software and Manual], 8th Edn. Los Angeles, CA: Muthén & Muthén.

Olineck, K. M., & **Poulin-Dubois, D.** (2005). Infants' ability to distinguish between intentional and accidental actions and its relation to internal state language. *Infancy*, **8**(1), 91–100, https://doi.org/10.1207/s15327078in0801_6

Olineck, K. M., & **Poulin-Dubois, D.** (2007). Imitation of intentional actions and internal state language in infancy predict preschool theory of mind skills. *The European Journal of Developmental Psychology*, **4**, 14–30. https://doi.org/10.1080/17405620601046931

Pascual, B., **Aguado, G.**, **Sotillo, M.**, & **Masdeu, J. C.** (2008). Acquisition of mental state language in Spanish children: A longitudinal study of the relationship between the production of mental verbs and linguistic development. *Developmental Science*, **11**, 454–6. https://doi.org/10.1111/j.1467-7687.2008.00691.x

Raines, E. (2014). Inhibitory control and internal state language in 30-month-old toddlers. *Unpublished doctoral thesis.* https://core.ac.uk/download/pdf/210595455.pdf

Raudenbush, S. W., **Bryk, A.**, & **Congdon, R.** (2011). *HLM 7: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Recchia, H. E., & **Howe, N.** (2008). Family talk about internal states and children's relative appraisals of self and sibling. *Social Development*, **17**, 776–94. https://doi.org/b3tx7c

Roger, K. M., **Rinaldi, C. M.**, & **Howe, N.** (2012). Mothers' and fathers' internal state language with their young children: An examination of gender differences during an emotions task. *Infant and Child Development*, **21**, 646–66. https://doi.org/10.1002/icd.1762

Ruffman, T., **Slade, L.**, & **Crowe, E.** (2003). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, **73**(3), 734–51. https://doi.org/10.1111/1467-8624.00435

Rumpf, A.-L., **Kamp-Becker, I.**, **Becker, K.**, & **Kauschke, C.** (2012). Narrative competence and internal state language of children with Asperger syndrome and ADHD. *Research in Developmental Disabilities*, **33**, 1395–1407. https://doi.org/10.1016/j.ridd.2012.03.007

Shatz, M., **Wellman, H. M.**, & **Silber, S.** (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, **14**(3), 301–321. https://doi.org/10.1016/0010-0277(83)90008-2

Siller, M., **Swanson, M. R.**, **Serlin, G.**, & **Teachworth, A. G.** (2014). Internal state language in the storybook narratives of children with and without autism spectrum disorder: Investigating relations to theory of mind abilities. *Research in Autism Spectrum Disorders*, **8**(5), 589–96. https://doi.org/f5zvz8

Singer, J. D., & **Willett, J. B.** (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195152968.001.0001

Tager-Flusberg, H., & **Sullivan, K.** (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics*, **16**, 241–256. https://doi.org/10.1017/S0142716400007281

Tardif, T., & **Wellman, H. M.** (2000). Acquisition of mental state language in Mandarin- and Cantonese-speaking children. *Developmental Psych.*, **36**, 25–43. https://doi.org/cpg2hv

Wiebe, S. A., **Espy, K. A.**, & **Charak, D.** (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, **44**(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575

Wiebe, S. A., **Sheffield, T.**, **Nelson, J. M.**, **Clark, C. A. C.**, **Chevalier, N.**, & **Espy, K. A.** (2011). The structure of executive function in 3-year-old children. *Journal of Experimental Child Psychology*, **108**(3), 436–452. https://doi.org/10.1016/j.jecp.2010.08.008

Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. Albany: State University of New York Press.

Cite this article: McFayden, T.C., & **Bruce, M.** (2024). Internal state language factor structure and development in toddlerhood: Insights from WordBank. *Journal of Child Language* **51,** 721–740, https://doi.org/10.1017/S0305000924000060