Maternal discourse continuity and infants' actions organize 12-month-olds' language exposure during object play.

https://escholarship.org/uc/item/6p20p233

Chang, LM
Deák, GO

2018-11-10

10.1111/desc.12770

Peer reviewed
Maternal discourse continuity and infants’ actions organize 12-month-olds’ language exposure during object play

Lucas M. Chang  |  Gedeon O. Deák

Department of Cognitive Science, University of California, San Diego, La Jolla, California

Correspondence
Gedeon O. Deák, Department of Cognitive Science, University of California, San Diego, La Jolla, CA.
Email: gdeak@ucsd.edu

Funding information
Division of Social and Economic Sciences, Grant/Award Number: SES-0527756; Division of Graduate Education, Grant/Award Number: DGE-1144086; Office of Multidisciplinary Activities, Grant/Award Number: NSF SMA-1005256

Abstract
Infant language learning depends on the distribution of co-occurrences within language—between words and other words—and between language content and events in the world. Yet infant-directed speech is not limited to words that refer to perceivable objects and actions. Rather, caregivers’ utterances contain a range of syntactic forms and expressions with diverse attentional, regulatory, social, and referential functions. We conducted a distributional analysis of linguistic content types at the utterance level, and demonstrated that a wide range of content types in maternal speech can be distinguished by their distribution in sequences of utterances and by their patterns of co-occurrence with infants’ actions. We observed free-play sessions of 38 12-month-old infants and their mothers, annotated maternal utterances for 10 content types, and coded infants’ gaze target and object handling. Results show that all content types tended to repeat in consecutive utterances, whereas preferred transitions between different content types reflected sequences from attention-capturing to directing and then descriptive utterances. Specific content types were associated with infants’ engagement with objects (declaratives, descriptions, object names), with disengagement from objects (talk about attention, infant’s name), and with infants’ gaze at the mother (affirmations). We discuss how structured discourse might facilitate language acquisition by making speech input more predictable and/or by providing clues about high-level form-function mappings.

KEYWORDS
attention, discourse, infant-directed speech, language input, object play, parent-child interaction

1 | INTRODUCTION

1.1 | Infant-directed discourse

From the earliest ages, infant-directed speech (IDS) constitutes a rich discourse including interconnected utterances of varied syntactic and semantic content. These utterances include but are not limited to object-naming, descriptive comments, attention-directing, imperatives, questions, and stereotyped routines such as greetings and games. Caregivers’ varied utterance types have interactive functions including regulating infants’ affective state and attention (e.g., Fernald, 1993; Ninio & Snow, 1996; Papoušek, Bornstein, Nuzzo, Papoušek, & Symmes, 1990). Distinct functions can be at least partially discriminated using low-level acoustic features (Fernald, 1989) and by linguistically naive observers (Bryant & Barrett, 2007). Yet, most research has focused on how infants’ environment supports learning object words (e.g., Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006; Suanda, Smith, & Yu, 2016; Trueswell, Lin, Armstrong, Cartmill, Goldin-Meadow, & Gleitman, 2016). The diversity of utterance types in natural discourse, however, exacerbates...
the inherent difficulty of mapping words to meanings: for example, there are frequent opportunities for spurious associations not only between words and objects, but between words or phrases and varied social purposes ranging from action-imperatives such as “don’t touch that!” to non-referential utterances such as “Hi, baby!”

Despite this potential difficulty, the amount of complex and diverse speech heard by infants predicts their vocabulary growth (Akhtar, Dunham, & Dunham, 1991; Bornstein, Haynes, & Painter, 1998; Hart & Risley, 1995; Hoff, 2003; Huttonlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Pan, Rowe, Singer, & Snow, 2005). This appears to be true even though a richer discourse environment would potentially complicate the task of a learner who relies on co-occurrences of words and referents to construct a lexicon (Yu & Ballard, 2007; Yu & Smith, 2012). Perhaps, however, infants also start to detect and discriminate patterns of occurrence of different discourse content types, even as they are learning their first object words. Different functions of IDS utterances such as attention-getting, prohibition, encouragement, and comforting might be indicated not only by prosodic differences (e.g., Fernald, 1989; Katz, Cohn, & Moore, 1996), but also by the frequency of utterance contexts or frames that associate words with phrase structures and syntactic roles. Such findings underscore the question: how does syntactically and pragmatically diverse discourse shape not only infants’ linguistic environment, but also their language acquisition? One possibility is that discourse diversity is merely incidental: parents simply produce complex language out of habit, but the diversity of words and discourse-functions does not directly affect early language acquisition. A second possibility is that caregivers use IDS to regulate infants’ attention, to make word-learning opportunities more predictable and therefore more effective (Benitez & Smith, 2012). A third possibility, not mutually exclusive with the second, is that complex discourse is itself an input for infant learning – not, perhaps, of word-object mappings, but of the different types of utterances that can accompany various social-interactive events. That is, perhaps a burgeoning awareness of discourse types, working in tandem with infants’ growing social-event knowledge (Cannon & Woodward, 2012; Tomasello, 1992) allows infants to classify or even predict adults’ communicative actions during social interactions. However, the relevant statistical patterns characterizing infant-directed discourse – its structure over time and correspondences with other observable events – are poorly characterized. Thus, it remains unclear what role the diversity of varied discourse content plays in language acquisition.

1.2 | Discourse continuity

Conversational analysis studies show that successive speakers’ turns show well-structured, predictable sequences of functional content types, such as question-answer pairs (Ninio, Snow, Pan, & Rollins, 1994; Schegloff, 2007). Conversations between adults and young children show developmental trajectories toward a variety of such functional conversational sequences, including question-answering (Gallagher, 1981) and repair (Tomasello, Ramsden, & Ewert, 2009). However, relatively few studies have focused on discourse structure in parents’ speech to prelinguistic infants, even though infant-directed discourse is presumably the developmental context for the emergence of early conversational patterning. These studies are reviewed below.

One proposed mechanism by which infant-directed discourse could facilitate learning is continuity of reference (Messer, 1980). Frank, Tenenbaum, and Fernald (2013) identified object references in a corpus of IDS and found that mothers’ consecutive utterances were more likely to refer to the same object than to other objects that were present. Moreover, a computational model was able to use this continuity of reference to more accurately infer which object was the topic of each utterance. However, sequential continuity might also characterize discourse functions other than object reference. For instance, if utterances that either name objects, describe objects, or refer to actions each tend to follow previous utterances of the same type, this continuity might be useful to a learner tasked with learning a lexicon of nouns, adjectives, and verbs, respectively. That is, sequences of utterances of the same type might provide not only repetition of topical content words and continuity of reference, but also correlated repetition of syntactic forms. Repetition of these parallel levels of content, function, and context variables might facilitate semantic and pragmatic mapping.

Another form of discourse continuity might be available to infants: short, formulaic sequences of different utterances that tend to occur in predictable order. For example, utterances that capture the infant’s attention might be followed by utterances that direct their attention to an object, and then by utterances that name the object, and finally by elaborating utterances that describe or evaluate it. By this hypothesis, IDS would be organized around stereotyped sequences (or “paths”) of content-types (“states”), with different probabilities associated with all of the possible transitions from one content type to the next. For example, Rohde and Frank (2014) found that mothers produced more utterance-final
object names early within runs of same-topic utterances; conversely, they produced more pronouns later in runs. However, the aforementioned studies analyzed only object-naming and/or object-referring utterances, and therefore did not address whether continuity over time or predictable transitions characterize other types of infant-directed utterances. Therefore, in the current study we recorded transition probabilities among different types of infant-directed utterances.

1.3 Contextual correlates of discourse types

In addition to sequential continuity from one utterance to the next, the discourse structure of IDS could facilitate language learning if different types of utterances appear preferentially in specific contexts. Relatively few studies have examined the contextual correlates of utterances other than object-referring utterances. Moreover, studies on the nonlinguistic contexts of object-referring utterances often measure only the degree to which the infant and/or parent's gaze or manual actions are directed to the particular object being spoken about (e.g. Frank et al., 2013; Yu & Smith, 2012). However, independent of object-specific associations, object-naming as a discourse function occurs coupled with infants' general object-directed gaze and developmentally advanced manual object exploration (Chang, de Barbaro, & Déak, 2016), and analogous coupling might exist between other discourse functions and other non-verbal actions or states within social interactions. If infants can learn to associate utterance-types with the contexts they most often occur in, this could form the basis for infants to develop expectations that are well coordinated with caregiver actions (Rochat, Querido, & Striano, 1999). These contexts can be studied at a macro level (e.g. the activity being engaged in and the participants present), or at a micro level that is, the moment-to-moment actions and attention-states of participants within an activity. Several studies have investigated the frequency of various types of linguistic content in IDS across different activities: book-reading and object play (Choi, 2000; Gros-Louis, West, & King, 2016; Tamis-LeMonda, Song, Leavell, Kahana-Kalman, & Yoshikawa, 2012; Yont, Snow, & Vernon-Feagans, 2003), as well as eating and dressing (Hoff-Ginsberg, 1991; Pan, Perlmutter, & Snow, 2000). These efforts have revealed significant variability in the proportions of types of maternal utterances, such as questions, directives, questions, and naming, produced in these varied social contexts.

Other studies have investigated caregivers' use of different types of utterance content in response to infants' moment-to-moment actions within an activity context. One series of studies found that mothers follow infants' social initiatives with responsive utterances, whereas they tend to ignore or redirect infants' object initiatives, and to redirect infants when they disengage from toys (Masur, Flynn, & Lloyd, 2013; Masur & Lloyd, 2014). Similarly, Tamis-LeMonda, Kuchirko, and Tafuro (2013) found that mothers contingently respond to 14-month-old infants' object exploration actions by handling the objects themselves and by producing more referential language. Another study found that mothers respond contingently to infants' object exploration and play with utterances of various types including questions, descriptions, affirmations, imitations, and prompts (Bornstein, Tamis-LeMonda, Hahn, & Haynes, 2008). Beyond social and object play, infants' walking and crawling influences parents' use of prohibitions and action directives (Karasik, Tamis-LeMonda, & Adolph, 2014; Zumbahlen, 1997), and infants' pointing encourages parents to ask questions (Wu & Gros-Louis, 2014). Collectively, these studies show that caregivers' utterance content types depend on infants' actions, however, each has limitations. In several studies, infants' activity was judged subjectively by different criteria, making comparisons across studies difficult. In others, parental utterances were only counted if they were judged by human observers to be appropriate responses to the infants' actions, or utterance types were defined based on how they related to infant actions. These designs do not allow for comparisons between observed utterance-action co-occurrence rates and chance, and are difficult to interpret in terms of informativeness to naive learners. Therefore, in the current study we comprehensively documented both infants' action primitives, specifically gaze target and object handling, and mothers' speech, which was then classified for different types of discourse content.

1.4 Utterance types

Several utterance content types have been studied in IDS and infant language, and the evidence points to predictable developmental trajectories. At the lexical level, researchers have been interested in infants' first-acquired words. These include many "routines" (e.g. thank you, all gone, hello), and an increasing proportion of object nouns in the first months of production (Caselli et al., 1995). As early as 6 months of age, infants show evidence of having formed associations with common nouns (Bergelson & Swingley, 2012) and their own name (Mandel, Jusczyk, & Pisoni, 1995). Other researchers have examined which types of discourse content in speech input correlate with developmental outcomes. A consistent finding has been that a higher proportion of responsive utterances — typically defined as utterances that refer to the infants' focus of attention, or non-referential affirmations of infants' communicative actions — correlate with positive outcomes in language development (Baumwell, Tamis-LeMonda, & Bornstein, 1997; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Masur, Flynn, & Eichorst, 2005; Rollins, 2003; Tamis-LeMonda, Bornstein, & Baumwell, 2001). In contrast, a higher proportion of directive utterances is unrelated to (Baumwell et al., 1997; Carpenter et al., 1998) or negatively correlated with vocabulary growth (Hughes, Dote-Kwan, & Dolendo, 1999; Nelson, 1973; Tomasello & Todd, 1983), perhaps because directives disrupt infants' attention (Akhtar et al., 1991; Lloyd & Masur, 2014; Masur et al., 2005).

We therefore sought to code maternal utterances in a way that captured the salient lexical, syntactic, and topical content types reviewed above. Standardized taxonomies of communicative acts have been proposed (Ninio et al., 1994; Searle, 1976); however, these schemes either did not cover the desired range of levels of description or did not match well with the set of content types that were...
developmentally and contextually appropriate to the current study. Therefore, we chose to create a coding system suitable for the population and context represented in the current study. We chose to tag infant-directed utterances at the syntactic level as declarative, imperative, or questions, and at the topical level as about attention, infant actions, or object-properties (i.e., descriptions). We also tagged utterances for several other lexical or functional content types with specific relevance to infant language: infant’s name, object names, social routines, and affirmations. Similar approaches to coding IDS have been used in studies of infant-parent interaction (Bornstein et al., 2008; Gros-Louis, West, Goldstein, & King, 2006; Toda, Fogel, & Kawai, 1990).

1.5 | Infant activity

While previous studies have extensively documented parents’ tendency to refer to specific objects of infant visual attention (e.g., Yu & Smith, 2012), we chose to analyze more general categories of gaze: at toys, at the mother’s face, or other. We expected that mothers would tailor their speech content to the infant’s focus of attention (Frank et al., 2013), preferentially using speech with referential, interpersonal, or attention-directing content, depending on infants’ gaze state. We also expected that simple versus complex object handling would elicit different speech patterns. Chang et al. (2016) and de Barbaro, Johnson, and Deák (2015) showed that infants’ manipulation of two objects is related to their cognitive and motor maturity, and that bimanual manipulation elicited more object naming by mothers of infants between 4 and 9 months. Handling of multiple objects might also be related to subjective measures of infant manual behavior such as the distinction between object exploration and play (Bornstein et al., 1998). Object handling was therefore classified as involving no objects, one object, or two objects.

1.6 | Current study

In the current study, we observed mother-infant dyads in a free-play in-home interaction with toys. We transcribed the mothers’ speech and tagged utterances as containing declaratives, imperatives, questions, talk about action, talk about attention, object descriptions, affirmations, social routines, object names, and the infant’s name. We also annotated infants’ gaze at toys and at the mother’s face, and infants’ handling of the toys. The first set of analyses describes the structure of mothers’ discourse over time. We measured the repetitiveness of all content types, and report transitions between the types that occur at above chance rates. The second set of analyses describes contingencies between infants’ actions and mothers’ discourse. We first investigated whether changes in infants’ gaze target or object handling affect mothers’ probability of repeating each content type in successive utterances. We also investigated whether each content type varies in frequency as a function of infants’ gaze target and/or object handling. Finally, we quantified the degree of predictability of each content type as a function of previous utterance content and of infant activity state. We expected that most content types would tend to follow utterances of the same type at above-chance levels, and that infant gaze or hand shifts would decrease this tendency. We also expected that mothers’ speech would include both significant sequential transitions between different utterance content types, and correlations between utterance content and infants’ gaze and object handling. Furthermore, the specific patterns should not be arbitrary, but should reflect the social, attentional, and referential functions of the utterances.

2 | METHODS

2.1 | Participants

The participants were 42 mother-infant dyads (20 female) from a longitudinal study of infant social development (Chang et al., 2016; Deák, Triesch, Krasno, de Barbaro, & Robledo, 2013). Participants were recruited as a sample of convenience from the greater San Diego area. Mothers’ mean age upon recruitment was 32.1 years (range = 21–42) and they had completed a mean of 16.1 years of formal education (range = 12–21). Twenty-nine infants were Caucasian, two were Asian, four were Hispanic, five were “other” or multiple races, and two parents did not report ethnicity. No infants had any neurological, cognitive, or sensory deficits, according to parental report.

An experimenter visited the participants’ home each month between 3 and 9 months, and again at 12 months, and participants also visited the laboratory to complete various tasks every month. All participants gave informed consent before participating in each session. For the current study, only the 12-month home session was used. Six additional participants were recruited but dropped out of the study before the 12-month session, and four participants were tested but excluded from analyses due to equipment failure or use of non-English language, resulting in a final sample of 38 dyads. Infants’ mean age was 371 days (range: 356–450; SD = 14.5 days); all but one infant was tested before 13 months of age, but excluding this participant did not affect the qualitative results.

2.2 | Procedure

An experimenter visited the participants’ home. Infants and mothers participated in a free-play task in which they were seated on the floor and experimenters provided three different sets of toys (switched out by the experimenter every ~3 min; Figure 1). Similar tasks have been widely used to study infant language learning (Song, Spier, & Tamis-LeMonda, 2014; Sosa, 2016; Suanda et al., 2016; Tamis-LeMonda et al., 2001; Yu & Smith, 2012). Object play tasks elicit rich parental speech with many opportunities for interaction, and are representative of “peaks” in infants’ daily language exposure rather than average periods (Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017). Multiple sets of toys reduced the possibility of spurious results due to specific toys. The first set of toys included a set of three colorful blocks, two wooden ladybugs (red and green), and a chain of plastic rings. The second set included
We analyzed mothers' utterance content types with mixed-effects logistic regression models, using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015) to model the probability of occurrence of each utterance type. These models fit the log-odds of occurrence of a content type as a linear function of the predictor variables, with random intercepts fit for each participant dyad. *p*-values for each coefficient are calculated using likelihood ratio tests between the full model and the model with that predictor removed.

The first set of analyses examined sequences of mothers' utterance content types, testing for which content types predicted the type(s) of the following utterance. We first selected all consecutive utterance pairs, so long as the following utterance occurred within 5 s of the end of the previous one (including all utterance pairs regardless of the delay yielded qualitatively similar results). We then fit separate mixed-effects logistic regression models for each content type as a function of all the types in the previous utterance. Because most content types were most strongly predicted by previous occurrence of the same type, we conducted an additional analysis of temporal and interactive factors affecting the probability of repetition. For each content type, we selected utterances that followed an utterance containing that content type. We then fit mixed-effects logistic regression models predicting the probability of repeating the content type as a function of three factors: time elapsed between utterances, infants' shifts in gaze target, and infants' shifts in object handling.

The next set of analyses examined whether the proportion of utterances with each content type differed as a function of infants' gaze target and infants' object handling. For each content type, we fit mixed-effects logistic regression models with infants' gaze target (FACE, TOY, OR OTHER) and previous utterance content types as predictors, and then with object handling (none, one, or two objects) and previous utterance type as predictors. These models were compared against the baseline models using previous utterance type alone (including these baseline predictors was necessary to remove serial correlation between consecutive utterances).

The final set of analyses examined how both factors—the previous utterance content, and the infant's gaze and object handling state—predicted the content of the next utterance. We fit mixed-effects logistic regression models using previous utterance content,
the infants’ gaze state, and the infant’s object-handling state. We then quantified how much each subset of factors (previous utterance, gaze and hands, both) predicts each content type in the next utterance. Performance was evaluated using subsets of fixed effects estimated from the joint model to predict the content types for participants held out under 10-fold cross validation (Zhang, 1993). Results were quantified using the area under the ROC curve.

To correct for the total number of comparisons, we adopted a stringent α = 0.01 criterion for statistical significance throughout the results (note that a correction for multiple comparisons such as Bonferroni would be inappropriate because correlations among tested variables are predicted by previous results; e.g., Chang et al., 2016; Frank et al., 2013; Suanda et al., 2016). Effects with 0.01 < p < 0.05 are indicated in figures and tables but not discussed in detail. The data and code used to generate the results may be found at: https://github.com/cogdevlabucsd/discourse-continuity-dev-sci.

3 | RESULTS

3.1 | Descriptive statistics

The frequencies of mother’s production of utterances with each content type are shown in Table 2, and the proportion of time that infants spent in each gaze and object-handling state are shown in Table 3.

3.2 | Continuity and sequential structure of discourse

We analyzed the sequential structure of infant-directed discourse by identifying the predictive relations between utterance content types in consecutive utterances (Figure 2). The data included 7,432 utterances, of which 6,535 were preceded by an utterance within 5 s. All utterance types showed significant repetitiveness: values greater than 0 on the main diagonal of Figure 2 indicate that all content types were more likely to occur when the previous utterance contained the same content type (ps < 0.01). Pairs of successive utterance content types that were significantly predictive are highlighted in Figure 2 and listed in Table 4. The presence of significant transitions between different content types indicates that sequential contingencies in content types are a source of regularity in mothers’ utterances, and above content repetition (Figure 3): however, the largest magnitude effects are observed for repetitiveness. Although exact repetitions were sometimes observed, the

<table>
<thead>
<tr>
<th>Content type</th>
<th>Definition</th>
<th>Example</th>
<th>Example also contains</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATIVE</td>
<td>Declarative statements, excluding one-word utterances</td>
<td>That’s pretty neat</td>
<td>–</td>
</tr>
<tr>
<td>IMPERATIVE</td>
<td>Imperative syntax, including “Let’s”</td>
<td>Go get it</td>
<td>ACTION</td>
</tr>
<tr>
<td>QUESTION</td>
<td>Question syntax, or utterance-final pitch rise in appropriate context</td>
<td>What’s over here?</td>
<td>–</td>
</tr>
<tr>
<td>DESCRIPTION</td>
<td>Adjectives or predicates referring to toys</td>
<td>It’s the blue one</td>
<td>DECLARATIVE</td>
</tr>
<tr>
<td>ATTENTION</td>
<td>Directing or commenting on infant’s focus of attention</td>
<td>Can you see mommy?</td>
<td>QUESTION</td>
</tr>
<tr>
<td>ACTION</td>
<td>Directing or commenting on infant’s non-perceptual actions</td>
<td>You wanna try and squeeze it?</td>
<td>QUESTION</td>
</tr>
<tr>
<td>OBJECT NAME</td>
<td>Contains name for physically present toys</td>
<td>We got a little rubber ducky</td>
<td>DECLARATIVE</td>
</tr>
<tr>
<td>INFANT’S NAME</td>
<td>Contains infant’s name or other term of address</td>
<td>Hi [name]</td>
<td>SOCIAL ROUTINE</td>
</tr>
<tr>
<td>AFFIRMATION</td>
<td>Contains a form of Yes or acknowledges a conversational turn without further content</td>
<td>Yeah</td>
<td>–</td>
</tr>
<tr>
<td>SOCIAL ROUTINE</td>
<td>Fixed expressions such as greetings or exclamations</td>
<td>Thank you</td>
<td>–</td>
</tr>
</tbody>
</table>

| TABLE 1 | Discourse content types, definitions of types, and a representative utterance. Because content types are not mutually exclusive, the right-most column indicates additional content types that were tagged in the example utterance, if any

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency/minute: mean (range), SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total utterances</td>
<td>18.1 (8.7–26.4), 4.0</td>
</tr>
<tr>
<td>DECLARATIVE</td>
<td>4.9 (1.9–8.8), 1.8</td>
</tr>
<tr>
<td>IMPERATIVE</td>
<td>3.5 (0.9–7.7), 1.9</td>
</tr>
<tr>
<td>QUESTION</td>
<td>4.9 (2.1–7.9), 1.6</td>
</tr>
<tr>
<td>DESCRIPTION</td>
<td>2.2 (0.2–8.2), 1.5</td>
</tr>
<tr>
<td>ATTENTION</td>
<td>2.3 (0.4–6.0), 1.4</td>
</tr>
<tr>
<td>ACTION</td>
<td>4.0 (1.0–7.5), 1.5</td>
</tr>
<tr>
<td>OBJECT NAME</td>
<td>4.5 (1.2–8.1), 1.6</td>
</tr>
<tr>
<td>INFANT’S NAME</td>
<td>1.0 (0.2–4.1), 0.8</td>
</tr>
<tr>
<td>AFFIRMATION</td>
<td>1.6 (0.3–3.6), 0.9</td>
</tr>
<tr>
<td>SOCIAL ROUTINE</td>
<td>2.0 (0.5–5.1), 1.1</td>
</tr>
</tbody>
</table>

| TABLE 2 | Frequency of utterance content types
repetition coefficients remained positive when exact repetitions were removed from the dataset for all content types except affirmation, and remained significant for all content types except affirmation, imperative, and infant’s name (Table S4). The proportion of exact repeats ranged from 0.08 (declarative) to 0.41 (affirmation). Partial repeats (i.e., at least one overlapping word, but not exact repeats) constituted the majority of repeats for all content types except affirmation and social routine (Table S5).

3.3 | Relations between maternal discourse and infants’ gaze and object handling

To further characterize the repetitiveness observed in the previous section, we investigated whether infants’ activity plays a role in the repetitiveness of maternal discourse. Specifically, we tested whether infants’ gaze shifts or object handling shifts affect the probability of repetition of each content type. Gaze shifts were defined as any difference in the gaze target at the start times of the two utterances, and handling shifts were defined as any difference in the object(s) being handled at the start times of the two utterances. For each content type, we first identified all utterances that followed an utterance with that same type. We then modeled the probability of repetition for each type using mixed-effects linear models with time (i.e., seconds between end of utterance and start of next utterance).

TABLE 3 Proportion of time spent in gaze/object-handling states

<table>
<thead>
<tr>
<th>Variable</th>
<th>Proportion: mean (range), SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaze:0</td>
<td>0.74 (0.27–0.92), 0.15</td>
</tr>
<tr>
<td>Gaze:1</td>
<td>0.03 (0–0.13), 0.03</td>
</tr>
<tr>
<td>Gaze:2</td>
<td>0.23 (0.04–0.68), 0.15</td>
</tr>
<tr>
<td>Handling:0 objects</td>
<td>0.31 (0.01–0.63), 0.13</td>
</tr>
<tr>
<td>Handling:1 objects</td>
<td>0.48 (0.21–0.70), 0.11</td>
</tr>
<tr>
<td>Handling:2 objects</td>
<td>0.21 (0.03–0.43), 0.11</td>
</tr>
</tbody>
</table>

FIGURE 2 Predictive relations between content types in consecutive utterances. Row indicates content types present in the preceding utterance. Column indicates content type being predicted. Values shown are coefficients representing the difference in log-odds when the previous content type is present. *p < 0.05, †p < 0.01, and **p < 0.001. Significance is only shown for positive predictors. Dec: declarative; Imp: imperative; Que: question; Dsc: description; Att: attention; Act: action; ONm: object name; INm: infant’s name; Aff: affirmation; Soc: social routine

3.4 | Predictability of utterance content

We next investigated how much infants’ concurrent gaze and object handling and the previous utterance content contributed to the predictability of utterance content type. Using the mixed-effects logistic regression, we fitted models predicting the probability of each utterance
content type based on the full set of predictors (gaze, handling and previous utterance content). We then used 10-fold cross-validation (assigning each dyad to one fold) to quantify how much fixed effects for gaze/handling, previous utterance content, or both improved predictability over chance levels in sessions not used for model fitting. We evaluated the predictability using the area under the ROC curve (AUC; Table 6). This measure is computed by plotting the true positive rate against the false positive rate as the decision boundary (i.e., the predicted probability of the content type necessary for a positive prediction) is varied. An AUC value of 0.5 indicates chance prediction (true positives do not outnumber false positives for any decision boundary), and 1.0 indicates perfect accuracy (i.e., some decision boundary detects all true positives and no false positives).

Consistent with the previous analyses, prediction was better than chance for most gaze/handling models (AUC from 0.49 to 0.60), for all discourse models (AUC from 0.56 to 0.69), and for all full models (AUC from 0.58 to 0.70). The previous utterance content was generally a stronger predictor than gaze and handling information. However, gaze/handling state was a stronger predictor than previous utterance content for infant’s name. The two factors performed approximately equally well for predicting affirmation (Figure 6). For most content types, the model using only previous content performed as well as (i.e., AUC within 0.01 of) the full model, indicating that gaze/handling information was redundant with previous content; however, for utterances with infant’s name and object name, both sources of information made independent contributions.

4 | DISCUSSION

The current results confirm that a combination of sequential discourse structure and concurrent infant gaze and manual activity predict a wide range of content types in naturalistic maternal speech to 12-month-old, English-learning infants. Unlike previous studies on the linguistic and action context of IDS (e.g., Frank et al., 2013; Trueswell et al., 2016; Yu & Smith, 2012), these effects were not measured in terms of relations between specific words and their referents. Nor were they limited to concrete utterances referring to observable objects or actions. Instead, sequential and cross-modal relations were found for a wide range of abstract content types that are prevalent in IDS and relevant to language acquisition.

Maternal speech exhibited pervasive repetitiveness, in that utterances with a given type of content were likely to be followed by one or more successive utterances with the same type of content. Notably, this pattern held across all the 10 syntactic, thematic, and lexical content types tested. For most content types, most repetitions showed only partial (not complete) overlap with the previous utterance. In addition, discourse type repetition was more likely when utterances were closely spaced in time.

Repetitiveness in IDS has been documented previously: it is known that caregivers frequently produce both exact and partial repetitions (Hoff-Ginsberg, 1990; Kaye, 1980; Snow, 1972). Infants, notably, prefer the expected level of repetitiveness of IDS (McRoberts, McDonough, & Lakusta, 2009). This repetition seems to have specific pragmatic functions: for example, parents’ consecutive utterances tend to refer to the same object even after taking into account infants’ and parents’ ongoing and coordinated gaze and object handling (Frank et al., 2013). Moreover, this structured repetition impacts language learning. Continuity in discourse has been found experimentally to facilitate word learning: 2-year-old learn new object words more effectively when they are repeated in consecutive

### TABLE 4
Significant sequential structure in mothers’ discourse. Effects are listed if their coefficients were positive with \( p < 0.01 \)

<table>
<thead>
<tr>
<th>Content type</th>
<th>Predicts</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATIVE</td>
<td>DECLARATIVE</td>
</tr>
<tr>
<td>IMPERATIVE</td>
<td>IMPERATIVE</td>
</tr>
<tr>
<td>QUESTION</td>
<td>QUESTION</td>
</tr>
<tr>
<td>DESCRIPTION</td>
<td>DESCRIPTION</td>
</tr>
<tr>
<td>ACTION</td>
<td>ACTION</td>
</tr>
<tr>
<td>OBJECT NAME</td>
<td>OBJECT NAME</td>
</tr>
<tr>
<td>INFANT’S NAME</td>
<td>INFANT’S NAME</td>
</tr>
<tr>
<td>AFFIRMATION</td>
<td>AFFIRMATION</td>
</tr>
<tr>
<td>SOCIAL ROUTINE</td>
<td>SOCIAL ROUTINE</td>
</tr>
</tbody>
</table>

**FIGURE 3** Graph of above-chance transitions in mothers’ discourse. Significant predictive relations are shown as bold arrows. Then, smaller dashed arrows were added for the highest remaining coefficients until the graph was fully connected.

### TABLE 5
Effects of time between utterances and infant attention shifts on probability of content type repeats

<table>
<thead>
<tr>
<th>Content type</th>
<th>Time (s)</th>
<th>Gaze shift</th>
<th>Hand shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATIVE</td>
<td>−0.36**</td>
<td>−0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>IMPERATIVE</td>
<td>−0.20**</td>
<td>−0.20</td>
<td>0.08</td>
</tr>
<tr>
<td>QUESTION</td>
<td>−0.15*</td>
<td>0.15</td>
<td>−0.10</td>
</tr>
<tr>
<td>DESCRIPTION</td>
<td>−0.54**</td>
<td>−0.08</td>
<td>−0.10</td>
</tr>
<tr>
<td>ATTENTION</td>
<td>−0.17†</td>
<td>−0.04</td>
<td>0.07</td>
</tr>
<tr>
<td>ACTION</td>
<td>−0.19**</td>
<td>−0.17</td>
<td>−0.30†</td>
</tr>
<tr>
<td>OBJECT NAME</td>
<td>−0.05</td>
<td>0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>INFANT’S NAME</td>
<td>−0.06</td>
<td>−1.21*</td>
<td>0.30</td>
</tr>
<tr>
<td>AFFIRMATION</td>
<td>0.02</td>
<td>0.01</td>
<td>−0.45</td>
</tr>
<tr>
<td>SOCIAL ROUTINE</td>
<td>−0.26*</td>
<td>−0.44†</td>
<td>0.09</td>
</tr>
</tbody>
</table>

\( \text{†} p < 0.05, \text{*} p < 0.01, \text{and **} p < 0.001. \)
utterances (Schwab & Lew-Williams, 2016), and toddlers can use continuity of reference to infer the referent of otherwise ambiguous utterances (Horowitz & Frank, 2015). Effects of repetition on learning might be more general: for example, repetitive exposure to the same individuals in time has been proposed to facilitate learning in other domains such as face processing (Jayaraman & Smith, 2018).

Our findings demonstrate that content repetition in IDS is more general than previously recognized, in that discourse continuity is not specific to one or a few types of linguistic content. If infants use the continuity of adjacent utterances to reduce the ambiguity of object-referential utterances, they could potentially use the same mechanism to more accurately identify the communicative function of a wide variety of infant-directed utterances.

In addition to repetition of content types, we also found significant sequential patterning of mothers’ use of different content types across utterances, demonstrating that continuity of infant-directed discourse also includes sequential transitions between content types. We identified sequences or paths of frequent transitions, such as infant’s name → attention → declarative and infant’s name → attention → description. These sequences might help to focus infants’ attention on the most informative utterances by first capturing the infant’s attention by saying their name, then orienting the infant to
Infants might initially experience such sequences while spontaneously exploring their environment or following overt cues. However, as they learn to predict and comply with caregivers’ attentional bids, infants might also increasingly choose actions that facilitate the emergence of more complex and reliable sequences.

Consistent with this idea, our results show systematic patterning of mothers’ discourse content depending on infants’ state of engagement with the physical and social environment. Although we did not find general effects of infant attention shifts on content type repetition, there was at least one specific effect: a lower rate of repetition of the infant’s name after infant gaze shifts, suggesting that infants’ reorientations in relation to maternal utterances might facilitate sequential transitions in discourse. Given that only two infant behavioral states were coded, a more thorough characterization of infant states might reveal more specific relations to the caregiver’s speech repetition. Additionally, several content types differed in frequency based on infants’ attention states. When infants gaze at toys, they hear more descriptive language and more object names; when they gaze at faces, they hear more affirmations, and when they gaze away from faces and toys, they hear more imperatives, their own name, and attention-related language. Similarly, when infants handle one or more toys, they hear more descriptive language; whereas when they do not handle objects, they hear more imperatives and their own name. These results complement and expand a sizable literature (reviewed above) on maternal speech about objects, and maternal responsiveness in dyadic and triadic interactions.

Some of these differences might be due to mothers responding to infants’ attentional engagement. Nevertheless, such associations could still help establish the pragmatic force of some utterances as regulatory and others as informative. Furthermore, some utterance types are observed to differ in frequency across “on-task” states. Notably, affirmations are associated with gaze at the mother’s face rather than at objects, reflecting their interpersonal rather than referential role. Also, whereas descriptions are associated with both gaze to and handling of toys, object names were associated only with gaze.

We also compared how much the infant’s gaze and object handling and/or the previous utterance content predicted occurrences of each utterance content type. One or both types of contextual factors predicted every content type above chance, although substantial uncertainty remained for all content types, suggesting that other predictive contextual factors have not yet been identified. Previous utterance content was generally a stronger predictor than concurrent gaze and object handling, but the relative importance of the cues differed substantially across content types. Thus, some discourse types were typically produced within a longer discourse sequence, whereas other functions were more immediately responsive to the infant’s actions. At one extreme, object names and descriptions were highly dependent on the previous utterance, but occurred frequently in all configurations of the infant’s gaze and object handling; at the other, the infant’s name occurred in response to infants’ gaze away from toys and the mother’s face but depended only weakly on previous utterances. These differences can reflect the interacting roles of different discourse types: for instance, mothers typically did not talk about attention as a direct consequence of an infant’s disengagement with the toys, but rather oriented the infant to an interesting object only after first capturing their attention by calling their name.

Previous research suggests how correspondences between infants’ actions and parents’ speech might facilitate language learning. In the well-studied case of object-name learning, 17- to 18-month-old infants are more likely to learn words that are uttered during episodes of joint attention (Baldwin, 1991; Tomasello & Farrar, 1986), and 18-month-old infants learn more object labels that are presented while the object dominates their visual field (Yu & Smith, 2012). The effectiveness of object naming, therefore, depends not only on the correspondence between object and label, but also on a supportive context in which the utterance is expected and appropriate. The current results suggest that nonverbal context could similarly support the usage of a range of referential (descriptions, object names, talk about action) and regulatory (attention, imperative, infant’s name, affirmations) discourse types. Our approach complements other empirical studies of structured relations between maternal speech and concurrent activities that are predictable within a variety of contexts (e.g., diaper-changing; Nomikou & Rohlfing, 2011). Characterizing language input at the level of associations between maternal discourse content and infants’ embodied behavioral states is also consistent with
the proposal that infants represent linguistic meanings not as specific objects or events, but as frames of routinized goal-directed activity (Rohlfing, Wred, Vollmer, & Oudeyer, 2016). As infants learn their role in interactive sequences, they might further improve their ability to participate in and learn from routines, eventually leading to the emergence of conversation. The current results characterize speech-activity relations experienced by infants at a particularly important developmental stage, as predictable, structured activities play an increasingly complex and integral role in 1-year-old social interactions (De Barbaro, Johnson, & Deák, 2013; Nelson, 1998), despite their still immature and incomplete linguistic knowledge and symbolic skills.

Discourse continuity in IDS might facilitate language acquisition in several ways. Partial repetitions of parental utterances correlate positively with children’s syntactic development (Baker & Nelson, 1984; Hoff-Ginsberg, 1985, 1986). Discourse continuity has been previously argued to facilitate syntax acquisition because partial overlap in adjacent utterances constrains word and phrase boundaries (Onnis, Waterfall, & Edelman, 2008), and facilitates lexical acquisition by clustering object-labeling events (Schwab & Lew-Williams, 2016; Vlach & Johnson, 2013). Structured sequences in discourse and interaction might cue infants to expect utterances with novel information such as object names or descriptions. In fact, these were the two most predictable content types in our data. Our finding of significant transitions from attention-regulating to declarative and descriptive utterances suggests that attention-directing talk might be partially responsible for this predictability. This raises the possibility that findings of negative correlations between directive maternal talk and child language outcomes (Hughes et al., 1999; Nelson, 1973; Tomasello & Todd, 1983) are related to the disruption of these sequences, perhaps due to infants’ failure to respond to bids for attention, rather than differences in speech style per se.

Sequential continuity and repetitiveness might further facilitate learning of syntactic and/or pragmatic categories if continuity effects are stronger for valid categories than for arbitrary ones. Repetition in discourse might highlight similarities among exemplars of a content type, whereas sequences emphasize contrastive features (Carvalho & Goldstone, 2017). Moreover, infants might better recognize repeated word tokens if prosody or other utterance-level features are consistent across utterances (Bortfeld & Morgan, 2010). Consistent sequences in discourse might also make it possible to infer similar meanings between utterances that occur in similar discourse contexts, analogous to inferring word meaning by familiar sentence context (Miller & Charles, 1991). Once children learn to detect discourse continuity, contextual utterances could constrain the interpretation of utterances that would be ambiguous in isolation. This effect has been called discourse bootstrapping. For example, by 2 years of age children can use discourse continuity to select a referent for an unfamiliar word suggested by a speaker’s previously expressed preferences (Sullivan & Barner, 2016), and parents use multiple cues such as syntactic frames, utterance-final position, and prosodic emphasis to introduce unfamiliar words to 2- to 5-year-old children (Clark, 2010). Although we cannot yet determine how much discourse bootstrapping facilitates meaning-inference relative to syntactic and semantic bootstrapping, the current results indicate that information supporting the former is available by 12 months.

Finally, discourse continuity might facilitate language acquisition via general information processing factors such as increased processing speed (see Conway, Bauernschmidt, Huang, & Pisoni, 2010) and reduced working memory load. These mechanisms might bridge correlational findings indicating that parental speech style influences language acquisition (Barnes, Gutfriend, Satterly, & Wells, 1983; Hart & Risley, 1995; Tamis-LeMonda et al., 2001) with mechanistic accounts of how infants find structure in linguistic input (Hoff & Naigles, 2002; see also Tamis-LeMonda, Kuchirko, & Song, 2014).

The current study has several limitations. We observed mother-infant dyads interacting in a restricted context that captured only a limited portion of their daily activities. Also, the dyads were a sample of convenience of mostly middle- to upper-class families that did not faithfully represent local population demographics with respect to education and ethnicity. This limits the generality of the findings—specifically, the patterns of experience found in this sample might not hold in different contexts or populations. However, context-specific structure in infants’ language exposure might facilitate language learning over and above regularities that are evident when aggregating over all contexts (Beals & Tabors, 1995; Roy, Frank, DeCamp, Miller, & Roy, 2015; Roy, Frank, & Roy, 2012). Another limitation of the current study is the necessarily somewhat arbitrary classification of verbal content types. Further research could attempt to derive a “bottom-up” taxonomy based on raw verbal data, and investigate how much the sequential and cross-modal structure of maternal speech affords unsupervised discovery of content types.

The current study simplified interactions that unfold over time by studying state transitions and concurrent co-occurrences between utterances and other events. Although a lagged analysis of relations between parental speech and infants’ past attention states is outside the scope of the study, concurrent relations between infants’ attention state and mothers’ speech can also be interpreted as contingent responses to infant state changes. However, mothers’ speech might also depend on sequences and durations of infant activities. Infants’ object handling, for example, tends to occur later in episodes of shared attention to objects, compared to object gaze (Chang et al., 2016; Rohde & Frank, 2014). Therefore, another layer of contingency remains to be revealed. Future work also could explore how parents’ discourse patterns are related to infants’ comprehension of specific words or constructions, as the current study and related studies have not disaggregated contingencies based on the familiarity or novelty of linguistic elements to the infant. In addition, an extension of the current study could examine discourse contingencies at multiple ages because parents integrate speech into play differently depending on the infant age or developmental status (Bornstein et al., 2008; Murphy, 1978). Finally, future research might improve models by adding individual differences in caregiver speech (e.g., level of discourse repetitiveness) as a predictor of infants’ emerging speech processing fluency or early linguistic knowledge.
5 | CONCLUSION

The current results demonstrate that mothers’ speech to 12- month-old infants is repetitive at multiple levels of analysis, is sequentially structured, and is contingently related to infants’ exploratory activity. All of the syntactic, thematic, and lexical content types that we examined tended to occur preferentially in specific positions in the discourse. Also, most types showed significant associations with infants’ ongoing gaze and/or manual activity, even when ignoring specific word-object pairings. Yet the specific predictors of each utterance content type varied: some types were more repetitive, others more sequential, others more responsive to infant actions. This richly structured input thus provides opportunities for infants to distinguish between pragmatic functions of different content types in caregiver speech. The patterns we identified might also reflect mothers’ efforts to ensure that informative utterances occurred at times when their infant was attentive to the referent object, and therefore might expect informative input to quickly follow. If infants are sensitive to these regularities, then the diversity of structured discourse contexts might constitute a source of information that facilitates, rather than complicates, learning. Future research on language acquisition should account for such information within connected, multimodal, interactive discourse, thereby more fully explicating the mechanisms by which social interaction endows infants’ linguistic experience with an organizing structure.

ACKNOWLEDGEMENTS

We would like to thank Jordan Danly for recording the sessions, Shirlene Wade for supervising the speech transcription, and Scott Ragland, Kimberlì Ng, Tiffany Ziebol, Vivian Zhang, Maria Kremljovsky, Christen Buetz, and Erin Maxwell for assisting with coding. This research was supported and by the National Science Foundation (NSF) Graduate Research Fellowship under grant no. DGE-1144086, NSF grant SES-0527756 to G. Deák and J. Triesch, and by the Temporal Dynamics of Learning Center (NSF SMA 1005256).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Chang LM, Deák GO. Maternal discourse continuity and infants’ actions organize 12-month-olds’ language exposure during object play. Dev Sci. 2018;e12770. https://doi.org/10.1111/desc.12770