Journal of Experimental Child Psychology 185 (2019) 128–147

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Toddlers' sensitivity to phonetic detail in child speech

Dana E. Bernier *, Katherine S. White

Department of Psychology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

A R T I C L E   I N F O

Article history:
Received 6 October 2018
Revised 26 April 2019

Keywords:
Child speech
Spoken word processing
Phonetic sensitivity
Disambiguation response
Language environments
Toddlers

A B S T R A C T

Young language learners acquire their first language(s) from the speech they are exposed to in their environment. For at least some children (e.g., those in daycare), this environmental speech includes a large quantity of speech from other children. Yet, we know little about how young learners process this type of speech and its status as a source of input. Across two experiments, we assessed 21- to 23-month-olds' processing of a child's speech using the preferential looking paradigm. We found that toddlers processed the child speaker's productions as well as those of an adult and with the same level of sensitivity to phonetic detail previously shown for adult speakers. Although the amount of experience toddlers had interacting with other children outside the home had little influence on their processing of familiar words, only toddlers with high levels of experience with other children outside the home showed a disambiguation response after hearing novel labels. Whether this is truly due to the number or variety of other child speakers or to other correlated aspects of toddlers' language environments is unclear and remain intriguing questions for future research. Overall, these findings demonstrate that child speech may represent useful input for young language learners.

* Corresponding author.
E-mail address: dbernier@uwaterloo.ca (D.E. Bernier).

https://doi.org/10.1016/j.jecp.2019.04.021

Introduction

Young language learners acquire their first language(s) through exposure to speech in their environment. Although this environment is typically thought of exclusively in terms of adult speech, language learners are also exposed to the speech of other children. In some environments, this child speech may even occur in an amount similar to that of adult speech (Bernier & Soderstrom, 2016). Yet, we know very little about how well young children process this kind of speech. Given that other children's speech potentially represents a substantial source of input for at least some children (e.g., those with siblings or attending daycare), this represents a significant gap in our understanding of early language processing.

Very young children's productions are characterized by a number of phonological deviations from adult targets. These changes include (among others) substitutions of one sound for another (e.g., *fumb* for *thumb*) and omissions of sounds (or syllables) altogether (e.g., *nake* for *snake*) and can lead to low intelligibility for naïve adult listeners (Flipsen, 2006; Hodson & Paden, 1981). The majority of these more significant phonological deviations have decreased by 4 years of age, and they are largely absent after 6 years (Dodd, Holm, Hua, & Crosbie, 2003). These later productions are (on average) quite intelligible for even naïve adult listeners, although adults with more experience listening to children are more accurate and reliable in their judgments of individual sounds (Munson, Johnson, & Edwards, 2012). However, even once these larger deviations are no longer present, children continue to show less accuracy and more variability in their productions than adults. For example, early school-aged children produce less accurate /l/ (Lin, Inkelas, McDonnaughey, & Dohn, 2016), more variable /s/ (Koenig, Lucero, & Perlman, 2008; Munson, 2004), and less distinguishable /s/ and /ʃ/ (Maas & Mailend, 2017; Nissen & Fox, 2005; Romeo, Hazan, & Pettinato, 2013) than adults, and they have a larger vowel space (Hillenbrand, Getty, Clark, & Wheeler, 1995; Lee, Potamianos, & Narayanan, 1999). These deviations and increased variability mean that child speech may be more difficult for young language learners to process than adult speech.

In an early study exploring young children's processing of child speech, Dodd (1975) tested 2- to 4-year-olds' comprehension of their own productions of object labels as well as those of another child and an adult. She found that children were less accurate in choosing the named object when hearing labels produced by a child (including their own productions) compared with labels produced by an adult. More recently, Cooper, Fecher, and Johnson (2018) used a looking paradigm to test the same question with 2½-year-olds and similarly found an advantage for adult speech. These studies suggest that young children have difficulty processing speech that deviates significantly from adult target forms. In other work more closely examining the precision with which child speech is processed, Strömbergsson, Wengelin, and House (2014) examined slightly older (4- to 6-year-old) children's processing of their own and other children's synthetically modified speech. They found that typically developing children were able to distinguish between correct and mispronounced versions of resynthesized speech that involved changes between /t/ and /k/ regardless of whether the speech was their own or that of another child. Although children with a phonological disorder were found to have some difficulty processing their own speech (at least without a time delay), they did not show this difficulty when processing the speech of other children. Thus, by 4–6 years of age, children are sensitive to even small pronunciation changes in the speech of similarly aged children.

These studies provide some indication of how young children process the speech of their age-matched peers. Namely, on average, toddlers have difficulty processing the speech of other toddlers, but preschool-aged children seem to fare much better in processing the speech of other preschool-aged children. However, age-matched peers are not the only type of child–child interaction possible. Given that child speech constitutes a large percentage of the input for some young language learners, it is important to continue exploring how toddlers process the speech of children of various ages, the factors that might affect ease of processing, and the extent to which speech from other children is used to guide language learning.

Although there is limited work to date on toddlers' processing of child speech, there is much more work on their processing of other forms of noncanonical speech, such as non-native accented speech. Although there are processing costs associated with initial exposure to an unfamiliar accent (e.g., Best, Tyler, Gooding, Orlando, & Quann, 2009; van Heugten & Johnson, 2014; White & Aslin, 2011), continued exposure to the accent can lead to successful word recognition (Schmale, Cristia, & Seidl, 2012; van Heugten & Johnson, 2014; van Heugten, Krieger, & Johnson, 2015; White & Aslin, 2011). These studies suggest that if child speech is initially difficult for toddlers to process, experience hearing other children speak may help.

It is clear that toddlers hear their own productions (and so have experience with child speech). However, the role of experience with *other* children has not previously been considered. Experience hearing other children's speech could facilitate processing in a number of ways. One way that experience may help is by allowing listeners to learn something about the sound categories children produce. Just as exposure to ambiguous sounds embedded in words alters adults' judgments of phoneme category boundaries (Norris, McQueen, & Cutler, 2003), exposure to the speech of other children could refine toddlers' perception of those speech productions. In this case, it is expected that toddlers with greater exposure to other children would show *more* sensitivity to phonetic detail in child speech.

Alternatively, experience with children could lead toddlers to pay less attention to the specifics of child speakers' productions and rely more on context to determine referential intent. It is clear that adult listeners rely heavily on context for comprehending very young children (whose extreme phonological deviations make them highly unintelligible). And recent work has shown that even young children will use top-down information during word processing when the acoustic input is unreliable (Yurovsky, Case, & Frank, 2017). If toddlers have knowledge that children misarticulate sounds, it is possible that they would adopt this strategy in the case of child speech. In this case, it is expected that toddlers with greater exposure to other children would be more tolerant of deviations (and, as a result, show *less* sensitivity to phonetic detail) in child speech.

Regardless of how experience might influence processing of children's speech (either by affecting perception directly, through learning of children's speech patterns, or by affecting expectations about child speech), environments where there is a high amount of child-to-child speech will be those most likely to provide toddlers with this experience. Environments such as daycare (whether home-based or center-based) regularly place toddlers in situations where there are more children than adults (e.g., 1 adult per 2–5 children), which is in sharp contrast to home environments where the number of adults typically ranges from 1 to 1.5 per child during the day (Soderstrom, Grauer, Dufault, & McDivitt, 2018). Moreover, Soderstrom et al. (2018) found that the amount of time toddlers spent with other children was far higher in daycare settings than in home environments (even for toddlers at home with siblings). This means that toddlers who spend time in group settings on a regular basis, such as those in daycare, are much more likely to encounter child speech than toddlers who spend the majority of their time at home (with or without siblings).

In addition to the amount of child speech, the variety of speakers toddlers hear may be important. Exposure to more variable input has been shown to improve phonological processing and learning. For example, greater variability has been shown to increase infants' attention to phonetic detail during word learning (Rost & McMurray, 2009, 2010) and their learning of a novel phonotactic pattern (Seidl, Onishi, & Cristia, 2014). Therefore, toddlers who regularly experience group settings (whether in daycare or in some other activity) may be those most likely to develop knowledge of the way that other children speak, regardless of whether or not they have siblings. There are, of course, other features that occur in many (but not all) group settings that could affect children's language knowledge or processing, such as the greater variety of adults encountered and the structured interactions typical of organized activities. However, because we were interested in examining differences in the processing of child speech, we used the amount of time spent with other children in daycare, informal playgroups, and organized activities as our measure of experience, despite these possible co-occurring features.

In the current study, we presented toddlers with visual displays containing two objects (either both familiar or one familiar and one novel) and recorded their looking behavior in response to instructions directing them to look at one of the objects. In Experiment 1, we compared their processing of productions from a child speaker and productions from an adult. In this experiment, words were produced only in their standard form. In Experiment 2, we examined the specificity of toddlers' processing by asking whether they show the same sensitivity to mispronunciations in child speech that they have shown previously for adult speech (e.g., White & Morgan, 2008). Given the research showing that toddlers have difficulty processing the speech of age-matched peers (Cooper et al., 2018; Dodd, 1975) and the lack of studies examining toddlers' processing of slightly older children's speech, we chose a female first-grade student as our child speaker. Although speech from a child this age does not typically contain large deviations from adult target forms, it does have the more subtle deviations described above and clearly has the voice qualities of a child.

Because experience could affect toddlers' processing of child speech, we included the amount of experience with other children as a factor in our analyses. If experience with a diversity of other children sharpens the perception of child speech, toddlers with more experience should show better processing (Experiment 1) and greater phonetic sensitivity (Experiment 2) than those with less experience. If, however, experience with other children primarily leads to more tolerance for deviations in child speech, toddlers with more experience should show equivalent or better processing than those with less experience (Experiment 1) but less phonetic sensitivity to mispronunciations (Experiment 2).

Experiment 1

In Experiment 1, we compared toddlers' processing of familiar and novel object labels produced by a female child and by a female adult. If child speech is challenging to process, toddlers may be less accurate for words produced by a child, particularly toddlers who have little experience interacting with other children. To determine whether experience is a factor in the processing of child speech, we compared toddlers who spent 8 h or less per week (via parent report) with other children in group settings to those with more experience.

We presented toddlers with three types of trials from each speaker. For the first two types of trials, a label was presented for a familiar target, but the type of distractor object differed (familiar vs. novel). In a situation where both pictured objects have known labels, it may be fairly easy to reject the distractor even if the target label is not pronounced exactly as expected. However, the same is not true when the target object is paired with a novel distractor (see White & Morgan, 2008, for discussion). With a novel distractor, toddlers must decide whether the match between their lexical representation and the label being heard is sufficient, making this a more sensitive test of word recognition. If child speech is harder to process than adult speech, toddlers may be slower or less likely to map familiar labels to familiar objects in the presence of a novel distractor.

It is also possible that there will be differences in processing for the two speakers as a function of the type of label (familiar vs. novel). For this reason, we also included a third type of trial where the novel object was labeled. The task of interpreting a novel label also requires toddlers to evaluate the match between the familiar lexical representation and the label being heard. However, in the case of a novel label, after evaluating (and rejecting) the familiar object as a potential referent, the label must then be mapped to the novel object. If child speech is more difficult to process, toddlers may perform worse on these trials for a child speaker. For example, toddlers could be slower to shift away from the familiar object and toward the novel object within the allotted time window (see Halberda, 2003, for a related finding with younger infants).

Method

Participants

A total of 48 monolingual English-learning toddlers (M_{age} = 22.1 months, range = 21.2–23.6) were recruited from the Kitchener/Waterloo region of Ontario, Canada. An additional 3 toddlers were tested but not included, due to completing fewer than half of the child speaker trials successfully using the criteria below (n = 2) or parental report indicating that they knew the labels of fewer than half of the familiar objects shown (n = 1).

Half of the participants were classified as High Experience (M_{age} = 22.2 months, range = 21.4–23.0) and half were classified as Low Experience (M_{age} = 22.0 months, range = 21.2–23.6). Approximately half of the participants in each experience group had older siblings. Experience groups were determined a priori by the average number of hours parents reported that they spent interacting with

Table 1
Hours per week participants in Experiment 1 spent with children other than siblings, reported separately for toddlers who did and did not attend daycare.

		n	Number of older siblings	Total hours[a] [*M* (range)]	Daycare [*M* (range)]	Other [*M* (range)]
High experience						
Daycare	First born	10		37.7 (20–48)	35.1 (20–45)	2.6 (0–8)
	Later born	8	1–2	32.4 (18–45)	30.3 (16–45)	2.1 (0–4)
Other	First born	3		15.0 (13–18)	0	15.0 (13–18)
	Later born	3	1	12.0 (10–16)	0	12.0 (10–16)
Low experience						
	First born	12		2.4 (0–6)	0	2.4 (0–6)
	Later born	12	1–3	1.8 (0–8)	0	1.8 (0–8)

[a] Total hours = Daycare + Other.

children other than siblings (see Table 1 for summary). Because it is likely that only a very small number of toddlers have absolutely no regular exposure to children in group settings, we set the cutoff between the experience groups at 8 h per week. Toddlers in the Low Experience group were those who were home with a caregiver (i.e., not in any type of formal childcare setting) and who regularly spent 8 h or less per week interacting with children other than siblings. Those in the High Experience group spent 10 h or more per week with other children outside the home (there were no children whose time in group settings fell between 8 and 10 h). Although most of these toddlers were in daycare, some were home with a caregiver who regularly provided them with activities where there were other children present. Toddlers with these two types of experience are presented separately in Table 1.

Design

There were 24 test trials in total, each with a unique object pair (a list of the familiar objects is presented in Appendix A; sample images are presented in Fig. 1). Half of these trials had labels produced by a child and half had labels produced by an adult (blocked, within participant). Of the 12 test trials for each speaker, 8 involved the labeling of familiar objects (4 of these trials had a familiar distractor and 4 had a novel distractor) and 4 involved the labeling of novel objects with novel labels. For the novel labels, the two speakers each produced four different labels (paired across speakers such that the labels within a pair were matched for length, syllable structure, and segment type, e.g., *tibble* for the child speaker and *boogle* for the adult speaker). Each pair of novel labels was assigned to a pair of novel objects; the speaker assigned to label each object was counterbalanced across participants.

Two versions of the experiment were generated, each with different target–distractor pairings for both familiar labeled trials (e.g., shoe paired with a bike vs. a novel object) and novel labeled trials (e.g., *tibble* paired with a doll vs. a spoon). Version, speaker order, and objects labeled by each speaker were counterbalanced, producing a total of eight between-participant counterbalancing conditions.

Stimuli

Stimuli were recorded in a sound-attenuated room using a Sennheiser e945 microphone connected to a laptop via a blue icicle USB adaptor. They were recorded into Praat (Boersma & Weenink, 2014). Target words were recorded in one of four sentence frames ("Where's the ____?", "Do you see the ____?", "Look! A ____!", and "Can you find the ____?"). The child stimuli were produced by a female first-grade student (7 years 5 months of age). To elicit these productions, the child sat with a female adult in a sound-attenuated room. For each target word, an image of the object appeared on a laptop while the adult produced a labeling sentence in infant-directed speech for the child to repeat. The adult stimuli were recorded in a separate session by the same female adult speaking in an infant-directed speech register. At least three sentence productions per speaker were elicited for each target word; the clearest token of each was selected. Stimuli were later adjusted as necessary by the first author in Praat to ensure that all tokens were of approximately equal perceived intensity (because

Fig. 1. Sample image pairs for trials with two familiar objects (left) and trials with one familiar object and one novel object (right).

Table 2
Mean acoustic values for child and adult speakers.

	Child speaker	Adult speaker	Child–Adult difference
Full sentence			
Duration (ms)	1592	1983	−392
F0[a] (Hz)	283	262	21.3
F0 variation (SD; Hz)	82.4	103.1	−20.7
Target word			
Duration (ms)	664	781	−117
F0 (Hz)	325.8	310.9	14.9
F0 variation (SD; Hz)	86.2	112.4	−26.1

Note. F0 is the acoustical parameter most closely related to perceived pitch. F0 values were calculated using a floor of 100 Hz and a ceiling of 600 Hz for both speakers.

of the difference in mean pitch noted in Table 2, the mean intensity of the child's productions was set slightly lower than the adult's once equated for perceived intensity). In the testing room, all stimuli were presented at a comfortable listening volume of 65–70 dB.

Table 2 presents the mean acoustic values for the adult and child productions. These measures show that, on average, the child speaker's productions were both shorter and less variable in pitch than the adult's and were also slightly higher in mean pitch. See the online supplementary material for additional information about the productions.

Procedure

Toddlers were tested using the intermodal preferential looking procedure (IPLP) in a sound-attenuated room. Each child sat on the parent's lap while the parent listened to music over circumaural headphones. In front of them was a 42-inch widescreen television and two hidden speakers located at the base of the television; both were connected to a computer in an adjacent room. Participants were monitored over a closed-circuit video feed that was recorded for later offline coding; the camera was centrally located beneath the television and hidden behind a black curtain.

Toddlers were presented with 24 trials in total (12 per speaker). For each speaker, there were four blocks of 3 trials (one trial type per block in pseudorandom order). The four blocks for each speaker were presented in random order. Each trial consisted of images of two objects presented on the left and right sides of the screen. The trial began with the objects shown in silence for 3 s; this was used as the baseline phase. Following the baseline phase, the audio stimulus (e.g., "Where's the shoe"?) began to play. The images remained on the screen for an additional 5 s from the start of the audio (for a total trial length of 8 s). We defined a priori a 3-s naming phase that commenced 267 ms (8 frames) after the start of the target word (based on the time needed to program an eye movement and previous convention; e.g., Swingley & Aslin, 2000; White & Morgan, 2008). The dependent measure was the change in the proportion of time toddlers spent looking at each object from the baseline phase to the naming phase.

Following the testing session, parents completed a questionnaire on their children's comprehension and production of the experimental items (both familiar and novel) and the amount of time their children interacted with other children (daycare and playgroups/organized activities) on a weekly basis.

Analysis

Looking behavior was coded offline by trained coders blind to condition using customized software at a rate of 30 frames per second (~33.33 ms/frame). For each trial, the proportion of time toddlers spent looking at the *labeled* object (out of the total time spent looking at the two objects) was calculated for both the baseline and naming phases. For exposition purposes, proportions for the naming phase as a function of time are provided in the supplementary material. However, our dependent measure was a difference score across the entire naming and baseline phase windows (naming minus baseline), which indicates how much toddlers changed their looking to the labeled object after hearing the label. A positive score indicates that they increased their looking to the labeled object following naming, whereas a negative score indicates that they decreased their looking to that object.

For a trial to be included in the analysis, participants needed to look at each of the objects for a minimum of 8 frames (267 ms) during the baseline phase. This criterion resulted in 16.2% of trials being discarded (High Experience: 15.6%; Low Experience: 16.8%). Participants also needed to attend to the objects for a minimum of 1 s total during each phase. This criterion resulted in 1% of trials being discarded (High Experience: 0.9%; Low Experience: 1%).

Data were analyzed using linear mixed effects regression (Baayen, Davidson, & Bates, 2008) with R's *lme4* package (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2018). Fixed effects of experience (high vs. low), speaker (adult vs. child),[1] and condition (familiar labeled/familiar distractor vs. familiar labeled/novel distractor vs. novel labeled) were contrast-coded. The first condition contrast compared the two conditions where the familiar object was labeled and the second condition contrast compared these first two conditions with the final condition where the novel object was labeled. The first condition contrast tests the effect of competitor type on familiar word recognition; the second tests the effect of familiar versus novel object labeling. The *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017), which uses Satterthwaite's method for approximating degrees of freedom, was used to assess the significance of intercepts as well as contrasts of simple models with no more than one fixed effect (Luke, 2017). Random effects error variance was used as an estimate of σ^2 to calculate Cohen's *d* for model estimates (Brysbaert & Stevens, 2018; Westfall, Kenny, & Judd, 2014).

Results

Baseline looking

We first assessed looking during the baseline phase against chance (.50). To do this, baseline looking was centered around 0 by subtracting .50 from each value, thereby making the intercept of the following models equivalent to the deviation from chance. Linear mixed effects regression models were generated separately for each type of display (two familiar objects or one familiar object and one novel object), with experience entered as a fixed effect along with random intercepts for participant and item (i.e., the specific object being labeled). Because baseline performance should be unaffected by speaker, condition, and test version,[2] by-participant and by-item random slopes were not included.

Considering first trials with two familiar objects, model estimates indicate that, overall, toddlers spent equivalent amounts of time on each object ($\beta_0 = 0.034$, β_0 $SE = 0.021$), $t(14.8) = 1.613$, $p = .128$, $d = 0.17$, and this did not differ across experience groups ($\beta_1 = 0.003$, β_1 $SE = 0.022$), $t(286.4) = 0.140$,

[1] Although toddlers in the High Experience group showed an overall increase in looking for the second speaker, whereas those in the Low Experience group did not, test order did not interact with condition for either experience group and, therefore, was not included in our main analyses.

[2] Recall that version refers to the two different target–distractor pairings generated for each target object (see "Design" section above).

p = .889, d = 0.02. For trials with one familiar object and one novel object, baseline looking was assessed as looks to the *familiar* object. Model estimates indicate that, overall, toddlers spent significantly more time on the familiar object (β_0 = 0.071, β_0 SE = 0.015), t(22.2) = 4.618, p < .001, d = 0.36. This is in keeping with other work showing a familiarity bias during baseline with this kind of display (e.g., White & Morgan, 2008). However, the Low Experience group spent significantly less time on the familiar object than the High Experience group (β_1 = −0.040, β_1 SE = 0.014), t(629.1) = −2.820, p = .005, d = −0.21.

Word recognition

Our main analyses involved the naming–baseline difference scores. Condition, speaker, and experience, as well as their interactions, were entered into the model as fixed effects. The maximum random effects structure that allowed for convergence included random intercepts for participant and item as well as by-participant random slopes for speaker and by-item random slopes for version. To assess whether the inclusion of each effect (main effect and interaction) significantly added to the model, maximum likelihood models were compared with and without the effect of interest, with all other possible effects included. These analyses revealed only a significant Condition × Experience interaction, X^2(2) = 14.193, p < .001, with model estimates indicating that experience had an effect on performance for the contrast testing the type of distractor object (β = 0.106, β SE = 0.049), t = 2.152, d = 0.33, as well as the contrast testing the type of target object (β = −0.128, β SE = 0.042), t = −3.056, d = −0.40). No other effects were significant (X^2s ≤ 2.15, ps ≥ .347), including all of those involving speaker.[3]

To further examine the Condition × Experience interaction (Fig. 2), separate models were generated for each experience group with condition entered as a fixed effect. The maximum random effects structure that allowed for convergence included random intercepts for participant and item as well as by-participant random slopes for speaker for the High Experience group and by-item random slopes for version for the Low Experience group. In addition, each condition's estimated difference from baseline (0 or no change; see Table 3) was assessed using intercept-only models (no fixed effects and only random intercepts for participant and item were included).

First, in examining the effect of distractor type, no difference between familiar and novel distractors was found for either the High Experience group (β_1 = −0.052, β_1 SE = 0.034), t(428.8) = −1.538, p = .125, d = −0.16, or the Low Experience group (β_1 = 0.059, β_1 SE = 0.043), t(22.7) = 1.377, p = .182, d = 0.17. However, the effects are in opposing directions for the two groups, thereby explaining the interaction for this contrast. Assessment of the estimated differences from baseline (0 or no change) show that toddlers in both experience groups significantly increased their looking to the target object for familiar labels regardless of speaker or distractor (ps ≤ .025, ds ≥ 0.15). Therefore, whether the distractor object was a familiar versus novel object did not affect familiar label processing for either experience group.

For the type of target object (familiar vs. novel), we found that, whereas toddlers in the High Experience group looked to the target object equally for familiar and novel labels (β_2 = 0.019, β_2 SE = 0.048), t(22.3) = 0.398, p = .694, d = 0.06, toddlers in the Low Experience group increased their looking to the target significantly less for novel labels than for familiar ones (β_2 = −0.108, β_2 SE = 0.039), t(22) = −2.74, p = .012, d = −0.32. Assessments against baseline show that when the novel object was labeled, toddlers in the High Experience group increased their looking to the target object for both speakers (ps ≤ .009, ds ≥ 0.23), whereas toddlers in the Low Experience group did not change their looking from baseline for either speaker (ps ≥ .318, ds ≤ 0.08). Supporting this difference across experience groups, models assessing effects of experience (with random intercepts for participant and item) revealed a significant difference across experience groups for novel labels (β_1 = −0.106, β_1 SE = 0.043), t(42.1) = −2.482, p = .017, d = −0.31, but not for familiar labels (β_1 = 0.028, β_1 SE = 0.029), t(45) = 0.928, p = .358, d = 0.09.

Discussion

When familiar objects were labeled, toddlers responded similarly regardless of the type of distractor or their experience with other children. However, this was not the case with novel labels. On these

[3] Model comparisons show that birth order had no effect on the Condition × Experience interaction (X^2s ≤ 2.00, ps ≥ .176).

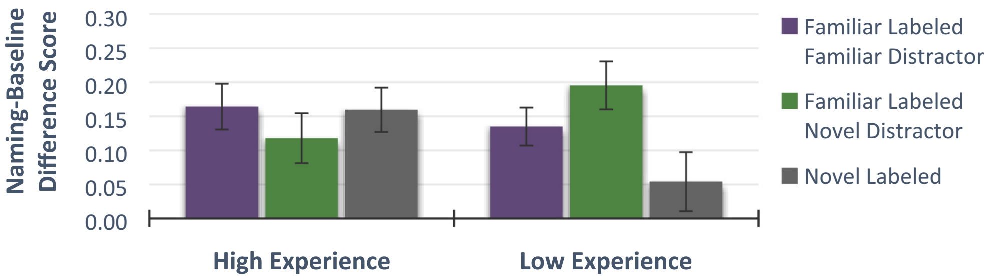

Fig. 2. Estimated change in looking to the *labeled* object for each condition as a function of experience group in Experiment 1. A positive score indicates an increase in looking at the labeled object following naming. Error bars represent standard errors.

Table 3
Estimated proportion changes (and standard errors) in looking to the labeled object for each speaker as a function of condition and experience group.

		Familiar labeled, Familiar distractor	Familiar labeled, Novel distractor	Novel labeled
High	Adult speaker	0.182 (0.046)	0.092 (0.037)	0.189 (0.012)
Experience	Child speaker	0.153 (0.044)	0.142 (0.053)	0.132 (0.035)
Low	Adult speaker	0.105 (0.040)	0.173 (0.037)	0.052 (0.028)
Experience	Child speaker	0.149 (0.047)	0.214 (0.045)	0.051 (0.047)

trials, only toddlers with higher levels of experience with other children showed a disambiguation response and increased their looking to the novel object. Interestingly, these patterns were true regardless of speaker. Therefore, our task did not reveal any differences in the processing of speech produced by an adult and by a 7-year-old child.

Although these results suggest that experience influences the mapping of novel words, one concern is that the effect of experience for novel label trials may have been driven by differences in baseline preference across the groups. In other words, perhaps the toddlers in the High Experience group were more likely to shift to the novel object during the naming phase because of their overall stronger preference for the familiar object during baseline. However, we have at least three arguments against this concern. First, looking patterns during the naming phase alone mirror the reported difference scores for both speakers. Namely, although both experience groups looked to the familiar object at above-chance levels after hearing a familiar label ($ps \le .009$, $ds \ge 0.44$), only the High Experience group trended toward looking at the novel object after hearing a novel label ($ps \le .059$, $ds \ge 0.29$ for the High Experience group vs. $ps \ge .615$, $ds \le 0.09$ for the Low Experience group). Second, if baseline preference determined the degree of naming preference, we would expect a negative correlation between the amount of time spent looking at the familiar object during the baseline and naming phases. This was not the case. Third, baseline preferences on adult speaker trials differed more as a function of experience than baseline preferences on child speaker trials ($d = -0.33$ vs. $d = -0.11$), yet toddlers changed their looking equivalently for both speakers. Therefore, differences in baseline preference do not explain the effect of experience for novel labels.

In summary, we found that toddlers, regardless of experience with other children, were equally likely to look to the familiar object when familiar labels were produced by a child speaker vs. an adult speaker. In other words, toddlers exhibited no difficulty processing the speech of an early school-aged child in this task. However, all familiar labels were pronounced correctly. This means that toddlers may have been able to successfully deduce the target for the child speaker through the use of a more general matching strategy that did not require them to pay close attention to the phonetic detail. Next, we turned to an approach that has been fruitful in testing toddlers' sensitivity to phonetic detail during processing—that of mispronunciation detection.

Experiment 2

Previous work has demonstrated that, for familiar words produced by adults, toddlers are very sensitive to phonetic changes, detecting even slight mispronunciations of those words (e.g., Swingley & Aslin, 2000). For example, toddlers look to an image of a baby less, and more slowly, when it is labeled as *vaby* compared to when it is labeled correctly.

Not only do toddlers show processing costs for mispronunciations, they also show graded sensitivity to the degree of mispronunciation (Mani & Plunkett, 2011; Ren & Morgan, 2011; Tamási, McKean, Gafos, Fritzsche, & Höhle, 2017; White & Morgan, 2008). In looking paradigms, toddlers progressively decrease their looking to the target object following labeling as the degree of mispronunciation increases (White & Morgan, 2008). For example, there may be a relatively small cost associated with a one-feature phonetic change, but there may be a larger cost associated with a two-feature change and an even larger cost associated with a three-feature change. This graded sensitivity has also been demonstrated using pupillometry, where the larger the degree of mispronunciation, the greater the change in pupil diameter (Tamási et al., 2017).

In Experiment 2, we used a mispronunciation procedure to tap toddlers' sensitivity to phonetic mismatch in speech produced by a child. Labels for familiar objects were pronounced either correctly or with an onset mispronunciation of one to three phonetic features. We envisioned at least two possible outcomes. The first was that toddlers would show graded sensitivity to the degree of mispronunciation, with a penalty for even single-feature mispronunciations (the same pattern of response for a child speaker as toddlers have shown in previous research for an adult speaker). This would indicate high sensitivity to phonetic detail in child speech. A second possible outcome was that at least a subset of toddlers would accept anything close to a correct label from a child speaker. This outcome could arise for two reasons. If it were due to difficulty resolving subtle differences (i.e., reduced sensitivity to phonetic detail) in child speech, we would expect to see reduced mispronunciation penalties in the group of toddlers with *less* experience with child speech. If, on the other hand, it were due to a greater tolerance for child mispronunciations, we would expect to see reduced mispronunciation penalties for toddlers with *more* experience with child speech.

Method

Participants

A total of 40 monolingual English-learning toddlers (M_{age} = 21.7 months, range = 20.9–23.2) participated in this experiment. An additional 12 toddlers were tested but not included due to technical issues with equipment (n = 3), fussiness/crying (n = 4), completing fewer than half of the trials successfully using the criteria below (n = 2), parental interference (n = 1), not successfully completing at least two correct label trials (n = 1), or performance on correct label trials more than 3.5 standard deviations below the mean (n = 1).

Half of the participants were classified as High Experience (M_{age} = 21.6 months, range = 20.9–22.6) and half were classified as Low Experience (M_{age} = 21.8 months, range = 21.0–23.2). Experience groups were determined in the same way as in Experiment 1. This time a quarter of the participants in each experience group had older siblings (see Table 4 for summary).

Design

There were 20 test trials in total, each with a unique familiar object–novel object pair (a list of the familiar objects is presented in Appendix A). Of these 20 trials, the familiar object was labeled correctly on 4 trials, the novel object was labeled with a novel word (e.g., *tibble*) on 4 trials, and the familiar object was labeled with one of three types of mispronunciations (4 trials for each type) on the remaining 12 trials. These mispronunciations consisted of onset consonant changes of one to three features that resulted in either a nonword or a word that toddlers are unlikely to know (see Table A2 for the full set of mispronunciations).

Each participant heard one type of pronunciation (correct, one-feature change, two-feature change, or three-feature change) for each familiar target object. The assignment of pronunciations to familiar

Table 4
Hours per week participants in Experiment 2 spent with children other than siblings, reported separately for those who did and did not attend daycare.

		n	Number of older siblings	Total hours[a] [*M* (range)]	Daycare [*M* (range)]	Other [*M* (range)]
High experience						
Daycare	First born	13		34.6 (18–50)	32.2 (16–45)	2.4 (0–8)
	Later born	5	1	27.4 (11–37)	26.0 (10–35)	1.4 (0–3)
Other	First born	2		12.5 (10–15)	0	12.5 (10–15)
Low experience						
	First born	15		3.0 (0–5)	0	3.0 (0–5)
	Later born	5	1–2	3.6 (0–7)	0	3.6 (0–7)

[a] Total hours = Daycare + Other.

objects was counterbalanced across participants. Novel label trials used different familiar object–novel object pairs and were the same across all participants.

There were four blocks of 5 trials (one type of trial per block in pseudorandom order). The blocks were presented in random order. An additional 4-s trial presented the image of a young girl before each block of trials.

Stimuli

Correctly pronounced familiar labels and novel labels were the same as those used in Experiment 1. Mispronounced labels were recorded by the child during the same recording session and were elicited in the same way. Stimuli were presented to participants at a comfortable listening volume of 65–70 dB. Acoustic measures for the stimuli are presented in Table 5, and show that the child speaker's productions of the mispronounced labels did not differ systematically from their familiar counterparts.

Procedure

The procedure was the same as in Experiment 1.

Analysis

As in Experiment 1, looking behavior was coded offline by trained coders blind to condition using customized software at a rate of 30 frames per second (~33.33 ms/frame). This time, however, the difference score (our dependent measure) was computed based on the proportion of time toddlers spent looking at the *familiar* object. Once again, for exposition purposes, proportions for the naming phase as a function of time are provided in the supplementary material. The same criteria for trial exclusion as in Experiment 1 were used. This resulted in 17.3% of trials being discarded for participants not attending to both objects during baseline (High Experience: 18.0%; Low Experience: 16.5%) and 1.0% of trials being discarded for not being on-task for at least 1 s in each of the baseline and naming phases (High Experience: 1.0%; Low Experience: 1.0%).

As in Experiment 1, data were analyzed using linear mixed effects regression. The fixed effect of experience (high vs. low) was contrast-coded, whereas the fixed effect of condition (correct vs. one-feature vs. two-feature vs. three-feature vs. novel) remained treatment-coded to enable a comparison of correct against each of the mispronunciation conditions.

Results

Baseline looking

We first assessed baseline preferences against chance (.50). To do this, baseline looking was centered around 0. Experience was entered as a fixed effect into a linear mixed effects regression model along with random intercepts for participant and item (i.e., the specific object being labeled). Model estimates indicate that, just like in Experiment 1, toddlers spent significantly more time looking at

Table 5
Mean acoustic values for the child speaker for each condition.

	Correct	One-feature change	Two-feature change	Three-feature change	Novel[a]
Full sentence					
Duration (ms)	1582	1665	1633	1613	1626
F0[b] (Hz)	281.2	279.0	273.0	277.8	289.0
F0 variation (SD; Hz)	82.8	76.8	72.6	76.9	76.2
Target word					
Duration (ms)	650	675	658	639	606
F0 (Hz)	326.3	322.7	302.7	321.3	301.4
F0 variation (SD; Hz)	83.9	90.7	84.3	80.3	98.4

[a] $n = 4$ novel labels compared with $n = 16$ for each of the other conditions.
[b] F0 is the acoustical parameter most closely related to perceived pitch. F0 values were calculated using a floor of 100 Hz and a ceiling of 600 Hz.

the familiar object ($\beta_0 = 0.097$, β_0 *SE* = 0.014), $t(19.9) = 7.162$, $p < .001$, $d = 0.53$, and toddlers in the Low Experience group spent marginally less time on the familiar object than those in the High Experience group ($\beta_1 = -0.031$, β_1 *SE* = 0.015), $t(34.7) = -2.013$, $p = .052$, $d = -0.17$.

Word recognition

As in Experiment 1, our main analyses involve the naming–baseline difference scores. The maximum random effects structure that allowed for convergence for each of the following analyses included only random intercepts for participant and item. We first considered only trials in which the familiar object was labeled correctly, with experience entered into the model as a fixed factor. This analysis showed that toddlers recognized these words, increasing their looking to the familiar object from baseline ($\beta_0 = 0.161$, β_0 *SE* = 0.030), $t(14.7) = 5.317$, $p < .001$, $d = 0.53$, with no difference across experience groups ($\beta_1 = -0.009$, β_1 *SE* = 0.063, $t(35.7) = 0.951$, $p = .857$, $d = -0.03$. Evaluation of the estimated difference from baseline (0 or no change) via intercept-only models showed that both the High Experience group ($\beta_0 = 0.169$, β_0 *SE* = 0.035) and the Low Experience group ($\beta_0 = 0.152$, β_0 *SE* = 0.049) significantly increased their looking to the familiar object (*ps* $\leq$ .007, *ds* $\geq$ 0.48).

We next assessed whether there was an effect of pronunciation type across the four conditions in which the familiar object was labeled (correct vs. one-feature vs. two-feature vs. three-feature). Condition, experience, and their interaction were entered into the model as fixed effects. Maximum likelihood model comparisons were performed to assess whether the inclusion of each effect significantly added to the model. This analysis revealed only a significant main effect of condition, $X^2(3) = 41.42$, $p < .001$, with model estimates indicating that toddlers were less likely to increase their looking to a familiar object for mispronounced labels than for correct ones and that this was true for all degrees of change (one-feature: $\beta_1 = -0.088$, β_1 *SE* = 0.040, $t = -2.186$, $d = -0.27$; two-feature: $\beta_2 = -0.149$, β_2 *SE* = 0.040, $t = -3.716$, $d = -0.45$; three-feature: $\beta_3 = -0.255$, β_3 *SE* = 0.040, $t = -6.349$, $d = -0.77$). No other effects reached significance (X^2s $\leq$ 1.29, *ps* $\geq$ .455).

Given our hypotheses about the role of child experience, we conducted exploratory analyses to examine the effects of experience despite the lack of interaction involving this factor. To do so, models of the effect of condition were generated separately for each experience group (Fig. 3). For the High Experience group, the model showed that all three types of mispronunciations significantly differed from the correct condition (one-feature: $\beta_1 = -0.114$, β_1 *SE* = 0.057, $p = .047$, $d = -0.35$; two-feature: $\beta_2 = -0.154$, β_2 *SE* = 0.056, $p = .007$, $d = -0.47$; three-feature: $\beta_3 = -0.297$, β_3 *SE* = 0.057, $p < .001$, $d = -0.90$). Evaluation of each condition's estimated difference from baseline showed that toddlers did not change their looking for either a one-feature change ($\beta_0 = 0.051$, β_0 *SE* = 0.044, $p = .258$, $d = 0.14$) or a two-feature change ($\beta_0 = 0.012$, β_0 *SE* = 0.043, $p = .777$, $d < 0.01$), but they significantly *decreased* their looking to the familiar object for a three-feature change ($\beta_0 = -0.124$, β_0 *SE* = 0.055, $p = .039$, $d = -0.35$). In contrast, the model for the Low Experience group showed that the smallest change was not treated differently from the correct condition (one-feature: $\beta_1 = -0.058$, β_1 *SE* = 0.056, $p = .305$, $d = -0.17$), but larger changes were (two-feature: $\beta_2 = -0.142$, β_2 *SE* = 0.057, $p = .012$, $d = -0.43$; three-feature: $\beta_3 = -0.211$, β_3 *SE* = 0.057, $p < .001$, $d = -0.63$). Estimated differ-

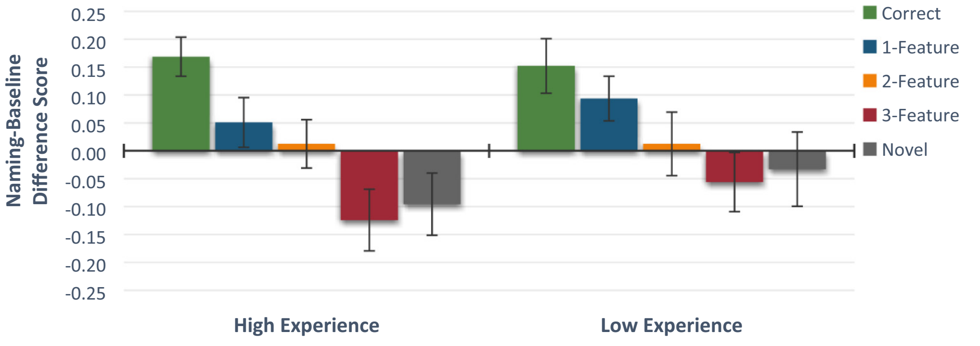

Fig. 3. Estimated change in looking to the *familiar* object for each condition as a function of experience group in Experiment 2. A positive score indicates an increase in looking at the familiar object following naming. Error bars represent standard errors.

ences from baseline showed that toddlers significantly increased their looking to the familiar object for a one-feature change (β_0 = 0.094, β_0 *SE* = 0.040, *p* = .034, *d* = 0.29) and showed no difference from baseline for a two-feature change (β_0 = 0.012, β_0 *SE* = 0.057, *p* = .831, *d* = 0.03) or a three-feature change (β_0 = −0.06, β_0 *SE* = 0.05, *p* = .658, *d* = −0.16). Despite these somewhat different patterns of significance for the two groups, models that include only the three types of mispronunciations coded to assess a polynomial fit reveal significant linear trends for both the High Experience group (β_1 = −0.128, β_1 *SE* = 0.043), *t*(184.1) = −3.016, *p* = .003, and the Low Experience group (β_1 = −0.109, β_1 *SE* = 0.040), *t*(171.8) = −2.749, *p* = .007. Together, these analyses indicate that, although toddlers in the Low Experience group showed somewhat reduced mispronunciation penalties compared with toddlers in the High Experience group, both groups were sensitive to the degree of mispronunciation in child speech.

Finally, we looked at performance on novel label trials with experience entered into the model as a fixed factor. Overall, toddlers did not significantly change their looking to the familiar object from baseline (β_0 = −0.065, β_0 *SE* = 0.037), *t*(3.3) = −1.827, *p* = .156, *d* = −0.19, with no significant difference across experience groups (β_1 = 0.060, β_1 *SE* = 0.064), *t*(35.7) = 0.951, *p* = .348, *d* = 0.18. However, as can be seen in Fig. 3, whereas the High Experience group trended toward the novel object (β_0 = −0.096, β_0 *SE* = 0.056), *t*(6.4) = −1.75, *p* = .134, *d* = −0.29, the Low Experience group did not change their looking from baseline (β_0 = −0.038, β_0 *SE* = 0.067), *t*(2.9) = −0.492, *p* = .658, *d* = −0.09. This pattern is similar to that in Experiment 1, where toddlers in the High Experience group looked to the target object equivalently for familiar and novel labels, but toddlers in the Low Experience group looked to the target object significantly less for novel labels than for familiar ones. This suggests that, as in Experiment 1, it is in the processing of novel labels that effects of experience appear. To explore this possibility, we assessed the similarity of these patterns acros the two experiments directly.

Novel label processing across experiments

To compare performance on familiar and novel label trials across experiments, we compared responses for the two trial types that appeared in both experiments (for Experiment 1, only trials with the child speaker were included)—trials involving familiar–novel object pairs with either correctly pronounced familiar labels or novel labels. To do this, we first recoded the novel label data for Experiment 2 so that they represented *looks to the labeled object* (as in Experiment 1). We then pooled the data for these two trial types across experiments, resulting in 44 toddlers in each experience group. Condition (familiar vs. novel), experience (high vs. low), and Experiment (1 vs. 2) were entered into the model as contrast-coded fixed effects along with their interactions. The maximum random effects structure that allowed for convergence consisted of random intercepts for participant and item,[4] as well as a by-item random slope for experiment. Maximum likelihood model comparisons revealed a main effect of condition, $X^2(1)$ = 4.59, *p* = .032, as well as a Condition × Experience interaction, $X^2(1)$

[4] Due to the use of different novel object–label pairings across participants, for this analysis item represented the label used rather than the object being labeled.

= 4.75, p = .029. The model showed that overall toddlers increased their looking to the target object less for a novel label than for a familiar one (β = −0.092, β SE = 0.041, t = −2.272, d = −0.28) and this difference was more pronounced for the Low Experience group than the High Experience group (β = −0.99, β SE = 0.051, t = −1.930, d = −0.31). No other effects reached significance (X^2s ≤ 0.93, ps ≥ .335) (Fig. 4).

To assess the Condition × Experience interaction (Fig. 4), separate models for each experience group were generated with condition as a fixed effect along with random intercepts for participant and item (the maximum random effects structure). These analyses showed that toddlers in the High Experience group increased their looking to the target object similarly for familiar and novel labels (β_1 = −0.040, β_1 SE = 0.046), t(4.8) = −0.861, p = .430, d = −0.13, whereas toddlers in the Low Experience group increased their looking to the target object for novel labels significantly less than they did for familiar labels (β_1 = −0.143, β_1 SE = 0.041), t(6.1) = −3.448, p = .013, d = −0.43. Intercept-only mixed effects models for each condition show that toddlers in both the High Experience and Low Experience groups significantly increased their looking to the target object for familiar labels (β_0s ≥ 0.156, ps < .001, d ≥ 0.54), but only those in the High Experience group did so for novel labels (β_0 = 0.117, β_0 SE = 0.023), t(38.8) = 4.045, p < .001, d = 0.38. Toddlers in the Low Experience group did not change their looking from baseline for novel labels (β_0 = 0.044, β_0 SE = 0.029), t(149) = 1.505, p = .134, d = 0.12. This indicates that, across experiments, all toddlers recognized the familiar objects, but only toddlers in the High Experience group showed a disambiguation response for the novel labels.

Discussion

When object labels were pronounced correctly, toddlers successfully mapped them to the appropriate referents. When these same familiar object labels were mispronounced, toddlers' looking behavior differed from the correct pronunciations, showing that overall they were sensitive to each level of change. However, subtle differences arose in the treatment of one-feature and three-feature changes across experience groups. Toddlers in the High Experience group demonstrated somewhat greater sensitivity to phonetic differences, showing a larger penalty for both one-feature and three-feature mispronunciations than toddlers in the Low Experience group. Although the difference across groups was not significant, this pattern suggests that toddlers with limited experience with other children may have some difficulty resolving subtle differences in child speech. That said, regardless of experience, toddlers showed graded sensitivity to the degree of change, with progressively larger mispronunciation penalties for one-feature, two-feature, and three-feature changes, a pattern that is similar to the one previously found for toddlers' processing of adult speech (White & Morgan, 2008). Thus, despite some differences between the two experience groups, toddlers were overall very sensitive to the phonetic content of the child speech, suggesting that high levels of child experience are not required to demonstrate this sensitivity. Finally, this experiment also provides additional evidence

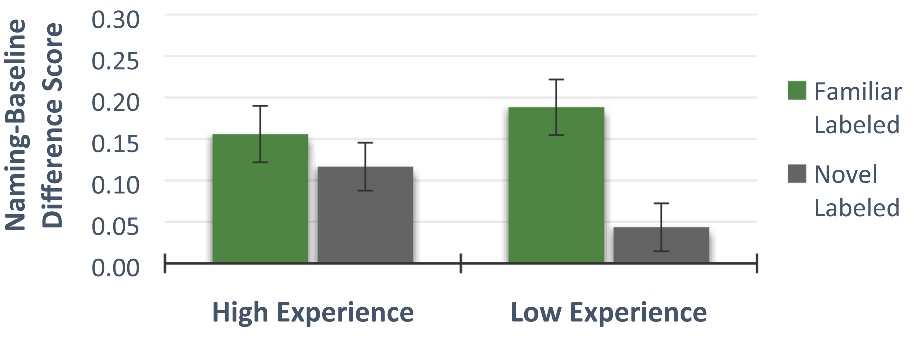

Fig. 4. Estimated change in looking to the *labeled* object for familiar labels in the presence of a novel distractor and novel labels as a function of experience group. A positive score indicates an increase in looking at the labeled object following naming. Estimates are for the child speaker pooled across both experiments. Error bars represent standard errors.

that toddlers with varying levels of child experience differ in their treatment of novel labels. We return to this topic in the General Discussion.

General discussion

Across two experiments, we have shown that toddlers process familiar labels from a 7-year-old speaker in our task as accurately, and with as much sensitivity, as they process labels from an adult. For the most part, this sensitivity was unaffected by toddlers' experience with other children. However, only toddlers with experience in interacting regularly with a variety of other children showed a disambiguation response for novel labels.

We predicted that the effects of increased experience might manifest in one of two ways: first, as increased accuracy in processing child speech (i.e., higher looking to the target in Experiment 1 and higher sensitivity to phonetic changes in Experiment 2) or, second, as increased tolerance for deviations in child speech (in Experiment 2). Our results were not strongly in line with either of these predictions. Instead, we found that toddlers were equally accurate in processing familiar labels regardless of their amount of experience with child speech, although there was a hint that toddlers with less experience showed reduced sensitivity to mispronunciations in Experiment 2 (consistent with some effect of experience on processing the fine details of child speech).

It may be that our child speaker was too old (and her productions too mature) for us to have seen significant effects on processing. In two previous studies (Cooper et al., 2018; Dodd, 1975), 2- to 4-year-old children were less accurate in processing their own speech, as well as the speech of another child, compared with that of an adult. One obvious difference between studies is the age of the child speaker (2–4 years vs. 7 years). Together, these findings suggest that toddlers are able to cope with the more subtle deviations that older children's speech may exhibit and that it is the larger distortions (e.g., substitutions, deletions) that impair comprehension. Nonetheless, we cannot rule out the possibility that other measures (e.g., pupil dilation) or more challenging listening conditions (e.g., with background noise) might reveal differences in the processing of speech from adults and older children. Future work could explore speech from a greater range of speaker ages, and using a greater range of tasks, to determine at what age children's speech is processed in a mature manner and whether experience with speech from other children is more important for processing younger children's speech.

Our findings suggest that toddlers are quite adept at processing the speech of a child speaker, but our study leaves open the question of what assumptions toddlers made about the age of the speaker. Adults are extremely sensitive to the acoustic markers of speaker age (Amir, Engel, Shabtai, & Amir, 2012; Cerrato, Falcone, & Paoloni, 2000) and would have recognized that the child speaker was young based on the acoustic properties of her voice (e.g., higher mean pitch). However, it is unclear at what age this ability develops. Because we were purely interested in toddlers' ability to process the speech itself, our task did not require that they be aware of the speakers' ages. Future work could explore when children begin to map acoustic properties to speaker age.

Although we did not find a difference in accuracy for the child and adult speakers, our results showed that toddlers who had little exposure to other children did not change their looking from baseline after hearing a novel label and that, unexpectedly, this was true for both the child and adult speakers. In other words, this group did not show the expected disambiguation response that is typically seen in monolingual toddlers as young as 17 months when a novel label is presented by an adult (Byers-Heinlein & Werker, 2009; Halberda, 2003; White & Morgan, 2008). Given that this was true for both speakers, it seems unlikely that this is due to difficulties in processing child speech in particular.

Why is it that only toddlers whose experience came from a variety of child speakers, such as occurs in a daycare setting, successfully mapped novel labels to novel objects? Although we focused on their amount of experience with other children, toddlers who are in group settings are also likely to hear a greater variety of adult speakers (although there are certainly some exceptions to this, e.g., home-based daycare settings with only one adult). Therefore, it is possible that there is an influence of speaker variability more generally. Previous work has demonstrated that speaker variability introduced during the learning process affects attention to phonetic detail and the generalizability of word representations (e.g., Rost & McMurray, 2009; Singh, 2008). If speaker variability played a role here, it would suggest that hearing speech from a greater variety of speakers (child and/or adult) affects the efficiency or robustness of *subsequent* word learning with new speakers.

Another possibility is that exposure to multiple speakers in the environment has effects on vocabulary size (perhaps via the influence of variability or as a result of having a wider variety of interactions in which different topics or objects are discussed) and that this in turn affects novel label processing. Indeed, the connection between number of speakers and vocabulary size has been demonstrated for bilingual children (Place & Hoff, 2011), as has the connection between vocabulary size and the use of a disambiguation strategy (Bion, Borovsky, & Fernald, 2013). Together, these studies are consistent with the possibility that toddlers with language exposure from a wider variety of individuals (all else being equal) may have larger vocabularies and that these toddlers with larger vocabularies will be more likely to disambiguate novel labels. However, this connection remains speculative at this point.

Although we did not test vocabulary levels directly, one aspect of our findings that is potentially relevant is the difference in baseline familiarity bias between the High Experience and Low Experience groups for trials with a familiar–novel object pair. In both Experiment 1 and Experiment 2, toddlers with more experience with other children showed a stronger baseline preference for the familiar object. This stronger preference could indicate that toddlers in the High Experience group were more familiar with the familiar object labels (Schafer, Plunkett, & Harris, 1999). Previous work has shown that children's degree of knowledge about specific familiar labels predicts their disambiguation response when a novel label is presented (Grassmann, Schulze, & Tomasello, 2015). Therefore, it is possible that weaker knowledge of individual familiar labels was responsible for the difficulty toddlers in the Low Experience group had with novel labels. It may also be that knowledge of individual familiar labels is correlated with vocabulary size, consistent with a link between processing efficiency and vocabulary size (Fernald, Perfors, & Marchman, 2006).

To more directly investigate the potential role of vocabulary in explaining our findings, we conducted a follow-up with a separate sample of same-aged children drawn from the same population ($N = 47$; 24 female, 20 male, and 3 undisclosed). In this follow-up, we administered parent questionnaires and the MacArthur Communicative Development Inventories short form (Fenson et al., 2000) to examine the relationship between vocabulary size and experience in group settings. However, inconsistent with a role for vocabulary size, we found no relationship between vocabulary size and the amount of time toddlers spent with other children outside the home ($r = -.08$). In other words, toddlers' vocabulary size was similar regardless of whether they attended daycare, engaged in regular playgroups, or spent the majority of their time at home. In addition, maternal education did not differ across these three groups, with more than 80% of mothers in each group having 3 or more years of postsecondary education. Although not conclusive, these survey results (drawn from the same population) suggest that neither the quality of maternal input (as approximated by maternal education) nor toddlers' vocabulary size was responsible for the difference in novel label processing in the current study.

The lack of a relationship between care setting and vocabulary in our follow-up survey is not entirely surprising given that previous literature has not found consistent effects of daycare on language outcomes. For example, Booth, Clarke-Stewart, Vandell, McCartney, and Owen (2002) observed no differences in 15-month-olds' vocabulary scores as a function of the children's daycare status, whereas Laing and Bergelson (2019) found that 17-month-olds with a combination of home care and daycare had larger vocabularies than toddlers with either home care or daycare alone. But in a large-scale NICHD ECCRN (2000) study, 3-year-olds in high-quality daycare did have larger vocabularies than children in other care situations. Thus, overall, it appears to be the quality of care that is most important (Burchinal, Roberts, Nabors, & Bryant, 1996; McCartney, 1984), although there is considerable debate about how quality of care should be assessed (see Falenchuk, Perlman, McMullen, Fletcher, & Shah, 2017, and Perlman et al., 2016, 2017, for meta-analyses of factors such as the education of staff, child–staff ratios, and child–staff interaction quality).

Although our discussion so far has focused on how the number of speakers and vocabulary size might contribute to novel word mapping, it is also possible that other factors differed across our experience groups. One possibility is that environments such as daycare settings provide more opportunities for structured word learning play, and that this in turn affected performance on our novel label trials. However, there are a wide variety of daycare settings, not all of which provide these very organized and deliberate kinds of learning experiences to the same degree. In addition, not all of the toddlers in our High Experience group had experience with formal childcare settings; some had experience with other children through informal playgroups and community-based activities such as swimming and gymnastics. Therefore, it remains possible that something about interacting with other children (or a variety of different individuals) may provide toddlers with the kinds of experience that boost novel label processing. Because this was not the primary goal of the current study, future work will be needed to examine more systematically this intriguing relationship among novel word processing, type of experience with other children, and other experiential factors.

In conclusion, toddlers are exposed not only to the speech of adults, but also to the speech of other children. The current study is the first demonstration that toddlers show considerable sensitivity to phonetic detail in the speech of a child, processing it as well as that of an adult. This was true (although to a somewhat lesser extent) even for toddlers with less experience hearing other children. Given this sensitivity, speech from children of this age could be useful input for young language learners. A full picture of the impact that this has on toddlers' language development will require additional research, especially because some studies of toddlers in other cultures (e.g., the Mayan Yucatec) indicate that their vocabulary size is best predicted by the quantity of speech directed at them from *adults* and not from the children with whom they spend most of their time (Shneidman & Goldin-Meadow, 2012). Finally, our findings also demonstrate that toddlers with more versus less exposure to speech from other children show a difference in their treatment of novel labels even when they are produced by adults. Whether this is truly due to the number or variety of other child speakers or to other correlated aspects of toddlers' language or broader environment is unclear. These are intriguing questions for future research.

Acknowledgements

The authors thank Emily McIntosh, Joel LeForestier, Shaquille Sealy, Erin Kim, Ashley Blayney-Hoffer, and Ayah Taji for help with recruitment and coding. We also thank the many families who volunteered their time to participate. This work was funded by an operating grant from the Natural Sciences and Engineering Research Council of Canada to KSW.

Appendix A. Familiar objects and novel labels used in each experiment

Table A1 and A2.

Table A1
Familiar objects and novel labels used in Experiment 1.

Block 1			Block 2		
Set	Labelled Object	Distractor Object	Set	Labelled Object	Distractor Object
1	monkey	flower	1	pants	bunny
	ball	*novel*		keys	*novel*
	murg/zurk	shirt		*murg/zurk*	hand
2	shoe	bike	2	phone	truck
	dog	*novel*		cookie	*novel*
	semp/neech	balloon		*semp/neech*	bird
3	vacuum	banana	3	mouth	brush
	book	*novel*		fish	*novel*
	gorp/tilk	horse		*gorp/tilk*	block
4	cat	doll	4	cup	hat
	foot	*novel*		sock	*novel*
	tibble/boogle	spoon		*tibble/boogle*	chair

Note: Within a set, distractor objects were counterbalanced across participants. The order of the 2 blocks, and the particular speaker (child, adult) assigned to each block, were counterbalanced across participants. Novel labels are presented as child/adult productions.

Table A2
Familiar, mispronounced, and novel labels used in Experiment 2.

Set	Correct/Novel Label	Mispronounced labels[a]		
		One-feature	Two-feature	Three-feature
1	monkey	bunky	gunky	tunky
	ball	gall	nall	sall
	mouth	nouth	pouth	kouth
	fish	pish	zish	gish
	tibble	familiar distractor: chair		
2	shoe	foo	voo	goo
	pants	tants	sants	nants
	dog	nog	vog	fog
	keys	gees	dees	zees
	gorp	familiar distractor: doll		
3	sock	zock	pock	bock
	book	gook	nook	sook
	vacuum	zacuum	pacuum	tacuum
	cup	gup	shup	vup
	semp	familiar distractor: jacket/coat		
4	cat	tat	dat	zat
	foot	soot	buut	guut
	phone	pone	tone	doan
	cookie	pookie	dookie	mookie
	murg	familiar distractor: block		

Note. Toddlers heard each familiar label produced with one of four pronunciations (correct or with a one-feature, two-feature, or three-feature onset mispronunciation). Within a set, the particular items assigned to each pronunciation were counterbalanced across participants.

[a] Rimes are the same as those of the correct productions regardless of spelling.

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2019.04.021.

References

Amir, O., Engel, M., Shabtai, E., & Amir, N. (2012). Identification of children's gender and age by listeners. *Journal of Voice, 26*, 313–321.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Bernier, D. E., & Soderstrom, M. (2016). [LENA recording study contrasting different child care environments]. Unpublished raw data.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science, 20*, 539–542.

Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition, 126*, 39–53.

Boersma, P., & Weenink, D. (2014). *Praat: Doing phonetics by computer [computer program]* http://www.praat.org/.

Booth, C. L., Clarke-Stewart, K. A., Vandell, D. L., McCartney, K., & Owen, M. T. (2002). Child-care usage and mother–infant "quality time". *Journal of Marriage and Family, 64*, 16–26.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*, 9.

Burchinal, M. R., Roberts, J. E., Nabors, L. A., & Bryant, D. M. (1996). Quality of center child care and infant cognitive and language development. *Child Development, 67*, 606–620.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science, 12*, 815–823.

Cerrato, L., Falcone, M., & Paoloni, A. (2000). Subjective age estimation of telephonic voices. *Speech Communication, 31*, 107–112.

Cooper, A., Fecher, N., & Johnson, E. K. (2018). Toddlers' comprehension of adult and child talkers: Adult targets versus vocal tract similarity. *Cognition, 173*, 16–20.

Dodd, B. (1975). Children's understanding of their own phonological forms. *Quarterly Journal of Experimental Psychology, 27*, 165–172.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics & Phonetics, 17*, 617–643.
Falenchuk, O., Perlman, M., McMullen, E., Fletcher, B., & Shah, P. S. (2017). Education of staff in preschool aged classrooms in child care centers and child outcomes: A meta-analysis and systematic review. *PLoS One, 12*(8), e183673.
Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21*, 95–116.
Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*, 98–116.
Flipsen, P. J. (2006). Measuring the intelligibility of conversational speech in children. *Clinical Linguistics & Phonetics, 20*, 303–312.
Grassmann, S., Schulze, C., & Tomasello, M. (2015). Children's level of word knowledge predicts their exclusion of familiar objects as referents of novel words. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.01200.
Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*, B23–B34.
Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*, 3099–3111.
Hodson, B. W., & Paden, E. P. (1981). Phonological processes which characterize unintelligible and intelligible speech in early childhood. *Journal of Speech and Hearing Disorders, 46*, 369–373.
Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: Results of functional data analysis. *Journal of the Acoustical Society of America, 124*, 3158–3170.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26.
Laing, C., & Bergelson, E. (2019). Mothers' work status and 17-month-olds' productive vocabulary. *Infancy, 24*, 101–109.
Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*, 1455–1511.
Lin, S., Inkelas, S., McDonnaughey, L., & Dohn, M. (2016). /l/ in clusters: An articulatory–acoustic study of children's productions. *Paper presented at The 15th Conference on Laboratory Phonology*. NY: Ithika.
Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*, 1494–1502.
Maas, E., & Mailend, M. L. (2017). Fricative contrast and coarticulation in children with and without speech sound disorders. *American Journal of Speech-Language Pathology, 26*, 649–663.
Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language, 38*, 606–627.
McCartney, K. (1984). Effect of quality of day care environment on children's language development. *Developmental Psychology, 20*, 244–260.
Munson, B. (2004). Variability in /s/ production in children and adults: Evidence from dynamic measures of spectral mean. *Journal of Speech, Language, and Hearing Research, 47*, 58–69.
Munson, B., Johnson, J. M., & Edwards, J. (2012). The role of experience in the perception of phonetic detail in children's speech: A comparison between speech-language pathologists and clinically untrained listeners. *American Journal of Speech-Language Pathology, 21*, 124–139.
National Institute of Child Health and Human Development Early Child Care Research Network (2000). The relation of child care to cognitive and language development: National Institute of Child Health and Human Development Early Child Care Research Network. *Child Development, 71*, 960–980.
Nissen, S. L., & Fox, R. A. (2005). Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective. *Journal of the Acoustical Society of America, 118*, 2570–2578.
Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238.
Perlman, M., Falenchuk, O., Fletcher, B., McMullen, E., Beyene, J., & Shah, P. S. (2016). A systematic review and meta-analysis of a measure of staff/child interaction quality (the Classroom Assessment Scoring System) in early childhood education and care settings and child outcomes. *PLoS One, 11*(12), e167660.
Perlman, M., Fletcher, B., Falenchuk, O., Brunsek, A., McMullen, E., & Shah, P. S. (2017). Child–staff ratios in early childhood education and care settings and child outcomes: A systematic review and meta-analysis. *PLoS One, 12*(1), e170256.
Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development, 82*, 1834–1849.
R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Ren, J., & Morgan, J. L. (2011). Sub-segmental details in early lexical representation of consonants. In W.-S. Lee & E. Zee (Eds.), *Proceedings of The 17th International Congress of Phonetic Sciences (ICPhS XVII)*. Hong Kong: City University of Hong Kong.
Romeo, R., Hazan, V., & Pettinato, M. (2013). Developmental and gender-related trends of intra-talker variability in consonant production. *Journal of the Acoustical Society of America, 134*, 3781–3792.
Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*, 339–349.
Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy, 15*, 608–635.
Schafer, G., Plunkett, K., & Harris, P. L. (1999). What's in a name? Lexical knowledge drives infants' visual preferences in the absence of referential input. *Developmental Science, 22*, 187–194.
Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science, 15*, 732–738.
Seidl, A., Onishi, K. H., & Cristia, A. (2014). Talker variation aids young infants' phonotactic learning. *Language Learning and Development, 10*, 297–307.
Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science, 15*, 659–673.
Singh, L. (2008). Influences of high and low variability on infant word recognition. *Cognition, 106*, 833–870.

Soderstrom, M., Grauer, E., Dufault, B., & McDivitt, K. (2018). Influences of number of adults and adult:child ratios on the quantity of adult language input across childcare settings. *First Language, 38*, 563–581.

Strömbergsson, S., Wengelin, Å., & House, D. (2014). Children's perception of their synthetically corrected speech production. *Clinical Linguistics & Phonetics, 28*, 373–395.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76*, 147–166.

Tamási, K., McKean, C., Gafos, A., Fritzsche, T., & Höhle, B. (2017). Pupillometry registers toddlers' sensitivity to degrees of mispronunciation. *Journal of Experimental Child Psychology, 153*, 140–148.

van Heugten, M., & Johnson, E. K. (2014). Learning to contend with accents in infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General, 143*, 340–350.

van Heugten, M., Krieger, D. R., & Johnson, E. K. (2015). The developmental trajectory of toddlers' comprehension of unfamiliar regional accents. *Language Learning and Development, 11*, 41–65.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143*, 2020–2045.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science, 14*, 372–384.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language, 59*, 114–132.

Yurovsky, D., Case, S., & Frank, M. C. (2017). Preschoolers flexibly adapt to linguistic input in a noisy channel. *Psychological Science, 28*, 132–140.