Journal of Child Language (2025), 1–24
doi:10.1017/S0305000925100317

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Flexible Use of Word Learning Strategies: Monolingual and Bilingual Children's Word Learning Under Different Language Contexts

Kin Chung Jacky Chan[1,2] and Padraic Monaghan[1]

[1]Department of Psychology, Lancaster University, Lancaster, UK and [2]Department of Psychology, Durham University, Durham, UK

Corresponding author: Kin Chung Jacky Chan; Email: kin.c.chan@durham.ac.uk

(Received 21 September 2024; revised 24 August 2025; accepted 08 September 2025)

Abstract

Monolingual children tend to assume that a word labels only one object, and this mutual exclusivity supports referent selection and retention of novel words. Bilingual children accept two labels for an object (lexical overlap) for referent selection more than monolingual children, but in these previous studies, information about speakers' language backgrounds was minimal. We investigated monolingual and bilingual 4-year-old children's ability to apply mutual exclusivity and lexical overlap flexibly when objects were labelled either by one or two speakers with the same or different language backgrounds. We tested referent selection and retention of word–object mappings. Both language groups performed similarly for mutual exclusivity, were more likely to accept lexical overlap in the two-language than one-language condition, and performance was similar for referent selection and later retention. Monolingual and bilingual children can adapt their word-learning strategies to cope with the demands of different linguistic contexts.

Keywords: bilingualism; word learning; mutual exclusivity; lexical overlap; speaker identity

1. Introduction

Word learning is an immensely complex task. One difficulty is due to the presence of infinitely many possible referents in the environment for a word (Quine, 1960). To address this difficulty, several researchers (e.g., Markman, 1994; Markman & Hutchinson, 1984; Markman & Wachtel, 1988; Mervis & Bertrand, 1994; Waxman, 1989) have argued that young children possess operating principles to guide their word learning by constraining the number of potential referents for a word (Golinkoff et al., 1994; Markman, 1990). One of these proposed constraints that children use to map words to their referents is mutual exclusivity (ME), a strategy that assigns a new word label to an unfamiliar rather than a familiar object, assuming that every object can only have one label (Kalashnikova et al., 2014; Markman & Wachtel, 1988). ME has been discovered in even very young children: studies with 10-month-olds (Mather & Plunkett, 2010) and 17-month-olds (Halberda,

2003) using looking time paradigms, and 17.5-month-olds (Mervis & Bertrand, 1994) and 2 years and older (Clark, 1990; Littschwager & Markman, 1994; Markman et al., 2003) using behavioural selection tasks.

Bilingual children, in contrast, tend to use ME differently than monolingual children. Byers-Heinlein and Werker (2009) compared the use of ME in monolingual and bilingual 17- and 18-month-olds in a looking time study and found that bilingual children relied less on ME than monolinguals. Extending this approach, Byers-Heinlein et al. (2014) investigated the use of ME when an English speaker labelled an object, then a Chinese speaker used a different label. Two-year-old monolingual and bilingual children used ME in the English condition, whereas only the monolingual children systematically used ME in the Chinese condition, suggesting that monolingual children assume that words are conventionally shared across speakers of all languages, while bilingual children are aware that speakers of one language are ignorant of words in another language. Similar to Byers-Heinlein et al.'s (2014) study of children's sensitivity to different language speakers in applying ME, Scott and Henderson (2013) found that 13-month-old monolingual children accepted an English and a French speaker using the same label for an object, whereas Henderson and Scott (2015) found that bilingual children of the same age were surprised by the overlap in labels.

These studies show that monolingual and bilingual children apply ME differently, but they have not explicitly tested how children respond to lexical overlap (LO), that is, when a referent has more than one label. Frank and Poulin-Dubois (2002) tested 27- and 35-month-old monolingual and bilingual children on their ability to accept LO. In their first experiment, they tested children's ability to accept two labels for the same object when spoken in the same language. They found similar levels of acceptance of LO by monolingual and bilingual children. In a second experiment, they tested bilingual children on LO when two labels were spoken in different languages by two experimenters, with both languages known to the children. The rationale for testing this is that bilingual children have more exposure to speakers of different languages using different labels for the same referents (Barron-Hauwaert, 2004; Lanza, 1997). Bilingual children performed similarly to the first experimental condition, where labels were spoken in the same language, though monolingual children were not tested, as familiarity with the two languages was a prerequisite of the study design.

Kandhadai et al. (2017) directly compared bilingual and monolingual children on a LO task. When a new label for a known object was provided, 17- to 18-month-old bilingual children were more likely to accept the label as another name for the object, whereas monolingual children of similar age were more likely to link the label to the colour of the object. However, older monolingual children have been shown to be able to accept two labels for the same object. Kalashnikova et al. (2016a) tested whether 4- to 5-year-old monolingual children could both apply ME and accept LO when two speakers labelled the same object with a different novel label. The monolingual children performed significantly better than chance level in both ME and LO conditions, indicating that children aged 3 to 4 years are able to accept two labels from the same language for one object (Au & Glusman, 1990; Waxman & Hatch, 1992). These results are in line with the Emergentist Coalition Model of word learning, whereby children are sensitive to multiple cues to learning the meaning of new words, and their reliance on different cues changes as they develop (Hollich et al., 2000; Pruden et al., 2006). Although children show an early reliance on lexical constraints, by 24 months, they rely more on socio-pragmatic cues over constraints (Akhtar et al., 1996; Diesendruck & Markson, 2001).

In a variation of Kandhadai et al.'s (2017) task, Kalashnikova et al. (2015) tested the extent to which children could flexibly apply ME or LO according to the communicative situation. Two puppets either both used the same label, or each used a different label, for an unfamiliar object. Children were then given four labels sequentially and asked to select from four objects: two familiar objects, one unlabelled unfamiliar object, and the labelled unfamiliar object. In the exclusivity condition, the four labels corresponded to the four objects that the children could choose from; whereas in the overlap condition, two of the labels referred to the same object – the labelled unfamiliar object, and the remaining two labels each referred to a familiar object. All children were able to apply ME and accept LO to a certain degree, but older bilingual children increased in acceptance of LO compared to younger bilingual children, and older monolingual children relied on ME in word learning more and accepted fewer LOs than their bilingual counterparts, with a smaller difference for the younger monolingual and bilingual children (see also Kalashnikova et al., 2019).

In summary, these studies indicate that monolingual and bilingual children apply ME differently, and that this difference is emphasised when labels are produced by speakers of different languages (Byers-Heinlein et al., 2014; Henderson & Scott, 2015; Scott & Henderson, 2013). However, in studies that test LO explicitly, older bilingual children appear to accept LO more readily than monolingual children (Kalashnikova et al., 2014; Kandhadai et al., 2017), but the effect of different speakers on this effect is less clear (Frank & Poulin-Dubois, 2002; Kalashnikova et al., 2015; Kalashnikova et al., 2016a), with studies not directly testing monolingual and bilingual children on LO when labels are spoken by speakers of the same or different languages. In natural language exchanges, it is unlikely for children to encounter two speakers of the same language labelling an object in two different ways, given that they both tend to name the object with its basic category label (Clark, 1987; Rosch et al., 1976). Hence, the design of Kalashnikova et al.'s (2015, 2016a) studies might have underestimated the ability of children in using socio-pragmatic cues in order to relax ME and accept LO. Consequently, we do not know the extent to which monolingual and bilingual children can use speaker identity to support LO. The first aim of our study was to provide this direct comparison.

These previous studies of ME and LO have focused on referent selection tasks – where very soon after being exposed to the labelling of an object, children are tested on their ability to distinguish between a set of objects with a similar or novel label. However, referent selection ability is not the same as word learning, and the latter can be measured by testing children's learning after a delay rather than immediately after being exposed to the label. Horst and Samuelson (2008) found that 24-month-old monolingual children show poor retention of words learned through the application of ME, even after a very short delay of a few minutes (see also Vlach & DeBrock, 2019). These studies point to the importance of distinguishing referent selection from retention for word learning (Horst & Samuelson, 2008; McMurray et al., 2012), and so investigating the extent to which LO as well as ME relate to the acquisition of word-referent mappings, as well as guiding referent selection, is a key test for understanding the role of these behaviours in language acquisition. Learning two-to-one word–object mappings (i.e., through accepting LO) presents a different and perhaps more complicated problem than learning through ME, and so there is the possibility that LO may support retention of words to a greater degree than ME. For instance, Fitneva and Christiansen (2011) found that learners acquire (one-to-one) word–object mappings better when their initial learning is difficult, and so it is possible that learning words from accepting LO may be more resilient to a delay than learning from applying ME. The second aim of our study was thus to test retention of learning as a consequence of LO.

For studies examining ME, there has been substantial variability in evidence across studies in terms of eliciting reliable ME responses in young children, particularly up to the age of 18 months (Bion et al., 2013; Halberda, 2003; Mather & Plunkett, 2010). One possible explanation for these inconsistent results is that the use of ME is related to vocabulary size (Graham et al., 1998; Mervis & Bertrand, 1994), as the use of ME requires children to have some vocabulary – knowing the name of the known object – and their experience with words may also lead them to be more proficient in using ME. Bion et al. (2013) found that 24- and 30-month-olds' ME performance was positively correlated with their vocabulary knowledge, and Kalashnikova et al. (2016b) found that 17- to 19-month-olds' vocabulary knowledge significantly predicted their use of ME in a looking time task. Vocabulary knowledge may also relate to LO, with greater skill at linking words to referents supporting application not only of ME but also of multiple labels to the same referent, and Kalashnikova et al. (2019) found there was a relation between bilingual, but not monolingual, vocabulary knowledge and LO. We thus included vocabulary level as an individual difference measure to relate to ME and LO performance in our study, in order to determine how our results related to these previous studies of children's vocabulary size and referent selection performance.

Studies of language development have conventionally adopted classic inferential statistics to test relations among measures or differences among groups. However, these approaches do not enable us to gather evidence for *similar* behaviour between groups, as it is not possible to determine whether finding no statistical difference between groups is due to the groups being similar in performance or due to noise in the sample preventing the difference from being observed. An alternative approach that is now seeing a resurgence in psychological research (e.g., Wagenmakers et al., 2018) uses Bayesian inference, which enables evidence for both similarities and differences to be appraised. Our analyses include both inferential statistical model building approaches and Bayes Factor analyses to determine evidence for similarity and differences between the monolingual and bilingual groups in their ME and LO behaviour.

1.1. The current study

To investigate whether the additional cue of linguistic background of speakers would differentially affect the use of ME and acceptance of LO in monolingual and bilingual 3- to 4-year-olds, we adapted Kalashnikova et al.'s (2016) puppet study but provided more salient information about speakers' linguistic identity, similar to the study by Frank and Poulin-Dubois (2002), by providing cues to the linguistic backgrounds of the speakers in the task. In one condition, both speakers spoke English, whereas in the other condition, the two speakers spoke English and Hungarian, respectively. The age range of the participants was selected based on the age effect (i.e., differences in the use of ME and acceptance of LO only evident between 4- to 5-year-old monolinguals and bilinguals) found in Kalashnikova et al.'s (2015) related study that used puppets instead of human speakers. Additionally, performance was measured not only in terms of referent selection ability, but also word learning after a 10-minute delay following ME and LO training conditions, and children's vocabulary knowledge was tested to relate to their use of ME and acceptance of LO.

2. Method

2.1. Participants

Twenty monolingual (M_{age} = 4.10 years, SD_{age} = 0.43, 12 females) and 20 bilingual children (M_{age} = 3.92 years, SD_{age} = 0.50, 12 females) took part in the present study between March and December, 2017. All monolingual children and four bilingual children were recruited from and tested at preschools and nurseries in the local area of Lancaster, UK. The remainder of the bilingual sample (*n* = 17) was recruited through and tested at Lancaster University Babylab, and these children were drawn from a similar general population in terms of socio-economic background and geographical location. All children in the monolingual group only spoke English, and none had experience with Hungarian. The bilingual group consisted of children who spoke English and an additional language as reported by their caregivers: Arabic (*n* = 1), Dutch (*n* = 1), French (*n* = 2), German (*n* = 5), Italian (*n* = 1), Malagasy (*n* = 1), Polish (*n* = 1), Russian (*n* = 1), Slovak (*n* = 1), and Spanish (*n* = 6)[1]. Three additional monolingual and three additional bilingual children were tested but excluded due to either testing being done in a noisy classroom, resulting in an inability to follow the experiment instructions (*n* = 3), very low English proficiency raw scores that were not convertible to age-standardised scores (*n* = 2), or experimenter error (*n* = 1).

2.2. Materials and apparatus

All children took part in two experimental conditions: English–English, where they saw video clips of two English speakers speaking and naming unfamiliar objects, and English–Hungarian, where one speaker in the video clips spoke English while the other spoke Hungarian. This was in line with Samara et al.'s (2017) experimental design that showed that children can associate distinct speech input with different speakers.

The two conditions were administered a week apart. Each condition contained two tasks: exclusivity and overlap, each of which consisted of four immediate test trials and eight delayed test trials. The order of conditions and tasks was counterbalanced (see Procedures for more details).

2.2.1. Experimental stimuli

Eight images of familiar objects were selected from the TarrLab Object Databank (1996) for use in the familiarisation trials (see Supplementary Appendix A for a list of the stimuli used). The images of familiar objects were placed in four two-object sequences for familiarisation.

Thirty-two images of unfamiliar objects and novel words were selected from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016) for the test trials. All selected novel words were phonotactically legal in both English and Hungarian, and pronunciations were aligned across English and Hungarian, such that the Hungarian version was matched to the phonology of the English version of the nonword. For the English–English condition, a male and a female English speaker were recorded. For the

[1]Our collection of participant demographic information followed that in Kalashnikova et al. (2015) and, therefore, did not include parental education, SES, and bilingual language exposure measures.

English–Hungarian condition, a male English native speaker (different than the speaker in the English–English condition) and a female Hungarian native speaker were recorded.

The images of the unfamiliar objects were ordered into 16 pairs, appearing on the screen with a video recording of one or two people. Four pairs were assigned to each task in each condition.

For the immediate test trials, two objects were shown on the screen, comprising the previously named object and an unnamed unfamiliar object.

For the delayed test trials, four objects were shown on the screen. The four objects in a delayed test trial were all either objects for which the children were presented the label ostensively in the immediate test trials or objects for which the children had been given the chance to learn the names through ME (during the immediate test trials).

All stimuli were presented on a Surface Pro 4 touchscreen using PsyScript 3 (Slavin, 2014), and children's responses were collected via touches on the screen. See Figure 1 for the experiment flow and screenshots of example trials in each task and phase.

2.2.2. Language proficiency

The British Picture Vocabulary Scale – Third Edition (BPVS III; Dunn et al., 2009) was administered to all children immediately after the second experimental condition. The monolingual ($M = 104.45$, $SD = 13.00$) and bilingual group ($M = 98.85$, $SD = 8.98$) did not differ significantly on their age standardised scores, $t(38) = 1.59$, $p = .121$, $d = 0.50$. Note that similar scores were found for monolingual and bilingual children aged between 3;6 and 4;5 years collected from a similar demographic by Kalashnikova et al. (2015). Whereas one might anticipate higher vocabulary scores for monolingual than bilingual children, this difference only emerged for slightly older children in Kalashnikova et al. (2015) – aged between 4;5 and 5;9 years. Parents of all but one bilingual child (exposure to English since 2;2 years) reported that their child had been exposed to English since birth.

2.3. Procedure

For the children tested at preschools and nurseries, information sheets and consent forms were handed to parents by the contacted preschools and nurseries. Visits to preschools and nurseries were arranged after obtaining parental consent. For the children tested at Lancaster University Babylab, parental consent was sought prior to the experiment on the first day of testing. The experiment took place on two separate days (one week apart), with one condition running on each day. The order of conditions was counterbalanced across participants. On the day of testing, the children were tested individually in a quiet area, under the supervision of a member of staff of the preschool or nursery (a parent in the case of testing at the Babylab). The children were either sitting at a table or on the floor.

The experimenter greeted the child by introducing his name and explaining what the child was expected to do, in terms of finding things for the people in the video. Then, the children completed the familiarisation trials. In each trial, pictures of two familiar objects were shown on the screen alongside a video clip featuring a female English speaker uttering a familiar word (e.g., *cup*). After hearing the label, the children were asked by the experimenter: "Which one is it?" and were encouraged to make their response by tapping an object on the screen. The target word was repeated if the touchscreen did not receive a response after 3.5 s, as pilot testing showed that if children had not responded within this

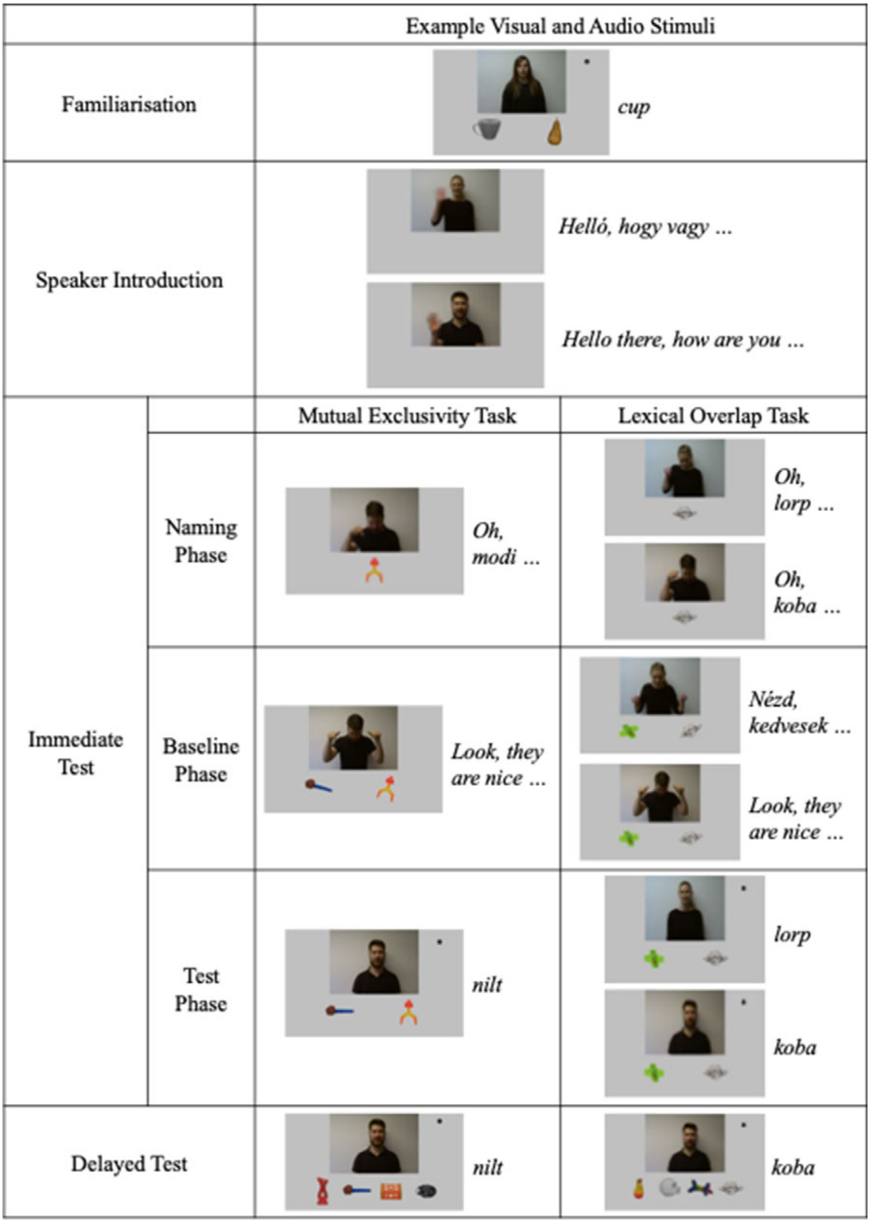

Figure 1. Example visual and audio stimuli for familiarisation, speaker introduction, and both the mutual exclusivity and lexical overlap tasks of the immediate and delayed tests.

time, they were unlikely to make a response without further prompting. In the event of the touchscreen failing to register the children's touch, the experimenter provided assistance.

The study only proceeded if a child had provided correct answers to all four familiarisation trials (five children failed to provide the correct answer on one trial at the first

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

instance, one failed on two trials at the first instance, but they all provided the correct answer on their second attempts). At this point, the experimenter repeated the instructions to the child. Children then took part in the English–English or the English–Hungarian condition and were then tested one week later on the other condition. The order of the language conditions (English–English first or English–Hungarian first) was counterbalanced.

2.3.1. English–English condition

Children were first shown a short introductory video clip of a male and a female English speaker, featured one at a time, saying: "Hello there, how are you? We are going to play a game. Would you like to play a game with me?" This provided socio-pragmatic information about the speakers' linguistic backgrounds. The order of appearance of the speakers was counterbalanced across participants. There were then two tasks: ME and LO, with a short pause between the tasks. The order of tasks was counterbalanced across participants.

ME Task. All children completed two sets of test trials: four immediate test trials consisting of three phases (naming, baseline, and test, in order of presentation), and eight delayed test trials. The immediate and delayed test trial sets were separated by a 10-minute break (see below for details). In the immediate test trial set, each test trial featured only one speaker. The presentation of test trials featuring the two speakers was alternated, and the speaker featured in the first test trial was counterbalanced across participants. For the ME task, the children were taught four novel words through explicit naming and were expected to learn four novel words through the application of ME during the immediate test trials.

Immediate Test Trials. In the naming phase, the speaker in the test trial labelled an unfamiliar object three times, each preceded by a short meaningless utterance ("oh", "hmm", or "ah"), while pointing at it and alternating gaze between the object and the children. In the baseline phase, two objects, the just-named object and an unnamed unfamiliar object, appeared on the screen and jittered to maintain the children's attention. At the same time, the speaker said: "Look! They are nice! Wow! They are pretty!", while pointing at both objects and alternating gaze between the objects and the children. This was to provide the children with an opportunity to view both objects that were going to be in the test phase to control for possible familiarity/novelty biases. The positions of the just-named and unnamed unfamiliar objects were randomised across test trials, but were the same in the baseline and test phases of a test trial. In the test phase, the speaker in the video looked at the children and uttered a novel label that was different from that in the naming phase. Then, the children were asked by the experimenter: "Which one is it?" The children were reminded to tap the screen if they only pointed to the object but did not touch the screen. Across the experiment, the children did not hear the label in the test phase more than two times. After all four trials, the children were told that they would be coming back to play some more of the game after 10 minutes. This pause of 10 minutes was included for practical reasons in the testing context of the preschool, but note that this exceeds other studies' retention interval (e.g., Horst and Samuelson (2008) investigated retention over a 5 minute delay).

Delayed Test Trials. On returning to the designated testing area, children were reminded of the instructions. They then saw four objects, the names of which either all occurred during ostensive teaching or all occurred in the ME immediate test trials

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

(i.e., they were either the targets in the naming or test phases of the ME immediate test trials). The objects appeared alongside a video clip of one of the two speakers uttering the name of one of the objects. The positions of the objects were randomised and were different across test trials. The speakers only uttered the labels that they used in the immediate test trials, in either a naming phase or a test phase.

LO Task. The procedure was identical to that of the ME task, with each child completing four immediate test trials and eight delayed test trials, with the exception that all video clips featured both speakers (in alternation). For the LO task, the children were taught eight words through explicitly naming.

Immediate Test Trials. All phases were identical to those in the ME task with the following exceptions: (1) the video clips in all phases featured two speakers, one after another; (2) in the naming phase, the two speakers named the same object with different names; and (3) in the test phase, each of the speakers spoke the same name that they used during the naming phase, rather than a different novel word, which was the case in the ME task. The order of appearance of the speakers was counterbalanced across trials and remained the same for all the phases of the same trial. The first speaker of the naming phase was counterbalanced across participants.

Delayed Test Trials. The procedure was identical to that in the ME task, with the exception that all four objects presented in any given trial were learned through exposure during the naming phase, as the children never had the opportunity to learn the names of the four unnamed objects that appeared in the immediate test trials.

2.3.2. English–Hungarian condition

The procedure was the same as that in the English–English condition, with the exception that whenever a video clip featured a Hungarian speaker, the sentences were spoken in Hungarian using an equivalent translation to the English version. The speakers used different languages in order to provide a stronger cue to different language backgrounds of the speakers than have been used in studies that varied accented English (e.g., Corriveau et al., 2013; Kinzler et al., 2011; Kinzler & DeJesus, 2013).

When children were distracted from the task by looking away from the screen, the experimenter redirected the child's attention. Trials where a child was distracted such that the stimuli were not observed were removed from analysis (n = 29).

3. Results

The datasets analysed for this study can be found in the Open Science Framework repository: https://osf.io/pyb2v/.

3.1. Accuracy analyses

Children's responses to test trials were scored as correct if they pressed the picture that was intended to be the answer or incorrect if they pressed any other pictures. See Table 1 for means and standard deviations of proportion correct.

In order to compare children's performance between the immediate and delayed tests (which differed in terms of the number of options available for selection), a likelihood score was computed for each child for their performance on each task in each condition in

Table 1. Means (and standard deviations) of proportion correct in the immediate and delayed tests by language group, condition, and task

	English–Hungarian		English–English	
	Monolingual	Bilingual	Monolingual	Bilingual
Immediate (chance = .50)				
Mutual exclusivity	.60 (.33)	.54 (.27)	.61 (.27)	.56 (.29)
Lexical overlap	.50 (.11)	.50 (.19)	.51 (.09)	.51 (.08)
Delayed (chance = .25)				
Mutual exclusivity	.26 (.19)	.30 (.15)	.27 (.19)	.31 (.18)
Lexical overlap	.34 (.13)	.21 (.17)	.29 (.16)	.29 (.15)

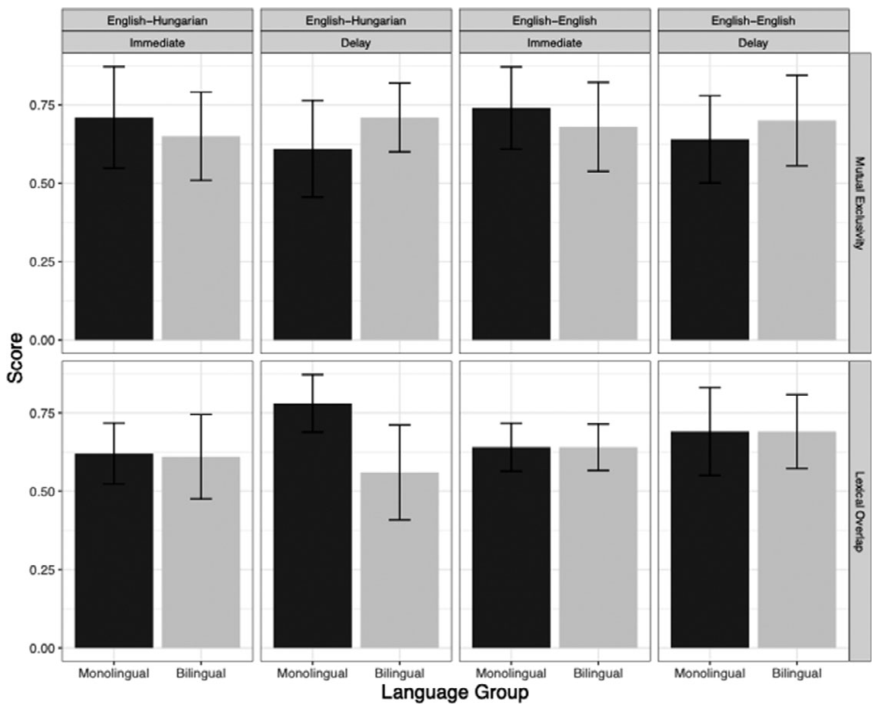

Figure 2. Children's likelihood scores in the immediate and delayed tests by language group, condition, and task. *Note.* Error bars represent standard errors.

the immediate and delayed tests. Each score was an indication of the likelihood of a child's performance on a given task in a given condition not being due to chance (likelihood score): Score = 1 - nCk $p^k(1\text{-}p)^{n\text{-}k}$, where n represents the number of trials, k the number of trials with a correct answer, and p the probability of success on each trial. See Figure 2 for means and standard errors of the likelihood scores.

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

Linear mixed-effects (LME) modelling (Baayen, 2008) was used to determine, with respect to our first aim, the learning of both monolingual and bilingual children, whether this differed by language context (one or two languages), and whether there were differences in flexibility in applying ME and LO for these groups in situations where speaker identity information was available. We also compared performance on immediate and delayed testing to see if that differently reflected learning for our second aim. In order to relate our sample to previous studies of vocabulary level relating to referent selection, we included the effect of children's English vocabulary (BPVS III) scores in the analysis in order to determine if learning words from ME or LO varied according to language proficiency. The effect of age was also tested to see if language proficiency or chronological development related more closely to the observed learning. The use of LME allows the investigation of both systematic and random individual differences (Jiang, 2007). There were a total of 320 observations. All likelihood scores were arcsine-root transformed prior to analysis to allow the bounded scores to be analysed using linear models, which assume dependent variables to be unbounded.

Intercorrelations (Pearson) between all predictor variables and the outcome variable (score) were examined and are shown in Table 2; note that many of the correlations are 0 due to the careful design of the study. Collinearity diagnostics indicated no possible risk of collinearity (condition number = 18, all $|r|$s $\leq .25$).

A series of LME models were fitted using the lmer function in the lme4 package in R, in order to determine the effect of task (LO or ME), language condition, one- versus two-speaker, immediate and delayed testing, and the role of vocabulary on performance. In all models, all predictors were entered simultaneously. First, assuming the same random effects of participants on intercepts, the following models differing in fixed effects, fitted using the *REML = FALSE* setting in lmer, were compared:

1. A model with just the intercept.
2. The final model obtained through the following backwards elimination steps: (a) A model with all participant and item attributes and additional predictor

Table 2. Intercorrelations (Pearson) between all predictor and outcome variables for immediate and delayed tests

	1.	2.	3.	4.	5.	6.
1. Score	-					
2. Delay (immediate vs. delayed)	.02	-				
3. Language group (monolingual vs. bilingual)	−.05	.00	-			
4. Age	.05	.00	−.19***	-		
5. English vocabulary (BPVS III) score	−.08	.00	−.25***	.20***	-	
6. Condition (English–Hungarian vs. English–English)	.04	.00	.00	.00	.00	-
7. Task (mutual exclusivity vs. lexical overlap)	−.05	.00	.00	.00	.00	.00

*** $p < .001$.

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

(extraneous) variables, including setting and gender of child, order of task (ME first vs. LO first), day of testing (day 1 vs. day 2), and which speaker appeared first in the introductory video clip. The drop1 function (test = "Chisq") was used to remove variables until the removal of all variables yielded a significant result from the likelihood ratio test. This was to ensure that the extraneous variables did not influence children's performance, and if any of these extraneous variables did influence children's performance, they were identified and included as a predictor in subsequent models. The final model was an empty model without any fixed effects.

3. The final model obtained through the following backwards elimination steps: (a) The most complex model with all six predictor variables and all interactions among them was first fitted to the data. (b) Then, the drop1 function (test = "Chisq") was used to determine whether dropping the highest order fixed effect would fit the data better. (c) The highest order fixed effect with the highest likelihood ratio test p-value once dropped was then removed from the model, and a model with the identified fixed effect removed was then fitted to the data. (d) Steps (b) to (c) were then repeated until all likelihood ratio test *p*-values between a more complex model and all simpler models with one of the highest order fixed effects removed from the complex model was smaller than .05 to obtain the final model.

The final model was the model with the following fixed effects: task, delay, English vocabulary (BPVS III) score; the two-way interactions between (a) delay and task, (b) English vocabulary score and delay, and (c) English vocabulary score and task; and the three-way interaction between English vocabulary score, delay, and task. Comparing a model with versus without the fixed effects was significant, $\chi^2(7) = 16.28$, $p = .022$.

To determine the random effects structure, the final model was then compared to the most preferred model (based on likelihood ratio tests) that included random effects of participants on the slopes of the fixed effects of delay, task, and/or the interaction between delay and task using the *REML = TRUE* setting in lmer. The inclusion of random effects of participants on the slopes of all the named fixed effects was not justified, $\chi^2(3) = 0.33$, $p = .953$, and so these slopes were not included. The final model is reported in Table 3.

Testing our first aim in terms of language group differences, the main effect of and any interactions involving language group were non-significant and removed in the model selection process. In addition, there was a significant main effect of task: the children's performance on the LO task was 1.14 times poorer than their performance on the ME task. Together, these suggest that the two language groups did not differ in their performance on both the ME and LO tasks for both immediate and delayed testing, and their performance on immediate and delayed testing was better for the ME than the LO learning conditions.

In terms of relations with the vocabulary level of the children, the three-way interaction between English vocabulary score, delay, and task was significant: In the ME task, the children's performance on both the immediate and delayed tests had a tendency to worsen with increasing English proficiency level, with their performance on the delayed test being affected more by their English proficiency. In contrast, in the LO task, while the children's performance on the immediate test had a tendency to worsen with increasing English proficiency level, their performance on the delayed test had a tendency to become better with increasing English proficiency (see Figure 3). Thus, the

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

Table 3. Summary of final model

Fixed Effects	Estimated Coefficient	SE	Wald Confidence Intervals 2.50%	Wald Confidence Intervals 97.50%	t	$pr(>\|t\|)$
(Intercept)	1.0555	0.0352	0.9864	1.1245	29.956	< .0001
Task (mutual exclusivity vs. lexical overlap)	−0.1326	0.0498	−0.2303	−0.0350	−2.662	.0082**
Delay (immediate vs. delay)	−0.0694	0.0498	−0.1671	0.0282	−1.394	.1644
English vocabulary score	−0.0011	0.0031	−0.0073	0.0050	−0.364	.7162
Delay × Task	0.1449	0.0705	0.0068	0.2830	2.057	.0405*
English vocabulary score × Delay	−0.0072	0.0044	−0.0159	0.0015	−1.632	.1037
English vocabulary score × Task	−0.0021	0.0044	−0.0108	0.0065	−0.483	.6298
English vocabulary score × Delay × Task	0.0139	0.0063	0.0016	0.0262	2.221	.0271*

Random effects	Name	Variance	*SD*	
Subject	(Intercept)	0	0	
	AIC	**BIC**	**logLik**	**Deviance**
	189.1	226.8	−84.5	169.1

* $p < .05$;; ** $p < .01$.

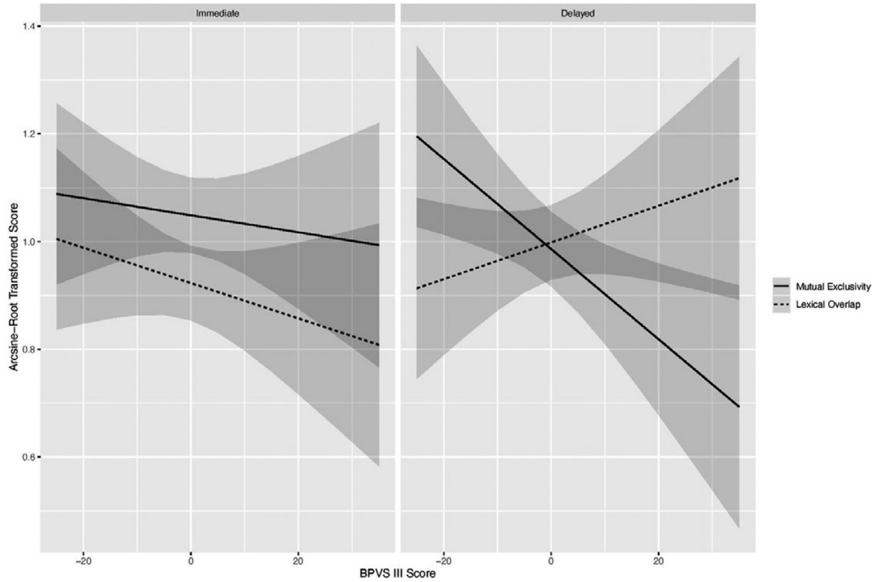

Figure 3. The interaction between English vocabulary (BPVS III) score, delay, and task on children's scores. *Note.* Shaded areas represent standard errors.

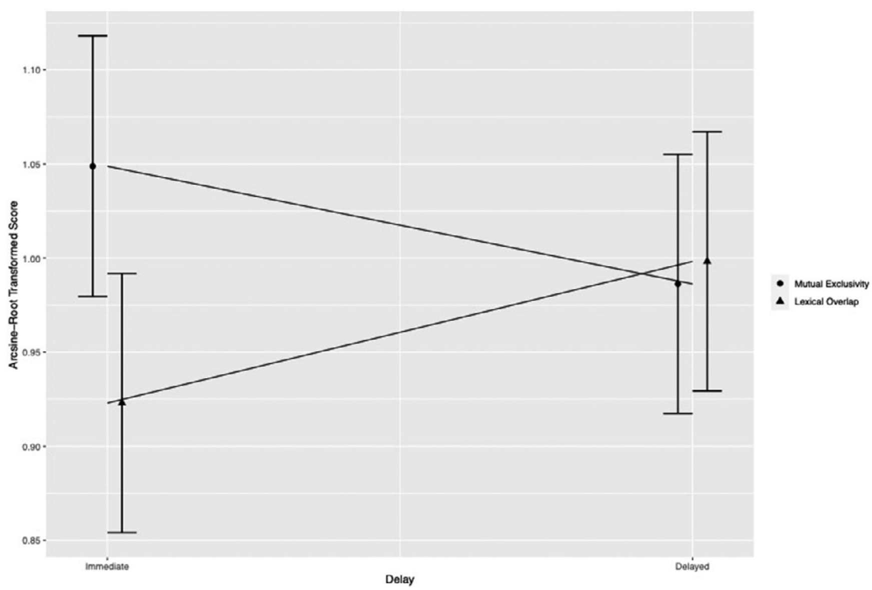

Figure 4. The interaction between delay and task on children's scores.
Note. Error bars represent standard errors.

relation between language proficiency and word learning from ME or LO tasks varies according to whether testing is immediate – a referent selection task – or after a delay, which reflects word learning performance. Children with varying language proficiency

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

levels are similar in terms of using ME for referent selection, but vary in their performance for retention.

Addressing our second aim, in determining influences on word learning after a retention period, the significant interaction between delay and task showed that the children's performance on the ME task had a tendency to worsen after the delay, whereas their performance on the LO task had a tendency to improve after the delay (see Figure 4). This suggests that immediate performance versus retention of learning varied according to the initial manner in which the word is learned – a difference observed between ME and LO in a referent selection task resolves to similar performance after a delay.

3.2. Bayes analyses on the effects of language group

The results of the mixed-effects analyses showed that monolingual and bilingual children did not perform significantly differently on the immediate and delayed tests across all tasks and conditions. However, in order to determine whether there was positive evidence that the language group did *not* have any influence on the children's performance, Bayes factors were computed for all tasks in all conditions of both the immediate and delayed tests. The Bayes factor is an indicator of whether the data support the research hypothesis, the null hypothesis, or neither. A value of 3 or higher indicates noticeable support for the research hypothesis, and a value of 1/3 or less indicates noticeable support for the null hypothesis (Dienes, 2014; Jeffreys, 1939/1961). Intermediate values between 3 and 1/3 indicate no substantial evidence for a difference or for no difference. For the present study, Bayes factors were computed based on arcsine-root transformed response accuracies. Bayes factors relating to the immediate test were computed with bounds of the difference being [0, 0.785], as the differences could range from no different from chance based on the null hypothesis (i.e., 0) to the maximum difference based on the research hypothesis: the difference between performance at chance level (.50; 0.785 when arcsine-root transformed) and 100% accuracy (1.00; 1.571 when arcsine-root transformed). By contrast, Bayes factors relating to the delayed test were computed with bounds of the difference being [0, 1.047], due to the chance level of the delayed test being .25 (0.524 when arcsine-root transformed). Additional Bayes factors based on arcsine-root transformed tendencies of accepting LO were computed for the LO tasks of the immediate test. The bounds of the difference for these Bayes factors were [0, 1.571], as the differences could range from not accepting LO at all (i.e., 0) to accepting LO for all pair trials (i.e., 1.00; 1.571 when arcsine-root transformed). When computing the Bayes factors, a uniform distribution was used, as the maximum plausible difference in all cases is known. No priors were drawn from previous studies as the manipulations of the present study differed substantially from those of previous studies (e.g., different number of distractors in a trial and inclusion of a two-language context). All computed Bayes factors are shown in Table 4.

The Bayes factors based on response accuracies demonstrate that the children's performance on all tasks in all conditions, except for the LO task in the English–Hungarian condition, provided evidence for the null hypothesis (i.e., evidence for no difference between the two language groups). In contrast, the monolingual and bilingual children performed differently on items in the LO task in the English–Hungarian condition in the delayed test. Interestingly, by examining the mean response accuracies of the LO task in the English–Hungarian condition of the delayed test, monolingual children were better than bilingual children in remembering word–object pairs that violate the one-label assumption in a two-language context.

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

Table 4. Bayes Factors comparing response accuracies of monolingual and bilingual children in all tasks in all conditions of the immediate and delayed tests

					Likelihood		
Test	Condition	Task	*M* difference	Difference *SE*	(data | difference)	(data | no difference)	Bayes Factor
Immediate	English–Hungarian	Mutual exclusivity	.065	.135	0.871	2.632	0.33*
		Lexical overlap	.009	.060	0.710	6.575	0.11*
	English–English	Mutual exclusivity	.052	.130	0.833	2.833	0.29*
		Lexical overlap	.004	.026	0.710	15.163	0.05*
Delayed	English–Hungarian	Mutual exclusivity	.073	.073	1.070	3.315	0.32*
		Lexical overlap	.194	.069	1.270	0.111	11.43**
	English–English	Mutual exclusivity	.037	.078	0.867	4.570	0.19*
		Lexical overlap	.028	.064	0.850	5.665	0.15*

*notable evidence for no difference between language groups; **notable evidence for difference between language groups.

4. Discussion

The present study investigated monolingual and bilingual preschoolers' ability to flexibly use and integrate different word-learning strategies when determining the meaning of new words: in particular, the application of ME, acceptance of LO, and use of speaker identity information in different language learning contexts. The first key aim of our study was to explore how monolingual and bilingual children learned from both ME and LO conditions, when words were spoken either by speakers of the same language or of different languages. Previous studies of LO, where an object is named with two names or a familiar object is named with a novel name, have tended to show a bilingual advantage – bilingual children are more likely to accept LO than monolingual children (Byers-Heinlein & Werker, 2009; Kalashnikova et al., 2015). However, in the only study that has directly tested children's acceptance of LO, speakers of the names for objects spoke the same language (Kalashnikova et al., 2015). We considered what would happen if the speakers of two labels for the same object are evidently speakers of different languages – could this reduce the bilingual advantage, converting monolinguals to also accept LO, by enabling the children to utilise socio-pragmatic information about different speakers producing distinct languages (Samara et al., 2017)? Our linear mixed-effects models suggest that this is the case. There was no significant effect of monolingual versus bilingual language background on learning from ME or from LO, and the Bayes Factor values indicate that there is evidence for similar performance between these language groups. Thus, monolingual and bilingual children are both able to accept LO during word learning.

Indeed, if anything, there was evidence of an advantage for the monolingual speakers in the two-language condition of the study: after a delay, monolingual children were more likely than bilingual children to have retained two labels for an object named in two languages. Thus, children around 4 years old are able to learn words when they are given one or two labels. Intuitively, the results could be anticipated to be in the opposite direction, as bilingual children should be more familiar with the idea that different languages can have different words for an object. However, it was possible that the monolingual children were particularly attentive to the speaker identity cue in the LO task in the English–Hungarian condition, as this was a situation which they were not used to (i.e., having to learn different names from different languages for an object). Alternatively, it was possible that, similar to Fitneva and Christiansen's (2011) finding that retention might not be optimal when the learning task is simple, the complexity of accepting LO and the presence of a novel language as socio-pragmatic information posed challenges to the monolingual children, which in turn boosted their retention performance.

Previous studies have found a bilingual advantage in the application of socio-pragmatic information to language, including speakers' language background (Akhtar et al., 2012) and to gestural communicative cues to referents (Byers-Heinlein et al., 2017; Yow et al., 2017; Yow & Li, 2018). However, there were key differences between these studies and ours, in that we focused on language comprehension, rather than production, as was tested in Akhtar et al. (2012), and comprehension may elicit more sensitivity to the use of cues by children than production tasks. The studies by Byers-Heinlein et al. (2017), Yow and Li (2018), and Yow et al. (2017) investigated a combination of communicative cues, such as eye-gaze, whereas our study focused on the relation between speaker identity and language and word learning.

The finding that both groups of children were more accurate in the ME than the LO task was not surprising. It has previously been documented that monolingual and

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

bilingual language learners are better at learning word–object mappings that adhere to ME (e.g., Benitez et al., 2016; Kachergis et al., 2009; Poepsel & Weiss, 2016). Although much of this evidence came from adult language learners, the results of Kalashnikova et al.'s (2015) study have provided support for this in young children. Kalashnikova et al. analysed cases whereby monolingual and bilingual children failed to learn two-to-one word–object mappings in their study and found that in those situations, the children reasoned the referent of a word by applying ME. This suggests that bilingual children also rely on ME to some extent when learning the meaning of words, as bilingual children, like monolingual children, also rely on strategies to help them reduce the number of potential referents when learning the meaning of a new word. Moreover, our analysis focusing on only the LO task (see additional analyses in Supplementary Appendix B) showed that the children were more likely to accept LO in the two-language condition. This finding is in line with that in the study by Samara et al. (2017), which showed that young children (and adults) are able to benefit from socio-pragmatic cues, in particular cues to speaker identity, in learning linguistic structures.

Together, these findings imply that monolingual and bilingual children are sensitive to the socio-pragmatic information present in their environment and can adjust their learning strategies – in the context of this study, relax ME and accept LO – to accommodate the demands of different learning contexts. Creel (2012) noted that ME is affected by similarities in accent between novel and known words, such that as the novel words become more distinct in accent from known words, novel objects have an increasing tendency to be selected. The bilingual population had a variety of different language backgrounds, and so similarity between the novel words and known words in their languages was not controlled. However, if this was adversely affecting the results, then we would expect this to result in differences in applying ME between the monolingual and bilingual children, and this was not the case. Rather, the key observation was a convergence of monolingual and bilingual children's performance on LO trials.

In terms of how learning from LO related to vocabulary development and whether this was similar to relations with ME, the results were complex. Previous studies of the application of ME have shown that, even for very young children, it appears to relate to vocabulary development (Bion et al., 2013; Kalashnikova et al., 2016b), and Kalashnikova et al. (2019) showed that bilingual vocabulary knowledge related to LO, but monolingual vocabulary level did not. In our results, we found a significant three-way interaction between vocabulary, whether the test was immediate or delayed, and task (ME or LO). In the delayed test, the higher a child's language proficiency score, the better they were at remembering words learned in the LO task, but worse at remembering those learned in the ME task.

This contrasting pattern could be explained in terms of the Emergentist Coalition Model of word learning, where older children are known to have a tendency to focus more on socio-pragmatic cues when learning the meanings of new words and less on basic constraints, such as ME (Pruden et al., 2006; Hollich et al., 2000). In the present study, the children's age and language proficiency score were significantly and positively correlated. Therefore, the contrasting pattern relating to language proficiency score and task in the delayed test could be due to the children's use of socio-pragmatic cues, in this case pointing and eye gaze, rather than ME to guide their learning of word–object mappings.

In contrast, the children's performance on both the ME and LO tasks in the immediate test worsened with increasing language proficiency. A possible explanation for this is that learning did not just take place during referent selection, but also during the delayed test via cross-situational statistical learning or associative learning, and that this learning was driven by the children's knowledge of or experience with the mapping between words and referents. This suggests that the children who were more proficient in English relied more on cross-situation statistical learning or associative learning, whereas the children who were less proficient in English relied more on ME. Note that these explanations are not inconsistent with the finding that the ME task was easier than the LO task for the children, as the children's performance on the ME task, averaged across the immediate and delayed tests, was still better than their performance on the LO task.

The second aim of our study was to investigate whether children could retain words as well as be able to select referents after being exposed to word–object mappings. Previous studies comparing ME and LO have tended to investigate only referent selection, yet the ability to retain mappings has been shown to be much more vulnerable in word learning studies (Horst & Samuelson, 2008). We thus investigated immediate and delayed learning from ME and LO conditions.

Though, overall, as seen in the accuracy analysis, learning from ME was easier than learning from LO, we found that this was affected by whether the testing was immediate or delayed. For referent selection, children were better able to identify word–object mappings when initial exposure was under ME conditions than LO. Hence, LO was a more difficult task for all the children in the study. However, the difference between ME and LO disappeared after a delay: now, children were similar in their learning of words under ME and LO conditions. Thus, the ME constraint – assuming a one-to-one mapping – was most evident as a referent selection advantage and dissipated after a delay in a measure of performance that more closely approximates children's word learning ability (Horst & Samuelson, 2008). Taking into account that children are sensitive to multiple cues of word learning (Pruden et al., 2006; Hollich et al., 2000), this finding implies that children may rely on the ME constraint as a referent-selection strategy for limited exposure to word–object pairings, but may turn to rely more on other, more reliable cues such as socio-pragmatic cues and cross-situational statistics or associative information with more exposure to the pairings (McMurray et al., 2012).

A limitation of the current study was that performance was not highly accurate in any condition. Given that children at a similar age are estimated to have a capacity of learning four new words per day (e.g., Bion et al., 2013; Mayor & Plunkett, 2011), our tasks may have been too demanding for the children. However, the Bayes Factor analyses indicate that null effects due to the different conditions were not due to noise or high error rates: there is positive evidence that language background had no effect under nearly all conditions. Repeating the study with older children, or with live rather than videoed presentations of word learning conditions, might increase children's attention further and enhance learning. Kalashnikova et al. (2015) presented their stimuli in a live puppet presentation, whereas in the present study, the presentation of stimuli was through a computer screen. The live presentation in Kalashnikova et al.'s (2015) study might be more effective in attracting and sustaining children's attention and focus on the task. For instance, the children in Kalashnikova et al.'s study were allowed to explore the objects in the task, whereas the children in the present study did not have the opportunity to do the same. In fact, in both the ME (based on likelihood scores) and LO (based on tendency to accept LO) tasks in the present study, the children's performance became more deviated from the expected performance as the experiment progressed, showing a possible fatigue effect.

In Kalashnikova et al.'s (2016a) study, a similar computerised paradigm was used, and monolingual children were able to apply ME and accept LO to greater accuracy than the children in the present study. Yet, the children in the present study were younger than

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

those in Kalashnikova et al.'s (2016a) study, who were between 4 and 5 years old. In Kalashnikova et al.'s (2015) study, when the children were divided into a younger and an older group, the performance of the younger group was significantly worse than that of the older group. Therefore, it was possible that repeating the study with older children might further increase the learning effects from ME and LO.

Another limitation of the present study was that, owing to Hungarian being a language that was unknown to all the children, the prompt to invite children to provide a response could not be provided by the speakers in the task (cf. Frank & Poulin-Dubois, 2002). In the present study, the prompt was provided by the experimenter, and the prompt was always in English. This could have an impact on the children's performance when the language in focus was Hungarian, as the prompt in English could have distracted them from concentrating on Hungarian being the language in focus. This could have reduced the sensitivity of the task in detecting whether monolingual and bilingual children would apply ME and accept LO differently in the two-language condition. Future studies could train children to provide a response upon a presentation of a visual cue on the screen, so that language would not be involved in the prompt for response. The similarity in performance of monolingual and bilingual children also contrasts with evidence for differences in LO between children according to their language background in referent selection tasks (Byers-Heinlein & Werker, 2009). In Byers-Heinlein and Werker's (2009) study, the expected difference between language groups could be observed even in 17- to 18-month-olds. Byers-Heinlein and Werker's (2009) study measured performance using eye-tracking, whereas our study (as well as that of Kalashnikova et al. (2015) relied on behavioural measures. Repeating the study with an implicit rather than an explicit behavioural measure may reveal processing differences between groups.

Further, we note that although we collected demographic information on what languages the bilingual children spoke and the number of years of exposure to each language as measures of language exposure, we treated the bilingual children in our study as a homogeneous group in terms of language experience and exposure. Indeed, all but one of the bilingual children had been exposed to two languages from birth, and so the duration of exposure to two languages was similar. However, information on the *quantity* of exposure to languages would also help us to ensure that the bilingual children are similar in terms of their language experience. Unfortunately, we did not collect these data, and future studies that measure the degree of exposure would ensure that our results are not dependent on particular patterns of bilingual language experience. Without these additional data, the generality of our results should be considered with caution.

Furthermore, the bilingual group were extremely heterogeneous in terms of which languages they spoke, limiting our ability to analyse the influence of the language distance between the two languages that the bilingual children spoke (e.g., English–German versus English–Arabic), which a close comparison of groups of bilinguals that speak the same two languages could provide. For instance, Byers-Heinlein and Werker (2013) found effects of individual languages on bilingual learners' application of ME – bilingual infants who knew more translation equivalents between their two languages were less likely to rely on ME, suggesting that the number of translation equivalents between the two languages that a bilingual learner speaks could influence their flexible use of ME. We also note that the language distance between the unfamiliar language used in the present study (i.e., Hungarian) and the different languages that the children in the present study spoke may also have influenced the children's responses. For instance, Hungarian is more phonotactically similar to Slovak than to Malagasy, so the socio-pragmatic cue may be more relevant/salient to some of the children than to others. Therefore, future studies

https://doi.org/10.1017/S0305000925100317 Published online by Cambridge University Press

should ideally take into account the different languages that bilingual children speak and their respective vocabulary knowledge.

5. Conclusion

In conclusion, the results of the present study showed that monolingual and bilingual preschoolers are sensitive to the socio-pragmatic cues, in terms of speakers' language identity, in their linguistic input and could integrate and alter their word learning strategies, for example, relax ME, with respect to their linguistic environment when learning the meanings of new words. In addition, it was also found that the children's use of ME and acceptance of LO were related to their vocabulary knowledge and timing of testing, showing a tendency to be more reliant on socio-pragmatic cues as their language developed, in line with the Emergentist Coalition Model (Hollich et al., 2000) of word learning. Nonetheless, although cross-situational statistical learning and associative learning were not directly investigated in the present study, our results provide some evidence that the children drew on cross-situational statistics or associative information to aid their learning of new words. Together, our findings suggest that monolingual and bilingual children can flexibly apply the same set of word learning strategies to different learning contexts, and that the ability to integrate speaker identity cues develops as a function of their vocabulary knowledge, highlighting the role of learning context and prior language experience in word learning.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0305000925100317.

Acknowledgements. This work was supported by a Leverhulme Trust Doctoral Scholarship (DS-2014-14) and the International Centre for Language and Communicative Development (LuCiD) at Lancaster University, funded by the Economic and Social Research Council (United Kingdom; ES/L008955/1).

Competing interests. The authors declare none.

References

Akhtar, N., **Carpenter, M.**, & **Tomasello, M.** (1996). The role of discourse novelty in early word learning. *Child Development*, **67**(2), 635–645. https://doi.org/10.2307/1131837.

Akhtar, N., **Menjivar, J.**, **Hoicka, E.**, & **Sabbagh, M. A.** (2012). Learning foreign labels from a foreign speaker: The role of (limited) exposure to a second language. *Journal of Child Language*, **39**(5), 1135–1149. https://doi.org/10.1017/S0305000911000481.

Au, T. K., & **Glusman, M.** (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development*, **61**(5), 1474–1490. https://doi.org/10.1111/j.1467-8624.1990.tb02876.x.

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge University Press.

Barron-Hauwaert, S. (2004). *Language strategies for bilingual families: The one-parent–one-language approach*. Multilingual Matters.

Benitez, V. L., **Yurovsky, D.**, & **Smith, L. B.** (2016). Competition between multiple words for a referent in cross-situational word learning. *Journal of Memory and Language*, **90**, 31–48. https://doi.org/10.1016/j.jml.2016.03.004.

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**(1), 39–53. https://doi.org/10.1016/j.cognition.2012.08.008.

Byers-Heinlein, K., **Behrend, D. A.**, **Said, L. M.**, **Girgis, H.**, & **Poulin-Dubois, D.** (2017). Monolingual and bilingual children's social preferences for monolingual and bilingual speakers. *Developmental Science*, **20**(4), e12392. https://doi.org/10.1111/desc.12392.

Byers-Heinlein, K., **Chen, K. H.**, & **Xu, F.** (2014). Surmounting the tower of babel: Monolingual and bilingual 2-year-olds' understanding of the nature of foreign language words. *Journal of Experimental Child Psychology*, **119**, 87–100. https://doi.org/10.1016/j.jecp.2013.09.011.

Byers-Heinlein, K., & **Werker, J. F.** (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, **12**(5), 815–823. https://doi.org/10.1111/j.1467-7687.2009.00902.x.

Byers-Heinlein, K., & **Werker, J. F.** (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition*, **128**(3), 407–416. https://doi.org/10.1016/j.cognition.2013.05.010.

Clark, E. V. (1987). The principle of contrast: A constraint on acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Lawrence Erlbaum Associates.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, **17**(2), 417–431. https://doi.org/10.1017/S0305000900013842.

Corriveau, K. H., **Kinzler, K. D.**, & **Harris, P. L.** (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology*, **49**(3), 470–479. https://doi.org/10.1037/a0030604.

Creel, S. C. (2012). Phonological similarity and mutual exclusivity: On-line recognition of atypical pronunciations in 3-5-year-olds. *Developmental Science*, **15**, 697–713. https://doi.org/10.1111/j.1467-7687.2012.01173.x.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, **5**, 781. https://doi.org/10.3389/fpsyg.2014.00781.

Diesendruck, G., & **Markson, L.** (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, **37**(5), 630–641. https://doi.org/10.1037/0012-1649.37.5.630.

Dunn, L. M., **Dunn, D. M.**, **Styles, B.**, & **Sewell, J.** (2009). *British picture vocabulary scale III*. NFER-Nelson.

Fitneva, S. A., & **Christiansen, M. H.** (2011). Looking in the wrong direction correlates with more accurate word learning. *Cognitive Science*, **35**(2), 367–380. https://doi.org/10.1111/j.1551-6709.2010.01156.x.

Frank, I., & **Poulin-Dubois, D.** (2002). Young monolingual and bilingual children's responses to violation of the mutual exclusivity principle. *International Journal of Bilingualism*, **6**(2), 125–146.

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**(1), 125–155. https://doi.org/10.1017/S0305000900008692.

Graham, S. A., **Poulin-Dubois, D.**, & **Baker, R. K.** (1998). Infants' disambiguation of novel object words. *First Language*, **18**(53), 149–164. https://doi.org/10.1177/014272379801805302.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, **87**(1), B23–B34. https://doi.org/10.1016/S0010-0277(02)00186-5.

Henderson, A. M. E., & **Scott, J. C.** (2015). She called that thing a mido, but should you call it a mido too? Linguistic experience influences infants' expectations of conventionality. *Frontiers in Psychology*, **6**, 332. https://doi.org/10.3389/fpsyg.2015.00332.

Hollich, G. J., **Hirsh-Pasek, K.**, **Golinkoff, R. M.**, **Brand, R. J.**, **Brown, E.**, **Chung, H. L.**, **Hennon, E.**, & **Rocroi, C.** (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, **65**(3), v–123. https://doi.org/10.1111/1540-5834.00090.

Horst, J. S., & **Hout, M. C.** (2016). The novel object and unusual name (NOUN) database: A collection of novel images for use in experimental research. *Behavior Research Methods*, **48**, 1393–1409. https://doi.org/10.3758/s13428-015-0647-3.

Horst, J. S., & **Samuelson, L. K.** (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, **13**(2), 128–157. https://doi.org/10.1080/15250000701795598.

Jeffreys, H. (1939/1961). *The theory of probability*, (1st/3rd ed.). Oxford University Press.

Jiang, J. (2007). *Linear and generalized linear mixed models and their applications*. Springer.

Kachergis, G., **Yu, C.**, & **Shiffrin, R. M.** (2009). Frequency and contextual diversity effects in cross-situational word learning. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual meeting of the cognitive science society* (pp. 2220–2225). Cognitive Science Society.

Kalashnikova, M., **Mattock, K.**, & **Monaghan, P.** (2014). Disambiguation of novel labels and referential facts: A developmental perspective. *First Language*, **34**(2), 125–135. https://doi.org/10.1177/0142723714525946.

Kalashnikova, M., **Mattock, K.**, & **Monaghan, P.** (2015). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Bilingualism: Language and Cognition*, **18**(4), 626–638. https://doi.org/10.1017/S1366728914000364.

Kalashnikova, M., **Mattock, K.**, & **Monaghan, P.** (2016a). Flexible use of mutual exclusivity in word learning. *Language Learning and Development*, **12**(1), 79–91. https://doi.org/10.1080/15475441.2015.1023443.

Kalashnikova, M., **Mattock, K.**, & **Monaghan, P.** (2016b). Mutual exclusivity develops as a consequence of abstract rather than particular vocabulary knowledge. *First Language*, **36**(5), 451–464. https://doi.org/10.1177/0142723716648850.

Kalashnikova, M., **Oliveri, A.**, & **Mattock, K.** (2019). Acceptance of lexical overlap by monolingual and bilingual toddlers. *International Journal of Bilingualism*, **23**(6), 1517–1530. https://doi.org/10.1177/1367006918808041.

Kandhadai, P., **Hall, D. G.**, & **Werker, J. F.** (2017). Second label learning in bilingual and monolingual infants. *Developmental Science*, **20**(1), e12429. https://doi.org/10.1111/desc.12429.

Kinzler, K. D., **Corriveau, K. H.**, & **Harris, P. L.** (2011). Children's selective trust in native-accented speakers. *Developmental Science*, **14**(1), 106–111. https://doi.org/10.1111/j.1467-7687.2010.00965.x.

Kinzler, K. D., & **DeJesus, J. M.** (2013). Northern= smart and southern= nice: The development of accent attitudes in the United States. *Quarterly Journal of Experimental Psychology*, **66**(6), 1146–1158. https://doi.org/10.1080/17470218.2012.731695.

Lanza, E. (1997). *Language mixing in infant bilingualism: A sociolinguistic perspective*. Oxford University Press.

Littschwager, J. C., & **Markman, E. M.** (1994). Sixteen- and 24-month-olds' use of mutual exclusivity as a default assumption in second-label learning. *Developmental Psychology*, **30**(6), 955–968. https://doi.org/10.1037/0012-1649.30.6.955.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, **14**(1), 57–77. https://doi.org/10.1207/s15516709cog1401_4.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, **92**, 199–227. https://doi.org/10.1016/0024-3841(94)90342-5.

Markman, E. M., & **Hutchinson, J. E.** (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, **16**(1), 1–27. https://doi.org/10.1016/0010-0285(84)90002-1.

Markman, E. M., & **Wachtel, G. F.** (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, **20**(2), 121–157. https://doi.org/10.1016/0010-0285(88)90017-5.

Markman, E. M., **Wasow, J. L.**, & **Hansen, M. B.** (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, **47**(3), 241–275. https://doi.org/10.1016/S0010-0285(03)00034-3.

Mather, E., & **Plunkett, K.** (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology*, **105**(3), 232–242. https://doi.org/10.1016/j.jecp.2009.11.004.

Mayor, J., & **Plunkett, K.** (2011). A statistical estimate of infant and toddler vocabulary size from CDI analysis. *Developmental Science*, **14**(4), 769–785. https://doi.org/10.1111/j.1467-7687.2010.01024.x.

McMurray, B., **Horst, J. S.**, & **Samuelson, L. K.** (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, **119**(4), 831–877. https://doi.org/10.1037/a0029872.

Mervis, C. B., & **Bertrand, J.** (1994). Acquisition of the Novel Name-Nameless Category (N3C) principle. *Child Development*, **65**(6), 1646–1662. https://doi.org/10.2307/1131285.

Poepsel, T. J., & **Weiss, D. J.** (2016). The influence of bilingualism on statistical word learning. *Cognition*, **152**, 9–19. https://doi.org/10.1016/j.cognition.2016.03.001.

Pruden, S. M., **Hirsh-Pasek, K.**, **Golinkoff, R. M.**, & **Hennon, E. A.** (2006). The birth of words: Ten-month-olds learn words through perceptual salience. *Child Development*, **77**(2), 266–280. https://doi.org/10.1111/j.1467-8642.2006.00869.x.

Quine, W. V. O. (1960). *Word and object*. MIT Press.

Rosch, E., **Mervis, C. B.**, **Gray, W. D.**, **Johnson, D. M.**, & **Boyes-Braem, P.** (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X.

Samara, A., **Smith, K.**, **Brown, H.**, & **Wonnacott, E.** (2017). Acquiring variation in an artificial language: Children and adults are sensitive to socially conditioned linguistic variation. *Cognitive Psychology*, **94**, 85–114. https://doi.org/10.1016/j.cogpsych.2017.02.004.

Scott, J. C., & **Henderson, A. M.** (2013). Language matters: Thirteen-month-olds understand that the language a speaker uses constrains conventionality. *Developmental Psychology*, **49**(11), 2102–2111. https://doi.org/10.1037/a0031981.

Slavin, S. (2014). *PsyScript 3 (version 3.2.1)* [Computer software]. Lancaster University. http://www.lancaster.ac.uk/psychology/research/research-software/psyscript3.

The Object Databank [Computer software]. (1996). Carnegie Mellon University. http://wiki.cnbc.cmu.edu/Objects

Vlach, H. A., & **DeBrock, C. A.** (2019). Statistics learned are statistics forgotten: Children's retention and retrieval of cross-situational word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **45**(4), 711–711. https://doi.org/10.1037/xlm0000611.

Wagenmakers, E. J., **Marsman, M.**, **Jamil, T.**, **Ly, A.**, **Verhagen, J.**, **Love, J.**, & **Morey, R. D.** (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, **25**(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3.

Waxman, S. R. (1989). Linking language and conceptual development: Linguistic cues and the construction of conceptual hierarchies. *Genetic Epistemologist*, **17**, 13–20. http://groups.psych.northwestern.edu/waxman/linkinglangandconceptualdev.pdf

Waxman, S. R., & **Hatch, T.** (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Psychology*, **19**, 153–166. https://doi.org/10.1017/S0305000900013672.

Yow, W. Q., & **Li, X.** (2018). The influence of language behavior in social preferences and selective trust of monolingual and bilingual children. *Journal of Experimental Child Psychology*, **166**, 635–651. https://doi.org/10.1016/j.jecp.2017.09.019.

Yow, W. Q., **Li, X.**, **Lam, S.**, **Gliga, T.**, **Chong, Y. S.**, **Kwek, K.**, & **Broekman, B. F.** (2017). A bilingual advantage in 54-month-olds' use of referential cues in fast mapping. *Developmental Science*, **20**(1), e12482. https://doi.org/10.1111/desc.12482.

Cite this article: Chan, K.C.J., & Monaghan, P. (2025). Flexible Use of Word Learning Strategies: Monolingual and Bilingual Children's Word Learning Under Different Language Contexts. *Journal of Child Language* 1–24, https://doi.org/10.1017/S0305000925100317