Preschoolers are more likely to direct questions to adults than to other children (or selves) during spontaneous conversational acts

Koeun Choi1 (koeunchoi0@gmail.com), Elizabeth Lapidow2 (elapidow@ucsd.edu), Jennifer Austin3 (jaustin@newark.rutgers.edu), Patrick Shafto4 (patrick.shafto@gmail.com),
Elizabeth Bonawitz1 (lbaraff@gmail.com)

1 Department of Psychology, Rutgers University, Newark, NJ 07102
2 Department of Psychology, UC San Diego, La Jolla, CA 92093
3 Department of Spanish and Portuguese Studies, Rutgers University, Newark, NJ 07102
4 Department of Math and Computer Science, Rutgers University, Newark, NJ 07102

Abstract

Question asking is a prevalent aspect of children’s speech, providing a means by which young learners can rapidly gain information about the world. Although past work demonstrates that children are sensitive to the knowledge state of potential informants (e.g., Koenig & Harris, 2005), less work has explored whether children spontaneously direct questions to adults over other children (who are less likely to be knowledgeable), and in particular if adult-directed questions focus on content that is more likely to support general learning. We recorded individual children’s spontaneous speech in 40-minute sessions during their preschool day; for every production we coded whether the speech was directed towards an adult, another child, or was stated to self. Our results (N = 30, totaling 2,232 utterances) showed that questions took up a greater proportion of children’s adult-directed speech as compared to the proportion of questions in child-directed and self-directed speech. Furthermore, although children asked many kinds of questions (including conversational clarifications, specific information questions, and questions intended for general learning), children more frequently asked the questions intended for learning when they spoke to adults than to the other groups. Analysis revealed a developmental effect, with results strongest for the older preschoolers. Our findings suggest that children discriminatingly choose “what” and “whom” to ask in daily conversations, and this ability improves over the course of development.

Keywords: question asking; selective, spontaneous speech

Introduction

Even young learners play an active role in their learning (Piaget, 1930). The benefits of self-directed learning in early childhood have been supported by a large body of literature on exploratory play (Bruner, Jolly, & Sylva, 1976; Singer, Golinkoff, & Hirsh-Pasek, 2006). Similarly, in cognitive psychology, active exploration has been considered as a critical component of learning, which allows interventions that help deconfound variables (Schulz, Kushnir, & Gopnik, 2007, 2007; Schulz, 2012; Schulz, Gopnik, & Glymour, 2007). Children’s exploration has been compared to scientific endeavors, including observation, hypothesis generation, and intervention to understand the world around them (Carey, 1985; Gopnik, Meltzoff, & Bryant, 1997; Wellman & Gelman, 1992). Question asking provides another means by which children can “explore” to learn, and may provide a particularly powerful way to quickly learn about the world (Gopnik et al., 1997; Harris, 2012).

Question asking is prevalent in young children’s day-to-day conversations (Callanan & Oakes, 1992; Chouinard, 2007). According to a daily diary study, children aged 1-5 years asked on average 76 questions per hour to acquire information during their conversation with adults (Chouinard, 2007). Children also make an effective use of questions to gain information from others (Greif, Kemler Nelson, Keil, & Gutierrez, 2006; Kemler Nelson, Egan, & Holt, 2004). That is, children tailor their questions to request domain-specific information (Callanan & Oakes, 1992; Greif et al., 2006), and monitor the answers they received to form follow up questions (Frazier, Gelman, & Wellman, 2009; Kemler Nelson et al., 2004). Results from these studies suggest that children actively use questions in their conversations to obtain information. However, less is known about whether children choose particular individuals over others as the targets of their question asking.

Epistemic trust and active learning

One reason to believe that even very young children may be able to consider relevant targets for questions is that children are relatively savvy reasoners about other’s knowledge states, intentions, and pedagogical goals. For example, children draw inferences about pedagogical goals of teachers from available information to guide their future exploration (Bonawitz et al., 2011; Gweon & Asaba, 2017; Gweon, Pelton, Konopka, & Schulz, 2014; Eaves & Shafto, 2012). Moreover, the large literature on epistemic trust has documented that children make inferences about sources of information based on the evidence provided to them (Koenig & Harris, 2005; Kushnir & Koenig, 2017; Pasquini, Corriveau, Koenig, & Harris, 2007; Sobel & Corriveau, 2010). This line of work has shown that children selectively learn from informants who provide accurate information or try to be helpful (Shafto, Eaves, Navarro, & Perfors, 2012; Koenig, Clément, & Harris, 2004; Mascaro & Sperber, 2009).

In fact, there is evidence that even young children exhibit the ability to adjust their questions based on the information about sources (Mills, Legare, Bills, & Mejias, 2010; Mills & Landrum, 2016; Mills, Legare, Grant, & Landrum, 2011). Mills et al. (2011) presented two puppet informants with varied knowledge state and allowed children to freely direct questions. Children asked more questions to knowledgeable informants over ignorant ones by age 3, suggesting that young children are equipped with a basic skill set to direct questions to the appropriate sources of informa-
tion. However, there were developmental differences as well. When knowledgeable informants were contrasted with inaccurate guessers, only 5-year-olds were able to selectively direct questions to knowledgeable sources, suggesting developmental progression in children’s question-asking strategies.

Do children actively seek out better informants in naturalistic social contexts? Although several studies have explored children’s questions depending on the knowledge state of information sources, the opportunities for asking questions in these studies were driven by the explicit information about the informants situated in environments that were structured by researchers (Mills et al., 2010; Mills & Landrum, 2016; Mills et al., 2011). Effective learners may make choices about whom to ask prior to observing explicit evidence. Such a learner must be able to weigh the probability of gains in the environment, consider the opportunity for learning in the current case, and even make choices among competing opportunities for investigation.

One possibility is that children seek out information selectively, depending on the assumptions about the knowledge state of others. That is, conversation with knowledgeable and fully fluent adults may provide opportunities to gain information by directly asking questions. In the absence of any prior information, preschoolers choose to learn from an adult over a child, suggesting that preschoolers assume that adults are better sources of information than children (Jaswal & Neely, 2006). However, it is not known how children play an active role in utilizing this understanding in their everyday inquiry. In querying others, children may use their understanding about others to evaluate potential information sources. Children may gauge the kinds of information others could provide depending on an informant’s expertise. Based on this judgment, children may ask for tailored information from differently skilled others.

**Study of preschooler’s spontaneous question asking**

The objective of this study was to explore naturally occurring questions in a pedagogical setting (preschools). Of particular interest was whether there are systematic differences in the kinds of speech acts (e.g. questions) preschoolers engage in, depending on the target of the conversation. In the current study, the speech productions of children ages 3 to 6 years were recorded while a live coder noted the contextual information and to whom each individual speech act was directed. This extensive coding allowed us to compile a detailed record of questions produced by children as well as the social contexts of the questions during the recording sessions.

Based on what is known about developmental changes in epistemic trust (Mills et al., 2010; Mills & Landrum, 2016) and preschoolers’ assumptions that adults are better information sources (Jaswal & Neely, 2006), we expect the preschoolers in our sample to identify teachers in the classroom as more knowledgeable than other students. If children can actively utilize this understanding to maximize their opportunities for learning, we would expect systematic differences in children’s questions by the target of speech. Specifically, we hypothesized three main effects. First, we predict children would be more likely to seek out information by asking questions when speaking with knowledgeable adults than when speaking to other (less knowledgeable) children or when speaking aloud to themselves. Second, the relationship of question to target should be particularly pronounced for “learning content” questions. That is, children are expected to direct more questions that support general learning to adults than the other groups; however, other kinds of questions (such as communication clarification questions or specific information questions) may not depend on the target being an adult. Third, given that older children are likely to have more experience asking questions, we expected developmental improvements in the efficiency of directing questions to information sources across early childhood. In other words, older children should be more proficient at directing more questions to adults than younger children.

**Methods**

**Participants** Participants were 30 preschoolers ($M = 4.3$ years, $SD = 0.92$ years, Range = 37-76 months, 50% Female) who participated in their respective classrooms during their free play time. Each classroom was led by a female teacher. The classroom size was an average of 15 children. The sample of participants spanned five classrooms all located in Essex County, NJ, which is one of the most racially and economically diverse counties in the United States1. We intentionally recruited from these sites to ensure the diversity of the sample. Prior to the study, consent from the sites, participating families, and the internal review board were obtained.

**Procedure** Participants were introduced to a clip-on microphone that attached to the child’s shirt and recorded continuously for the full 40-minute session in the child’s classroom. During the session, an experimenter observed surreptitiously from the corner of the classroom. Using time-stamps to align with the audio recordings, the experimenter independently recorded every action and speech partner the participant was engaged with during the session. After the session ended, audio data were transcribed by another researcher. These transcriptions were linked with the written data using the synced time-stamps. This produced a master data set which included the transcribed utterances, contextual cues, and targets of each speech act. Coding and analysis proceeded from the master data set.

**Coding** We coded a total of 3,341 constructions ($M = 111.37$ per child). Of the total constructions, 499 were excluded because of one of the following: non-words ($n = 22$), sound ($n = 361$), and unidentified constructions ($n = 116$). The remaining 2,842 constructions were coded as utterances ($M = 94.73$ per child).

For every utterance, we coded whether the speech was di-

1 As measured by the Gini Index (U.S. Census Bureau, 2016)
Table 1: Examples of questions for each coding type

<table>
<thead>
<tr>
<th>Question type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning</td>
<td>Do you know what this is?</td>
</tr>
<tr>
<td>Question</td>
<td>Is this a sponge?</td>
</tr>
<tr>
<td></td>
<td>How do you write F?</td>
</tr>
<tr>
<td>Communication</td>
<td>The boy’s table?</td>
</tr>
<tr>
<td>Question</td>
<td>Wait, what you mean?</td>
</tr>
<tr>
<td></td>
<td>What is that again?</td>
</tr>
<tr>
<td>Information</td>
<td>Can I get some water?</td>
</tr>
<tr>
<td>Question</td>
<td>Is my mom gonna be here soon?</td>
</tr>
<tr>
<td></td>
<td>This is my nap?</td>
</tr>
</tbody>
</table>

rected towards an adult, another child, or self. Of 2,842 utterances, 610 were excluded because either the speech partner was unidentifiable \( n = 417 \) or child speech was directed to a mixed group \( n = 193 \). The remaining 2,232 utterances were included in analyses \( M = 74.4 \) per child.

Two independent coders coded each of the 2,232 utterances. The coders first identified whether each utterance was a question. The interrater reliability between the two independent coders was high; Krippendorf \( \alpha \) was .93.

 Coders were also trained to identify different kinds of questions, dividing questions into one of three subcategories based on speakers’ intention (Yu, Bonawitz, & Shafto, 2017): (a) Learning questions can be interpreted as intended to learn, where learning is broadly defined as acquiring general or specific knowledge about objects, people, or events, (b) Communication questions can be interpreted as intended to be rhetorical (giving commands in the form of questions) or simply asking for clarification by repeating what others said, (c) Information questions can be interpreted as intended to get situation-specific information such as seeking permission, checking the physical status of object, or checking the physical, emotional, epistemic status of others. The examples for each question type are included in Table 1. For the three question types, agreement was high; Cohen’s \( k \) was .78. Discrepancies in codes were reviewed and resolved by a third coder.

Results

Evaluating the speech partner of question asking Children’s speech was directed to other children (59%) more frequently than to adults (22%) or themselves (19%). To control for the difference in the amount of speech by speech partner and an uneven child to adult ratio in each classroom, we examined the proportion of questions over other possible speech acts based on the target of speech production.

Overall, about 19% \( n = 414 \) of the preschoolers’ total utterances \( n = 2,232 \) consisted of questions. Questions took up a greater proportion (24%) of children’s adult-directed speech, compared to only 19% towards other children and 11% towards themselves, \( \chi^2(2) = 25.11, p < .001 \). As in Figure 1, and consistent with our hypothesis, children were more likely to direct their questions to adults than the other groups \( \chi^2(1) = 5.24, p = .022; \chi^2(1) = 12.52, p < .001 \). The error bars represent the standard error.

To examine individually calculated proportion, we conducted a mixed effect logistic regression (Level 1: Speech Partner, Level 2: Participant) with participants as random effects and with speech partner as a fixed effect to predict the probability of utterance being a question, using lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). Consistent with the finding based on the aggregated data, the model showed a significant effect of speech partner: Children directed relatively more questions to adults than their peers (\( \beta = -.31, p = .039 \)) or themselves (\( \beta = -.89, p < .001 \)).

Examining the types of questions posed also revealed significant differences. Overall, of all the questions \( n = 414 \), the majority were classified as information questions (58%, \( n = 241 \)) compared to 28% conversation questions \( n = 117 \) and 14% learning questions \( n = 56 \). Critically, we examined the proportion of each of these types of questions by recipient. Learning questions took up a greater proportion (28%) of children’s adult-directed questions, compared to only 5% for other children and 4% for themselves (Fisher’s exact, \( p < .001 \); Figure 2). In other words, children were more likely to direct their learning questions to adults than other children \( (\chi^2(1) = 57.32, p < .001) \) or themselves \( (\chi^2(1) = 14.71, p < .001) \). A mixed effect logistic regression model revealed a consistent pattern: The probability of question intended for learning (1 = Learning Question; 0 = Communication or Information Question) was significantly predicted by the recipient of questions. That is, children directed more learning questions to adults than their peers \( (\beta = -2.39, p = .001) \) or themselves \( (\beta = -2.74, p = .002) \).

Developmental changes in questioning asking Finally, we explored the degree to which question asking strategies might differ across the preschool years. We took a median split of our population and compared younger preschoolers \( 37\text{-}56\text{-month-olds, } M = 48.4, SD = 6.13 \) to older
Children can systematically direct questions to speech partners. Learning questions consisted of greater proportion of adult-directed questions than child- or self-directed questions. Error bars denote SE.

Preschoolers (57-month-olds to 76-month-olds, $M = 67.2$, $SD = 5.52$). Younger and older children spent about the same proportion of their utterances on questions (18% vs 19% respectively, $\chi^2(1) = 0.40, p = .528$). However, older children were more likely to direct those questions to a knowledgeable adult than younger children (27% vs 17%, $\chi^2(1) = 5.42, p = .020$; Figure 3). Consistent with this finding, a separate mixed effect model for each age group showed that the probability of utterance being a question was higher when questions were directed to adults than children ($\beta = -0.47, p = .011$) or themselves ($\beta = -1.08, p < .001$) only for the older group, but not for the younger group ($ps > .250$). These results suggest that children’s question asking strategies improve over the course of development.

We also compared whether the type of questions differed by age. Replicating our main finding, in the younger age group, children were more likely to ask learning content questions to adults (62%) than to other children (6%) or themselves (0%; Fisher’s exact, $p < .001$). The same pattern was found in the older group such that older preschoolers directed more learning content questions to adults (28%) than to the other groups (4% and 6%; Fisher’s exact, $p < .001$). However, there was no significant age differences in learning questions ($\chi^2(1) = 0.46, p = .497$). A separate mixed effect model for each group confirmed this finding: The probability of question intended for learning was higher for adult-directed questions than child-directed questions ($\beta_{young} = -2.24, p = .023$; $\beta_{old} = -1.92, p = .003$) for both age groups. This suggests that when asking learning questions, even young children can systematically direct questions to speech partners.

**Discussion**

Question asking is widespread in children’s spontaneous conversations. The findings of this study suggest that children systematically direct questions to actively seek out information from others. Consistent with research on epistemic trust, we found that children exhibit the sensitivity to weigh information gains from different social partners. The proportion of questions in adult-directed speech was higher than those in child-directed and self-directed speech. Critically, children generated more questions with the intention of learning when talking to adults than to the other groups. Furthermore, older children directed more questions to adults relative to younger children, suggesting that this ability increases with age.

Prior work has demonstrated that children can use principles such as information gain to effectively choose which questions to ask (Ruggeri & Lombrozo, 2015). In our study, children directed more questions to adults than other children or themselves. In particular, adult-directed questions focused on content that is more likely to support general learning. These findings suggest that children are sensitive to who is informative and likely to provide helpful feedback. This initial result provides evidence to support the claim that children are not only being guided by these inferences about others (Koenig & Harris, 2005; Sobel & Corriveau, 2010), but also actively using this knowledge to decide who to seek out for information.

In our study, older children directed relatively more questions to adults than younger children. This developmental progression in children’s question asking coincides with the developmental trajectory of epistemic trust (Koenig & Harris, 2005; Pasquini et al., 2007). Computational work on epistemic trust explains developmental differences as a result of increasing understanding that not everyone is helpful (Shafto et al., 2012; Eaves & Shafto, 2012). That is, these models suggest that younger children have a stronger default expectation that other people are helpful. Research on infants’ pointing provides support for selective information seeking from early on (Begus & Southgate, 2012; Franco, Perucchini, & March, 2009). However, our finding and those of other
studies on question asking (Mills et al., 2011; Mills & Landrum, 2016) showed developmental changes in early childhood. Thus, verbally communicating intentions via questions may require additional experiences to overcome younger children’s strong assumptions.

Our data were collected during the free play time, and there are potential contextual factors that may have contributed to children’s conversations. For example, the presence of adults during free play was relatively infrequent compared to that of children. Thus, it is not surprising that the majority of the conversations were directed to other children. Furthermore, teachers may intentionally encourage children to communicate with other children during this period rather than engaging with children directly. We further explored whether speech partners provided explicit cues prompting children to ask questions (e.g., "Do you have any questions for me?" "Go ask Mrs. [Teacher]"). However, only one of the 414 questions were preceded by these explicit cues. Thus, children appear to selectively orient their questions without direct guidance. Future work could explore whether the presence of implicit cues in the context such as pedagogical questions (e.g., "What are the colors of the rainbow?") is related to the frequency and type of questions.

It is possible that other demographic or social factors may influence the quantity and content of questions. Regarding socioeconomic status (SES), Kurkul and Corriveau (2018) reported that low-SES children asked only half the number of questions to their caregivers compared to mid-SES children despite no difference in question type. It would be an important next step to examine the extent to which socioeconomic backgrounds contribute to children’s question asking. In addition, when choosing from whom to learn, children tend to rely on an informant who shares the same culture or gender (Kinzler, Corriveau, & Harris, 2011; Wood, Kendal, & Flynn, 2013). In our study, all of the teachers and researchers were female, but children’s questioning patterns did not differ by child’s gender. We think that future research that includes the diverse speech partners can further explore the culture or gender match between children and their speech partners.

We find it interesting that a relatively large portion of children’s spontaneous speech, and questions in particular, were self-directed. When and why might children spontaneously ask questions of themselves during play? Future work could explore the factors surrounding this phenomenon, which would be related to self-explanation effects (Chi, De Leeuw, Chiu, & LaVancher, 1994; Rittle-Johnson, 2006) or working memory and meta-cognitive monitoring (Schneider, 2008).

A contribution of this work is a database that extends beyond simple transcriptions of naturalistic speech acts to include time-linked contextual cues and partners of speech acts. We believe this database will provide an important resource for language and development researchers. For example, because we also coded information such as mean length of utterance (MLU) as well as conjugation errors and awkward constructions, we believe this resource could be valuable to linguists who may be interested in the relationship between these production types and the speech partner.

Most prior studies on naturalistic question asking behaviors focused solely on adult-child interactions. To our knowledge, our study is one of the first attempts to quantify children’s questioning involving not only adults but also other speech partners in a preschool setting. Children spend most of their waking hours during the week in preschools; these classrooms provide the potential for important pedagogical opportunities. Thus, it remains an important area of study for our understanding of the environmental factors influencing early childhood development.

Children ask abundant questions in early childhood (Callanan & Oakes, 1992; Chouinard, 2007). We observed children’s selective information seeking even in their unstructured and spontaneous social interactions. This data provides support for the claim that children also consider who to ask discriminately in their day-to-day conversations.

Acknowledgments

This research has been supported by NSF-SES 1627971 to L.B. and the Rutgers University SEED award to L.B., P.S., J.A. We thank teachers, parents, and children of Little Scholar Preschool for their participation and to Courtney Bell, Sara Tariq, Mary Manansala, and Priscilla Mejia for help in conducting the study.

References


---


