Children Do Not Overuse "the" in Natural Production

Yuanfan Ying, Valentine Hacquard, Alexander Williams, and Jeffrey Lidz

1. Introduction

In adult English, when using a singular definite description "the N", a speaker presupposes the existence of an object that is either the only object satisfying N in the domain of the conversation (Russell, 1905) or already "familiar" in the discourse or common ground (Strawson, 1950; Heim, 1982; Roberts, 2003). In acquisition studies, young children have been claimed to produce a great proportion of definite phrases in indefinite-eliciting contexts where intended referents are neither unique (Wexler, 2011) nor in the common ground (Maratsos, 1976). For instance, an indefinite description might have been more appropriate for children to use in (1) ("Get a window"), as the investigator's response implies that there is perhaps not a uniquely identifiable referent in the shared context.

(1) Eve: (Have a) get the window.
Investigator: Which window? (Brown, 1973: Eve, 02;00;00)

This work investigates whether young children overuse definites in natural production. We examine young children's use of singular definites in natural production using two corpus studies (Study 1 and Study 2) and a behavioral study (Study 3) and find no clear evidence for their systematic overuse of "the". We discuss how we should be interpreting reports of *the*-overuse in the literature as well as its implications for young children's knowledge of definite expressions.

While focusing on certain elicited production studies may give us the impression that children overuse "the", we should be cautious in interpreting the overuse data. First, a closer examination of children's vs. adults' production data across studies reveals two crucial caveats: 1) children give widely variable responses and don't consistently overuse "the" (van Hout et al., 2010), and 2) adult controls in these production studies also exhibit "inappropriate" uses of definite expressions under certain experimental manipulations (Maratsos, 1976; Schafer and de Villiers, 2000; van Hout et al., 2010). Second, it has been observed that children show adult-like responses to "the" in comprehension studies. Children aged 3 to 5 understand that the use of definite expressions (e.g., "the red ball") is, in the relevant circumstances, infelicitous when the intended referent is

* Yuanfan Ying, Valentine Hacquard, Alexander Williams, and Jeffrey Lidz, University of Maryland, College Park. Corresponding author: jackying@umd.edu.

Proceedings of the 48th annual Boston University Conference on Language Development, ed. Hayat Abdullah Ali AlThagafi and Jupitara Ray, 685-698.
Somerville, MA: Cascadilla Press.

not unique (Syrett et al., 2010). Additionally, younger children at 19 months show looking responses that presumably signal their early grasp of the definite ("the") vs. indefinite ("a") distinction and their sensitivity to speaker knowledge in interpreting singular definites (Choi et al., 2018). Taken together, there is a lack of definite evidence for children's systematic overuse of "the". Accordingly, there is no strong evidence that children lack either the right meanings for "the" or the sensitivity to the perspectives of interlocutors that normal use of "the" requires.

The observed conflict between findings from children's elicited production and those from comprehension studies gives rise to two hypotheses. Hypothesis 1 is that children's actual understanding of "the" is revealed in production and their success in comprehension is an illusion. For some reason, their error is revealed only in particular circumstances. Hypothesis 2 is that their actual understanding of definites is masked in artificial production setups, where certain contexts may encourage children as well as adults to misrepresent the restriction of the domain and boost "inappropriate" uses. To address this, examining natural production is crucial. Hypothesis 1 predicts that children's overuse of "the" should be present not only in elicited production but also in natural production. Following Hypothesis 2, children's production errors should decrease significantly when experimental artifacts are removed in natural production. To test these hypotheses, we conducted three studies to look at children's use of "the" in natural production.

2. Study 1: Overall distribution of definites in natural production

We first report a corpus study examining children's natural production of singular definites compared to their mothers' in terms of quantity. We start at a very coarse level of analysis to see how easy it would be to find symptoms of having the wrong meaning or an insufficient perspective-taking ability. Specifically, we examine the distribution of children's vs. mothers' use of singular definites of the form "the N". Do children in fact produce singular definites more frequently than their mothers? Do children and mothers produce these definites in the same environments? To obtain a fuller picture of children's use, we present their proportions of singular definites overall and across contexts.

2.1. Corpus information and coding scheme

The distributional analysis of use of singular definites with children and mothers is based on several CHILDES corpora including Brown (1973), Soderstrom et al. (2008), Suppes (1974), and Valian (1991). The CHILDES corpora examined in this study include a large sample of 27 children from a broad age range (1;0-3;11), consisting of a total of 912,530 words and 254,753 sentences.

For distributional analyses in the study, we looked specifically at singular definites produced by children and mothers in the corpora, as it requires the choice of either a definite or an indefinite determiner in production. We looked at singular definites in the form of 'the N_{SG}' and excluded counts of plural definites, as they either require "the" for definite referents or require none for indefinite referents.

To examine whether children show an overall overuse of singular definites, we use mothers' input as the baseline and ask whether the proportion of 'the N' is higher for children than for mothers. Here, the proportion of 'the N' is defined as the count of 'the N_{SG}' divided by the count of 'a/an N_{SG}' and 'the N_{SG}'. We excluded other alternatives for the determiner (e.g., demonstratives like "this" and "that") in the denominator, given how "the" and "a" are often pitted against each other in elicited production studies, and we examine whether the overuse of "the" found in elicited production is replicated in natural production.

2.2. Results

In terms of the general pattern, children use a lower proportion of singular definites than mothers do for all age groups of children, as shown in Figure 1.

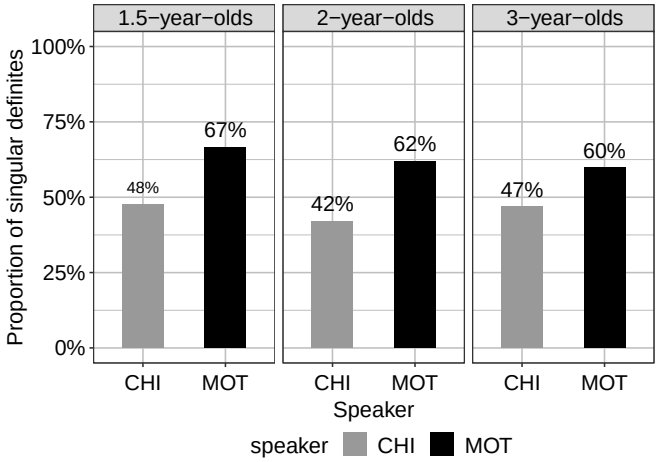

Figure 1: Overall proportion of singular *the*-definites by age

Specifically, children use the definite determiner (as opposed to the indefinite determiner) around half of the time, while mothers use it more often, i.e., for more than 60% of the utterances. The use of more "the N" by mothers at these stages might reflect the fact that mothers play a more dominant role in directing the attention of the other interlocutor (i.e., their child) to already mentioned or familiar objects in conversations with their children. The lack of "the"-overuse is also confirmed when we examine the individual data for our two-year-old corpora, as plotted in Figure 2.

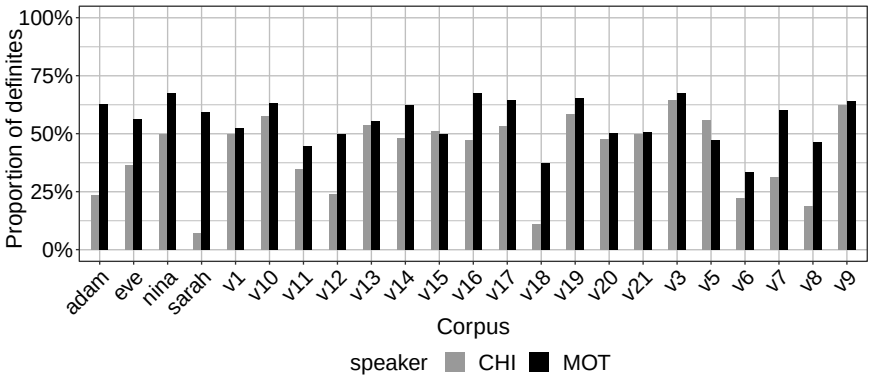

Figure 2. Overall proportion of singular *the*-definites by 2-year-old corpus

With the potential pragmatic differences across clause types and syntactic environments in mind, we ask 1) if children are sensitive to these differences and show different proportions of definite usages in the relevant contexts, and 2) if they use definites in an adult-like way. To do this, we break down usages of definites into different clause types, i.e. declaratives vs. interrogatives, and into different syntactic environments, i.e. subjects vs. objects vs. fragments.

First, children follow mothers in their pattern of determiner use across clause types (i.e., declaratives/statements vs. interrogatives/questions; Figure 3).

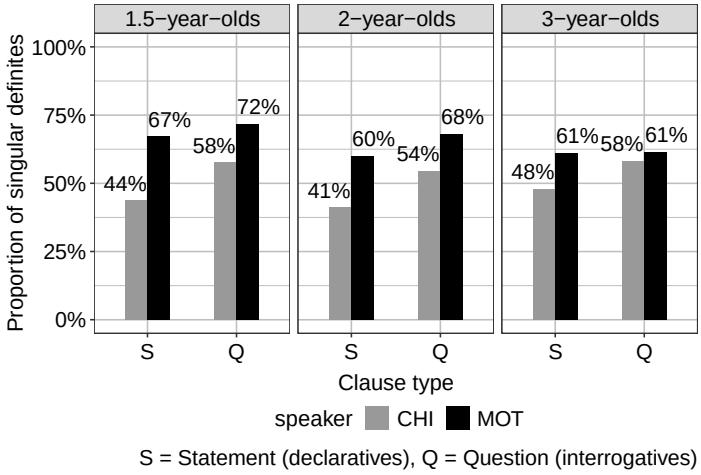

Figure 3. Proportion of singular *the*-definites across clause types (S for statements/declaratives versus Q for questions/interrogatives) by age

We do not observe children overusing definites against the baseline by mothers. Second, turning to children's pattern of determiner use in declarative clauses

across different syntactic positions (subjects vs. objects vs. fragments), we do not find *the*-overuse, either (Figure 4).

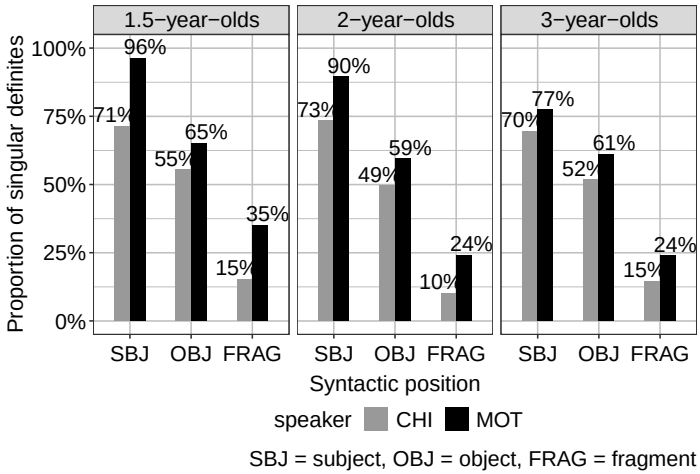

Figure 4. Proportion of singular *the*-definites across syntactic positions

Moreover, children use more singular definites for subjects than for objects. This is in line with their mothers' use, which reflects the generalization that in English, subjects tend to refer to things mentioned in prior discourse (Prince, 1992).

One concern is that children's rate of definites might be underestimated relative to their mothers', given that young children frequently produce NPs lacking determiners (e.g., Hyams, 1996). That is, if it is natural for children to drop determiners more for definites than for indefinites, not including these null determiners could drastically affect children's rate of definites overall. To examine this possibility, we calculated children's cases of null determiners (Table 1).

Table 1: Revised rate of definites (excluding fragment NPs)

Age (year)	#Def	#Ind	#Null	(Def+Null)%
1	209	166	550	82.1
2	1,364	1,342	2,599	74.7

If we treat all null determiners as potential cases of definites, children's revised rates of definites become 82.1% and 74.7% for 1-year-olds and 2-year-olds respectively, which are higher than the rates of definites by their mothers in the respective corpora, i.e., 66.5% and 60.2%. While children would be overusing definites if we assume the worst for their null determiners, we also notice that some of their null determiners in object positions are grammatical (2), and some seem to be a case of indefinite drop (3).

(2) I want coffee. (Valian, 1991: 01;09;25)
(3) I take walk. (Brown, 1973: Eve, 01;09;00)

To obtain a closer estimate of potential definites among children's cases of null determiners, we need to factor in children's distribution of definites in different syntactic environments. To do so, we calculated children's revised rate of definites that includes in the numerator null determiners that could have been a definite ($N_{null\ as\ def}$) probabilistically (5), i.e., based on the rates of produced definites for subject ($rate_{def\text{-}sub}$) and object ($rate_{def\text{-}obj}$) positions (4).

(4) $N_{null\ as\ def} = N_{null\text{-}sub} \times rate_{def\text{-}subj} + N_{null\text{-}obj} \times rate_{def\text{-}obj}$
(5) $rate_{revised\ def} = (N_{def} + N_{null\ as\ def}) / (N_{def} + N_{ind} + N_{null})$

The revised rates of definites by 1- and 2-year-olds become 58.2% and 52.1% respectively, both of which are lower than the rates of definites by their mothers in the respective corpora (i.e., 66.5% and 60.2%). Thus, children do not seem to overuse definites after we use a finer-grained estimate of children's null determiners that are potentially definite for subjects and objects.

In short, we do not find clear support for *the*-overuse in children's natural production quantity-wise. For the overall distribution of singular definites, we do not observe *the*-overuse globally, across clause types or syntactic positions. Rather, children's overall distributional pattern of definites follows their mothers'.

3. Study 2: Rate of *the*-driven miscommunications in natural production

In this section, we examine the quality of children's use of definites. If children had the wrong meanings for "the", we would expect them to use it in inappropriate contexts, which would lead to communication breakdowns. If children's definites were to occur in contexts where their presupposition is not supported or cannot be accommodated, their interlocutors would be confused with the referent of the definite or ask for clarifications.

To examine whether *the*-driven miscommunications is common in natural production, we compare the rates of *the*-driven miscommunications for mother-child interactions and adult-adult interactions. To do so, we calculate the rate of children's misuse of "the" for 1- and 2-year-olds and that of adults' misuse of "the". If children's knowledge of "the" is incomplete and non-adult-like, we would expect to see a great proportion of instances for miscommunications. Specifically, we would expect children's misuse of "the" causing befuddlement in listeners or eliciting responses from listeners suggestive of unfamiliar reference.

3.1. Corpus information and coding scheme

The rate of *the*-driven miscommunications was based on the same CHILDES corpora including Brown (1973), Soderstrom et al. (2008), Suppes (1974), and

Valian (1991), whereas the same rate for adult-adult interactions was calculated based on the CALLHOME corpus (Kingsbury et al., 1997).[1]

For children, we calculated the rate of *the*-driven miscommunications with singular definites, which requires a choice between "the" and "a/an" for the determiner. For adults, we did not limit our search to singular definites, as usages with both singular and plural definites could capture a fuller picture of their miscommunication scenarios. For the analysis of mother-child interactions in CHILDES, sentences containing 'a/an N_{SG}' or 'the N_{SG}' were extracted using commands and functions in CLAN, and for the analysis of adult-adult interactions in the CALLHOME corpus, sentences containing referential expressions were extracted using a Python script[2]. Cases of *the*-driven miscommunications were identified through the search for question marks occurring in 1-3 lines of texts following the target line with the definite expression, and we manually examined the cases in the output to ensure that these were genuine miscommunications by expanding the context and checking if they truly feature clarification questions.

3.2. Results

We first looked at 'the N' produced by 1-year-olds and found only 1 out of 205 (0.49%) cases of misuse but no case of miscommunication. We also noticed that quite a few instances of 'the N' produced by 1-year-olds are repetitions of mothers' utterances or instances following mothers' first mention of 'N', so we removed these cases (14 of them) from the denominator. The frequency of misuse remains low at 0.52% (1 out of 191).

We then examined the 2-year-old corpora and found only 0.39% (10 out of 2575) cases of *the*-driven miscommunications. Specifically, we identified two types of *the*-misuse in the corpora, namely ones associated with unfamiliar reference (8 cases) and others with non-unique reference (2 cases). (6) shows a case of unfamiliar reference, and (7) a case of non-unique reference.

(6) CHILD: What's the people doing?
MOTHER: What people?
CHILD: The people there.
MOTHER: There aren't any people out there.
(Brown, 1973: Eve, 02;02;00)

[1] The CALLHOME corpus includes 120 unscripted 30-minute telephone conversations between native speakers of English who were mostly family members or close friends. This is crucial, as the corpora we chose for analyzing children's miscommunication rate include interactions between mothers and children who know each other well. The CALLHOME corpus contains 301,805 words and 28,967 sentences.

[2] Many thanks to Weihang Wang and Sathvik Nair for helping with the script.

(7) CHILD: Hey, where's the truck? I need …
INVESTIGATOR: Where's the what?
MOTHER: The truck?
CHILD: Yeah.
MOTHER: There's two trucks.
(Brown, 1973: Eve, 02;02;00)

Turning to adult-adult interactions in the CALLHOME corpus, we found 0.13% (4 out of 3121) cases of *the*-driven miscommunications. All the cases identified in the corpus were due to unfamiliar reference. (8) illustrates such a case.

(8) A: You know, spent panting and spending every second of the day, w- with him only. You know what I'm saying and that probably part of the reason that it would be nice to actually stay with him is that **you could also go and visit the sights** or whatever and g- and...
B: Where? What are you talking about?
A: In Japan.

In brief, Study 2 does not support the claim that children overuse "the" in natural production. Children's rate of *the*-driven miscommunications in mother-child interactions is as low as that in adult-adult interactions, which is well below 1%. That is, the rate of misuse of "the" in children's corpora is not higher than what we see in adult-adult conversations, providing no support for *the*-overuse.

4. Study 3: Guessing determiners

The lack of miscommunication observed in Study 2 may result from mothers' accommodation of children's definites. To assess children's determiners more objectively, we need to involve adults who were not in selected conversations with children and are hence less able to accommodate presuppositions.

In Study 3, we probe whether children's productions of determiners are adult-like through conducting a determiner-guessing experiment, following Dieuleveut et al.'s (2022) adaptation of the Human Simulation Paradigm (Gillette et al., 1999). Dieuleveut et al. (2022) showed that this method could be used to assess whether children's use of a certain class of words is adult-like and capture nuances of adult-like-ness for children's use of different subclasses of words.

We presented adult participants with excerpts of mother-child conversations and asked them to choose a definite vs. an indefinite for a missing determiner used either by a child or a mother. We first ask whether adults can correctly guess mothers' use of definites. If yes, it implies that the context, together with adult knowledge of determiners, provides sufficient information for determiner choice by adults (mothers in this case). Similarly, if children tend to misuse definites, adults should be worse at guessing their definites, as the contexts where children used definites inappropriately will elicit indefinite determiners from adults.

4.1. Predictions

If children truly overuse definite determiners, we expect adults to be worse at predicting children's definite usages than predicting mothers'. On the contrary, if children do not overuse definite determiners, we expect no difference between adults' rate of correct predictions for children's definites and that for mothers'.

4.2. Designs and materials

Participants were recruited over Amazon Mechanical Turk (AMT). The experiment was run on PCIbex farm (Zehr & Schwartz, 2018). The duration of the task was roughly 15 to 20 minutes. Each participant received 40 trials, i.e., 10 definites, 10 indefinites, and 20 fillers.[3] One fifth of the trials (i.e., 8 trials) were followed by comprehension questions (i.e., simple memory questions) to ensure that they were reading the full conversations.

The materials consisted of excerpts of mother-child conversations from 23 mother-child pairs (age range of children: 2;1;4-4;11;2). Those excerpts consisted of 10-line conversations randomly drawn from Gleason's corpora in CHILDES (80,347 words; Menn & Gleason, 1986).[4] As we wanted both "the" and "a" to be possible determiner options for our test items, thereby making it more likely for us to capture cases of *the*-misuse if any, we removed items biased towards either "the" or "a" from the test list according to exclusion criteria established prior to our random sampling of conversations. Items were removed if the target determiner occurred in 1) cases of repetition, e.g., echoing, speaker disfluency, and finishing another speaker's sentence, 2) cases of idiomatic expressions (e.g., *wait a minute*) or common collocations, or 3) cases of plural or uncountable nouns.

4.3. Procedure

Before starting the actual experiment, participants were given two practice trials to learn that 1) they would have to press the spacebar for each conversation to unfold line by line, 2) choose an answer out of two options for the blank by the end of each conversation, and 3) that they would occasionally receive comprehension questions testing if they had paid attention to the context. Figure 5 shows a sample test item, where the final blank required a determiner to be chosen from two options between target determiners (i.e., "the" vs. "a"), filler determiners (e.g., "this" vs. "that"), or mixed determiners (e.g., "the" vs. "this").

---

[3] Many thanks to Weihang Huang and Sathvik Nair for helping with debugging the script for item randomization.

[4] The conversations took place in recorded lab sessions, where children and mothers interacted with new objects. This is good for our purpose, as we are interested in exploring how children register new referents in a dynamic common ground.

CHILD: Yeah..
MOTHER: Yeah?
MOTHER: I'll go get it.
MOTHER: Here, let's put the car back where it belongs.
MOTHER: Okay.
CHILD: No!
MOTHER: You don't wanna put the car there?
MOTHER: Okay.
CHILD: No.
CHILD: Wanna play with it and read __ book, okay?

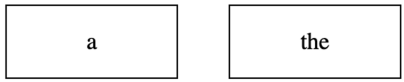

Figure 5. Sample test items during experiment

4.4. Participants

240 participants (127 males and 105 females; mean age: 41.3) were recruited over AMT. Data from 211 participants (109 males, 94 females, and 2 gender-unspecified; mean age: 41.9) were analyzed (those who scored less than 75% for comprehension questions were excluded).

4.5. Results

We looked at adult participants' rate of correct guesses for mothers' definites as our baseline for evaluating children's definites. They were above chance, hence good in predicting mothers' definites, at a rate of 0.9 (N =36, SD = .096), 0.895 (N = 39, SD = .102), and 0.765 (N = 31, SD = .236) for the 2- (V = 666, p < .001), 3- (V = 780, p < .001), and 4-year-old (V = 389, p < .001) corpora respectively.

Adults were also above chance in predicting children's definites, at a rate of 0.844 (N = 36, SD = .144), 0.857 (N = 35, SD = .127), and 0.894 (N = 34, SD = .110) for the 2- (V = 595, p < .001), 3- (V = 595, p < .001), and 4-year-old (V = 595, p < .001) groups respectively, as they did in guessing mothers' definites (Figure 6). Moreover, in no age group were adults significantly better at predicting mothers' definites than children's (2-year-olds W = 516, p = .124; 3-year-olds W = 571.5, p = .212; 4-year-olds W = 726.5, p = .007**).[5]

[5] The Wilcoxon test was used to see whether adults were above chance level at guessing definites for each condition, and the Mann-Whitney U test was for testing whether adults performed better with either child or mother utterances. These non-parametric tests were used, as the data did not follow normal distribution.

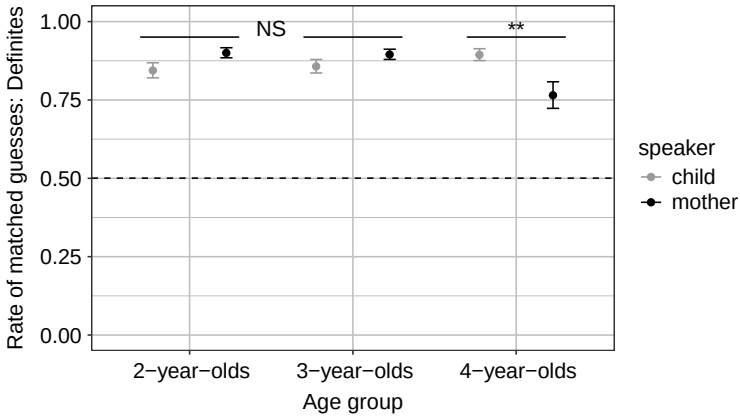

Figure 6. Rate of correct guesses for definites by speakers by child age

One possibility for adults' correct guesses for definites might be that they have a definite bias. This predicts worse performance with indefinites, which is not borne out (Figure 7). Adults were above chance at predicting indefinites used by both children (2-year-olds $V = 593.5$, $p < .001$; 3-year-olds $V = 527$, $p < .001$; 4-year-olds $V = 511.5$, $p < .001$) and adults (2-year-olds $V = 587$, $p < .001$; 3-year-olds $V = 768$, $p < .001$; 4-year-olds $V = 385$, $p < .001$). They were never significantly better at predicting mothers' indefinites than children's (2-year-olds $W = 526$, $p = .163$; 3-year-olds $W = 703.5$, $p = .819$; 4-year-olds $W = 498$, $p = .704$).

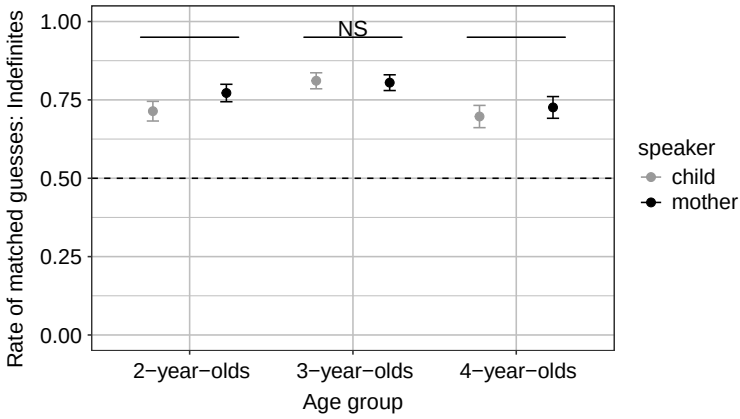

Figure 7. Rate of correct guess for indefinites by speakers by child age

Furthermore, participants in the definite condition show overall agreement on their determiner choice for most test items with either children's or mothers' utterances (Figure 8). Here, we use gray to mark regions with an accuracy range between 40% and 60% to capture items with wide disagreement (i.e., close to a

50-50 determiner split among the participants). As seen in the figure, very few items incur wide disagreement. We also examined items with the highest mismatch rates and found only one possible case of *the*-misuse by children.

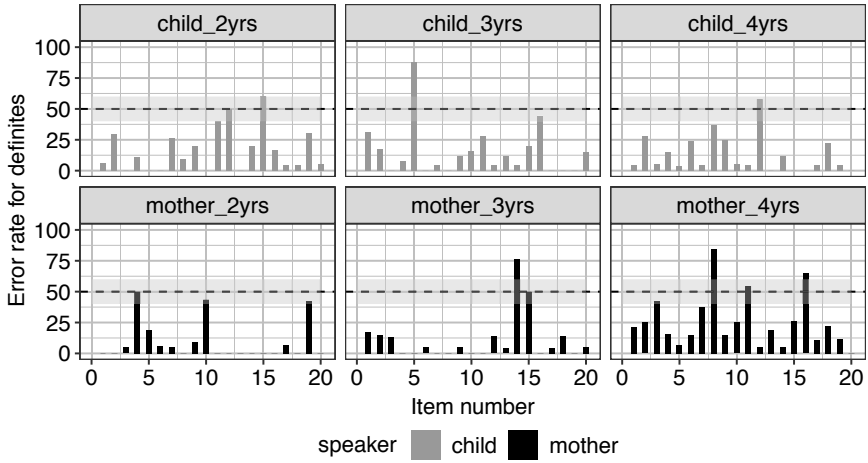

Figure 8. Adults' error rates for definites by item, speaker, and child age

In sum, our results suggest that children use singular definites in an adult-like way in natural production. This is reflected in adults' high rates of correct guesses for definites used by both children and their mothers.

5. Discussion

Our studies offer converging evidence that children do not overuse "the"; they exhibit adult-like uses of singular definites in natural production. Their usage patterns align with their mothers' from 1.5 years of age; the rate of miscommunication due to "the" is low among 1- and 2-year-olds; from two years of age, their use of definites is as guessable (by adults) as their mothers'.

These findings have two implications. First, children's knowledge of "the" grows fast. As early as children start to produce definites, they show adult-like patterns in usage. As proper use of "the" hinges on children's ability to consider the information state of their interlocutor, our results imply that children develop early pragmatic competence in discerning backgrounded presuppositions. Our results also align with findings on children's early ability in perspective-taking for non-verbal (e.g., Luo and Baillargeon, 2007) and verbal tasks (e.g., Saylor and Ganea, 2007). Second, our findings suggest that the overuse of definites found in elicitation studies may be due to experimental artifacts. While occasional usage errors occur in children's computing of the domain of reference or quantification, our findings support children's adult-like uses of definites in natural conversations. This suggests that children's knowledge of definites is in place but may be liable

to interference from artificial contexts. Alternatively, it may be that the contexts that lead to overuse are more prevalent in experiments than in real life.

We propose two general explanations for children's tendency to overuse the definite determiner "the" in some elicited production tasks. One is that some of the production tasks singled out a specific referent from several others through deictic gestures like pointing (Karmiloff-Smith, 1979) or contextual focus (e.g., Schafer & de Villiers, 2000), which likely prompted children to perceive the chosen referent as uniquely identifiable, thereby boosting their rate of *the*-usages. For instance, in a scenario featuring three identical balls, physically interacting with one of them could elevate its salience, making it the only referent within the domain of evaluation, which may incline children to use a definite expression (i.e., "the ball") to refer to it. Another explanation is that certain artificial contexts may interfere with children's reasoning about the knowledge of interlocutors. Children's rate of *the*-overuse could spike up in tasks with joint attention between the experimenter and the child, which may tempt children to assume shared perspectives. Cases of *the*-overuse could also increase in tasks lacking sufficient buildup of common ground between the experimenter and the child prior to production trials, which may hinder children's perspective-taking of interlocutors.

While natural production provides rich insights into children's use of referring expressions, there are limitations to what it can probe. There might be less dynamics involved for the common ground in mother-child interactions, and children's overall correct use may be owing to a more predictable common ground. Whether this is true is yet to be investigated for future studies.

6. Conclusion

In this paper, we showed that 1) there is a lack of clear evidence for systematic *the*-overuse in children's natural production and 2) their use of "the" is adult-like. Children's determiner use follows their mothers' overall and across clause types and syntactic environments; their rate of *the*-driven miscommunications is low; adults are equally successful at guessing definites used by children and their mothers. Our results suggest that children follow the discourse well, and *the*-overuse reported in elicited production studies may reflect performance issues due to experimental artifacts.

References

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

Choi, You-jung, Song, Hyun-joo, & Luo, Yuyan. (2018). Infants' understanding of the definite/indefinite article in a third-party communicative situation. *Cognition*, *175*, 69-76.

Dieuleveut, Anouk, van Dooren, Annemarie, Cournane, Ailís, & Hacquard, Valentine. (2022). Finding the force: How children discern possibility and necessity modals. *Natural Language Semantics*, *30*(3), 269-310.

Gilette, Jane, Gleitman, Henry, Gleitman, Lila, & Lederer, Anne. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135-176.

Heim, Irene R. (1982). *The semantics of definite and indefinite noun phrases* [Doctoral dissertation, University of Massachusetts Amherst].

Hyams, Nina. (1996). The underspecification of functional categories in early grammar. In Harald Clahsen (Ed.), *Generative perspectives on language acquisition* (pp. 91-128). John Benjamins Publishing Company.

Karmiloff-Smith, Annette. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge University Press.

Kingsbury, Paul, Strassel, Stephanie, McLemore, Cynthia, & McIntyre, Robert. (1997). Callhome American English transcripts, LDC97T14. *Philadelphia: Linguistic Data Construction*.

Luo, Yuyan, & Baillargeon, Renée. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition*, *105*(3), 489-512.

Maratsos, Michael P. (1976). *The use of definite and indefinite reference in young children: An experimental study of semantic acquisition*. Cambridge University Press.

Menn, Lise, & Gleason, Jean B. (1986). Babytalk as a stereotype and register: Adult reports of children's speech patterns. In Joshua A. Fishman (Ed.), *In honor of Charles A. Ferguson on the occasion of his 65th birthday. Volumn 2: Sociolinguistics and the sociology of language* (pp. 111-126).

Prince, Ellen F. (1992). The ZPG letter: Subjects, definiteness, and information status. In William C. Mann & Sandra A. Thompson (Eds.), *Discourse descriptions: Diverse linguistic analyses of a fund-raising text* (pp. 295-325). John Benjamins.

Roberts, Craige (2003). Uniqueness in definite noun phrases. *Linguistic and philosophy*, *26*, 287-350.

Russell, Bertrand (1905). On denoting. *Mind*, *14*(56), 479-493.

Saylor, Megan M., & Ganea, Patricia. (2007). Infants interpret ambiguous requests for absent objects. *Developmental Psychology*, *43*(3), 696.

Schafer, Robin, & de Villiers, Jill. (2000). Imagining articles: What *a* and *the* can tell us about the emergence of DP. In S. Catherine Howell, Sarah A. Fish, & Thea Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development* (pp. 609-620, Vol. 2). Cascadilla Press.

Soderstrom, Melanie, Blossom, Megan, Foygel, Rina, & Morgan, James L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, *35* (4), 869–902.

Strawson, Peter F. (1950). On referring. *Mind*, *59* (235), 320–344.

Suppes, Patrick. (1974). The semantics of children's language. *American psychologist*, *29* (2), 103–114.

Syrett, Kristen, Kennedy, Christopher, & Lidz, Jeffrey. (2010). Meaning and context in children's understanding of gradable adjectives. *Journal of semantics*, *27*(1), 1–35.

Valian, Virginia. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*(1-2), 21–81.

van Hout, Angeliek, Harrigan, Kaitlyn, & de Villiers, Jill. (2010). Asymmetries in the acquisition of definite and indefinite NPs. *Lingua*, *120*(8), 1973–1990.

Wexler, Ken. (2011). Cues don't explain learning: Maximal trouble in the determiner system. In Edward Gibson & Neal J. Pearlmutter (Eds.), *The processing and acquisition of reference* (pp. 15–42). MIT Press.

Zehr, Jeremy, & Schwarz, Florian. (2018). Penncontroller for Internet based experiments (IBEX).

Proceedings of the 48th annual Boston University Conference on Language Development

edited by Hayat Abdullah Ali AlThagafi and Jupitara Ray

Cascadilla Press Somerville, MA 2024

ISSN 1080-692X
ISBN 978-1-57473-097-5 (2 volume set, paperback)

Ordering information

To order a copy of the proceedings or to place a standing order, contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, sales@cascadilla.com, www.cascadilla.com