Received: 24 June 2020 | Accepted: 29 June 2021

DOI: 10.1111/infa.12422

RESEARCH ARTICLE

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

Symbolic play provides a fertile context for language development

Noëlie Creaghe[1,2] | **Sara Quinn**[1] | **Evan Kidd**[1,2,3,4]

[1]Research School of Psychology, The Australian National University, Canberra, ACT, Australia

[2]ARC Centre of Excellence for the Dynamics of Language, Canberra, ACT, Australia

[3]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[4]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

Correspondence
Noëlie Creaghe, Research School of Psychology, The Australian National University, Canberra, ACT 2601, Australia.
Email: noelie.creaghe@anu.edu.au

Evan Kidd, Language Development Department, Max Planck Institute for Psycholinguistics, P.O. Box 310, Nijmegen, 6500AH, The Netherlands.
Email: evan.kidd@mpi.nl

Funding information
Australian Research Council, Grant/Award Number: CE140100041

Abstract

In this study, we test the hypothesis that symbolic play represents a fertile context for language acquisition because its inherent ambiguity elicits communicative behaviors that positively influence development. Infant–caregiver dyads (N = 54) participated in two 20-minute play sessions six months apart (Time 1 = 18 months, Time 2 = 24 months). During each session, the dyads played with two sets of toys that elicited either symbolic or functional play. The sessions were transcribed and coded for several features of dyadic interaction and language; infants' linguistic proficiency was measured via parental report. The two contexts elicited different communicative and linguistic behaviors. Notably, the symbolic play condition resulted in significantly greater conversational turn-taking than functional play, and also resulted in the greater use of questions and mimetics in infant-directed speech (IDS). In contrast, caregivers used more imperative clauses in functional play. Correlational and regression analyses showed that frequent properties of symbolic play (i.e., turn-taking, yes–no questions, mimetics) were positively related to infants' language proficiency, whereas frequent features of functional play (i.e., imperatives in IDS) were negatively related. The results provide evidence supporting the hypothesis that symbolic play is a fertile context for language development, driven by the need to negotiate meaning.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. *Infancy* published by Wiley Periodicals LLC on behalf of International Congress of Infant Studies.

1 | INTRODUCTION

> The development of language, then, involves two people *negotiating*
>
> (Bruner, 1983, p. 39, italics added)

Symbolic play and language acquisition have long been linked (for reviews, see Hirsh-Pasek et al., 2009; Lillard et al., 2013; Quinn et al., 2018). Piaget (1962) attributed the relationship to the fact that both depend on the child's emerging understanding of symbols (the "semiotic function"), a suggestion that has framed much of the research on the topic (Bates et al., 1979; McCune, 1995; Werner & Kaplan, 1963). On this interpretation, the symbolic play–language relationship is an epiphenomenon of the symbolic capacity of our species (Deacon, 1997). On the other hand, sociocultural approaches to development, beginning with Vygotsky (1967, 1978; see also Bruner, 1983; Rakoczy, 2006), identify symbolic play as a rich and challenging context that may bootstrap communicative development. In the current paper, we test a crucial prediction of the sociocultural approach—that the representational nature of symbolic play elicits behaviors that foster language and communicative development.

1.1 | Symbolic play, development, and language

Play is a complex behavior that defies neat categorization across the biological and psychological sciences (Smith, 2009). While play behaviors such as object manipulation and play fighting are common across different species, symbolic play—the non-literal use of objects, action, or attributes—is likely an evolved trait unique to humans (Lillard, 2017). Consistent with this assertion, symbolic play appears to develop on a relatively fixed schedule and is present in every culture in which it has been studied (Callaghan et al., 2011), although its content and frequency are culturally mediated (Lillard, 2017).

Because symbolic play, broadly construed, captures a diverse range of behaviors, attempts to categorize its function have proven difficult, and it is therefore likely it has multiple functions (Bateson & Martin, 2013). One hypothesis is that, as a particularly prominent feature of childhood, it provides an especially framed situation within which children can hone or even acquire important socio-cognitive skills (Lillard, 2001; Lillard et al., 2011). Thus, with respect to language, the broad hypothesis is that engaging in symbolic play is *in some way* beneficial for development, although pinning down the exact details has proved elusive.

What we can be certain about is that the two domains are coupled across early development. In a meta-analysis of 35 studies investigating the correlational evidence for an association between symbolic play and language development, Quinn et al. (2018) reported a robust small-to-medium association ($r = .35$), which held across the ages of 1–6 years, and which was not largely affected by how language was measured (i.e., through comprehension or production) or by study design (i.e., concurrent or longitudinal). However, the inherent ambiguity of correlational evidence requires a complementary approach, which we attempt in this paper. Under the assumption that early play is jointly determined with a competent other (Bruner, 1983; Göncü & Gaskins, 2011; Haight & Miller, 1993; Vygotsky, 1978), we analyze infant–caregiver interaction during symbolic and non-symbolic play and measure how notable features of interaction influence infants' language.

Central to this investigation is the theoretical analysis of early symbolic play as the first unambiguous instance of an infant's ability to engage in *collective intentionality* (Rakoczy, 2006, 2008; Tollefsen, 2005). For example, object substitution, such as pretending a block is an airplane, indicates

two, arguably uniquely human capacities. Firstly, it indicates that the child understands that, in this context, the block can act as a symbol for something else, thus demonstrating representational capacity (Leslie, 1987). Secondly, the joint action indicates a "meeting of the minds": The transformation is successful only insofar as each participant (implicitly) understands that their interlocutor has the same mental representation. Thus, symbolic play indicates an emerging socio-cognitive sophistication regarding the understanding of symbols and others. Indeed, because the context involves inherent ambiguity of reference (Searle, 1995; Sutton-Smith, 2001), it necessarily *depends* upon these skills.[1]

The hypothesis we pursue in the current paper is that symbolic play provides a fertile context for language development because its inherent ambiguity elicits behaviors that facilitate the negotiation of meaning (Bruner, 1983; Trawick-Smith, 1998). Evidence in support of this hypothesis comes from research that shows both socio-communicative behaviors and infant-directed speech (IDS) are different in symbolic play when compared to non-symbolic contexts in ways that may positively influence language development. With respect to socio-communicative development, symbolic play contexts elicit greater amounts of joint attention and gesture use in infant–caregiver dyads (Lillard & Witherington, 2004; Lillard et al., 2007; Nishida & Lillard, 2007; Quinn & Kidd, 2019), both of which are important for early language acquisition and infant communicative expression (Bates et al., 1979; Bruner, 1983; Tomasello & Farrar, 1986).

With respect to IDS, early work showed that infant–caregiver symbolic play is highly interactive, with caregivers using instances of pretence to draw infants into the interaction (Dunn & Wooding, 1977; Kavanagh et al., 1983). Similarly, Reissland (1998) reported that caregivers used different discursive styles in pretend versus non-pretend activities with their infants, showing that they used an *interactive* style in pretence (pretending to feed a doll) but an *instructive* style in non-pretence activities (learning to feed themselves with a spoon). Such styles are categorized by the use of distinct linguistic frames that, in the case of pretence, serve to draw infants into conversation, or in the case of real activities, direct behavior.

How does a caregiver adopting an interactive style typical of symbolic play influence language acquisition? We suggest that it may do so both by influencing the language caregivers' use (and do not use) and by the way in which language draws infants into the to and fro of communicative interaction. Within conversation, caregivers' use of language can either promote or hamper language acquisition (Furrow et al., 1979; Huttenlocher et al., 2007; Newport et al., 1977). Interrogatives (i.e., questions), in particular, have been found to promote language growth (Barnes et al., 1983; Furrow et al., 1979; Hoff-Ginsberg, 1985; Ninio, 1980; Rowe et al., 2017). One hypothesis is that language development improves when caregivers use questions because they require and often necessitate a response from the child (Rowe et al., 2017), who must then recall, organize, and/or express information (Cristofaro & Tamis-LeMonda, 2012; Tamis-LeMonda et al., 2001). Thus, infants may benefit from hearing interrogatives because they are interactionally demanding and draw infants into a conversational exchange, and because symbolic play is more interactive, infants are likely to hear more interrogatives in this context.

Conversational turn-taking has also been positively linked to language development over and above the quantity of input that children hear. In an analysis of conversational interactions between caregivers and 2-year-old infants from low-income families, Hirsh-Pasek et al. (2015) found that ratings of conversation fluency and connectedness, which incorporated turn-taking, significantly predicted language ability one year later. Converging evidence comes from studies that have used automated day-long recordings, where evidence suggests that conversational turn-taking, as measured by the LENA®

[1]We follow Rakoczy (2006, 2008) in using "collective intentionality," but note that the core of our point is that symbolic play requires an intersubjective connection between infant and interlocutor (which could just as easily be framed as joint or "shared" intentionality).

technology (Greenwood et al., 2011), positively predicts language proficiency (Donnelly & Kidd, 2021; Gilkerson et al., 2014, 2018; Romeo et al., 2018; for meta-analytic evidence see Wang et al., 2020). If symbolic play is more interactive, then it may naturally increase conversational turn-taking, which may also contribute to symbolic play's positive influence on language development.

Not all language promotes interaction. Notably, Newport et al. (1977) found that imperatives (e.g., *eat your dinner!*) in IDS were negatively associated with acquisition (see also Furrow et al., 1979), a finding they linked to the tendency for imperatives to refer to concepts outside the current referential frame. However, an alternative interpretation is that caregivers who use imperatives tend to engage their child in conversation less (McDonald, 1979; see Hoff-Ginsberg & Shatz, 1982), since a primary function of imperatives is behavioral control. Thus, imperatives serve to establish a hierarchical asymmetry between child and caregiver, and as such do not communicatively challenge language learners. This is interesting in the context of Reissland's (1998) distinction between interactive and instructive styles, since instructive styles involving the use of imperatives are more typical of realistic interactions.

1.2 | The current study

The aim of the present study was to investigate language use in infant–caregiver dyads during symbolic play and compare it with an equivalent but non-symbolic play context (i.e., functional play, defined below). In a longitudinal design, we asked infants and caregivers to play with two sets of toys across a 20-min play session (10 min with each toy set) when infants were aged, on average, 18 months, and again when they were aged 24 months. Unbeknownst to the participants, the toy sets were chosen to elicit either symbolic or functional play. We then analyzed the participants' language use, focusing predominantly on the dynamics of conversational interaction, as measured by turn-taking, and infants' and caregivers' speech.

Following the idea that symbolic play requires the establishment of collective intentionality, we expected to observe linguistic behaviors that reflect the greater necessity for the negotiation of meaning and thus draw infants into the interaction (Kavanagh et al., 1983; Reissland, 1998). Specifically, we expected a greater use of interrogatives in IDS in symbolic as compared to functional play, but in contrast expected fewer imperatives. Given the likely interactive nature of symbolic play, we also hypothesized that we would observe significantly greater conversational turns in symbolic versus functional play. In addition, there is some evidence that suggests that play contexts may foster greater complexity in older children's speech (Fekonja et al., 2005; Garvey & Kramer, 1989); thus, we also hypothesized that infants' language would be more grammatically complex in symbolic compared with functional play.

Finally, we measured the infants' language proficiency at each time point using the MacArthur–Bates Communicative Development Inventory (Fenson et al., 2007). In a series of regressions, we analyzed how the variables that we found to differ across play contexts were related to language outcomes, with the aim of identifying the individual contribution of those features that are prevalent in symbolic and functional play to development.

2 | METHOD

2.1 | Participants

The research was approved by the Human Ethics Committee at the Australian National University and followed the ethical guidelines of the National Health and Medical Research Council of Australia.

Fifty-four primary caregivers (50 mothers) and their biological infants (31 girls) were recruited through opportunity sampling in Canberra, Australia. Written informed consent was obtained from one caregiver for each child before data collection. At Time 1, the infants were, on average, 18 months of age (M = 18.32, SD = 0.98, range = 16.58–20.26 months). Fifty-two of the original 54 participated in another session 6 months later (30 girls, M = 24.29, SD = 1.01, range = 22.73–26.45 months; time between sessions M = 183.44 days, SD = 5.76, range = 159–192 days). Participants were recruited within this age range because it marks the point at which infants begin to (i) regularly engage in symbolic play (Fein, 1981; McCune, 1995; Nielsen & Dissanayake, 2004; Rubin & Howe, 1985), and (ii) rapidly acquire language after a preceding period of slow developmental gains (see Fenson et al., 1994). All infants were acquiring Australian English as their only language and were typically developing with no known or suspected developmental delay. The majority were firstborn (70%, n = 38), 67% did not have any siblings (n = 36), and 65% attended childcare (n = 35; $M_{days/week}$ = 1.73, $SD_{days/week}$ = 1.51). Socioeconomic status was estimated from caregiver education as high: 78% of mothers (n = 42) and 69% fathers (n = 37) had bachelor's degrees or higher. Ethnicity was not recorded; however, the sample was representative of the local population, which is predominantly of white, Anglo-Celtic origin (approx. 90%), and contains a range of other ethnicities based on different waves of migration since the mid-20th century (Australian Bureau of Statistics, 2016).

2.2 | Materials

The study was designed to create a naturalistic play setting in which infant–caregiver dyads engaged in symbolic play and a comparable but non-symbolic play context, which we call "functional" play. Play is influenced by the number and type of toys and materials available to children, the form of the toy (e.g., shape, size, material, complexity), and the child's knowledge of its function (Morrissey, 2014; Rubin & Howe, 1985). In order to elicit functional and symbolic play separately, two different sets of toys were selected (see Figure 1).

Toys were selected according to their similarity to toys used in past research investigating symbolic (Bigham & Bourchier-Sutton, 2007; Brown et al., 2001; Fekonja et al., 2005; Largo & Howard, 1979; O'Brien & Nagle, 1987; Taylor et al., 1993) or functional play (Fenson et al., 1976; Laplante et al., 2007), as well as in standardized measures of play (e.g., Test of Pretend Play; Lewis & Boucher, 1997). As gender-stereotyped toys influence the nature of the parent–child interaction irrespective of the gender of the parent or child (Caldera et al., 1989), toys were selected to be relatively gender-neutral.

Each condition contained four sets of toys that lent themselves to four different activities. In the symbolic play condition, the toys comprised both representational and non-representational toys. The representational toys included a set of cooking equipment (saucepan with its lid, a wooden spoon), a tea set (comprising a teapot, two teacups, a teaspoon), a toy mobile phone, and a set of non-representational toys (a piece of red cloth, a small yellow cylinder, and a small white cube). Finally, there was also a teddy bear. The representational toys were selected because toy household items (e.g., tea set, saucepan, spoons) tend to elicit symbolic play (e.g., pretending to cook spaghetti or drink tea), toy mobile phones tend to elicit pretend conversations (see Taylor et al., 1993), and the teddy bear is an item upon which infants may project these behaviors and conversations (e.g., feeding; Brown et al., 2001; Lewis & Boucher, 1997). Non-representational objects (e.g., piece of red cloth, small yellow cylinder, and small white cube) were selected because they are more abstract and do not immediately represent real-world artifacts, thus encouraging object substitution (e.g., the cloth is a "picnic rug," a "blanket," or a "cape" for teddy).

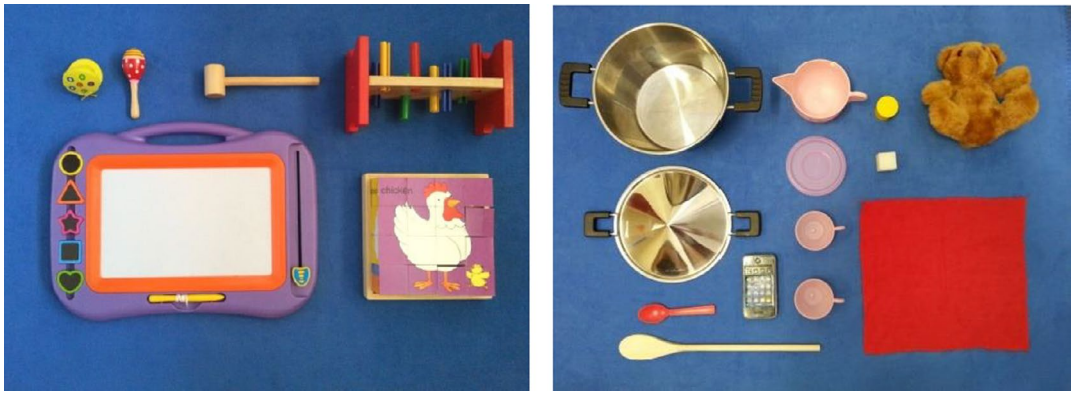

FIGURE 1 Functional (left) and symbolic (right) condition toys

Functional play was defined as object play during which the toy was used for its intended purpose in an *adult-defined* manner (Fenson et al., 1976; Laplante et al., 2007; Quinn & Kidd, 2019). These again lent themselves to four different activities, namely, drawing (a magnetic drawing board with magnetic stamps), hammering (a wooden peg and hammer set), puzzle completion (a wooden animal block puzzle containing 16 blocks; on each side of the block was a piece of a different puzzle, making one of 6 different animals when completed), and music (a wooden maraca and castanets). In this set, all toys do not immediately lend themselves to symbolic play. Instead, they are goal-oriented or "rule-based" (e.g., the castanet makes music, the stamps are for stamping on the drawing board).

As a manipulation check, infant–caregiver interactions were first assessed for the level of symbolic play in both play conditions at the two time points. The highest level of play of both infants and caregivers was recorded using the Pretend Play Observation Scale (Brown et al., 2001; see Appendix A, see also Quinn & Kidd, 2019). A subset of the play sessions was coded independently by two coders for reliability. At 18 months, 12 of 54 (22%) sessions were double-coded with 92% agreement. At 24 months, 10 of 52 (19%) were double-coded, with 93.4% agreement. At both time points, there were significantly greater levels of symbolic actions in the symbolic play as compared to the functional play condition, for both infants and caregivers (Time 1: infants: $M_{Symbolic} = 5.74$, SD = 2.19, $M_{Functional} = 0.19$, SD = 1.03, $t(53) = 17.80$, $p < .001$, $d = 3.59$, CI_{95} [2.98, 4.19]; caregivers: $M_{Symbolic} = 7.26$, $SD = 1.91$, $M_{Functional} = 0.52$, SD = 1.85, $t(53) = 15.78$, $p < .001$, $d = 3.24$, CI_{95} [2.67, 3.82]; and Time 2: infants: $M_{Symbolic} = 8.05$, SD = 1.22, $M_{Functional} = 1.21$, SD = 1.5, $t(51) = 25.86$, $p < .001$, $d = 3.59$, CI_{95} [2.84, 4.33]; caregivers: $M_{Symbolic} = 8.77$, SD = 1.31, $M_{Functional} = 1.75$, SD = 1.1, $t(51) = 30.44$, $p < .001$, $d = 4.22$, CI_{95} [3.36, 5.08]). Thus, the toys used in the symbolic play condition elicited higher levels of symbolic play from infants and their caregivers than did the toys in the functional play condition. Note that we cannot claim that dyads *never* engaged in symbolic behavior during functional play, although the low means in the functional condition suggest symbolic acts were exceedingly rare.

2.3 | Procedure

Testing sessions were conducted in the dyad's home. Caregivers were asked to sit on a play mat and play with their infants as they normally would for 20 min. Unlike in previous studies (Lillard & Witherington, 2004), caregivers were not primed to engage in pretence, thus ensuring that play was spontaneous and ecologically valid. Both the functional and the symbolic play conditions were presented consecutively as a continuous play session. Each dyad was randomly assigned their first

set of toys at each session, which they played with for approximately 10 min, at which point the experimenter exchanged the set of toys. Each play session lasted approximately 20 min. At 18 months, play conditions (i.e., functional or symbolic) averaged 10 min (M = 602s, SD = 52.62s) and ranged in duration from 7 min and 30 s to 12 min and 23 s. At 24 months, play conditions averaged just under 11 min (M = 648s, SD = 52.62s) and ranged in duration between 9 and 12 min and 45 s. Although they are referred to as separate play conditions in this paper, caregivers were not aware that the conditions were distinct. The session allowed for breaks as required. All sessions were video-recorded for later transcription and coding. The 18-month sessions were transcribed into CHAT format using Child Language Analysis (CLAN) software (MacWhinney, 2013). The 24-month sessions were transcribed into ELAN linguistic annotation software (The Language Archive, Max Planck Institute for Psycholinguistic, 2016), which has complete interoperability with CLAN to allow for quantitative analysis.

2.4 | Coding

Infant–caregiver interactions were coded for the following variables.

2.4.1 | IDS

In line with past research (Cameron-Faulkner et al., 2003; Graf Estes et al., 2016; Ninio & Snow, 1999), IDS was analyzed at the level of the sentence type. Firstly, four types of utterance were distinguished that differ in their function: *declaratives*, *interrogatives*, *exclamatives*, and *imperative*s (Hoff-Ginsberg, 1985; Ninio & Snow, 1999; Quinn, 2016; Wu & Gros-Louis, 2014; see Table 1). A declarative is a statement (e.g., *I am drawing a house*). Imperatives express an order and, within IDS, often serve as attention-getters or commands for a specific behavior (e.g., *Draw a house*). Exclamatives give emphasis to specific statements and tend to convey emotion (e.g., *Well done!*). Interrogatives, or questions, require a response from the conversation partner; they aim to elicit additional information from them (e.g., *Did you draw a house?*; Huddleston & Pullum, 2002). Interrogatives were further divided into: (a) *wh* questions: open-ended questions requiring complex and constructive responses and which in English begin with one of nine interrogative words: *who*, *whom*, *which*, *whose*, *what*, *where*, *when*, *why*, and *how* (Ninio & Snow, 1999); and (b) *yes–no* (Y/N) questions: close-ended questions that require a binary single-word response, usually either *yes* or *no*. Y/N questions can be subdivided further into canonical yes–no questions or tag questions, in which a declarative or an imperative statement is directly followed by an interrogative fragment that turns the original statement into a question (e.g., *It's a phone*, *isn't it?*). We only analyzed canonical Y/N questions, since the status of tag questions as true questions is unclear; they are not typically used to elicit information unknown to the speaker, but are instead used to negotiate common ground between speakers, often due to some level of uncertainty (e.g., *This is a cup of tea*, *isn't it? You like tea*, *don't you?* vs. *Is this a cup of tea? Do you like tea?*; see Huddleston & Pullum, 2002).

In addition to the four initial sentence types, three additional categories were used. An utterance was coded as *naming* when it referred solely to the label of an object or its parts (e.g., *It's a block*; Ninio, 1980; Ninio & Bruner, 1978; Wu & Gros-Louis, 2014). *Mimetics* (sound symbolic utterances) were also coded; these are utterances mimicking the sounds associated with the referent object (e.g., iconic sounds such as "bang," "crash"; symbolic sounds such as "numnum" to mimic eating; and animal sounds such as "bak bak," "woof woof"). Utterances that did not fit in any of the categories

TABLE 1 Types of utterances in infant-directed speech

Classification	Description	Example
Declaratives	Statements that convey information; requires a verb or implied verb	*Let's bake a cake* *You drew a cat*
Interrogatives		
Wh questions	Questions beginning with what/where/when/who/why/how, etc.	*Who are you calling?* *What should we do next?* *Where is the castanet?* *How are you?*
Yes/no questions	Questions eliciting a yes or no response	*Is this a cape?* *Should we have a cup of tea?* *Will you call daddy?*
Tag questions	Declarative, imperative, or naming statement ending with an interrogative tag	*It's a phone, isn't it?* *You like pasta, don't you?* *Teddy has had enough, has he?*
Imperatives	Directives, attempts to direct infant's attention or to perform an action	*Look at this* *Come back* *Stop*
Exclamatives	Statements that emphasize an idea typically with emotive meaning	*Delicious tea!* *Great job!* *Awesome!*
Naming	Identifying an object or feature of an object	*That's a puzzle* *It is yellow*
Mimetics	Words forms/sounds that mimic and/or symbolize the sound associated with the referent (e.g., animal, iconic, or pretense sound)	*Slurp* *Bang* *Crash* *Woof woof*
Other	Incomplete/inaudible sentences Social routines Discourse particles	*Unintelligible* *Thank you, bye, yum* *Oh, hey, oi, eh, hey*

described above were coded as *other*. This category included incomplete, inaudible, or unintelligible utterances, as well as conventional social routines (e.g., *yes*, *no*, *thank-you*, *hi*), singing, and routines or games involving language (e.g., peek-a-boo).

The frequency of each classification was coded manually and converted into a proportion of the total utterances produced by the caregiver per condition. To ensure reliability, 6 dyads were randomly chosen at each time point (11.11% of 18-month data; 11.5% of 24-month data) and their transcripts were recoded by a second coder. The overall agreement was almost perfect at both time points (18 months: 96.6% agreement, $\kappa = .93$ (SE$\kappa = 0.01$, CI95, κ [0.91, 0.95]); 24 months: 95.1% agreement, $\kappa = .91$ (SE$\kappa = 0.01$, CI95, κ [0.88, 0.92])).

2.4.2 | Conversational turns (CTs)

The frequency of turn-taking during each play condition was used as a measure of interactional complexity (Golinkoff & Ames, 1979; Ninio & Snow, 1999). A conversational turn was defined as "the time during which a single participant speaks, within a typical, orderly arrangement in which

participants speak with minimal overlap and gap between them" (Levinson, 1983, pp. 295–296). At this stage of development, expressive language is limited. For this reason, turns including nonwords and other communicative vocalizations were also included in the analysis. Nonverbal vocalizations such as moaning or sighing as a physical consequence of the infant's action (for instance, "ugh") were excluded, as were as gestures, because it was not always possible to easily determine whether such expressions were meaningful (i.e., communicative).

CT number was calculated automatically using the MLT function in CLAN. A turn was counted only when a new speaker took the floor. For example, consider the following exchange:

MOT: what's next?
MOT: oh, you've got Teddy.
CHI: yeah.

This exchange contains one CT, where the floor changed between the mother (MOT) and infant (CHI); the within-turn pause between the mother's two utterances not counting as a separate turn. Thus, the CT measure can be taken as an index of how often the conversation switched between infant and caregiver. We did not require that a speaker take the floor within a predetermined length of time (e.g., 2 or 5 s) in order to count the exchange in speakers as a turn. Developmental research shows that turn length time varies across a number of dimensions, including infant age, language proficiency, and the complexity of the utterance being produced (Casillas et al., 2016; Hillbrink et al., 2015). While the time frame in which a CT is defined has been shown to have some influence on aspects of interaction within turns in younger infants (Pretzer et al., 2019), our focus was not on what happened within turns. Regardless, we found that turn gaps were fairly short and comparable to similar past work. For example, an analysis of gaps at the 24-month stage showed that the average gap was 630.93 ms (median = 627.62 ms, SD = 166.33 ms), which is consistent with results reported by Hillbrink et al. (2015).

2.4.3 | Mean length of utterances (MLU)

MLU served as a measure of grammatical complexity for both infants and caregivers. It was calculated by dividing the total number of morphemes by the total number of utterances produced (Brown, 1973), using CLAN. Morphemes were counted when an utterance was both complete and intelligible. Irregular plurals and past tense and past participle verbs (e.g., *children*, *was*, *eaten*) and diminutives (e.g., *blankie*) were counted as one morpheme, because they were assumed to be stored as a whole word. MLU is typically a measure of grammatical development in early language acquisition, but caregiver MLU was also calculated in this analysis to measure language complexity in IDS and to compare it across different contexts.

2.4.4 | Type–token ratio (TTR)

Type–token ratio is a measure of flexibility and variability in vocabulary use. It was calculated by dividing the total number of unique words by the total number of words used by each speaker in the dyad, using CLAN. The ratio is reported as a number between 0 and 1. We note that TTR is particularly sensitive to important variables such as context of use (i.e., different language learning contexts

use different types) and how much language children hear (i.e., corpus size; Montag et al., 2018). We report TTR here to see whether there is equivalence in vocabulary use across contexts.[2]

2.4.5 | MB-CDI

The MacArthur–Bates Communicative Development Inventories (MB-CDI; Fenson et al., 2007) were used to provide a parental-report measure of the infant's language proficiency. At 18 months, parents completed the *Words and Gestures* form, which measures both infants' comprehension and production of words, in addition to their inventory of communicative and symbolic gestures. The form is standardized for American infants aged 8–18 months. Some minor changes in wording were used to reflect dialect differences (e.g., "biscuit" for "cookie"), but otherwise, the form was used as per the standardized instructions. From this form, we used total vocabulary comprehension, production, and gesture production as variables in our analyses. At 24 months, parents completed the *Words and Sentences* form, which measures children's productive vocabulary and emerging grammatical knowledge. The form is standardized for American children aged 16–30 months. Once again, minor changes in wording were made to reflect differences between American and Australian English. From this form, we used the total vocabulary production score and total grammatical complexity scores in our analyses, the latter giving us an estimate of children's emerging morphosyntactic knowledge. Both MB-CDI forms show excellent validity and reliability (see Fenson et al., 2007).

3 | RESULTS

Our data and analyses are publicly available on the Open Science Framework (https://osf.io/fnd58/). Our first set of analyses tested whether there were differences in language and communicative behavior across the two play contexts. In analytical terms, this aim is complicated by the fact that many of the measures are almost certainly interdependent. For instance, for every *wh* question produced in the finite time of the play session there is less chance of an imperative being produced. For this reason, for each measure we report pairwise *t* tests alongside their effect sizes and confidence intervals, the latter providing an unbiased and standardized measure of the magnitude of the observed effects. This approach is consistent with recommendations made by researchers who criticize null hypothesis significance testing (Nickerson, 2000; Schmidt & Hunter, 2002), and is recommended in fields such as behavioral ecology, where multiple comparisons are regularly made on interdependent measures (Garamszegi, 2006, 2016; Nakagawa, 2004). The difference between two means in terms of the size of the standard deviation is reported using Cohen's *d*.[3] We report analyses of the variables for which we have a priori hypotheses and the variables we coded but for which we had no hypotheses. We did this in the interest of having a full understanding of the linguistic ecology of symbolic versus non-symbolic play; knowing where the contexts do and do not differ is of interest beyond our specific hypotheses.

[2]To preface our results, we find no difference in TTR for either caregivers or infants across the two time points. However, we did find strong evidence that infants used more types and tokens in symbolic play at 24 but not 18 months (see OSF materials), which we interpret to derive from the fact that the infants play a greater role in interaction in this context. The caregivers never differed across contexts in their use of types and tokens.

[3]Following Cohen (1988, 1992), an effect size of less than or equal to 0.2 was considered small, an effect size of greater than 0.2 and less than or equal to 0.5 was considered medium, and an effect size greater than 0.5 was considered large.

TABLE 2 Descriptive statistics and pairwise comparisons for each dependent measure at 18-month time point

	Hypothesis	M_{sym} (SD)	M_{Fnl} (SD)	*t*	*df*	*p*	*d*	CI_{95} (*d*)	BF	$BF_{interpret}$
CTs	S > F	42.13 (21.58)	33.22 (19)	4.899	53	<.001	0.667	[0.416, ∞]	3959.46	Extreme (A)
YN-Qs	S > F	0.143 (0.051)	0.099 (0.049)	5.784	53	<.001	0.787	[0.527, ∞]	76,392.71	Extreme (A)
Wh-Qs	S > F	0.232 (0.067)	0.199 (0.069)	3.998	53	<.001	0.544	[0.302, ∞]	239.73	Extreme (A)
Imperatives	S < F	0.07 (0.042)	0.127 (0.063)	−7.3	53	<.001	−0.993	[−∞, −0.715]	$1.53e^{7}$	Extreme (A)
Other IDS										
Declaratives	N/A	0.145 (0.054)	0.171 (0.06)	−4.091	53	<.001	−0.557	[−0.841, −0.267]	157.66	Extreme (A)
Exclamatives	N/A	0.029 (0.022)	0.031 (0.026)	−0.588	53	.559	−0.08	[−0.347, 0.187]	0.175	Moderate (N)
Mimetics	N/A	0.051 (0.047)	0.037 (0.033)	2.069	53	.043	0.282	[0.008, 0.552]	1.061	Anecdotal (A)
Naming	N/A	0.084 (0.064)	0.095 (0.056)	−1.263	53	.212	−0.172	[−0.44, 0.098]	0.314	Moderate (N)
MLU_{parent}	N/A	4.1 (0.686)	4.206 (0.697)	−1.649	53	.105	−0.224	[−0.493, 0.047]	0.527	Anecdotal (N)
TTR_{parent}	N/A	0.074 (0.015)	0.079 (0.022)	−1.738	53	.088	−0.237	[−0.506, 0.035]	0.604	Anecdotal (N)
Other child										
MLU_{child}	S > F	0.934 (0.342)	0.844 (0.415)	1.979	53	.027	0.269	[0.04, ∞]	1.751	Anecdotal (A)
TTR_{child}	N/A	0.34 (0.311)	0.31 (0.289)	0.638	53	.527	0.087	[−0.181, 0.354]	0.180	Moderate (N)

Note: $N = 54$.

Abbreviations: BF, Bayes factor; $BF_{interpret}$, interpretation of BF; Extreme (A/N), extreme evidence in favor of the alternative/null hypothesis; F, functional play condition; S, symbolic play condition; TTR, type–token ratio.

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES —WILEY

TABLE 3 Descriptive statistics and pairwise comparisons for each dependent measure at 24-month time point

	Hypothesis	M_{sym} (SD)	M_{Fnl} (SD)	*t*	*df*	*p*	*d*	$CI_{95}(d)$	BF	$BF_{interpret}$
CTs	S > F	84.33 (21.61)	73.56 (30.54)	2.85	51	.003	0.396	[0.157, ∞]	11.11	Strong (A)
Wh-Qs	S > F	0.125 (0.049)	0.107 (0.056)	2.27	51	.014	0.315	[0.080, ∞]	3.10	Moderate (A)
YN-Qs	S > F	0.264 (0.083)	0.205 (0.081)	4.48	51	<.001	0.621	[0.369, ∞]	1000.19	Extreme (A)
Imperatives	S < F	0.041 (0.039)	0.080 (0.047)	−8.08	51	<.001	−1.12	[−∞, −0.824]	$1.85e^{8}$	Extreme (A)
Other IDS										
Declaratives	N/A	0.292 (0.086)	0.292 (0.106)	0.014	51	.989	0.002	[−0.27, 0.27]	0.151	Moderate (N)
Exclamatives	N/A	0.074 (0.065)	0.082 (0.053)	−0.995	51	.324	−0.138	[−0.41, 0.136]	0.241	Moderate (N)
Mimetics	N/A	0.062 (0.045)	0.03 (0.03)	4.151	51	<.001	0.576	[0.279, 0.867]	183.63	Extreme (A)
Naming	N/A	0.023 (0.018)	0.058 (0.03)	−8.664	51	<.001	−1.201	[−1.56, −0.84]	$6.84e^{8}$	Extreme (A)
MLU_{parent}	N/A	4.54 (0.55)	4.44 (0.73)	1.13	51	.264	0.157	[−0.12, 0.429]	0.276	Moderate (N)
TTR_{parent}	N/A	0.29 (0.243)	0.289 (0.266)	0.01	51	.922	0.01	[−0.27, 0.273]	0.151	Moderate (N)
Other child										
MLU_{child}	S > F	1.967 (0.609)	1.87 (0.471)	2.1	51	.020	0.291	[0.057, ∞]	2.22	Anecdotal (A)
TTR_{child}	N/A	0.409 (0.075)	0.416 (0.087)	−0.525	51	.602	−0.073	[−0.345, 0.2]	0.172	Moderate (N)

Note: $N = 52$.

Abbreviations: BF, Bayes factor; $BF_{interpret}$, interpretation of BF; Extreme (A/N), extreme evidence in favor of the alternative/null hypothesis; F, functional play condition; S, symbolic play condition; TTR, type–token ratio.

We also report Bayes factors (BF_{10}) derived from Bayesian paired sample *t* tests (using a default Cauchy prior = 0.707), which allow us to determine the evidence in favor of the null and alternative hypotheses. We report BF_{10} from directional tests where we had specific hypotheses, and BF_{10} from non-directional tests where we did not have hypotheses. Values greater than 1 provide evidence in favor of a difference across the conditions (see tables for interpretations, which follow Anandraszewicz et al. (2015)), and values less than 1 provide evidence in favor of no difference (i.e., the null hypothesis). All pairwise comparisons were computed in JASP v.0.12.1 (JASP Team, 2018).

3.1 | Time 1 results (18 months)

Table 2 presents the descriptive statistics and pairwise comparisons for all variables at 18 months. There were several differences across the contexts. With respect to our hypothesized differences, we saw that, as predicted, CTs and interrogatives (both *wh* and Y/N questions) were more frequent in symbolic play, whereas imperatives were more frequent in functional play. While infants' MLU was significantly higher in symbolic play than in functional play (with a medium effect size), the evidence in favor of the alternative hypothesis was anecdotal. There were also some differences that were not predicted. Notably, caregivers produced significantly more declaratives in functional play and significantly more mimetics in symbolic play, although the evidence in favor of a difference for the latter was anecdotal.

3.2 | Time 2 results (24 months)

We repeated the same analyses for Time 2 (see Table 3). Our hypothesized differences replicated at Time 2: There were more CTs and interrogatives in symbolic play and more imperatives in functional play. Infants' MLU was also significantly higher in symbolic play, although the evidence for the difference was again anecdotal. There were once again significantly more mimetics in symbolic play, and at this time point, caregivers produced more naming in the functional play condition.

3.3 | Interim discussion

Across both time points, we found evidence in support of our hypotheses. Consistent with symbolic play inducing an interactive conversational style (Reissland, 1998), we found that caregivers used significantly more interrogatives (i.e., *wh* and Y/N questions) with their infants in this context when compared to functional play. Additionally, we also observed a significantly greater number of conversational turns between infant–caregiver dyads. Finally, consistent with the idea that functional play would constitute an instructional and therefore less interactive context, we found that caregivers produced significantly more imperatives in functional when compared to symbolic play. Overall, these results are consistent with our argument that, because symbolic play requires the establishment of collective intentionality (Rakoczy, 2006, 2008), it elicits a rich set of socio-communicative behaviors that enable the negotiation of meaning (Quinn & Kidd, 2019; Reissland, 1998). We discuss these core results further in the main Discussion.

We observed some additional results that we did not predict. Firstly, caregivers produced significantly more mimetics in symbolic play than in functional play. Past research has reported that mimetics are used frequently during free play (Fekonja et al., 2005) and are used by caregivers to signal

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES — WILEY

TABLE 4 Bivariate correlations between 18-month measures of sentence types in infant-directed speech, conversational turns, word tokens and types, and language outcome measures at 18 and 24 months (significant correlations are in bold font).

		CTs (r_s)	Tokens	Types	Wh-Qs	YN-Qs	Dec.	Imp.	Mim. (r_s)
		M = 75.35 (38.4)	M = 626.93 (150.97)	M = 45.22 (5.32)	M = 0.12 (0.04)	M = 0.22 (0.06)	M = 0.16 (0.05)	M = 0.10 (0.04)	M = 0.04 (0.03)
18 months (N = 54)									
Gestures	M = 44.38 (9.16)	.189	.075	−.035	.098	.073	.043	−.198	**.370****
Vocab. Comp.	M = 224.73 (83.62)	**.414****	.126	−.005	.217	.083	−.124	−.174	**.313****
Vocab. Prod (r_s)	M = 75.5 (70.84)	**.471 ** **	.120	.101	.105	−.018	−.099	**−.342***	**.370****
24 months (N = 52)									
Vocab Prod.	M = 359.56 (155.24)	**.501****	.236	.226	.069	.117	−.004	**−.427****	**.245**#**
Grammatical Complexity (r_s)	M = 11.46 (9.99)	**.503****	.056	.112	.175	**.307***	−.048	**−.489****	.079
MLU (r_s)	M = 1.92 (0.52)	**.389***	.155	**.38****	−.140	.206	.036	**−.418****	.107

Note: #.10 < p < .05, * p < .05, **p < .01.

Abbreviations: CTs, number of conversational turns; Dec., proportion of declaratives; Imp., proportion of imperatives; M, mean (standard deviation in brackets); Mim., proportion of mimetics; r_s, Spearman's rank–order correlation; Tokens, number of tokens produced by caregivers at 18 months; Types, number of types produced by caregivers at 18 month; Wh-Qs, proportion of *wh* questions; Y/N-Qs, proportion of yes–no questions.

pretence (Lillard & Witherington, 2004). We did not expect to see this difference in our data, because we thought that our toy-based manipulation of context might be too subtle (i.e., both contexts were free play, and the notion of pretence was not explicitly primed). Mimetics may serve to cue children to the pretend and thus symbolic nature of the interaction, in much the same way as do other social referencing cues such as eye gaze, smiling, and modulations in pitch height (Nishida & Lillard, 2007; Reissland & Snow, 1996). Motamedi et al. (2021) presented data to suggest that mimetic devices such as onomatopoeia, where a word's phonological form evokes the sound properties of its meaning (e.g., *bang*), provide a crucial link between language and sensory experience, enabling young infants to use iconicity to break into the linguistic system (see also Laing, 2019). Their use in symbolic play, along with other linguistic features such as reduplication (e.g., *bang! bang!* and *woof woof*; see Ota et al., 2018), may thus have an influence on language development, which we explore in our regression analyses.

3.4 | Regression analyses

We next conducted a series of correlational and regression analyses that investigated how the variables that differentiated between symbolic and functional play were related to the infants' language both concurrently and longitudinally. Our strategy was as follows: (i) We correlated those variables that differentiated symbolic from functional play at 18 months with our MB-CDI measures at 18 (gestures, vocabulary comprehension, and production) and 24 (vocabulary production and grammatical complexity score) months, as well as children's MLU at 24 months, as measured within the entire play session.[4] Note that the measures we take from the play session represent the overall proportion of use of particular utterance types, rather than from either the symbolic or functional play components of the session. The logic behind this is that, because the behavior is defined at the functional level, their effect on language should not vary across contexts (i.e., a *wh* question is the same regardless of the context in which it is used, even if its frequency of use varies across contexts). We also included the number of types and tokens produced by caregivers at 18 months. MLU represents the total MLU of the infant at 24 months across both the functional and symbolic play contexts.

Table 4 shows the correlational matrix. Several variables deviated from a normal distribution; in those cases, Spearman rank–order correlations are reported.

Table 4 shows several trends in the data. Firstly, we note that CTs were associated both concurrently and longitudinally with vocabulary and longitudinally with grammatical knowledge. Secondly, several properties of IDS were also associated with language outcomes. At 18 months, the proportion of mimetics in IDS was significantly and positively associated with both vocabulary comprehension and production. The proportion of imperatives in IDS was negatively associated with vocabulary production both concurrently and longitudinally. With respect to grammar, the proportion of Y/N questions in IDS was positively associated with 24-month MB-CDI grammatical complexity score, whereas the number of types in caregiver speech at 18 months was positively associated with 24-month MLU only. The proportion of imperatives in IDS was negatively associated with 24-month vocabulary production and grammatical proficiency.

We next report on a series of regressions that modeled vocabulary and grammatical knowledge both concurrently and longitudinally. We report on separate regressions that predict language outcomes from (a) CTs and (b) IDS, for the following reasons. Firstly, whereas IDS represents the speech directed to infants, CTs represent the dynamic *interplay* between infant and caregiver. Secondly,

[4]We thus follow recent recommendations to use multiple outcome measures in infant research (LoBue et al., 2020).

modeling these separately allows us to compare our data with past research on both topics (Donnelly & Kidd, 2021; Gilkerson et al., 2018; Newport et al., 1977; Rowe et al., 2017; Zimmerman et al., 2009). Finally, doing so ensured that we did not overcomplicate our models with too many predictors. We used generalized linear models (GLMs) estimated using maximum-likelihood estimation in SPSS (Mac version; build 1.0.0.1347). We use GLMs because several of the outcome measures were not normally distributed, and the technique provides more flexibility than commonly used ordinary least squares regression, allowing the specification of different distribution types and link functions.

3.5 | Predicting concurrent vocabulary knowledge

We report here the results of GLMs predicting Time 1 infant vocabulary comprehension and production from (a) CTs, and (b) IDS: mimetics and imperatives, the latter only for the analysis predicting vocabulary production. We include both infants' age at Time 1 and their MB-CDI gesture score as additional variables because (a) the infants varied in their age enough to justify controlling for any additional variance associated with it, and (b) gesture was significantly correlated with both 18-month comprehension ($r = .56$, $p < .001$) and positively correlated with production ($r_s = .19$, $p = .17$). For the CT analyses, we included the number of caregiver tokens to control for caregiver talkativeness because caregiver tokens and CTs at 18 months were weakly correlated ($r_s = .261$, $p = .057$). For the IDS analyses, differences in caregiver speech are accounted for because each variable was computed as proportion of total caregiver utterances.[5]

3.5.1 | Vocabulary comprehension at 18 months

Vocabulary comprehension at 18 months was normally distributed (Shapiro–Wilk = 0.984, $df = 54$, $p = .69$). We thus ran two generalized linear models specifying a normal distribution with an identity link function. For CTs, the fitted model was significantly different from an intercept-only model ($\chi^2 = 29.72$, $df = 4$, $p < .001$), as was the IDS model ($\chi^2 = 25.16$, $df = 3$, $p < .001$). Table 5 shows that both infants' gesture repertoire and the number of CTs independently and significantly predicted their comprehension, but that the proportion of mimetics did not predict vocabulary comprehension in the IDS analysis.

3.5.2 | Vocabulary production at 18 months

Vocabulary production was not normally distributed (Shapiro–Wilk = 0.791, $df = 54$, $p < .001$), showing a right skew that is typical of this age group. Two cases deviated considerably from the rest of the sample, having standardized residuals of $z = 2.957$ and $z = 4.496$. We modeled the data using a gamma distribution with a log-link function to account for the skewness in the data.[6] We ran separate models with and without the outliers, and although removing one or both of them increased model

[5]Additional analyses showed that caregiver tokens were not associated with any outcome variables.

[6]Since scores on the MB-CDI could also be considered count data, we also ran models specifying a negative binomial distribution for this analysis and for the analyses of grammatical complexity. For 18-month productive vocabulary, the gamma models fit better, as measured by AIC and Bayes Information Factor. For 24-month MB-CDI grammatical complexity, the fit was equivalent, and the pattern of results did not change. See OSF materials for comparison models.

TABLE 5 Model parameters of generalized linear models predicting 18-month vocabulary comprehension

	B	SE (*B*)	CI_{95} (Wald)	Wald χ^2	*p*
CT model[a]					
Intercept	−180.91	168.84	[−511.83, 150.01]	1.15	.284
Age	9.31	9.13	[−8.58, 27.19]	1.04	.308
Gesture	4.25	0.98	[2.34, 6.17]	18.98	<.001
Tokens	−0.007	0.06	[−0.12, 0.11]	0.015	.902
CTs	0.68	0.25	[0.19, 1.6]	7.33	.007
Scale	3957.59	761.64	[2714.05, 5770.92]		
IDS model[b]					
Intercept	−327.99	173.14	[−667.33, 11.36]	3.59	.06
Age	18.45	9.37	[0.09, 36.81]	3.88	.049
Gesture	4.34	1.05	[2.29, 6.39]	17.14	<.001
Mimetics	510.15	309.55	[−96.55, 1116.86]	2.72	.099
Scale	4306.12	828.71	[2953.05, 6279.12]		

[a]Log-likelihood = −300.274, Akaike's information criterion (AIC) = 612.548.

[b]Log-likelihood = −302.553, Akaike's information criterion (AIC) = 615.106.

fit, it did not change the qualitative pattern of results; we therefore report the results from the full dataset. For CTs, the fitted model was significantly different from an intercept-only model (χ^2 = 21.33, $df = 4$, $p < .001$), as was the IDS model ($\chi^2 = 19.658$, $df = 4$, $p = .001$), although Table 6 shows that only CTs and the proportion of mimetics predicted vocabulary production over and above control variables.

3.6 | Longitudinal analyses

We next report on the results of three analyses predicting infants' longitudinal language outcomes at 24 months from the 18-month data.

3.6.1 | Vocabulary production at 24 months

Modeling vocabulary development longitudinally is complicated by the fact that development in this age range is non-linear (for a recent demonstration, see Donnelly & Kidd, 2020). This was also the case in our data (see Appendix B). In order to account for the non-linearity, we first modeled 24-month vocabulary production from 18-month vocabulary production using linear and high-order polynomial terms. Vocabulary production at 24 months was normally distributed (Shapiro–Wilk = 0.976, df = 52, p = .388), so we specified a normal distribution and an identity link function. A model containing linear and quadratic terms for 18-month vocabulary production fit better than a model containing only the linear term (AIC linear-only model = 649.811; AIC linear + quadratic = 645.65), with both fixed effects significantly predicting 24-month vocabulary (linear term: $B = 2.68$, SE(B) = 0.547, χ^2 = 24.04, $p < .001$; quadratic term: $B = -0.004$, SE(B) = 0.0016, $\chi^2 = 6.55$, $p = .011$). However, the addition of a cubic term did not further improve model fit (AIC = 644.81, cubic term: $B = 3.26e^{-5}$, SE(B) = $1.91e^{-5}$, $\chi^2 = 2.92$, $p = .087$). We thus modeled the data with linear and quadratic terms. The

TABLE 6 Model parameters of generalized linear models predicting 18-month vocabulary production

	B	SE (*B*)	CI_{95} (Wald)	Wald χ^2	*p*
CT model[a]					
Intercept	−0.832	2.282	[−5.31, 3.64]	0.13	.715
Age	0.196	0.122	[−0.04, 0.43]	2.60	.107
Gesture	0.019	0.014	[−0.01, 0.05]	1.78	.182
Tokens	0.000	0.001	[−0.002, 0.001]	0.24	.623
CTs	0.011	0.003	[0.004, 0.017]	10.81	.001
Scale	0.555	0.099	[0.392, 0.786]		
IDS model[b]					
Intercept	−2.51	2.28	[−6.97, 1.95]	1.21	.270
Age	0.33	0.12	[0.09, 0.56]	7.35	.007
Gesture	0.02	0.014	[−0.007, 0.046]	2.10	.148
Mimetics	8.15	3.94	[0.43, 15.86]	4.29	.038
Imperatives	−4.94	2.66	[−10.16, 0.28]	3.45	.063
Scale	0.57	0.101	[0.403, 0.807]		

[a]Log-likelihood = −275.788, Akaike's information criterion (AIC) = 563.575.

[b]Log-likelihood = −276.624, Akaike's information criterion (AIC) = 565.247.

CT model was significantly different from an intercept-only model ($\chi^2 = 42.63$, $df = 5$, $p < .001$), as was the IDS model ($\chi^2 = 38.68$, dif = 5, $p < .001$). Table 7 shows that CTs at 18 months significantly positively predicted infant vocabulary production at 24 months over and above all control variables, whereas the proportion of imperatives in IDS was a negative predictor.

3.6.2 | Regressions predicting grammatical knowledge at 24 months

Just as early vocabulary development is non-linear, so is the relationship between vocabulary development and early grammatical knowledge (Bates & Goodman, 1997). This was the case in the current dataset (see Appendix B). Therefore, we first modeled the relationship between vocabulary and grammar using higher order polynomial terms. The outcome measures were not normally distributed and right-skewed, so we modeled the data using a gamma distribution with a log-link function (Shapiro–Wilk tests: MB-CDI grammatical complexity: Shapiro–Wilk = 0.915, $df = 52$, $p < .001$; MLU (24 months) = 0.952, $df = 52$, $p = .035$). For grammatical complexity as measured by the MB-CDI, the model containing a linear and quadratic term fit the data better than the model containing only a linear term (AIC linear-only model = 357.97; AIC linear + quadratic = 355.13), with both fixed effects significantly predicting MB-CDI grammatical complexity (linear term: $B = 0.015$, SE(B) = 0.004, $\chi^2 = 14.14$, $p < .001$; quadratic term: $B = -2.72e^{-5}$, SE(B) = $1.11e^{-5}$, $\chi^2 = 5.91$, $p = .015$).[7] The addition of a cubic term did not further improve model fit (AIC = 355.92, cubic term: $B = 1.53e^{-7}$, SE(B) =$1.34e^{-7}$, $\chi^2 = 1.3$, $p = .254$). We thus modeled the data with linear and quadratic terms. The CT model was significantly different from an intercept-only model ($\chi^2 = 25.413$, df =5, $p < .001$), as was the IDS model ($\chi^2 = 30.071$, $df = 5$, $p < .001$). Table 8 shows that CTs at 18 months

[7]Note that a constant of 1 was added to all infants' MB-CDI grammatical complexity scores because some infants had zero scores at this age and the gamma distribution requires non-zero values.

TABLE 7 Model parameters of generalized linear models predicting 24-month vocabulary production

	B	SE (*B*)	CI_{95} (Wald)	Wald χ^2	*p*
CT model[a]					
Intercept	212.924	284.597	[−344.88, 770.723]	0.56	.454
Age	−9.376	16.174	[−41.08, 22.32]	0.336	.562
Vocab 18mths	2.501	0.534	[1.46, 3.55]	21.98	<.001
Vocab $18mths^2$	−0.004	0.002	[−0.007, −0.002]	8.84	.03
Tokens	0.14	0.10	[−0.05, 0.34]	2.04	.154
CTs	1.068	0.453	[0.18, 1.96]	5.57	.018
Scale	10,412.304	2042.021	[7089.45, 15,292.60]		
IDS model[b]					
Intercept	356.23	325.86	[−282.45, 994.91]	1.20	.274
Age	−3.52	17.19	[−37.22, 30.18]	0.04	.838
Vocab 18mths	2.28	0.57	[1.16, 3.40]	15.87	<.001
Vocab $18mths^2$	−0.003	0.002	[−0.006, $-6.43e^{-5}$]	4.00	.045
Mimetics	168.81	531.95	[−873.79, 1211.4]	0.101	.751
Imperatives	−852.38	400.05	[−1636.47, −68.30]	4.54	.033
Scale	11,234.95	2203.36	[7649.57, 16,500.82]		

[a]Log-likelihood = −314.304, Akaike's information criterion (AIC) = 642.608.

[b]Log-likelihood = −316.281, Akaike's information criterion (AIC) = 646.562.

significantly positively predicted infants' 24-month grammatical complexity score over and above all control variables, whereas the proportion of imperatives in IDS was a negative predictor.

For MLU at 24 months, we again first modeled the non-linear relationship between vocabulary at 18 months and grammar at 24 months. The model containing a linear and quadratic term fit the data better than the model containing only a linear term (AIC linear-only model = 70.51; AIC linear + quadratic = 68.49), with both fixed effects significantly predicting MB-CDI grammatical complexity (linear term: $B = 0.004$, SE(B) = 0.001, $\chi^2 = 10.26$, $p = .001$; quadratic term: $B = -7.04e^{-6}$, SE(B) = $3.37e^{-6}$, $\chi^2 = 4.37$, $p = .037$). The addition of a cubic term did not further improve model fit (AIC = 69.71, cubic term: $B = 3.49e^{-8}$, SE(B) = $3.92e^{-8}$, $\chi^2 = 0.79$, $p = .37$). We thus modeled the data with linear and quadratic terms. Since caregiver vocabulary types were significantly correlated with 24-month MLU but tokens were not, we included types as a control variable for the CT model. The CT model was significantly different from an intercept-only model ($\chi^2 = 27.057$, $df = 5$, $p < .001$), as was the IDS model ($\chi^2 = 24.204$, df = 5, $p < .001$). Table 9 shows that CTs at 18 months did not significantly predict infant 24-month MLU over and above all control variables, whereas for the IDS analysis, the proportion of imperatives was a significant negative predictor.

4 | DISCUSSION

In the current paper, we reported on a longitudinal study of play and its relationship to early language development. Unlike much past research on the topic, which has often focused on how children's symbolic acts are associated with language outcomes (see Quinn et al., 2018), we first measured how language and communicative interaction differed in infant–caregiver dyadic interaction across

TABLE 8 Model parameters of generalized linear models predicting 24-month MB-CDI grammatical complexity

	B	SE (*B*)	CI_{95} (Wald)	Wald χ^2	*p*
CT model[a]					
Intercept	−0.948	2.26	[−5.69, 3.49]	0.175	.676
Age	0.114	0.123	[−0.13, 0.36]	0.856	.355
Vocab 18mths	0.013	0.004	[0.006, 0.021]	11.876	.001
Vocab $18mths^2$	$-2.93e^{-5}$	$1.09e^{-5}$	[$-5.07e^{-5}$, $-7.92e^{-6}$]	7.218	.007
Tokens	0.000	0.001	[−0.002, 0.001]	0.071	.79
CTs	0.008	0.003	[0.002, 0.014]	5.96	.015
Scale	0.565	0.102	[0.397, 0.805]		
IDS model[b]					
Intercept	1.53	2.25	[−2.83, 5.89]	0.471	.492
Age	0.002	0.123	[−0.239, 0.243]	0.000	.998
Vocab 18mths	0.013	0.004	[0.006, 0.021]	12.64	<.001
Vocab $18mths^2$	$-2.42e^{-5}$	$1.02e^{-5}$	[$-4.42e^{-5}$, $-4.10e^{-6}$]	5.57	.018
Imperatives	−7.51	2.70	[−12.81, −2.21]	7.72	.005
YN-Qs	3.29	1.77	[−0.18, 6.77]	3.45	.063
Scale	0.52	0.095	[0.367, 0.747]		

[a]Log-likelihood = −170.031, Akaike's information criterion (AIC) = 354.062.

[b]Log-likelihood = −167.702, Akaike's information criterion (AIC) = 349.404.

symbolic and functional play. Following the suggestion that symbolic play requires the establishment of collective intentionality (Rakoczy, 2006, 2008; Tollefsen, 2005), we predicted that infant–caregiver interactions would be qualitatively different in comparison with functional play because the ambiguity associated with symbolic play necessitates the joint establishment of meaning and is thus more interactive (Quinn & Kidd, 2019; Reissland, 1998). Consistent with our hypotheses, we found that caregivers produced significantly greater interrogatives and significantly fewer imperatives in symbolic versus functional play and that there were more CTs between dyads in symbolic versus functional play. In addition, the observational nature of our design allowed us to explore other differences in language use across the two contexts and their association with concurrent and subsequent language proficiency. We discuss these results in more detail below.

The communicative complexity of symbolic play and its relationship to infants' language proficiency was most evident in our results concerning infant–caregiver CTs, which were higher in symbolic play at both time points, and at 18 months was significantly associated with vocabulary both concurrently and longitudinally, in addition to predicting grammatical knowledge at 24 months, as measured by the MB-CDI (but not in-session MLU). CTs appear to be a reliable predictor of child language (Donnelly & Kidd, 2021; Gilkerson et al., 2014, 2018; Romeo et al., 2018; Wang et al., 2020; Zimmerman et al., 2009), and it is therefore notable that we observed more CTs in symbolic play. Consistent with our theoretical framework, we attribute the result to the greater need in symbolic play to jointly establish meaning.

The finding that CTs predicted language development is consistent with socio-pragmatic approaches to language acquisition (Bruner, 1983; Nelson, 2007; Tomasello, 2003; Vygotsky, 1978), which identify social interaction as the foundation upon which the linguistic system is built. Toward the end of their first year, infants become increasingly capable of engaging in joint attention (Carpenter

TABLE 9 Model parameters of generalized linear models predicting infants' mean length of utterance (MLU) at 24 months

	B	SE (*B*)	CI_{95} (Wald)	Wald χ^2	*p*
CT model[a]					
Intercept	−0.738	0.586	[−1.89, 0.41]	1.586	.208
Age	0.011	0.032	[−0.05, 0.073]	0.132	.716
Vocab 18mths	0.003	0.001	[0.001, 0.005]	9.064	.003
Vocab 18mths^2	$-6.79e^{-6}$	$2.99e^{-6}$	[$-1.27e^{-5}$, $9.18e^{-7}$]	5.138	.023
Types	0.020	0.005	[0.01, 0.031]	13.878	<.001
CTs	0.001	0.001	[−0.001, 0.003]	1.140	.286
Scale	0.041	0.008	[0.028, 0.06]		
IDS model[b]					
Intercept	0.627	0.607	[−0.562, 1.86]	1.07	.301
Age	−0.009	0.034	[−0.076, 0.058]	0.067	.795
Vocab 18mths	0.003	0.001	[0.001, 0.005]	6.88	.009
Vocab 18mths^2	$-5.57e^{-6}$	$3.18e^{-6}$	[$-1.18e^{-5}$, $6.70e^{-7}$]	3.06	.08
Imperatives	−2.04	0.769	[−3.55, −0.533]	7.04	.008
YN-Qs	0.926	0.499	[−0.053, 1.90]	3.44	.064
Scale	0.044	0.009	[0.030, 0.064]		

[a]Log-likelihood = −23.354, Akaike's information criterion (AIC) = 60.708.

[b]Log-likelihood = −24.78, Akaike's information criterion (AIC) = 63.56.

et al., 1998), which affords competent others opportunities to scaffold language development via engagement in objects and shared communicative routines (Bruner, 1983). These interactions constitute a zone of proximal development for infants (Vygotsky, 1978), where learning is grounded within temporally contingent scaffolded social interaction (Csibra & Gergely, 2009, Leong et al., 2017; Rogoff, 1990; Tomasello, 1999; Vygotsky, 1978).

Caregivers produced significantly greater interrogatives in symbolic than in functional play. While there was no association between caregivers *wh* questions and infants' language (cf. Rowe et al., 2017), we did observe a significant longitudinal correlation between Y/N questions and grammatical knowledge at 24 months, although the relationship did not meet conventional standards of statistical significance in the regression analyses. The finding of a positive association is consistent with past research (Barnes et al., 1983) and likely comes from multiple sources. Notably, questions actively engage children in conversation (Golinkoff & Ames, 1979; Hoff-Ginsberg, 1985; Kavanaugh & Jirkovsky, 1982; Kruper & Užgiris, 1987; Rowe et al., 2017) and, because they require a response, place performance demands on their developing linguistic system. The specific association between Y/N questions and grammar may derive from the fact that English Y/N questions require challenging language-specific syntactic operations, including subject-auxiliary inversion (*This is a duck →Is this a duck?*) and *do*-support (*I know the teacher → Do you know the teacher?*). Accordingly, following classic research on the topic (Hoff-Ginsberg, 1985; Hoff-Ginsberg & Shatz, 1982), we suggest that greater exposure to questions supports the acquisition of auxiliary and modal verbs, and provides a key source of variety in word–order relations that help the learner identify form–function correlations via distributional analysis.

Caregivers also used consistently more mimetics in symbolic than in functional play, which at 18 months was positively correlated with infants' gesture use and vocabulary comprehension and production, but which only significantly predicted vocabulary production in our regression analysis. We did not expect this link between mimetic use and infants' language, although in light of recent work by Motamedi et al. (2021), a plausible explanation is that mimetics exert an influence on early vocabulary via their iconic link to the concept they denote. Interestingly, this phenomenon may coincide with another prominent feature of symbolic play—the use of "in-hand" iconic gestures (Quinn & Kidd, 2019; Vigliocco et al., 2019). Taken together, these two effects show how symbolic play may scaffold early vocabulary via its tendency to draw infants into the context of communicative and embodied action (Bates et al., 1983).

In contrast to linguistic behaviors that actively drew infants into the interaction, imperatives were used significantly more often in functional than in symbolic play and were negatively associated with infants' language proficiency. All except the association with concurrent productive vocabulary remained significant in the regression analyses. This result is consistent with past research with similarly high SES samples as ours (Furrow et al., 1979; Newport et al., 1977), and in the current sample, it is notable that the effect held longitudinally once controlling for infants' language proficiency at 18 months. We suggest that the negative effect is indicative of how communicative behaviors that direct and thus do not draw infants into sustained interaction may not provide the optimum conditions for acquisition. However, we caution in making this generalization beyond the cultural context of our sample, since past research on more socioeconomically diverse samples has found the opposite effect (Barnes et al., 1983), and because imperatives are often prominent features of IDS and language socialization in other cultures (where children still acquire the language proficiently, e.g., Abels et al., 2021).

The general pattern of our data supports suggestions for a strong social basis to language development (Bruner, 1983; Nelson, 2007; Tomasello, 2003). While the results of the current study reveal consistent support for the argument that symbolic play constitutes a fertile context for language development, there are several limitations that merit comment. Firstly, our sample was homogenous and not representative of the entire socioeconomic spectrum. This no doubt had some influence on how caregivers interacted with their infants (Hoff, 2003; Rowe, 2008, 2012), although it is important to also acknowledge that within SES strata there are still large and meaningful individual differences (see Kidd & Donnelly, 2020), which we also observed here. Equally, our sample is drawn from a so-called WEIRD (Western Educated Industrial Rich Democratic) society, which reflects a general sampling bias in developmental research (Nielsen et al., 2017). Although symbolic play appears to be a universal human behavior, its prevalence is moderated by culture (Lillard, 2017). Thus, following work such as Fletcher et al. (2020), it will be important to test our hypotheses in more diverse samples. Finally, our manipulation of play context was dependent on the toys we selected for each condition. Although our play coding demonstrated that the manipulation was successful—symbolic play largely occurred only in the symbolic play context—our data are open to the criticism that the results are toy-dependent. While this is a possibility, we point out that our toy sets were quite diverse and representative of typical play items in Australian households and pre-schools and that dyads were free to choose to interact with any toy within a set.

5 | CONCLUSION

Symbolic play has long been linked to language development, but the explanation for the link is unclear. In the current study, we tested a hypothesis derived from the sociocultural approaches to development; namely, that symbolic play has an influence on infants' language because the inherent

ambiguity of the context creates particularly fertile ecology that fosters communicative exchange (Bruner, 1983; Quinn & Kidd, 2019; Rakoczy, 2006, 2008; Reissland, 1998; Tomasello, 2003). The hypothesis was supported: Most notably, we found that symbolic play elicited greater conversational interaction between infant–caregiver dyads than did functional play, and was associated with patterns of IDS that foster language development. The results suggest that symbolic play is one potentially important context that can foster infants' language development, and supports those theories of language development that identify social processes as a crucial driving force in this process.

ACKNOWLEDGEMENTS

The research reported in this paper formed part of the PhDs of the first and second author (Creaghe, 2019; Quinn, 2016) and was supported by the Australian Research Council (CE140100041, CI: Kidd). The authors declare no conflicts of interest regarding the study's funding source. We thank Robin Panneton and two anonymous reviewers for helpful feedback. Open access funding enabled and organized by ProjektDEAL.

ORCID

Noëlie Creaghe https://orcid.org/0000-0002-2611-3228
Evan Kidd https://orcid.org/0000-0003-4301-2290

REFERENCES

Abels, M., Kilale, A., & Vogt, P. (2021). Speech act addressed to Hadza infants in Tanzania: Cross-cultural comparison, speaker age, and camp livelihood. *First Language*, *41*, 294–313. https://doi.org/10.1177/0142723720972000

Anandraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E. (2015). An introduction to Bayesian hypothesis testing for management research. *Journal of Management*, *41*, 521–543. https://doi.org/10.1177/0149206314560412

Australian Bureau of Statistics. (2016). *2016 census community profiles: Australian Capital Territory*. Canberra: Australian Bureau of Statistics. Retrieved from https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/communityprofile/8ACTE

Barnes, S., Gutfreund, M., Satterly, D., & Wells, G. (1983). Characteristics of adult speech which predict children's language development. *Journal of Child Language*, *10*, 65–84. https://doi.org/10.1017/s0305000900005146

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. Academic Press.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*, 507–584. https://doi.org/10.1080/016909697386628

Bates, E., Shore, C., Bretherton, I., & McNew, S. (1983). Names, gestures, and objects: Symbolization in infancy and aphasia. In K. E. Nelson (Ed.), *Children's language* (Vol. *4*, pp. 59–123). Lawrence Erlbaum.

Bateson, P., & Martin, P. (2013). *Play, playfulness, creativity, and innovation*. Cambridge University Press.

Bigham, S. A., & Bourchier-Sutton, A. (2007). The decontextualization of form and function in the development of pretence. *British Journal of Developmental Psychology*, *25*, 335–351. https://doi.org/10.1348/026151006x153154

Brown, P. M., Rickards, F. W., & Bortoli, A. (2001). Structures underpinning pretend play and word production in young hearing children and children with hearing loss. *Journal of Deaf Studies and Deaf Education*, *6*, 15–31. https://doi.org/10.1093/deafed/6.1.15

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Bruner, J. S. (1983). *Child's talk: Learning to use language*. Oxford University Press.

Caldera, Y. M., Huston, A. C., & O'Brien, M. (1989). Social interactions and play patterns of parents and toddlers with feminine, masculine, and neutral toys. *Child Development*, *60*, 70–76. https://doi.org/10.2307/1131072

Callaghan, T., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., Behne, T., & Collins, W. A. (2011). Early social cognition in three cultural contexts. *Monographs of the Society for Research in Child Development*, i-142. https://doi.org/10.1111/j.1540-5834.2011.00603.xo

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, *27*, 843–873. https://doi.org/10.1207/s15516709cog2706_2

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs for the Society of Research in Child Development*, *63*, 1–143. https://doi.org/10.2307/1166214

Casillas, M., Bobb, S. C., & Clark, E. V. (2016). Turn taking, timing, and planning in early language acquisition. *Journal of Child Language*, *43*, 1310–1337. https://doi.org/10.1017/S0305000915000689

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Creaghe, N. V. (2019). *Symbolic play and language acquisition: The dynamics of infant-caregiver communication during symbolic play*. Retrieved from https://openresearch-repository.anu.edu.au/handle/1885/201973

Cristofaro, T. N., & Tamis-LeMonda, C. S. (2012). Mother-child conversations at 36 months and at pre-kindergarten: Relations to children's school readiness. *Journal of Early Childhood Literacy*, *12*, 68–97. https://doi.org/10.1177/1468798411416879

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*, 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Deacon, T. (1997). *The symbolic species*. W. W. Norton & Co.

Devescovi, A., Casselli, M. C., Marchione, D., Pasqualetti, P., Reilly, J., & Bates, E. (2005). A crosslinguistic study of the relationship between grammar and lexical development. *Journal of Child Language*, *32*, 759–786. https://doi.org/10.1017/S0305000905007105

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

Donnelly, S., & Kidd, E. (2021). The longitudinal relationship between conversational turn-taking and vocabulary growth in early language development. *Child Development*, *92*(2), 609–625. https://doi.org/10.1111/cdev.13511

Dunn, J., & Wooding, C. (1977). Play at home and its implications for learning. In B. Tizard & D. Harvey (Eds.), *Biology of play* (pp. 45–58). Heinemann.

ELAN (Version 4.9.4) [Computer software]. (2016). *Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive*. Retrieved from https://archive.mpi.nl/tla/elan

Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development*, *52*, 1095–1118. https://doi.org/10.2307/1129497

Fekonja, U., Umek, L., & Kranjc, S. (2005). Free play and other daily preschool activities as a context for child's language development. *Studia Psychologica*, *47*, 103–117.

Fenson, L., Bates, E., Dale, P. S., Marchman, V. A., Reznick, J. S., & Thal, D. J. (2007). *MacArthur-Bates communicative development inventories* (2nd ed). Brookes Publishing.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), i. https://doi.org/10.2307/1166093

Fenson, L., Kagan, J., Kearsley, R. B., & Zelazo, P. R. (1976). The developmental progression of manipulative play in the first two years. *Child Development*, *47*, 232–236. https://doi.org/10.2307/1128304

Fletcher, K. K., Cates, C. B., Mendelsohn, A. L., & Tamis-LeMonda, C. S. (2020). Play in Mexican-American mothers and toddlers is frequent, multimodal, and rich in symbolic content. *Infancy*. https://doi.org/10.1111/infa.12339

Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language*, *6*, 423–442. https://doi.org/10.1017/s0305000900002464

Garamszegi, L. Z. (2006). Comparing effect sizes across variables: Generalization without the need for Bonferroni correction. *Behavioral Ecology*, *17*, 682–687. https://doi.org/10.1093/beheco/ark005

Garamszegi, L. Z. (2016). A simple statistical guide when data are constrained due to practical or ethical reasons. *Animal Behavior*, *120*, 223–234. https://doi.org/10.1016/j.anbehav.2015.11.009

Garvey, C., & Kramer, T. L. (1989). The language of social pretend play. *Developmental Review*, *9*, 364–382. https://doi.org/10.1016/0273-2297(89)90035-X

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, K., Hansen, J. H. L., & Paul, T. D. (2014). Mapping the language environment using all-day and automated analysis. *American Journal of Speech-Language Pathology*, *26*, 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language experience in the second year of life and language outcomes in late childhood. *Pediatrics*, *142*, e20174276. https://doi.org/10.1542/peds.2017-4276

Golinkoff, R. M., & Ames, G. J. (1979). A comparison of fathers' and mothers' speech with their young children. *Child Development*, *50*, 28–32. https://doi.org/10.2307/1129037

Göncü, A., & Gaskins, S. (2011). Comparing and extending Piaget's and Vygotsky's understandings of play: Symbolic play as individual, sociocultural, and educational interpretation. In P. Nathan & A. D. Pellegrini (Eds.), *The Oxford handbook of the development of play* (pp. 48–57). Oxford University Press.

Graf Estes, K., Gluck, S. C. W., & Grimm, K. (2016). Finding patterns and learning words. *Journal of Experimental Child Psychology*, *146*, 34–49. https://doi.org/10.1016/j.jecp.2016.01.012

Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., & Gilkerson, J. (2011). Assessing children's home language environments using automatic speech recognition technology. *Communication Disorders Quarterly*, *32*, 83–92. https://doi.org/10.1177/1525740110367826

Haight, W. L., & Miller, P. J. (1993). *Pretending at home: Early development in a sociocultural context*. SUNY Press.

Hillbrink, E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: A longitudinal study of mother-infant interaction. *Frontiers in Psychology*, *6*, Article 1492. https://doi.org/10.3389/fpsyg.2015.01492

Hirsh-Pasek, K., Adamsdon, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K. S., & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*, 1071–1089. https://doi.org/10.1177/0956797615581493

Hirsh-Pasek, K., Golinkoff, R., Berk, L. E., & Singer, D. (2009). *A mandate for playful learning in preschool: Applying the scientific evidence*. Oxford University Press.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368–1378. https://doi.org/10.1111/1467-8624.00612

Hoff-Ginsberg, E. (1985). Some contributions of mothers' speech to their children's syntactic growth. *Journal of Child Language*, *12*, 367–385. https://doi.org/10.1017/s0305000900006486

Hoff-Ginsberg, E., & Shatz, M. (1982). Linguistic input and the child's acquisition of language. *Psychological Bulletin*, *92*, 3–26. https://doi.org/10.1037//0033-2909.92.1.3

Huddleston, R., & Pullum, G. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Huttenlocher, J., Vasilyeva, M., Waterfall, H. R., Vevea, J. L., & Hedges, L. V. (2007). The varieties of speech to young children. *Developmental Psychology*, *43*, 1062–1083. https://doi.org/10.1037/0012-1649.43.5.1062

JASP Team. (2018). JASP (version 0.12.1).

Kavanagh, R. D., Whittington, S., & Cerbone, M. J. (1983). Mothers' use of fantasy speech to young children. *Journal of Child Language*, *10*, 45–55. https://doi.org/10.1017/S0305000900005122

Kavanaugh, R. D., & Jirkovsky, A. M. (1982). Parental speech to young children: A longitudinal analysis. *Merrill-Palmer Quarterly*, *28*, 297–311.

Kidd, E., & Donnelly, S. (2020). Individual differences in first language acquisition. *Annual Review of Linguistics*, *6*, 319–340. https://doi.org/10.1146/annurev-linguistics-011619-030326

Kruper, J. C., & Užgiris, I. C. (1987). Fathers' and mothers' speech to young infants. *Journal of Psycholinguistic Research*, *16*, 597–614. https://doi.org/10.1007/bf01067087

Laing, C. (2019). A role for onomatopoeia in early language: Evidence from phonological development. *Language and Cognition*, *11*, 173–187. https://doi.org/10.1017/langcog.2018.23

Laplante, D. P., Zelazo, P. R., Brunet, A., & King, S. (2007). Functional play at 2 years of age: Effects of prenatal maternal stress. *Infancy*, *12*, 69–93. https://doi.org/10.1111/j.1532-7078.2007.tb00234.x

Largo, R. H., & Howard, J. A. (1979). Developmental progression in play behavior of children between nine and thirty months. I: Spontaneous play and imitation. *Developmental Medicine & Child Neurology*, *21*, 299–310. https://doi.org/10.1111/j.1469-8749.1979.tb01622.x

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 13290–13295. https://doi.org/10.1073/pnas.1702493114

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review*, *94*, 412–426. https://doi.org/10.1037/0033-295X.94.4.412

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Lewis, V., & Boucher, J. (1997). *The test of pretend play*. Psychological Corporation.

Lillard, A. (2001). Pretend play as twin earth: A social-cognitive analysis. *Developmental Review*, *21*, 495–531. https://doi.org/10.1006/drev.2001.0532

Lillard, A. S. (2017). Why do the children (pretend) play? *Trends in Cognitive Sciences*, *21*, 826–834. https://doi.org/10.1016/j.tics.2017.08.001

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin*, *139*, 1–34. https://doi.org/10.1037/a0029321

Lillard, A., Nishida, T., Massaro, D., Vaish, A., Ma, L., & McRoberts, G. (2007). Signs of pretense across age and scenario. *Infancy*, *11*, 1–30. https://doi.org/10.1207/s15327078in1101_1

Lillard, A. S., Pinkham, A. M., & Smith, E. (2011). Pretend play and cognitive development. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (pp. 285–311). Wiley-Blackwell. https://doi.org/10.1002/9781444325485.ch11

Lillard, A. S., & Witherington, D. C. (2004). Mothers' behavior modifications during pretense and their possible signal value for toddlers. *Developmental Psychology*, *40*, 95–113. https://doi.org/10.1037/0012-1649.40.1.95

LoBue, V., Reider, L. B., Kim, E., Burris, J. L., Oleas, D. S., Buss, K. A., Pérez-Edgar, K., & Field, A. P. (2020). The importance of using multiple outcome measures in infant research. *Infancy*, *25*, 420–437. https://doi.org/10.1111/infa.12339

MacWhinney, B. (2013). Child Language Analysis (CLAN). (Version 18).

McCune, L. (1995). A normative study of representational play at the transition to language. *Developmental Psychology*, *31*, 198–206. https://doi.org/10.1037/0012-1649.31.2.198

McDonald, L. O. (1979). *A functional analysis of individual differences in conversational style among mothers*. Unpublished doctoral dissertation, University of Oregon.

Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and diversity: Simulating early word learning environments. *Cognitive Science*, *42*(Supp. 2), 375–412. https://doi.org/10.1111/cogs.12592

Morrissey, A. M. (2014). Scaffolding, analysis and materials: Contributing factors in an unexpected finding of advanced infant/toddler pretend play? *Journal of Early Childhood Research*, *12*, 195–213. https://doi.org/10.1177/1476718X13515428

Motamedi, Y., Murgiano, M., Perniss, P., Wonnacott, E., Marshall, C., Goldin-Meadow, S., & Vigliocco, G. (2021). Linking language to sensory experience: Onomatopoeia in early language development. *Developmental Science*, *24*(3), e13066. https://doi.org/10.1111/desc.13066

Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology*, *15*, 1044–1045. https://doi.org/10.1093/beheco/arh107

Nelson, K. (2007). *Young minds in social worlds*. Harvard University Press.

Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 109–149). Cambridge University Press.

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, *5*, 241–301. https://doi.org/10.1037/1082-989x.5.2.241

Nielsen, M., & Dissanayake, C. (2004). Pretend play, mirror self-recognition and imitation: A longitudinal investigation through the second year. *Infant Behavior and Development*, *27*, 342–365. https://doi.org/10.1016/j.infbeh.2003.12.006

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Ninio, A. (1980). Picture-book reading in mother-infant dyads belonging to two subgroups in Israel. *Child Development*, *51*, 587–590. https://doi.org/10.2307/1129299

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language*, *5*, 1–15. https://doi.org/10.1017/S0305000900001896

Ninio, A., & Snow, C. (1999). The development of pragmatics: Learning to use language appropriately. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of child language acquisition* (pp. 347–383). Academic Press.

Nishida, T. K., & Lillard, A. S. (2007). The informative value of emotional expressions: "Social referencing" in mother-child pretense. *Developmental Science*, *10*(2), 205–212. https://doi.org/10.1111/j.1467-7687.2007.00581.x

O'Brien, M., & Nagle, K. J. (1987). Parents' speech to toddlers: The effect of play context. *Journal of Child Language*, *14*, 269–279. https://doi.org/10.1017/S0305000900012927

Ota, M., Davies-Jenkins, N., & Skarabela, B. (2018). Why Choo-Choo is better than train: The role of register-specific words in early vocabulary growth. *Cognitive Science*, *42*(6), 1974–1999. https://doi.org/10.1111/cogs.12628

Piaget, J. (1962). *Play, dreams, and imitation in childhood*. Norton.

Pretzer, G. M., Lopez, L. D., Walle, E. A., & Warlaumont, A. S. (2019). Infant-adult vocal interaction dynamics depend on infant vocal type, child-directedness of adult speech, and timeframe. *Infant Behavior and Development*, *57*, 101325. https://doi.org/10.1016/j.infbeh.2019.04.007

Quinn, S. J. (2016). *Learning to play and playing to learn: the role of symbolic play in language acquisition*. Retrieved from http://hdl.handle.net/1885/105179

Quinn, S., Donnelly, S., & Kidd, E. (2018). The relationship between symbolic play and language acquisition: A meta-analytic review. *Developmental Review*, *49*, 121–135. https://doi.org/10.1016/j.dr.2018.05.005

Quinn, S., & Kidd, E. (2019). Symbolic play promotes non-verbal communicative exchange in infant–caregiver dyads. *British Journal of Developmental Psychology*, *37*, 33–50. https://doi.org/10.1111/bjdp.12251

Rakoczy, H. (2006). Pretend play and the development of collective intentionality. *Cognitive Systems Research*, *7*, 113–127. https://doi.org/10.1016/j.cogsys.2005.11.008

Rakoczy, H. (2008). Pretence as individual and collective intentionality. *Mind & Language*, *23*, 499–517. https://doi.org/10.1111/j.1468-0017.2008.00357.x

Reissland, N. (1998). Context dependency in parental speech. *British Journal of Developmental Psychology*, *16*, 365–373. https://doi.org/10.1111/j.2044-835X.1998.tb00758.x

Reissland, N., & Snow, D. (1996). Maternal pitch height in ordinary and play situations. *Journal of Child Language*, *23*, 269–278. https://doi.org/10.1017/S0305000900008795

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*, 700–710. https://doi.org/10.1177/0956797617742725

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*, 185–205. https://doi.org/10.1017/S0305000907008343

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*, 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L., Leech, K. A., & Cabrera, N. (2017). Going beyond input quantity: Wh-questions matter for toddlers' language and cognitive development. *Cognitive Science*, *41*, 162–179. https://doi.org/10.1111/cogs.12349

Rubin, K. H., & Howe, N. (1985). Toys and play behaviors: An overview. *Topics in Early Childhood Special Education*, *5*, 1–9. https://doi.org/10.1177/027112148500500302

Schmidt, F., & Hunter, J. (2002). Are there benefits from NHST? *American Psychologist*, *57*(1), 65–66. https://doi.org/10.1037/0003-066X.57.1.65

Searle, J. (1995). *The social construction of reality*. Allen Lane.

Smith, P. K. (2009). *Children and play: Understanding children's worlds*. Wiley-Blackwell.

Sutton-Smith, B. (2001). *The ambiguity of play*. Harvard University Press.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development*, *72*, 748–767. https://doi.org/10.1111/1467-8624.00313

Taylor, M., Cartwright, B. S., & Carlson, S. M. (1993). A developmental investigation of children's imaginary companions. *Developmental Psychology*, *29*, 276–285. https://doi.org/10.1037/0012-1649.29.2.276

Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences*, *35*, 75–97. https://doi.org/10.1177/0048393104271925

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

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

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*, 1454–1463. https://doi.org/10.2307/1130423

Trawick-Smith, J. (1998). A qualitative analysis of metaplay in the preschool years. *Early Childhood Research Quarterly*, *13*, 433–452. https://doi.org/10.1016/S0885-2006(99)80049-0

Vigliocco, G., Motamedi, Y., Murgiano, M., Wonnacott, E., Marshall, C. R., Maillo, I. M., & Perniss, P. (2019). Onomatopoeias, gestures, actions, and words in the input to children: How do caregivers use multimodal cues in their communication to children? In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the cognitive science society* (pp. 1171–1177). Cognitive Science Society.

Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Soviet Psychology*, *5*(3), 6–18. https://doi.org/10.2753/RPO1061-040505036

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Harvard University Press.

Wang, Y., Willimas, R., Dilley, L., & Houston, D. M. (2020). A meta-analysis of the predictability of LENA™ automated measures for child language development. *Developmental Review*, *57*, 100921. https://doi.org/10.1016/j.dr.2020.100921

Werner, H., & Kaplan, B. (1963). *Symbol formation*. John Wiley.

Wu, Z., & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language*, *34*, 72–90. https://doi.org/10.1177/0142723714521925

Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yapanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, *124*, 342–349. https://doi.org/10.1542/peds.2008-2267

How to cite this article: Creaghe, N., Quinn, S., & Kidd, E. (2021). Symbolic play provides a fertile context for language development. *Infancy*, 26, 980–1010. https://doi.org/10.1111/infa.12422

APPENDIX A

Pretend Play Observation Scale

Symbolic play was coded according to the Pretend Play Observation Scale (Brown et al., 2001), a coding scheme that describes the typical developmental sequence of pretend play according to ten stages. For a description of these stages and their examples, see Table A1.

Following Morrissey (2014), there were two modifications to the coding scheme. Firstly, Stage 7 (object transformations) of Brown's scale was separated into two stages: *simple* (7.0) to represent early forms of object transformations (e.g., using a saucepan as a hat, cylinder as a cup) and *complex* (7.5) forms of object transformations (e.g., transforming more than one object at a time, transforming within a combinatorial sequence, and transformations involving greater dissimilarity to the represented object). Secondly, the coding of Stage 6 *planned play* was not applied to caregivers' play as their play activity was considered "modeling" and was frequently accompanied by verbalized intention (planning). If all planned caregiver play was coded at Stage 6 (*planned action*), the contribution of caregiver modeling of play between earlier Stages 2 and 5 (*autosymbolic*, *decentered*, *linear*, and *combinatorial sequences*) would have been lost. Therefore, removing Stage 6 as a coding stage for caregiver play allowed for nuances of their earlier play activity to be captured descriptively within the coding scheme.

TABLE A1 Pretend Play Observation Scale

Stage	Age (months)	Descriptor and example
1	>12	Pre-symbolic (closes eyes and pretends to sleep)
2	12–15	Autosymbolic (feeds self with empty spoon)
3	13–18	Decentered (feeds doll/partner with empty spoon)
4	16–19	Linear sequence (feeds self and doll in any order)
5	18–24	Combinatorial sequence with single recipient (feeds and bathes doll in any order)
6[a]	18–26	Planned action (searches for, requests, offers materials incorporated into play)
7.0	>20	Simple object transformation (uses saucepan as hat, aerosol lid as cup)
7.5	>20	Complex object transformation (more than one object at a time, within a combinatorial sequence, and/or involving greater dissimilarity to the represented object)
8	21–30	Agency attribution (adopts vocal or physical attributes of another, for example, cat, driver)
9	>30	Ordered sequences (mixes cake, bakes it, eats it, retaining logical order)
10	>30	Imaginary transformation (places imaginary cake on plate, interacts with imaginary character)

Note: Adapted from Brown et al. (2001).

[a] Stage 6 was not coded for caregivers.

(Continues)

APPENDIX B

Non-linear relationship between 18-month productive vocabulary and 24-month language outcome measures

Figures B1–B3 reveal non-linear relationships between productive vocabulary at 18 months and productive vocabulary, grammatical complexity (as measured by the MB-CDI), and MLU (as measured in the play session) at 24 months, respectively.

Such non-linear relationships are common in early language development, although it may differ depending on the domain and language in question (Devescovi et al., 2005; Donnelly & Kidd, 2020).

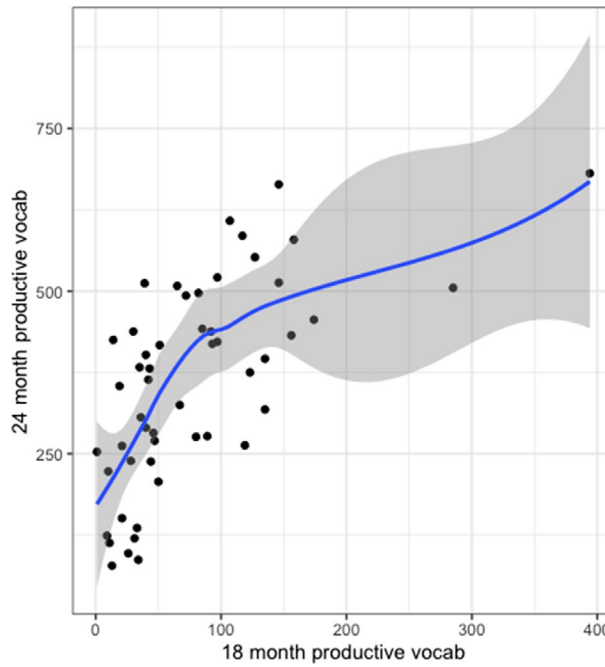

FIGURE B1 Relationship between 18- and 24-month productive vocabulary

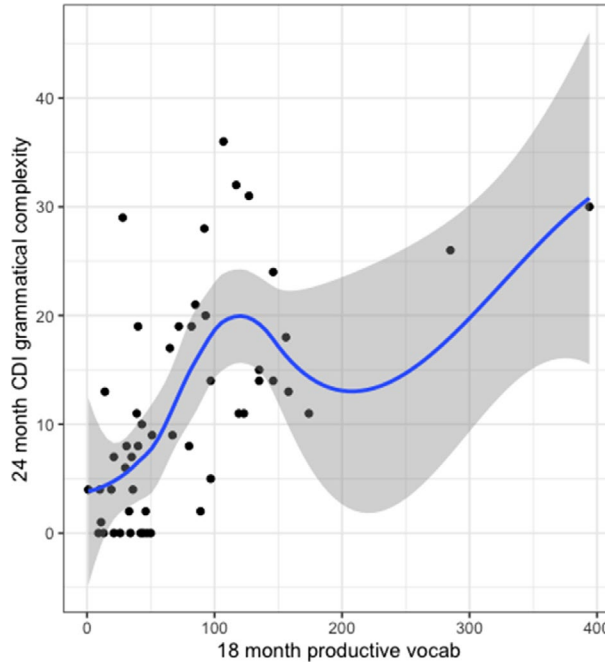

FIGURE B2 Relationship between 18-month productive vocabulary and 24 month grammatical complexity

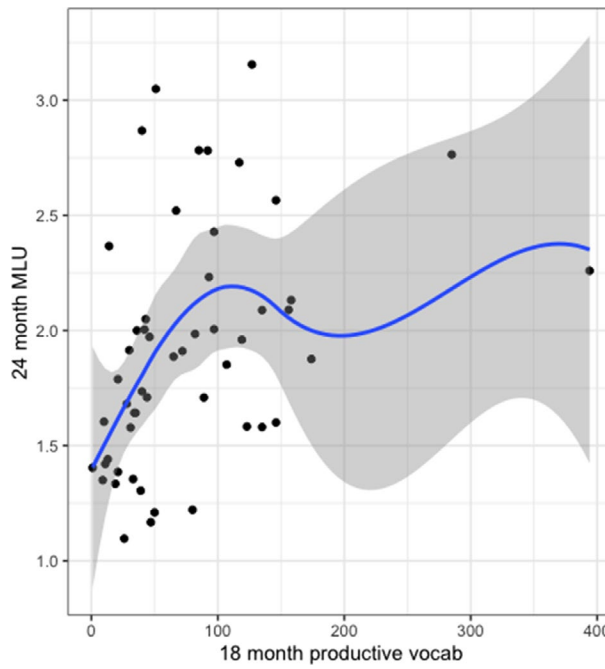

FIGURE B3 Relationship between 18-month productive vocabulary and 24 month MLU