Received: 15 March 2018 | Revised: 25 May 2019 | Accepted: 5 June 2019
DOI: 10.1111/desc.12881

PAPER

Developmental Science WILEY

Actions speak louder than words: Differences in memory flexibility between monolingual and bilingual 18-month-olds

Rachel Barr[1] | **Sylvia N. Rusnak**[1] | **Natalie H. Brito**[2] | **Courtney Nugent**[1]

[1]Georgetown University, Washington, District of Columbia
[2]New York University, New York, New York

Correspondence
Rachel Barr, Georgetown University, Washington, DC.
Email: rfb5@georgetown.edu

Funding information
Division of Behavioral and Cognitive Sciences, Grant/Award Number: BCS-1551719

Abstract

Bilingual infants from 6- to 24-months of age are more likely to generalize, flexibly reproducing actions on novel objects significantly more often than age-matched monolingual infants are. In the current study, we examine whether the addition of novel verbal labels enhances memory generalization in a perceptually complex imitation task. We hypothesized that labels would provide an additional retrieval cue and aid memory generalization for bilingual infants. Specifically, we hypothesized that bilinguals might be more likely than monolinguals to map multiple perceptual features onto a novel label and therefore show enhanced generalization. Eighty-seven 18-month-old monolingual and bilingual infants were randomly assigned to one of two experimental conditions or a baseline control condition. In the experimental conditions, either no label or a novel label was added during demonstration and again at the beginning of the test session. After a 24-hr delay, infants were tested with the same stimulus set to test cued recall and with a perceptually different but functionally equivalent stimulus set to test memory generalization. Bilinguals performed significantly above baseline on both cued recall and memory generalization in both experimental conditions, whereas monolinguals performed significantly above baseline only on cued recall in both experimental conditions. These findings show a difference between monolinguals and bilinguals in memory generalization and suggest that generalization differences between groups may arise from visual perceptual processing rather than linguistic processing. A video abstract of this article can be viewed at https://youtu.be/yXB4pM3fF2k

KEYWORDS
bilingual, imitation, infant, labels, memory flexibility, memory generalization

1 | INTRODUCTION

Bilinguals are exposed to wide variations in language, and these experiences are known to alter perceptual trajectories during infancy (Graf Estes & Hay, 2015; Kovács & Mehler, 2009a, 2009b; Sebastián-Gallés, 2010; Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012; Singh, Loh, & Xaio, 2017; Weikum et al., 2007; Werker, 2012). Bilinguals show enhanced auditory and visual discrimination; French-English (Weikum et al., 2007) and Spanish-Catalan (Sebastián-Gallés et al., 2012) bilingual 8-month-olds discriminate French and English on the basis of visual cues, but monolingual infants do not. In contrast, monolinguals show perceptual narrowing earlier than bilingual infants do across multiple domains during the first year of life. Monolinguals more rapidly perceptually tune to the phonemes of their native language than bilinguals do and are more likely to process audiovisual streams that match both their own race and native language (for review, see Pascalis et al., 2014). Singh et al. (2017) examined 11-month-olds on both face discrimination and perceptual auditory discrimination tasks. They found that monolinguals and bilinguals exhibited similar face discrimination patterns, but bilinguals exhibited more flexible auditory perceptual discrimination than monolinguals. Bilingual theorists (Byers-Heinlein, 2014; Kovács & Mehler, 2009b; Sebastián-Gallés et al., 2012; Singh et al., 2017; Werker & Byers-Heinlein, 2008) argue that bilingual language acquisition is boot-strapped by early perceptual discrimination of surface language features in both auditory and visual domains.

It is highly likely that perceptual processing capacities not only influence how infants learn language, but also how they process, encode, and remember information non-verbally. Therefore, bilingual language acquisition may have implications for general memory processing and cognitive outcomes as well. Although perceptual effects of bilingual exposure during infancy have been extensively researched, Barac, Bialystok, Castro, and Sanchez (2014) noted that there has been little investigation of how bilingualism might alter early memory processing trajectories. This is despite the fact that memory is essential for knowledge acquisition across all domains. Furthermore, because perceptual, linguistic, and memory systems are less specialized early in development, early modifications in one system are likely to affect the developmental trajectories and the underlying neural architecture of other systems as well (Costa & Sébastian-Gallés, 2014; D'Souza & Karmiloff-Smith, 2011; Newcombe, 2011). Consistent with this view, non-linguistic cognitive differences between monolinguals and bilinguals have been reported in infants as young as 6- and 7-months of age (Brito & Barr, 2014; Kovács & Mehler, 2009a; Singh et al., 2015). For example, using a standard habituation paradigm, 6-month-old bilingual infants showed better visual recognition memory for familiar items than monolingual infants did (Singh et al., 2015). These findings suggest that language exposure, and not language expression, contributes to early emerging differences in memory processing between monolinguals and bilinguals and that this divergence in processing in multiple domains begins during infancy.

There is a direct link between perceptual processing and nonverbal memory retrieval (Gerhardstein, Lui, & Rovee-Collier, 1998). Infants are faced with the daily mnemonic challenge of learning about their world from a variety of sources, and they must then apply what they learned to diverse problems, in different contexts, and sometimes after long delays. The ability to retrieve memories, despite changes in perceptual cues, allows learning to be generalized to novel situations. This ability to generalize information to novel situations has been referred to as memory flexibility (Barr & Brito, 2014; Eichenbaum, 1997; Hayne, 2006; Karmiloff-Smith, 1992). Indexed via memory generalization paradigms, memory flexibility emerges gradually during development. Early in development, successful memory performance is contingent on an exact match between the cues at the time of encoding and the cues available at retrieval. A mismatch at learning and test can decrease memory performance, but with age, toddlers can increasingly tolerate differences between conditions at encoding and retrieval (for review, see Barr & Brito, 2014; Karmiloff-Smith, 1992). Memory flexibility is needed to extend learning beyond individual exemplars. For example, in a memory generalization deferred imitation task, infants are shown how to construct a rattle. At test, they are given a different set of materials that can also be used to make a rattle. If infants recognize that it is possible to assemble a rattle with new, perceptually different stimuli, and thus successfully generalize across perceptually different materials, they are demonstrating memory flexibility. That is, memory becomes more flexible across perceptual changes during development.

Research Highlights

- Bilingual infants showed a performance difference in a complex memory generalization task but not in cued recall compared to monolingual infants.
- Bilinguals generalized across cues after a 24-hour delay under high levels of cognitive load.
- Labels did not aid or interfere with generalization performance in 18-month-old bilinguals.

Several studies have examined the effects of bilingualism on performance on memory generalization imitation paradigms at a variety of ages. In the puppet paradigm, children are first exposed to one puppet (e.g., a gray rabbit) and are shown three target actions performed on that puppet. After a delay, children are exposed to another puppet (e.g., a pink mouse). Using this paradigm, 6-month-old bilingual infants were able to generalize across perceptual changes in shape and color but monolingual infants were not (Brito & Barr, 2014). When cognitive load was reduced and infants were presented with a less complex task of generalizing across only one perceptual feature, color (i.e., pink mouse to gray mouse), monolingual 6-month-olds were then also able to generalize (Brito & Barr, 2014). At 18 months, bilinguals generalize across more complex perceptual changes in puppet stimuli (e.g., yellow duck to black and white cow) than monolinguals do—again showing a shift in the developmental trajectory toward greater memory flexibility (Brito & Barr, 2012). In a replication study, Brito, Sebastián-Gallés, and Barr (2015) showed that the rhythmic class of the two languages (English-Spanish vs. Spanish-Catalan) did not change the pattern of results. All bilingual groups generalized, whereas monolingual English- and monolingual Spanish-speaking 18-month-olds did not.

When perceptual features differ, additional retrieval cues may reduce cognitive load associated with memory generalization and enhance performance (e.g. Brito & Barr, 2014). Labels are often provided to young children as an additional retrieval cue to support learning. Labels are symbolic, age-dependent cues. They draw attention to perceptual similarities, such as shape (Golinkoff, Mervis, & Hirsh-Pasek, 1994; Graham, Kilbreath, & Welder, 2004; Henderson & Graham, 2005; Samuelson & Smith, 2005), which promotes more generalizable representations (Barr & Brito, 2014; Hayne, 2004, 2006). Studies with monolinguals have demonstrated that labels can facilitate memory generalization performance for 15- to 24-month-olds when narrative is added to the demonstration and provided as a retrieval cue at test (Hayne & Herbert, 2004; Herbert, 2011; Simcock, Garrity, & Barr, 2011). Novel labels facilitated two-year-old monolinguals' generalization performance (Barr & Wyss, 2008; Herbert & Hayne, 2000), but not 15- (Zack, Gerhardstein, Meltzoff, & Barr, 2013) or 18-month-old monolinguals' generalization performance (Herbert & Hayne, 2000).

For labels to be an effective retrieval cue, children need to conjointly map novel perceptual cues to a linguistic cue, which may be cognitively challenging for younger children. If the task is already perceptually complex, an additional cue to encode may increase cognitive load. Herbert and Hayne (2000) tested the effect of adding novel labels on 18- and 24-month-old monolingual infants' cued memory recall and memory generalization. They tested 18-month-old infants on two perceptually different but functionally equivalent stimulus sets either immediately or after a 24-hr delay. When tested immediately, 18-month-olds performed above baseline regardless ofwhethertheyweretested with the same or a different stimulus set, demonstrating evidence of both cued recall and memory generalization, respectively. After a 24-hr delay, however, 18-month-olds continued to show evidence of cued recall but not memory generalization. In a follow-up experiment, the authors added novel labels ("meewa" and "thornby") as an additional retrieval cue during the demonstration and again at the time of test. When tested after a 24-hr delay with the same stimulus set, infants at both ages showed cued recall and the addition of novel labels did not influence their performance on cued recall. In the memory generalization condition, however, when tested with a perceptually different stimulus set, neither 18- nor 24-month-old monolinguals generalized in the no label condition, demonstrating that both the memory delay and perceptual changes in the stimuli made this a complex generalization task. However, the novel labels facilitated generalization performance by the 24-month-olds but not the 18-month-olds. Herbert and Hayne (2000) concluded that the label bound together many of the features of the stimulus set, facilitating memory generalization by the 24-month-olds but not the 18-month-olds. Taking a more linguistic interpretation, the 24-month-old word learner assumed that the symbolic label ("meewa") was a word that could be applied not only to an individual exemplar but also to the other perceptually dissimilar, but functionally similar, items and utilized this linguistic cue to generalize across stimuli. The 18-month-olds were not able to do so. Replicating Herbert and Hayne's protocol, Brito, Grenell, and Barr (2014) found that at 24 months, bilinguals were able to generalize between perceptually different stimulus sets without the addition of novel labels but the monolinguals were not (Brito et al., 2014). That is, 24-month-old bilinguals did not need the additional label cue to generalize across stimulus sets. These findings demonstrated that bilinguals showed an advanced memory flexibility trajectory.

Just as there are constraints on memory flexibility during the second year of life, there are also constraints on label learning. For example, 12-month-olds will map non-verbal mouth noises (e.g., "psst") to objects using an interactive preferential looking paradigm (Hollich, Golinkoff, & Hirsh-Pasek, 2007). Between 12 and 24 months, infants begin to limit what they will accept as a symbol-object pairing; by 2 years, they will only accept words and no longer map non-symbolic sounds or gestures to objects (e.g., Graf Estes, Antovich, & Hay, 2018; Graham & Kilbreath, 2007; Namy & Waxman, 1998). For example, Graf Estes et al. (2018) found that 18- to 19-month-old infants became more constrained in label learning. Unlike 14-month-olds, they did not associate non-word labels with objects and only applied labels when they met typical language-specific properties of the language(s) the infants were learning. Further, label learning was maximized for 18-to 19-month-olds when it occurred in a social interaction that included joint attention (Graf Estes et al., 2018). Finally, LaTourette and Waxman (2019) demonstrated that providing even a few labels enhanced categorization by 2-year-olds.

The ability to make label-object associations may be easier for bilingual infants. Singh (2018) found that bilingual infants were better able to learn a novel word in a non-native third language perhaps because of their enhanced phonological flexibility (Singh et al., 2017). In Singh's (2018) study, 18- to 20-month-old English-Mandarin bilingual infants were more likely to learn a click consonant from the Southern African language, Ndebele, than monolingual infants were. They did not, however, learn to associate non-words, such as a handclap, to objects. This pattern of results is similar to constraints observed in monolinguals in this age range (Graf Estes et al., 2018). The transitions in associating labels and objects in the latter half of the second year of life is not surprising given rapid increases in both vocabulary production and comprehension at this age. Tsui, Byers-Heinlein, and Fennell (2019) conducted a meta-analysis on word-object associative learning in 12- to 20-month-olds using the switching task. In the switching task, infants are repeatedly shown two word-object pairs until they habituate and then are tested with the word-object combinations switched. Longer looking time at test indicated that they had associated the word and the object and detected the change at the time of the switch. Specifically, the authors reported that bilingual infants who were exposed to two languages in the home performed better than their monolingual counterparts on the switching task (Tsui et al., 2019).

Studies have not yet, however, examined how labeling influences memory flexibility in bilingual infants. There are multiple competing constraints on learning that may operate differently for monolinguals and bilinguals. Differences between monolinguals and bilinguals may have implications for how labels are utilized in memory tasks. Therefore, testing the effects of labels on 18-month-old memory flexibility using an imitation paradigm provides a unique opportunity to examine an intersection between perceptual, linguistic, and memory flexibility constraints on learning. In the present study, we replicated the Herbert and Hayne (2000) experimental design and tested the effects of novel labels on cued recall and memory generalization performance with 18-month-old monolingual and bilingual infants after a 24-hr delay. Although 18-month-old bilingual infants were able to generalize on the deferred imitation puppet task (Brito & Barr, 2012; Brito et al., 2015), the rattle and animal tasks used in the present study involve many more perceptual changes between stimuli, making the tasks more challenging. In fact, the test conditions for the present study are the most challenging ones created for

18-month-olds to date. Not only were the memory demands high (6 actions total), but also the long delay (24-hr) and novel labels placed additional demands on cognitive resources at the time of encoding and retrieval. Based on Herbert and Hayne's (2000) findings, we hypothesized that both language groups would be unable to generalize in the no label condition. Due to differences in word learning constraints between monolinguals and bilinguals, we predicted that bilinguals may be more receptive to novel linguistic information and therefore be more likely to learn a novel label-object association in the present study. Specifically, we hypothesized that by 18 months of age, bilinguals would be more likely than monolinguals to utilize additional retrieval cues and therefore show enhanced memory generalization in the label condition. Alternatively, novel labels may lead to cognitive overload, resulting in poor generalization performance by both monolingual and bilingual infants (Zack et al., 2013).

2 | METHOD

2.1 | Participants

Our final sample included 87 18-month-olds (75 recruited and tested for this study and 12 participants drawn from a similar pool of participants who had previously participated, details reported below). The 75 (37 female) 18-month-olds, *M* age = 18 months, 19 days (*SD* age = 20 days) were recruited in a metropolitan area. Within this sample, 41 were classified as monolingual and 34 classified as bilingual (see language exposure interview below). Additional infants were excluded from the analyses due to parental/sibling interference (*n* = 8), no exposure to English in home (*n* = 3), child refusal to interact with stimuli (*n* = 2), or experimenter/stimulus error in either labeling or presentation of the stimuli (*n* = 23). This is a higher rate of attrition than normal due to experimenter error. This occurred because one experimenter consistently ran one condition incorrectly. Because trilingual infants, defined as being exposed to three languages with L3 > 10%, perform more similarly to monolinguals on similar memory generalization tasks, we also excluded trilingual infants (*n* = 7; Brito et al., 2014; Brito et al., 2015). These infants did not differ from those included in the sample on any demographic variables (rank socioeconomic index [SEI] or education) or on raw scores from the MacArthur Communicative Development Inventory (MCDI), and attrition occurred in all conditions. Parents were primarily Caucasian (79%) or mixed race (17%), and in addition 23% of participants also identified as Latino. The families were middle- to high-income (SEI = 80.28 based on Nakao & Treas, 1992, where higher numbers reflect higher economic and occupational status, 81% families reporting), and well-educated (*M* = 17.35 years, *SD* = 10 months, 100% families reporting), with no differences between the monolingual, bilingual, or baseline groups on mean parental educational attainment, $F(2, 72) = 0.47$, $p = 0.63$, or mean rank SEI, $F(2, 58) = 0.27$, $p = 0.77$ (see Table 1). Infants were randomly assigned to the label (*n* = 32), no label (*n* = 33), or baseline

TABLE 1 Means (standard deviations) for demographic variables

	Infant Age in Months	Parental Education in Years	Rank SEI
Monolingual	18 months 15 days (18 days)	16.78 (0.83)	80.78 (9.56)
Bilingual	18 months 21 days (20 days)	16.60 (1.14)	85.11 (7.54)

control (*n* = 10) condition. There were 16 monolinguals and 16 bilinguals in the label condition, 17 monolinguals and 16 bilinguals in the no label condition, and 8 monolinguals and 2 bilinguals in the baseline condition. Because the baseline condition with these stimuli has been replicated on multiple prior occasions (see Barr & Brito, 2014 for review), we included 12 (9 boys) additional 18-month-old baseline participants drawn from a similar pool of participants who had previously participated (Barr, Muentener, Garcia, Fujimoto, & Chavez, 2007). Their mean age was 18 months and 20 days (*SD* = 10.8 days). They were Caucasian (*n* = 6), Asian (*n* = 1), African American (*n* = 1), mixed race (*n* = 1), and Latino (*n* = 3). The families were also middle- to high-income (SEI = 79.18, *SD* = 17.50 based on Nakao & Treas, 1992, with 91.6% families reporting), and well-educated (*M* = 15.5 years, *SD* = 18 months, 100% families reporting).

2.1.1 | Language exposure interview and demographics questionnaire

The caregiver was asked to complete a general information questionnaire (assessing parental occupation, parental education, and language). A child's language exposure was measured by an adapted version of the Language Exposure Questionnaire (Bosch & Sebastián-Gallés, 2001), used by Brito et al. (2014), Brito et al. (2015), and Brito and Barr (2012, 2014), to obtain specific estimates of the child's exposure to each language from all possible language partners. Parents were interviewed to find out who spoke to the child (i.e., mother, father, sibling, grandparent, nanny, day care, etc.) in which language, for how many hours per day, for each day of the week. The percentage of time exposed to each language was calculated for each infant from the interview.

Bilingual infants were defined as those who had been exposed to two languages on a daily basis from birth and who had exposure to a second language 20% or greater of the time that they were awake and hearing any language (see also Brito & Barr, 2012; Tsang, Atagi, & Johnson, 2018). See Table 2 for a description of languages and the range of percent exposure for each group. Past studies examining the influence of multilingualism on memory generalization have found that bilingual advantages are not dependent on exposure to specific language pairs (Brito & Barr, 2012, 2014; Brito et al., 2015), so language exposure type was not controlled.

TABLE 2 Description of language pairs (L1-L2) in monolingual and bilingual groups

Monolingual		Bilingual	
English only English-Spanish English-Arabic English French-Spanish English-German English Korean-Spanish English-Portuguese		English-Spanish English-Chinese English-Portuguese English-German English-Korean English-Polish English-French English-Ga English-Arabic	Spanish-English Chinese-English Portuguese-English German-English Korean-English Swedish-English
L1 Avg. Percent	98.53% (range = 83.05–100)	63.06% (range = 50.24–77.78)	
L2 Avg. Percent	1.47% (range = 0–16.95)	36.94% (range = 22.22–49.77)	

2.2 | Apparatus

The stimuli used in the present experiment were identical to those used by Herbert and Hayne (2000) and Brito et al. (2014). The stimuli were constructed specifically for the present study and were not commercially available. There were two types of stimuli (a rattle and an animal) and two versions of each type. The stimuli were constructed in such a way that exactly the same target actions could be performed with each version (see Table 3).

2.2.1 | Rattle stimulus sets

The stimuli for the green rattle consisted of a green stick (12.5 cm long) attached to a white plastic lid (9.5 cm in diameter) with Velcro attached to the underside of the lid, a green cylindrical bead (3 cm in diameter x 2.5 cm in height), and a clear plastic square cup with Velcro around the top (5.5 cm in diameter x 8 cm in height). The opening of the plastic cup (3.5 cm in diameter) was covered with a 1 mm black rubber diaphragm, with 16 cuts radiating from the center. The stimuli for the red rattle consisted of a D-shaped handle (gap between stick and handle = 1.5 × 8 cm) attached to a red wooden stick (12.5 cm long) with a plug on the end, which fit into a blue plastic ball with a hole cut in the top (4 cm in diameter), and a red cylindrical wooden bead (2 cm in diameter).

2.2.2 | Animal stimulus sets

The stimuli for the rabbit toy consisted of two plastic eyes (3 cm x 2 cm) attached to a 9 cm x 6 cm piece of plywood with Velcro on the back, a 12 cm orange wooden carrot with green string attached to the top, a white circle of wood (the head, 15 cm in diameter) mounted horizontally on a white rectangular wooden base (30 cm x 20 cm). A 3 cm (in diameter) hole was drilled at the bottom of the head and a 5 cm x 15 cm piece of white Velcro was attached to the top of the head. Two white ears (20 cm x 5 cm) decorated with stripes of pink felt were hidden behind the head. A 10 cm wooden stick attached to the top of the right ear allowed the ears to be pulled up from behind the head in a circular motion to a point above the head. The stimuli for the monkey toy consisted of two plastic eyes (2.5 cm in diameter) that were attached to a piece of brown plywood in the shape of two diamonds joined at the center (11.5 cm in width, 6.5 cm in height), with brown Velcro on the back, a 20.5 cm yellow plastic banana, and a brown wooden head and shoulders shape mounted horizontally on a brown rectangular wooden base (22 cm x 38 cm). A 4 cm hole was drilled at the bottom of the head and a 5 cm x 18 cm piece of white Velcro was attached to the top of the head. Two brown ears (3.5 cm x 7 cm) decorated with a piece of yellow felt were hidden behind the head. A 3 cm lever with a wooden button (3.5 cm in diameter) on the top, attached to the right ear, allowed for the ears to be pulled up from behind the head in a circular motion to the side of the head.

2.2.3 | MacArthur Communicative Development Inventory

Words and Sentences Short Form (MCDI) was used to measure children's productive vocabulary (Fenson et al., 2000). From a list of 100 words, parents filled which of these words their child produced at the time of the study. Due to the wide variety of languages, language specific vocabulary measures were not feasible. For the bilingual children, the caregiver was asked to fill out the same form for all languages, marking the words the child could produce and in which language (e.g., for a Spanish-English bilingual child: English, Spanish, or both). The MCDI could then be scored for English (or other dominant language, see Table 5) vocabulary and a combined vocabulary score (see Hoff et al., 2012; Song, Tamis-LeMonda, Yoshikawa, Kahana-Kalman, & Wu, 2012).

2.3 | Procedure

All protocols were approved by the university IRB. All stimuli and deferred imitation procedures were identical to Herbert and Hayne (2000, Exp. 1A). Infants were tested in their homes on a date and time chosen by the parents. The children were seen on two consecutive days, with the demonstration of target actions for the deferred imitation tasks occurring on the first day and children's ability to recall target actions tested on the following day (24-hr delay; range 23 to 28 hr between visits). Before the demonstration, the experimenter interacted with the infants for approximately 5 min or until a smile was elicited. On the first day, the caregiver was asked to complete the general information questionnaire, language exposure information, and the vocabulary measure. All sessions were video recorded for coding purposes.

2.4 | Demonstration session

During the demonstration of the target actions, children sat on the floor with the caregiver, across from the experimenter. The experimenter performed the three target actions with one version of each stimulus type, and the entire demonstration lasted approximately 60 s. For infants in the label condition, for the rattle, infants were told, "We can use these things to make a meewa." The experimenter then placed/pushed the ball inside the cup, placed the handle on the cup, and shook the rattle. This demonstration was repeated two more times. For infants in the label condition, for the animal, infants were told, "We can use these things to make a thornby." The experimenter then used the handle to pull the ears up, attached the Velcro eyes onto the animal, and placed the carrot/banana in the rabbit's/monkey's mouth. The animal stimulus was tilted at a 45-degree angle throughout the duration of the demonstration so that the child could best see the three target actions. This demonstration was repeated two more times. The child was not allowed to touch the stimuli. The order of presentation of the stimulus sets (rattle, animal) was counterbalanced across participants.

For infants in the no label condition for both the rattle and animal, infants were told, "We can use these things to make something." The same label ("something") was provided in each condition (cued recall and generalization), so it would not allow infants to use a specific label to discriminate between conditions. That is, the label "something" does not provide infants with a novel, discriminative label for the stimulus set (see Table 4 for prompts for each condition).

2.5 | Test session

On the second day, children were first tested on the deferred imitation task. Children were tested with one set of stimuli that had been used in the original demonstration (cued recall) and one set of stimuli that was perceptually different from the one seen during demonstration (generalization) but that required the same target actions. The two stimulus sets (rattle or animal) and the order of presentation at test (cued recall or generalization) were counterbalanced across children. The rattle was always called the meewa, and the animal was always called the thornby. For example, in the label condition, when the rattle did not differ from that of the previous day, the experimenter said, "Yesterday I showed you how to make a meewa. These were the things we used to make a meewa. Can you show me how we can use these things to make a meewa?" When the rattle differed from that of the previous day, the experimenter said, "Yesterday I showed you how to make a meewa. These are some other things we can use to make a meewa. Can you show me how we can use these things to make a meewa?" The identical procedure was performed for the animal task, following the same script as listed above except using the label "thornby" instead of "meewa" for the label condition. In the no label condition, these labels (meewa and thornby) were replaced with "something" for both the rattle and animal stimulus sets. During the test, children were given the first set of stimuli and the verbal prompt, and the experimenter encouraged the child to interact with the stimuli for 60 s from the time the child first touched the stimuli. Children were then given the second set of stimuli and the verbal prompt and then given another 60 s to interact with the stimuli. The animal was tilted at a 45-degree angle throughout the duration of the test so that the 18-month-olds could motorically perform the target

TABLE 3 Stimulus sets used and target actions demonstrated

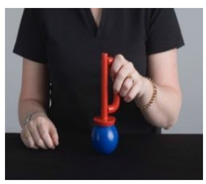
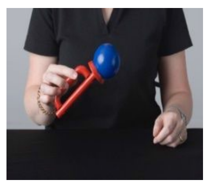
Stimulus Set	Target Action 1		Target Action 2		Target Action 3	
Monkey or Rabbit 		Pull lever in circular motion to raise ears		Attach eyes to face		Put carrot in the rabbit's mouth
Green or Red Rattle 		Drop ball into cup		Attach stick to jar		Shake stick

TABLE 4 Verbal prompts for each test type and label condition

	Day 1 – Demonstration		Day 2 – Test	
	Cued Recall	Generalization	Cued Recall	Generalization
No label	We can use these things to make *something!*	We can use these things to make *something!*	Yesterday I showed you how to make *something*. These were the things we used to make *something*. Can you show me how we can use these things to make *something?*	Yesterday I showed you how to make *something.* These are some other things we can use to make *something*. Can you show me how we can use these things to make *something?*
Label	We can use these things to make a *thornby/meewa!*	We can use these things to make a *thornby/meewa!*	Yesterday I showed you how to make a *thornby/meewa*. These were the things we used to make a *thornby/meewa*. Can you show me how we can use these things to make a *thornby/meewa?*	Yesterday I showed you how to make a *thornby/meewa*. These are some other things we can use to make a *thornby/meewa*. Can you show me how we can use these things to make a *thornby/meewa?*
Baseline	N/A		Can you show me how to make something?	

actions. The test procedure was identical for the experimental and baseline control groups; however, children in the baseline control group were not shown the demonstration of the target actions on the first day. Rather, the baseline group was only seen for one session and simply shown each stimulus type, one at a time, at test to assess the spontaneous production of the target actions. Because there was no demonstration, we slightly modified the prompt. Children in the baseline group were asked: "Can you show me how to make something?"

After both the cued recall and generalization tests were completed, the experimenter performed a manipulation check with each stimulus set (rattle, animal) to ensure that infants had the motoric ability to complete the target actions. During the manipulation check, infants first watched the experimenter demonstrate the three target actions one time. Infants were then given the opportunity to reproduce the target actions. The manipulation check was conducted to assess whether infants who did not complete the target actions were fatigued or ill on the day of the test. All infants participated in the manipulation check and were retained in the sample.

3 | RESULTS

3.1 | Coding

3.1.1 | Deferred imitation

For both cued recall and generalization, one coder scored each videotaped test session for the presence of the three target actions during the 60*s* test period for each stimulus type. The number of individual target actions produced during the 60*s* after the child first touched one of the stimuli (e.g., the stick of the rattle) was summed to calculate the imitation score (range = 0–3) for each stimulus set (rattle, animal). Each child had an imitation score for the stimulus set that was identical to the demonstration session (cued recall) or perceptually different from the demonstration session (generalization). A baseline estimate was calculated by averaging the scores of the baseline control group. A second independent coder scored 27% of the videos to determine reliability of the ratings; interrater reliability was above the acceptable level of 0.70 (Landis & Koch, 1977), with a kappa of 0.92.

3.1.2 | Language coding MCDI

As recommended by studies measuring vocabulary scores using the MCDI with bilingual populations (Hoff et al., 2012; Song et al., 2012), the raw MCDI scores were analyzed instead of the percentile scores. The raw scores for each individual language were tallied, and for the bilingual children, the raw scores were combined.

TABLE 5 Means (standard deviations) for MCDI Vocabulary Raw Scores

	Dominant language	Combined
Monolingual	25.05 (19.53)	N/A
Bilingual	N/A	17.71 (12.30)

3.1.3 | Preliminary analyses

A preliminary analysis examining associations between parental education, family rank SEI, child gender, or stimuli order and imitation performance yielded no main effects or interactions on either cued recall or memory generalization. Therefore, the data were collapsed across these variables in the following analyses.

3.1.4 | MCDI

We examined differences in productive vocabulary scores by language groups. Controlling for gender, there was no significant difference between groups (*baseline*, *monolingual label*, *monolingual no label*, *bilingual label*, and *bilingual no label*) on raw MCDI scores, $F(4, 67) = 1.54$, $p = 0.20$. Although the use of one vocabulary inventory standardizes the measurement of productive vocabulary across languages, it is worth noting that language specific inventories vary by the acquisition of common words in that specific language and only using the English form may underestimate the productive language skills of the multilingual children and monolingual non-English speaking children. In this case, the reported levels of productive vocabulary were relatively low for both groups (see Table 5).

3.2 | Data analysis plan

Deferred imitation is operationally defined as performance by the experimental group that significantly exceeds baseline. First, we examined whether experimental groups performed significantly above baseline on cued recall and generalization. In the deferred imitation paradigm, imitation can only be inferred if a group's performance is greater than the performance of infants in the age-matched control group (Barr & Hayne, 1999, 2000; Herbert & Hayne, 2000; Meltzoff, 1985). To assess the test conditions under which infants exhibited imitation, the data were subjected to two separate one-way Analyses of Variance (ANOVAs), one for the cued recall test (e.g., label cued recall monolingual, label cued recall bilingual, no label cued recall monolingual, no label cued recall bilingual, baseline) and one for the generalization test. As null effects were also hypothesized, Bayesian analyses were also conducted.

3.3 | Experimental groups Versus baseline

First, we conducted a one-way (group: *baseline*, *monolingual label*, *monolingual no label*, *bilingual label*, *bilingual no label*) Analysis of Covariance (ANCOVA) for cued recall scores, controlling for raw

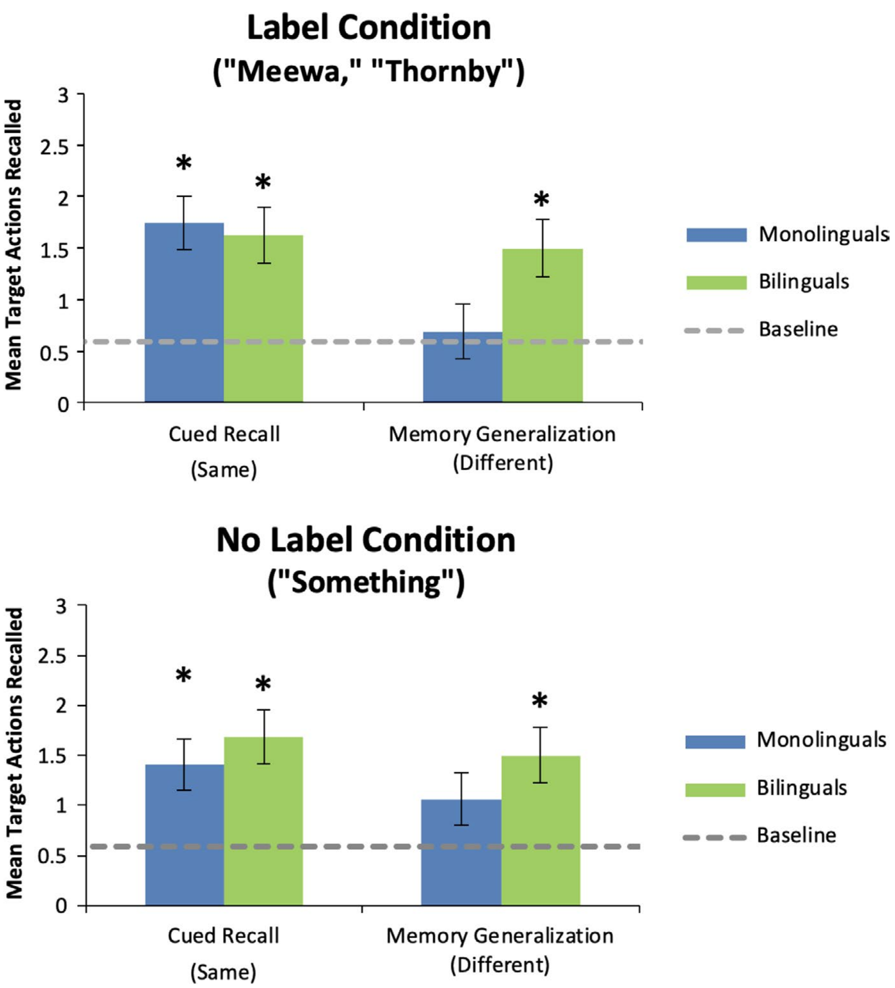


FIGURE 1 Top panel. Label condition. The mean number of target actions (± 1 SE) recalled as a function of language status on cued recall and memory generalization test conditions. An asterisk indicates group performance significantly above the baseline control. Bottom panel. No label condition. The mean number of target actions (± 1 SE) recalled as a function of language status on cued recall and memory generalization test conditions

MCDI scores. This revealed a main effect of group, $F(4, 67) = 4.16$, $p = 0.01$, $\eta^2 = 0.20$. As raw MCDI scores did not account for the effect of group and did not allow us to include the baseline participants, we removed this variable from the analysis and reran an ANOVA so that we could conduct post-hoc analyses to examine which experimental groups differed from baseline. A one-way ANOVA conducted for cued recall scores indicated a main effect of group, $F(4, 82) = 5.10$, $p = 0.001$, $\eta^2 = 0.20$ indicating a medium effect size. Dunnett's post-hoc tests indicated that *monolingual label* ($p = 0.001$), *monolingual no label* ($p = 0.032$), *bilingual label* ($p = 0.005$), and *bilingual no label* ($p = 0.003$) scores were all significantly different from *baseline*, showing that all four groups were able to recall the target actions after a 24-hr delay when the perceptual features of the stimuli did not change between demonstration and test (see Figure 1).

Again, we conducted a one-way (group: *baseline*, *monolingual label*, *monolingual no label*, *bilingual label*, *bilingual no label*) ANCOVA for memory generalization scores, controlling for raw MCDI scores. This revealed a main effect of group, $F(4, 67) = 3.88$, $p = 0.01$, $\eta^2 = 0.19$. As raw MCDI scores did not account for the effect of group, we removed it from the analysis and reran an ANOVA so that we could conduct post-hoc analyses to examine which experimental groups differed from baseline. A one-way ANOVA conducted for memory generalization scores indicated a main effect of group, $F(4, 82) = 5.54$, $p = 0.001$, $\eta^2 = 0.21$ indicating a medium effect size. Dunnett's post-hoc tests indicated that the mean differences between *baseline* and *bilingual label* ($p = 0.002$), and *bilingual no label* ($p = 0.002$), scores were significantly different. However, the mean differences between *baseline* and *monolingual label* ($p = 0.91$), and *monolingual no label* ($p = 0.13$), scores were not significantly different (see Figure 1). These results indicate that only the bilingual children in both the label and no label condition were able to recall the target actions after a 24-hr delay when the stimuli were not perceptually identical, but were functionally equivalent, from demonstration to test.

Next, we conducted confirmatory analyses to examine whether or not there were differences between monolinguals and bilinguals on the cued recall and generalization tasks. These analyses did not include baseline participants. A Bayes factor less than 1 is consistent with a significant $p < 0.05$ but the larger the Bayes factor, the more consistent it is with the probability that a null finding is correct (Jarosz & Wiley, 2014). Using the Rouder method, a Bayes independent *t*-test indicated that there was moderate support for no difference between the monolinguals and bilinguals on the cued recall task, $t(63) < 1$, $p = 0.75$, Bayes factor = 5.08, but a difference between groups for the generalization task, $t(63) = 2.77$, $p = 0.01$, Bayes factor = 0.188. This is consistent with the post-hoc *t*-tests showing a difference between monolinguals and bilinguals on the generalization test. Finally, we conducted two paired Bayes *t*-tests to confirm that this pattern held within subjects as well. For monolinguals, there was a difference between the cued recall and generalization tasks, $t(32) = 2.55$, $p = 0.02$, Bayes factor = 0.41, but for the bilinguals there was not, $t(31) < 1$, $p = 0.55$, Bayes factor = 6.11, which is moderate support for the null.

In sum, bilingual children in all groups—across test type (cued recall, generalization) and label condition (label, no label)—outperformed the baseline group; however, only monolingual children (across both no label and label groups) who were tested with the same test stimuli in the cued recall test significantly outperformed the baseline group. Follow-up Bayesian *t*-tests confirmed that the bilingual children outperformed the monolinguals on the generalization task but not the cued recall task.

4 | DISCUSSION

Bilinguals performed significantly above baseline on both cued recall and memory generalization, whereas monolinguals performed significantly above baseline only on cued recall, regardless of label condition. Our results were not in line with our predictions, and we were surprised that bilinguals were able to succeed on this cognitively challenging generalization task, both with and without the addition of the label. We had hypothesized that bilinguals might be better able to utilize the novel label to bridge the gap between the perceptual differences in stimuli in order to enhance generalization, but the label neither enhanced nor interfered with generalization performance. Because performance was enhanced in both the label and no label conditions in the bilinguals compared to the monolinguals, these findings both conceptually replicate and extend results from past studies (Brito & Barr, 2012, 2014; Brito et al., 2014, 2015). Few studies examine how action learning is influenced by labels in bilingual infants. This study provides a novel contribution showing that under conditions where monolinguals fail on the memory generalization test, bilinguals perform above baseline, and suggests that bilingual differences in generalization may arise from visual perceptual, rather than linguistic differences.

The findings were surprising because the memory generalization test conditions in this study were particularly challenging. The finding that 18-month-old monolinguals fail under these challenging conditions (Herbert & Hayne, 2000) was replicated in the present study. Yet, 18-month-old bilinguals in this study showed the same pattern of results as 24-month-old bilinguals in previous research who showed robust memory generalization on this imitation task in the no label condition as well (Brito et al., 2014). In this study, the memory demands were high (6 actions), the delay was long (24 hr), and novel labels may have placed an additional demand on cognitive resources at the time of encoding and retrieval for both monolingual and bilingual infants (see also Zack et al., 2013 for a similar argument). The developmental trajectory of coordination of labels and actions by bilingual infants is currently unknown.

The present study did not directly test word learning but not surprisingly, bilingual exposure also changes the trajectory of linguistic development. Word learning is constrained by a number of assumptions (Markman, 1991). For example, the mutual exclusivity

(ME) assumption means that the word learner assumes that only one word applies to the individual object and class of objects. Once the word learner has a word for dog, the word learner does not apply a second label to the object dog. It is thought that ME also allows children to learn about other object properties. Because ME constrains learning of one word to one object, a second word would not be learned as a second object label, but it may describe another feature of the object. The ME assumption poses a problem for bilingual infants, however, because the child is frequently provided with more than one label for an individual object. Monolingual infants adopt the ME principle (mapping one word to one object) beginning at 17 to 22 months, but bilingual infants have not been shown to adopt the ME principle in this age range (Byers-Heinlein & Werker, 2009; Houston-Price, Caloghiris, & Raviglione, 2010) and have demonstrated different label use strategies based on prior word learning experience (Kandhadai, Hall, & Werker, 2017). Due to the differences in the utilization of the ME assumption by bilinguals, bilinguals may need to rely more on the visual perceptual system to both group together and disambiguate objects than monolinguals do. In the present study, this may have been indexed by better generalization performance.

How might these linguistic constraints intersect with memory flexibility? The present study suggests that bilingual infants may be better able to integrate multiple perceptual cues during early infancy and thus may develop hierarchical mental representations earlier than monolinguals. Capitalizing on such memory processing, a label may help bilingual infants form a hierarchical memory representation and generalize across dissimilar stimuli. Hierarchical learning is associated with the frontostriatal pathway and some research has examined this in infants. For example, activation of the right dorsal lateral prefrontal cortex is associated with better memory generalization performance in monolingual infants (Werchan, Collins, Frank, & Amso, 2016). Further, infants' eye blink rate, a possible physiological index of striatal dopamine activity, was positively associated with generalization performance. The addition of neuroimaging techniques in memory generalization paradigms with monolingual and bilingual infants would further elucidate how language exposure may be shaping neural development across multiple domains. Specifically, an experiment could test whether there are differences between monolinguals and bilinguals in the formation of hierarchical representations and whether these differences are associated with differences in neural activation patterns in the frontostriatal pathway.

The present findings lead to more questions than answers regarding potential mechanisms to explain differences between monolinguals and bilinguals in memory flexibility across the infancy period. Differences between monolinguals and bilinguals in memory flexibility may also intersect with linguistic constraints on learning (e.g., Kandhadai et al., 2017) and provide promising avenues for future research. For example, in addition to the ME assumption, key word learning assumptions, such as whole object and taxonomic assumptions (Markman, 1991) require further investigation to examine how they interact with memory flexibility. The taxonomic assumption means that the word learner assumes that the word applies not only to an individual exemplar, but also to the whole taxonomy (e.g., the word "dog" will apply not only to one dog, but also to the class of dogs). Differences in learning constraints may result in different visual-auditory mapping of distinct labels to objects and differences in utilization of linguistic principles may tune the visual perceptual system as well.

Future studies should adopt Singh et al. (2017) approach of testing two paradigms within the same study (e.g., perceptual face processing and phonetic discriminations in their study) to examine the interaction between different learning constraints in monolinguals and bilinguals. Such studies should include measures of both word learning and memory generalization. Future research should also examine how processing of familiar versus novel labels impacts generalization performance for both monolinguals and bilinguals. Retention of novel word-object labels is notoriously short-lived when they are presented in the context of multiple new word-object pairings (Kucker, McMurray, & Samuelson, 2018). Had labels been familiar (e.g., "rattle" or "animal"), rather than our nonsense labels, performance may have been enhanced. In addition, these studies could manipulate exposure to the novel labels to test the strength of the label-object association on subsequent word learning and memory generalization.

These changes in memory flexibility may be precursors to later well-documented differences in cognitive flexibility between monolinguals and bilinguals. Cognitive flexibility is defined as the ability to adjust to changes in task demands and to switch between different rules and goals (Mahy & Munakata, 2015). Switching requires the ability to selectively attend to, as well as to integrate and adapt to, multiple cues in the environment (Deák & Wiseheart, 2015). Bilingual language acquisition is associated with accelerated cognitive flexibility emerging in preschool (Adi-Japha, Berberich-Artzi, & Libnawi, 2010; Bialystok & Martin, 2004; Bialystok & Senman, 2004; Carlson & Meltzoff, 2008) and continuing throughout the lifespan (Bialystok, Craik, & Ryan, 2006; Costa, Hernández, & Sebastián-Gallés, 2008). It is our contention that memory flexibility as demonstrated across multiple studies, along with enhanced perceptual processing capacity, are key candidate precursors to emerging advantages in cognitive flexibility. Longitudinal studies examining the trajectory of memory flexibility and cognitive flexibility are needed to trace this process developmentally and to assess how language and memory processes intersect.

5 | DATA SHARING STATEMENT

The data that support the findings of this study will be uploaded to databrary.org and are available from the corresponding author (rfb5@georgetown.edu) to be shared upon reasonable request.

ACKNOWLEDGEMENTS

We are grateful to all the families who participated in this research, to Amanda Grenell and Laura Zimmermann for their invaluable help in collecting data, and to Julia Tonnessen for her invaluable help in coding data and who was supported by the Georgetown Undergraduate Research Opportunity Program. This research was funded by the NSF BCS-1551719. *From memory flexibility to cognitive flexibility: Examining precursors to bilingual advantages during early childhood* to Barr.

ORCID

Rachel Barr https://orcid.org/0000-0002-5855-9718

Natalie H. Brito https://orcid.org/0000-0003-1016-0094

REFERENCES

Adi-Japha, E., Berberich-Artzi, J., & Libnawi, A. (2010). Cognitive flexibility in drawings of bilingual children. *Child Development*, *81*(5), 1356–1366. https://doi.org/10.1111/j.1467-8624.2010.01477.x

Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*, 699–714. https://doi.org/10.1016/j.ecresq.2014.02.003

Barr, R., & Brito, N. (2014). From specificity to flexibility: Early developmental changes in memory generalization. In P. J. Bauer & R. Fivush (Eds.), *The Wiley Handbook on the Development of Children's Memory* (pp. 453–479). Chichester, UK: Wiley Blackwell.

Barr, R., & Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Development*, *70*, 1067–1081. https://doi.org/10.1111/1467-8624.00079

Barr, R., & Hayne, H. (2000). Age-related changes in imitation: Implications for memory development. In C. Rovee-Collier, L. P. Lipsitt & H. Hayne (Eds.), *Progress in Infancy Research*, Vol. *1* (pp. 21–67). Mahwah, NJ: Erlbaum.

Barr, R., Muentener, P., Garcia, A., Fujimoto, M., & Chavez, V. (2007). The effect of repetition on imitation from television during infancy. *Developmental Psychobiology*, *49*, 196–207. https://doi.org/10.1002/dev.20208

Barr, R., & Wyss, N. (2008). Reenactment of televised content by 2-year olds: Toddlers use language learned from television to solve a difficult imitation problem. *Infant Behavior and Development*, *31*, 696–703. https://doi.org/10.1016/j.infbeh.2008.04.006

Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1341.

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, *7*, 325–339. https://doi.org/10.1111/j.1467-7687.2004.00351.x

Bialystok, E., & Senman, L. (2004). Executive processes in appearance–reality tasks: The role of inhibition of attention and symbolic representation. *Child Development*, *75*, 562–579. https://doi.org/10.1111/j.1467-8624.2004.00693.x

Bosch, L., & Sebastián-Gallés, N. (2001). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy*, *2*, 29–49. https://doi.org/10.1207/S15327078IN0201_3

Brito, N., & Barr, R. (2012). Influence of bilingualism on memory generalization during infancy. *Developmental Science*, *15*, 812–816. https://doi.org/10.1111/j.1467-7687.2012.1184.x

Brito, N., & Barr, R. (2014). Flexible memory retrieval in bilingual 6-month-old infants. *Developmental Psychobiology*, *56*, 1156–1163. 12/2013; https://doi.org/10.1002/dev.21188

Brito, N. H., Grenell, A., & Barr, R. (2014). Specificity of the bilingual advantage for memory: Examining cued recall, generalization, and working memory in monolingual, bilingual, and trilingual toddlers. *Frontiers in Psychology*, *5*, 1369. https://doi.org/10.3389/fpsyg.2014.01369

Brito, N., Sebastián-Gallés, N., & Barr, R. (2015). Differences in language exposure and its effects on memory generalization in monolingual, bilingual, and trilingual infants. *Bilingualism Language and Cognition*, *18*, 670–682. https://doi.org/10.1017/S1366728914000789

Byers-Heinlein, K. (2014). Languages as categories: Reframing the "One Language or Two" question in early bilingual development. *Language Learning*, *64*(s2), 184–201.

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. https://doi.org/10.1111/j.1467-7687.2009.00902.x

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*, 282–298. https://doi.org/10.1111/j.1467-7687.2008.00675.x

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, *106*, 59–86. https://doi.org/10.1016/j.cognition.2006.12.013

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews: Neuroscience*, *15*, 336–345. https://doi.org/10.1038/nrn3709

Deák, G. O., & Wiseheart, M. (2015). Cognitive flexibility in young children: General or task-specific capacity? *Journal of Experimental Child Psychology*, *138*, 31–53. https://doi.org/10.1016/j.jecp.2015.04.003

D'Souza, D., & Karmiloff-Smith, A. (2011). When modularization fails to occur: A developmental perspective. *Cognitive Neuropsychology*, *28*(3–4), 276–287. https://doi.org/10.1080/02643294.2011.614939

Eichenbaum, H. (1997). Declarative memory: Insights from cognitive neurobiology. *Annual Review of Psychology*, *48*, 547–572. https://doi.org/10.1146/annurev.psych.48.1.547

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. https://doi.org/10.1017/S0142716400001053

Gerhardstein, P., Liu, J., & Rovee-Collier, C. (1998). Perceptual constraints on infant memory retrieval. *Journal of Experimental Child Psychology*, *69*, 109–131. https://doi.org/10.1006/jecp.1998.2435

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, *21*, 125–155. https://doi.org/10.1017/S0305000900008692

Graf Estes, K., Antovich, D. M., & Hay, J. F. (2018). Intersecting constraints on label learning: Effects of age, label properties, and referential context. *Journal of Cognition and Development*, *19*, 532–551. https://doi.org/10.1080/15248372.2018.1523172

Graf Estes, K., & Hay, J. F. (2015). Flexibility in Bilingual Infants' Word Learning. *Child Development*, *86*, 1371–1385. Article first published online: 7 JUL 2015. https://doi.org/10.1111/cdev.12392

Graham, S. A., & Kilbreath, C. S. (2007). It's a sign of the kind: Gestures and words guide infants' inductive inferences. *Developmental Psychology*, *43*, 1111–1123. https://doi.org/10.1037/0012-1649.43.5.1111

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development*, *75*, 409–427. https://doi.org/10.1111/j.1467-8624.2004.00683.x

Hayne, H. (2004). Infant memory development: Implications for childhood amnesia. *Developmental Review*, *24*, 33–73. https://doi.org/10.1016/j.dr.2003.09.007

Hayne, H. (2006). Age-related changes in infant memory retrieval: Implications for knowledge acquisition. Processes of brain and cognitive development. *Attention and Performance XXI*, *209*–231.

Hayne, H., & Herbert, J. S. (2004). Verbal cues facilitate memory retrieval during infancy. *Journal of Experimental Child Psychology*, *89*, 127–139. https://doi.org/10.1016/j.jecp.2004.06.002

Henderson, A. M. E., & Graham, S. A. (2005). Two-year-olds' appreciation of the shared nature of novel object labels. *Journal of Cognition and Development*, *6*, 381–402. https://doi.org/10.1207/s15327647jcd0603_4

Herbert, J. S. (2011). The effect of language cues on infants' representational flexibility in a deferred imitation task. *Infant Behavior and Development*, *34*, 632–635. https://doi.org/10.1016/j.infbeh.2011.06.007

Herbert, J., & Hayne, H. (2000). Memory retrieval by 18–30-month-olds: Age-related changes in representational flexibility. *Developmental Psychology*, *36*, 473–484. https://doi.org/10.1037/0012-1649.36.4.473

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*, 1–27. https://doi.org/10.1017/S0305000910000759

Hollich, G., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology*, *43*, 1051–1061. https://doi.org/10.1037/0012-1649.43.5.1051

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, *15*, 125–150. https://doi.org/10.1111/j.1532-7078.2009.00009.x

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting bayes factors. *The Journal of Problem Solving*, *7(1)*, 2. https://doi.org/10.7771/1932-6246.1167

Kandhadai, P., Hall, D. G., & Werker, J. F. (2017). Second label learning in bilingual and monolingual infants. *Developmental Science*, *20*, n/a, e12429. https://doi.org/10.1111/desc.12429

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental approach to cognitive science*. Cambridge, MA: MIT Press.

Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556–6560. https://doi.org/10.1073/pnas.0811323106

Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science*, *325*(5940), 611–612.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2018). Too much of a Good Thing: How Novelty Biases and Vocabulary Influence known and Novel Referent Selection in 18-Month-Old Children and Associative Learning Models. *Cognitive Science*, *42* Suppl 2, 463–493. https://doi.org/10.1111/cogs.12610

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. https://doi.org/10.2307/2529310

LaTourrette, A., & Waxman, S. R. (2019). A little labeling goes a long way: Semi-supervised learning in infancy. *Developmental Science*, *22*(1), e12736. https://doi.org/10.1111/desc.12736

Mahy, C. E., & Munakata, Y. (2015). Transitions in executive function: Insights from developmental parallels between prospective memory and cognitive flexibility. *Child Development Perspectives*, *9*, 128–132. https://doi.org/10.1111/cdep.12121

Markman, E. M. (1991). The whole object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In J. P. Byrnes & S. A. Gelman (Eds.), *8* Perspectives on language and cognition: Interrelations in development (pp. 72–106). Cambridge: Cambridge University Press.

Meltzoff, A. (1985). Immediate and deferred imitation in fourteen- and twenty-four-month-old infants. *Child Development*, *56*, 62–72. https://doi.org/10.2307/1130174

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*, *69*, 295–308. https://doi.org/10.1111/j.1467-8624.1998.tb06189.x

Nakao, K., & Treas, J. (1992). *The 1989 socioeconomic index of occupations: Construction from the 1989 occupational prestige scores* General Social Survey Methodological Reports. Chicago, IL: NORC.

Newcombe, N. S. (2011). What is neoconstructivism? *Child Development Perspectives*, *5*(3), 157–160. https://doi.org/10.1111/j.1750-8606.2011.00180.x

Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W., & Lee, K. (2014). On the links among face processing, language processing, and narrowing during development. *Child Development Perspectives*, 1–6. https://doi.org/10.1111/cdep.12064

Samuelson, L. K., & Smith, L. B. (2005). They call it like they see it: Spontaneous naming and attention to shape. *Developmental Science*, *8*, 182–198. https://doi.org/10.1111/j.1467-7687.2005.00405.x

Sebastián-Gallés, N. (2010). Bilingual language acquisition: Where does the difference lie? *Human Development*, *53*(5), 245. https://doi.org/10.1159/000321282

Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, *23*(9), 994–999. https://doi.org/10.1177/0956797612436817

Simcock, G., Garrity, K., & Barr, R. (2011). The effect of narrative cues on infants' imitation from television and picture books. *Child Development*, *82*, 1607–1619. https://doi.org/10.1111/j.1467-8624.2011.01636.x

Singh, L. (2018). Bilingual infants demonstrate advantages in learning words in a third language. *Child Development*, *89*(4), e397–e413. https://doi.org/10.1111/cdev.12852

Singh, L., Fu, C. S. L., Rahman, A. A., Hameed, W. B., Sanmugam, S., Agarwal, P., … the GUSTO Research Team (2015). Back to Basics: A Bilingual Advantage in Infant Visual Habituation. *Child Development*, *86*, 294–302. https://doi.org/10.1111/cdev.12271

Singh, L., Loh, D., & Xaio, N. (2017). Bilingual Infants Demonstrate Perceptual Flexibility in Phoneme Discrimination But Perceptual Constraint in Face Discrimination. *Frontiers in Psychology*, *8*, https://doi.org/10.3389/fpsyg.2017.01563

Song, L., Tamis-LeMonda, C. S., Yoshikawa, H., Kahana-Kalman, R., & Wu, I. (2012). Language experiences and vocabulary development in Dominican and Mexican infants across the first 2 years. *Developmental Psychology*, *48*(4), 1106. https://doi.org/10.1037/a0026401

Tsang, T., Atagi, N., & Johnson, S. P. (2018). Selective attention to the mouth is associated with expressive language skills in monolingual and bilingual infants. *Journal of Experimental Child Psychology*, *169*, 93–109. https://doi.org/10.1016/j.jecp.2018.01.002

Tsui, A. S. M., Byers-Heinlein, K., & Fennell, C. T. (2019). Associative word learning in infancy: A meta-analysis of the switch task. *Developmental Psychology*, *55*, 934–950. https://doi.org/10.1037/dev0000699

Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, *316*(5828), 1159–1159.

Werchan, D. M., Collins, A. G. E., Frank, M. J., & Amso, D. (2016). Role of prefrontal cortex in learning and generalizing hierarchical rules in 8-month-old infants. *Journal of Neuroscience*, *36*, 10314–10322. doi/10.1523/JNEUROSCI.1351-16.2016.

Werker, J. F. (2012). Perceptual foundations of bilingual acquisition in infancy. *Annals of the New York Academy of Sciences*, *1251*(1), 50–61. https://doi.org/10.1111/j.1749-6632.2012.06484.x

Werker, J. F., & Byers-Heinlein, K. (2008). Bilingualism in infancy: First steps in perception and comprehension. *Trends in Cognitive Sciences*, *12*(4), 144–151. https://doi.org/10.1016/j.tics.2008.01.008

Zack, E., Gerhardstein, P., Meltzoff, A. N., & Barr, R. (2013). 15-month-olds' transfer of learning between touch screen and real-world displays: Language cues and cognitive loads. *Scandinavian Journal of Psychology*, *54*, 20–25. https://doi.org/10.1111/sjop.12001

How to cite this article: Barr R, Rusnak SN, Brito NH, Nugent C. Actions speak louder than words: Differences in memory flexibility between monolingual and bilingual 18-month-olds. *Dev Sci*. 2020;23:e12881. https://doi.org/10.1111/desc.12881