


ARTICLE

# The effect of masks on infants' ability to fast-map and generalize new words

Siying LIU<sup>1</sup> , Xun LI<sup>1</sup> and Renji SUN<sup>2</sup>

<sup>1</sup>Institute of Linguistics, Shanghai International Studies University, Shanghai, China

<sup>2</sup>East China University of Political Science and Law, China

**Corresponding author:** Siying Liu; Email: 2019616@shisu.edu.cn

(Received 05 May 2023; revised 09 November 2023; accepted 25 November 2023)

## Abstract

Young children today are exposed to masks on a regular basis. However, there is limited empirical evidence on how masks may affect word learning. The study explored the effect of masks on infants' abilities to fast-map and generalize new words. Seventy-two Chinese infants (43 males,  $M_{\text{age}} = 18.26$  months) were taught two novel word-object pairs by a speaker with or without a mask. They then heard the words and had to visually identify the correct objects and also generalize words to a different speaker and objects from the same category. Eye-tracking results indicate that infants looked longer at the target regardless of whether a speaker wore a mask. They also looked longer at the speaker's eyes than at the mouth only when words were taught through a mask. Thus, fast-mapping and generalization occur in both masked and not masked conditions as infants can flexibly access different visual cues during word-learning.

**Keywords:** infants; word learning; fast-mapping; generalization; visual cues; masks

## Introduction

Ever since the outbreak of COVID-19, wearing masks in public has become a common practice, especially for people living in Asian countries such as China, Japan and South Korea (Feng et al., 2020). Although not compulsory now, masks are still worn mainly for reasons of protection from respiratory illnesses and air pollution (Crimon et al., 2022; Greenhalgh et al., 2020). Consequently, young language learners today are exposed to masks on a regular basis and sometimes must acquire language from adults who wear masks (Singh et al., 2021). Masks are predicted by researchers to be present in young children's learning environment for the long term. It is therefore worth exploring their effects on language development (Yeung et al., 2020).

Past studies indicate that masks may alter the quality of speech transmission (Corey et al., 2020; Fecher & Watt, 2011; Saeidi et al., 2016; Saigusa, 2017). Cloth masks have been found to reduce speech transmission by 3 to 4% and the percentage for N95 masks is 13 to 17% (Palmiero et al., 2016). Another study found that common types of masks (N95, surgical, and cloth) affected the acoustic measures of speech (e.g., N95 masks impacted the power

distribution in frequencies above 3kHz) while other acoustic features such as measures of voice quality remain unaffected (Magee et al., 2020).

Masks also obscure a speaker's mouth which provides visual cues for language understanding (Flom & Bahrick, 2007; Kuhl & Meltzoff, 1982, 1984; Lalonde & Werner, 2019; Lewkowicz, 2010; Lewkowicz & Flom, 2014; Lewkowicz & Hansen-Tift, 2012). Starting from 6-months, infants can observe lip movements to discriminate between different phonemes (Teinonen et al., 2008). Four- to 10-month-olds are sensitive to the temporal synchrony between speech sound and the corresponding mouth movements (Lewkowicz, 2010). Twelve-month-olds differentiated familiar and unfamiliar words only under conditions in which the mouth movements are generally consistent with the words pronounced (Weatherhead & White, 2017). Starting from 18-months of age, infants begin to attend more to a speaker's mouth when exposed to both infant-directed and adult-directed utterance (de Boisferon et al., 2018). Thus, infants rely on both auditory and visual cues for language comprehension (Cohn et al., 2021) and masks could potentially hinder communication.

Studies with adults suggest that masks in some circumstances may affect language comprehension. For example, Magee et al. (2020) found that adults' ability to complete word and sentence translations was affected by all types of masks. Cohn et al. (2021) discovered that cloth masks affected adults' ability to repeat speech only when sentences were produced in a positive-emotional style. For sentences that were spoken in a casual style, repetition was unaffected by masks. There are also studies that indicate speech perception is unaffected by masks even in a noisy environment (Atcherson et al., 2017; Mendel et al., 2008).

Studies with children also produced mixed findings regarding the effect of masks on communication. For example, Tronick and Snidman (2021) found that mask wearing did not disrupt mother-child interaction for 5- to 19-month-olds. Lalonde et al. (2022) tested consonant recognition in children with bilateral hearing loss, children with normal hearing (between 7 to 19 years of age) and adults with normal hearing. It was found that speech recognition was similarly impacted by masks for all groups. Schwarz et al. (2022) asked 8- to 12-year-olds to repeat the last word of sentences presented in audio-visual format. Children made more mistakes and showed slower processing speed when the sentences were produced with a mask. In another study by Crimon et al. (2022), French and Japanese nursery school educators reported that masks reduced their language quantity and also children's verbal communication. Nonetheless, educators also reported increases in the use of non-verbal cues which could compensate for the reduction in language quantity.

Apart from the aforementioned studies, there is an emerging line of research that specifically investigates the impact of masks on word learning. Word segmentation (extracting wordforms from running speech), for example, is an ability that facilitates word learning (Bergmann & Cristia, 2016). In a study by Frota et al. (2022), 7- to 9-month-olds completed an auditory and an audiovisual word segmentation task in two conditions: without and with an FFP2 mask. It was found that unlike those from a pre-pandemic study (Butler & Frota, 2018), infants born during the pandemic showed no evidence for word segmentation both in the presence and absence of a mask. The authors argued that COVID-related changes in everyday communication could explain these findings. That is, masks are only one of the factors within an overall effect of the pandemic that disturbs word learning. In another study by Sfakianaki et al. (2021), the impact of masks on familiar word recognition was closely examined. Adults and children aged 6- to 7-years were presented a low frequency word recognition test in either quiet and noisy environments. Half of the

words were produced with a mask. The results revealed that word recognition was compromised when produced with masks regardless of participants' age and noise level. Kwon and Yang (2023) investigated the combined effects of mask usage (no mask, surgical masks and KF94 masks) and room acoustics on familiar word recognition in preschool children in real classroom settings. It was found that 4- and 5-year-olds' performance was disturbed by masks more than 6-year-olds. Using the Intermodal preferential paradigm (IPL), Singh et al. (2021) investigated 24-month-olds' ability to recognize familiar words spoken through different types of masks (no masks, opaque masks and clear masks). Infants saw two objects (one target and one distractor) while hearing a speaker producing a familiar word that was supposed to match the target. They then had the opportunity to visually identify the target. Results showed that 24-month-olds could recognize words spoken without a mask and through opaque masks, but not through clear masks. The authors argued that when encountering speakers with masks, infants can access cues other than those from the mouth to recover information. This finding was supported by studies suggesting that linguistic cues such as whole-head movement and eye gazes can also facilitate word learning (Langton et al., 2000; Munhall et al., 2004).

Taken together, previous research produced mixed results regarding the impact of masks on word learning in young children. However, apart from word segmentation and familiar word recognition, there are several other fundamental abilities that underlie word learning (Horváth et al., 2016; Schafer & Plunkett, 1998). For example, fast-mapping refers to the ability to learn a word (i.e., mapping word forms to an object or an action) after minimal exposure (Eyer et al., 2002; Houston-Price et al., 2005; Wilkinson & Mazzitelli, 2003). Using the IPL paradigm, Schafer and Plunkett (1998) taught 15-month-olds two novel words (e.g., *bard* and *sarl*) as names of two novel objects. After each object was named 12 times, infants saw both objects together and heard one word. It was found that infants consistently looked longer at the target objects than distractor object, revealing that they could rapidly "fast-map" a novel word to an object. Houston-Price et al. (2005) found that for 18-month-olds, three repetitions of a word were sufficient for learning to occur. Using a preferential pointing task, Spiegel and Halberda (2011) suggest that 24-month-olds can fast-map novel words after a brief single exposure and also retain the word after a one-minute break.

Another ability that facilitates word acquisition is word generalization across objects of the same category and across different people (Buresh & Woodward, 2007; Graham et al., 2006; Henderson & Woodward, 2012). This ability helps infants to save cognitive effort in word learning. That is, after acquiring a new word, infants do not have to learn the word repeatedly every time they encounter a new speaker or another object of the same category (Liu & Sun, 2019). Horváth et al. (2016) taught 16-month-olds two novel word-object pairs. Later in the tests, the objects changed colour and infants heard only one of the trained words. It was found that when provided enough time to consolidate the information (i.e., after taking a nap), infants were able to associate the words with the correct objects even when the objects changed colour. Using both the habituation and IPL paradigm, Liu and Sun (2019) repeatedly presented 13-month-olds with novel word-object pairs. The objects then changed colour and infants viewed the same speaker and a new speaker using the taught word to name either the different coloured target or distractor. They were also asked by both speakers to visually locate the correct referents. Results showed that infants expected both speakers to use the words for objects of the same category, thus demonstrating the ability to generalize words across people and across objects of the same category. These studies were supported by the finding that infants tend to rely on an object's shape rather than colour when identifying category (Graham &

Poulin-Dubois, 1999). This is because shape is more stable than other features such as colour and size and thus is more relevant for category identification and word generalization (Xu et al., 2004; Yoon et al., 2008).

To date, past studies only explored the effect of masks on word segmentation and familiar word recognition in young children, but there had been little empirical evidence about whether masks would influence fast-mapping and word generalization. Targeting this question, the present study adopted the IPL paradigm which has been reported to be suitable for pre-verbal infants (Ballem & Plunkett, 2005; Mani & Plunkett, 2008; Singh et al., 2015). Eighteen-month-olds were chosen as participants because previous literature indicates that, at this age, infants begin to fast-map words after very few exposures (Houston-Price et al., 2005). In the training trials of the study, infants were taught two novel word-object pairs for three times each by a speaker who either wore a cloth mask or did not wear a mask. They then completed the fast-mapping tests during which they saw the two objects being placed together, heard one matching word and had to visually identify the correct target object. Later in the generalization tests, the objects changed colour and a different speaker tested infants using the same procedure. Based on the mixed findings by Singh et al. (2021) and Frota et al. (2022), we asked whether mask wearing would impact infants' fixation on the objects.

## Methods

### Participants

The participants were 72 full term infants (43 males, mean age = 18 months, 8 days, range = 17 months, 22 days to 18 months, 18 days) recruited from three nurseries in Shanghai, China. No participant had known health-related issues and familial or other risks for language impairment. To detect an effect of partial  $\eta^2 = .06$  with 80% power in a repeated-measures, within-between interaction ANOVA (two groups,  $\alpha = .05$ , non-sphericity correction = 1), a priori GPower 3.1 analysis suggested that 30 participants would be sufficient (Faul et al., 2007). However, since a greater number of parents volunteered to take part in the study, the final sample consisted of 72 participants. Past studies suggest that larger sample sizes contribute to greater power (e.g., Desmond & Glover, 2002; Oakes, 2017).

All infants were Chinese and exposed to Mandarin as their first language. They were randomly assigned into either the No Mask Condition ( $n = 36$ , 24 males, mean age = 18 months, 6 days, range = 17 months, 26 days to 18 months, 18 days) or With Mask condition ( $n = 36$ , 19 males, mean age = 18 months, 11 days, range = 17 months, 22 days to 18 months, 17 days). Five additional infants were tested but not included in the final sample due to insufficient eye-tracking data (had over 25% data loss at calibration or less than 30% data during recordings at test trials) ( $n = 4$ ) and technical errors ( $n = 1$ ). The criteria for insufficient data were based on infant eye-tracking studies by Johnson et al. (2003) and Hessels et al. (2015). All parents gave informed consent for the infants' participation prior to their inclusion in the study.

The present study was conducted in the post-pandemic period. However, all infants were born during the Covid-19 pandemic period and all parents reported that their infants were exposed to masks on a daily basis since birth. The estimated mean time of infants' mask exposure from January 2022 to January 2023 was 1.86 hours per day ( $SD = .87$ , range = 0.50 to 4.00 hours).

### Stimuli

Four familiar objects which were a rabbit puppet, monkey puppet, a book and a ball were used. Four novel objects (two objects from each object category that differed only in colour) were also selected (see Figure 1). Infants did not show any biases towards a particular category of objects (for a detailed analysis see Appendix A).

One object category was named using the novel Mandarin pseudo-word ‘mi (1) dou (4)’ and another was named ‘ding (1) ge (2)’. The four syllables ‘mi’ ‘dou’, ‘ding’ and ‘ge’ were chosen based on previous findings that 1) /m/, /d/, /g/ were among the first and most frequent consonant sounds that infants as young as 6-months of age acquire and produce (Hua & Dodd, 2000; Nathani et al., 2006; Swingley, 2021) and 2) /i/, /ou/, /ing/, /e/ were also among the most common vowels in Chinese that could be produced by infants (Du, 2010; Hua & Dodd, 2000). The consonants and vowels were combined into four syllables and then the syllables were combined into two words. The first, second and fourth tones were then randomly assigned to the syllables. The above procedure was repeated to produce two words that did not coincide with existing Chinese lexicons or bear any inherent meaning. Note that the third tone was not used because it is significantly longer than the other tones (Cao & Sarmah, 2007). There was another pseudo-word ‘niu xing’ which was not taught as an object label in the training trials but was presented in the test trials. The two syllables were selected based on the finding that the consonants (/n/ and /x/) and vowels (/iu/ and /ing/) are also common in Chinese and could be articulated by 18-month-old Chinese infants (Hua & Dodd, 2000). The syllables were combined into a word using the same procedure as for ‘mi dou’ and ‘ding ge’.

Twenty adult native-speakers rated the level of resemblance of each pseudo-word to real words on a five-point scale (0 = no resemblance to real words, 5 = high resemblance to real words) and the mean rating were 4.92 ( $SD = .31$ ) for ‘mi dou’, 4.95 ( $SD = .22$ ) for ‘ding ge’ and 4.89 ( $SD = .28$ ) for ‘niu xing’.



Figure 1. Novel objects and speakers in the No Mask (left) and With Mask (right) Conditions.

The word-object pairs were taught and tested by two speakers (one male and one female, see [Figure 1](#)) in the form of videos. Acoustic analyses were conducted for all videos in the training phase and test phases. Results suggest that masks did not affect the acoustic features of the target words. See Appendix B for detailed analyses.

### *Eye-tracking*

Tobii Pro Fusion eye tracker with a sampling rate of 250Hz was used. Accuracy was about 0.3 degrees for binocular eye movements ( $M = .29$ ,  $SD = .07$ ,  $range = .19$  to  $.43$ ). The eye-tracker was connected to a laptop (controlled by the experimenter) that presented video stimuli on the screen and monitored children's visual attention using the Tobii Prolab software. The display monitor resolution was 1920 x 1080. The Areas of Interest (AOIs) were the speaker's eyes, mouth and hand in the training phase and the objects in the test phases (see [Figure 2](#)). The AOIs were dynamic that changed with movements of the head and hands. This was achieved using the dynamic AOI tool in the Tobii Prolab software (Tobii AB, 2023), which readjusted the positions of the AOIs with every movement of the corresponding areas. Areas in pixels were the same across videos (eyes: 161, 51; mouth: 93, 68; hand: 255, 164; objects: 155, 186).

### *Procedure*

The experiment consisted of a training phase, a fast-mapping test phase and a generalization test phase. In all phases, infants sat approximately 60cm from the eye-tracker in a quiet classroom at the nurseries. They sat in a baby chair with safety straps which restrained body movements. The surrounding was curtained so children could only see the screen. A nine-point calibration was first conducted with individual calibration points repeated until all calibration points were obtained for each participant. The 3D eye models in Tobii eye trackers can compensate for drift and are robust against changes in head position, therefore it is suggested that calibration is conducted only once prior to the experiment and needs not to be adjusted during recording (Tobii AB, 2023).

After calibration, the training phase began. There were 10 trials in total. Infants first saw two speaker familiarization trials. In the first trial, one speaker (e.g., the male speaker) appeared from behind a table and said "ni hao, wo zai zhe li" 'hi, here I am' in Mandarin, hid behind the table, then appeared again and smiled. Another speaker (e.g., the female speaker) repeated the same actions in the second trial. Each trial lasted for 8000ms. The order of speaker presentation was randomized.

Infants then saw two familiar word training trials presented in a random order. Each trial began with an attention getter (a ringing sound for 1000ms) and one speaker appeared on screen sitting behind a table with a puppet (e.g., a rabbit) in front of him or her. The speaker first attracted infants' attention by making direct eye contact for 2000ms. He or she then gazed at the puppet for 2000ms, named it "tuzi" 'rabbit', pointed at the puppet for another 2000ms and named it the second time by saying "yi zhi tu zi" 'a rabbit'. After that the speaker maintained his/her final position (in a still frame) for 3000ms before the trial ended. The second trial involved the same speaker using the same procedure to name a ball. These two trials demonstrated to infants that the speaker was explicitly naming the objects.

The familiar word training trials were followed by six novel word training trials. In a typical trial, the speaker had one of the novel objects in front of him or her and named the

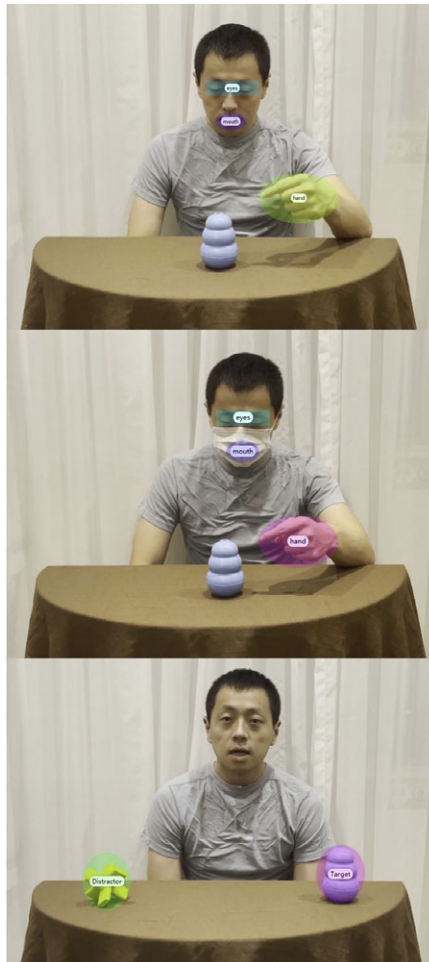


Figure 2. Examples demonstrating Areas of Interest in the training trials (top two) and test trial (bottom).

object “mi dou” for the first time and “yi ge mi dou” ‘a mi dou’ for the second time. The naming procedure and timing were the same as the familiar word training trials (see Figure 3 for the timeline of a training trial). This trial was repeated three times as previous studies suggest that three exposures were enough for 18-month-olds to learn word-object pairs (Houston-Price et al., 2005). The speaker then used the word “ding ge” to name the other object in the next three trials before the training phase ended. Note that half of the infants in each condition learned the word “mi dou” first and the other half was taught “ding ge” first.

Infants then underwent the fast-mapping test phase which consisted of eight test trials. The first two trials were familiar word test trials (presented in a random order) for the words “rabbit” and “ball”. In a typical trial, after a 1000ms attention getter, the speaker from the training phase appeared on screen with two familiar objects (e.g., rabbit as the target and monkey as the distractor) in front of him or her. He or she first established eye

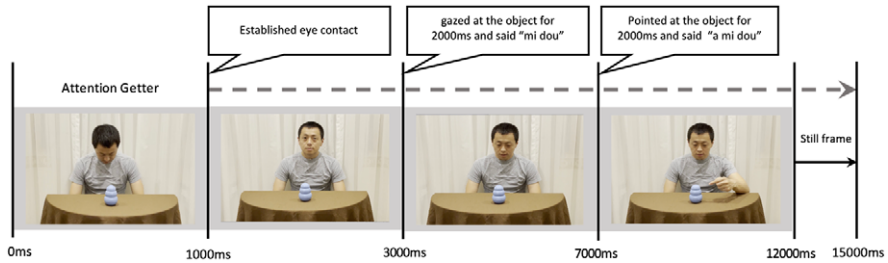


Figure 3. Timeline of a training trial in the No Mask Condition.

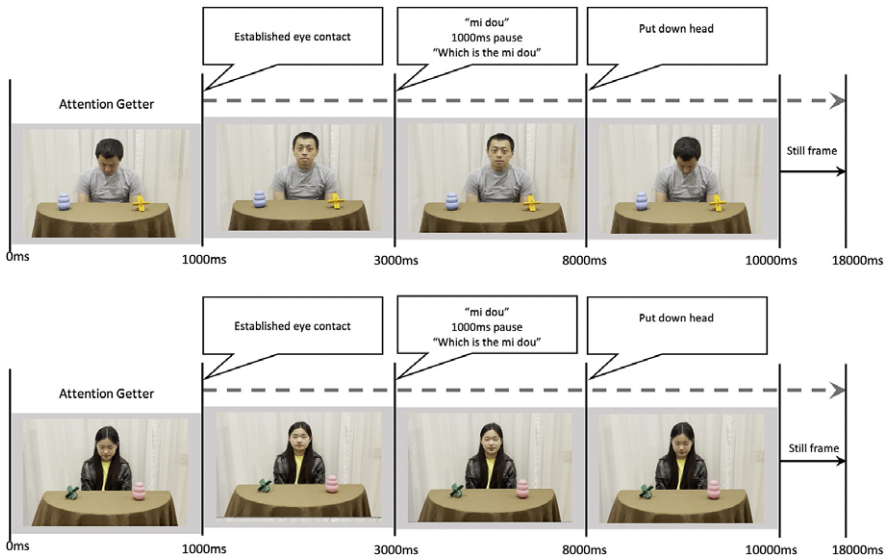
contact for 2000ms and then said “Tu zi. Tu zi shi na yi ge?” ‘Rabbit. Which is the rabbit?’. The two sentences were separated by a 1000ms pause. After finishing the questions, the speaker put down his or her head to break eye-contact so infants were more encouraged to look at the objects rather than the speaker’s face. Infants were then allowed 8000ms to look at the screen before the test trial ended. The second trial involved the speaker asking infants to find the ball in the presence of a book as the distractor. The familiar word test trials assessed whether infants could respond to the paradigm by looking at the target object upon hearing a taught word.

Infants then watched six fast-mapping test trials which were divided into two blocks. The first block contained three trials (presented in a random order). One test trial involved the word “mi dou”, one involved “ding ge” and one involved the word “niu xing” (which was not taught in the training phase). The procedure in these test trials was identical to the familiar word test trials (see figure 4 for the timeline of a fast-mapping test trial). The second block was a repetition of the first block. Note that the word “niu xing” was tested to further confirm that infants learned the association between the taught words and the two corresponding objects. If infants established the word-object associations, they were expected to show random looking at the objects (thus not identifying a target) or not look at the objects upon hearing the word “niu xing”.

After completing the fast-mapping test phase, infants last underwent the generalization test phase. This phase began with a new-speaker trial during which a different speaker (who was in the speaker familiarization trial but not in the training trials) appeared with two novel objects which changed colours. The speaker established eye-contact for 3000ms and maintained a neutral expression for 5000ms before the trial ended. This trial familiarized infants with the new speaker and the different coloured objects so infants’ looking responses in the subsequent generalization test trials would not be affected by sudden changes in speaker and object colours. After the new-speaker trial, infants saw six generalization test trials (arranged into two blocks). The procedures were identical to those in the fast-mapping test phase except that the speaker and objects’ colours were changed (see Figure 4).

In the above phases, infants in the With Mask Condition saw the two speakers wearing white surgical cloth masks with ear loops throughout the experiment; and for those in the No Mask condition, the speakers did not wear masks. If an infant did not look at the screen for a particular trial, it was repeated. The following factors were counterbalanced across infants in each condition: the target object, the positions (left and right) of the target and distractor and the gender of the speaker who presented the training trials, fast-mapping test trials and generalization test trials (i.e., infants who watched the male speaker during





**Figure 4.** Timeline of a fast-mapping test trial (top) and a generalization test trial (bottom) in the No Mask Condition.

training and fast-mapping test phases watched the female at generalization test phase and vice versa). The study procedure was approved by the Research Ethics Committee in accordance with the Declaration of Helsinki.

## Results

### *Preliminary analyses*

#### *Total attention analyses*

An independent samples t-test was first conducted to compare mean total fixation time (towards all trials) between the No Mask and With Mask Conditions. The analysis showed no significant difference in fixation time,  $t(70) = .44$ ,  $p = .663$ , indicating that infants were equally attentive to all videos in the two conditions. For the training trials involving the two novel words, infants demonstrated a tendency to look more to the With Mask Condition ( $M = 81.96$ ,  $SD = 6.53$ ) than to the No Mask Condition ( $M = 79.50$ ,  $SD = 6.12$ ), but the difference did not reach significance,  $t(70) = 1.65$ ,  $p = .052$ . For the test trials involving the novel words, looking time did not differ significantly between the two conditions,  $t(70) = -1.05$ ,  $p = .296$ .

#### *Familiar word test trials analyses*

To explore whether infants identified the target object upon hearing a familiar word, a 2 (objects: target, distractor)  $\times$  2 (conditions: no mask, with mask) mixed-design ANOVA was first conducted on infants' mean total fixation time in the familiar word test trial for the word "rabbit", with objects as the within-subject factor. Fixation time was recorded immediately after the speaker finished his or her questions (e.g., after the sentence "tu zi shi na yi ge?" 'which is the rabbit?'). Results showed a significant main effect of objects,  $F(1, 70) = 14.40$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .17$ .

To follow up on this main effect and further explore how infants differentiated between the target and distractor, a paired samples t-test was performed to compare fixation time to the target and distractor. Results indicated that infants (in all conditions) looked significantly longer at the target ( $M = 3.06$ ,  $SD = 2.23$ ) than at the distractor ( $M = 1.80$ ,  $SD = 1.49$ ),  $t(71) = 3.72$ ,  $p < .01$ ,  $d = .66$ ,  $r = .32$ , 95 % CI [.58, 1.93]. Longer looking time towards the target was observed regardless of condition as there were no interaction effects between objects and conditions,  $F(1, 70) = 3.72$ ,  $p = .058$ .

A similar ANOVA was then conducted on the second familiar word test trial for the word “ball” and revealed a significant effect of objects ( $F(1, 70) = 13.43$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .16$ ) but no significant interaction between objects and conditions,  $F(1, 70) = .84$ ,  $p = .364$ . A follow-up paired samples t-test showed that infants looked significantly longer at the target ( $M = 3.13$ ,  $SD = 2.23$ ) than at the distractor ( $M = 1.67$ ,  $SD = 1.89$ ),  $t(71) = 3.67$ ,  $p < .001$ ,  $d = .71$ ,  $r = .33$ , 95 % CI [.66, 2.25].

The above findings indicate that 18-month-olds in the present study could respond to the paradigm by visually identifying the target object.

## Main analyses

### Test trials analyses

The focal question of the present study was whether mask wearing would affect infants’ ability to 1) fast-map the two novel words and 2) generalize words across speakers and objects of the same category. To answer this question, infants’ mean total fixation time towards the target and distractor (after the speaker finished his or her questions) in the fast-mapping test phase and generalization test phase were compared within each condition.

For the No Mask Condition, a 2 (objects: target, distractor)  $\times$  2 (test types: fast mapping test, generalization test)  $\times$  3 (words: mi dou, ding ge, niu xing) repeated measures ANOVA was conducted on mean fixation time towards the target and distractor in the test trials. The ANOVA showed significant main effects of objects ( $F(1, 105) = 109.91$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .51$ ), test types ( $F(1, 105) = 15.50$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .13$ ) and words ( $F(2, 105) = 9.79$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .16$ ). The main effects of object and words were qualified by a significant interaction between objects and words, ( $F(2, 105) = 29.50$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .36$ ). No interactions were found for test types.

To obtain more information about the main effect of test types, a paired samples t-test was performed to compare infants’ mean total fixation time (on the objects) between the two types of tests. The analysis showed that fixation time was longer in the fast-mapping tests ( $M = 5.73$ ,  $SD = 2.81$ ) than in the generalization tests ( $M = 4.16$ ,  $SD = 2.28$ ),  $t(71) = 3.85$ ,  $p < .001$ ,  $d = .61$ ,  $r = .29$ , 95 % CI [.76, 2.39]. Past studies indicate that infants’ attention tends to decline after repeated presentations of similar test events (Henderson et al., 2008; Liu & Sun, 2019). Thus the decrease in looking time in the generalization tests could be a result of general decline of attention because the generalization test trials were always presented last in each block of test trials.

To follow-up on the interaction effect between objects and words, paired samples t-tests were performed on fixation time for the target and distractor objects in the fast-mapping tests and also in the generalization tests (see Figure 5). Results revealed that when asked to find the “mi dou”, infants looked significantly longer at the target ( $M = 4.60$ ,  $SD = 2.66$ ) than at the distractor ( $M = 1.59$ ,  $SD = 1.34$ ) in the fast-mapping tests,  $t(35) = 5.75$ ,  $p < .001$ ,  $d = 1.43$ ,  $r = .58$ , 95 % CI [1.95, 4.09]. This pattern was observed in

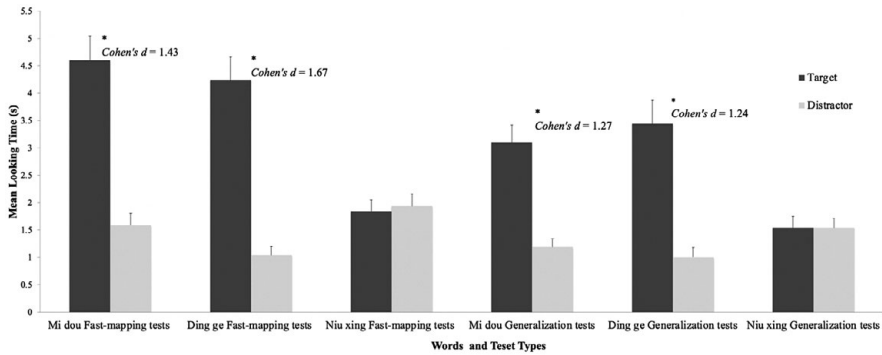


Figure 5. Infants' mean looking time towards the target and distractor in the fast-mapping and generalization tests in the No Mask Condition (\* indicates significant differences,  $p < .05$ ).

86% of the infants ( $n = 31$ ). Infants also looked significantly longer at the target ( $M = 3.10$ ,  $SD = 1.94$ ) than at the distractor ( $M = 1.19$ ,  $SD = .89$ ) in the generalization tests,  $t(35) = 4.81$ ,  $p < .001$ ,  $d = 1.27$ ,  $r = .52$ , 95 % CI [1.48, 2.95]. This was observed in 89% of the infants ( $n = 32$ ).

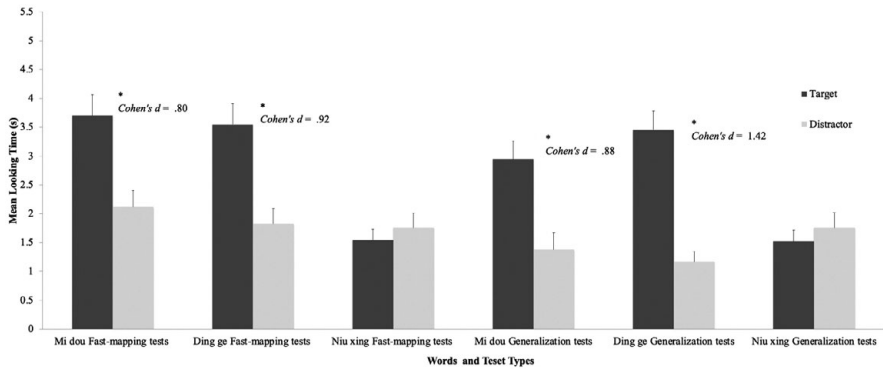
For the word “ding ge”, infants looked significantly longer at the target ( $M = 4.24$ ,  $SD = 2.53$ ) than at the distractor ( $M = 1.04$ ,  $SD = .98$ ) in the fast-mapping tests,  $t(35) = 7.27$ ,  $p < .001$ ,  $d = 1.67$ ,  $r = .64$ , 95 % CI [2.31, 4.11]. This was observed in 92% of the infants ( $n = 33$ ). Infants also looked significantly longer at the target ( $M = 3.44$ ,  $SD = 2.56$ ) than at the distractor ( $M = 1.00$ ,  $SD = 1.07$ ) in the generalization tests,  $t(35) = 4.96$ ,  $p < .001$ ,  $d = 1.24$ ,  $r = .53$ , 95 % CI [1.44, 3.44]. This was also observed in 92% of the infants ( $n = 33$ ).

Infants did not discriminate between the target and distractor upon hearing the word “niu xing” (which was never associated with the objects in the training phase) in both the fast-mapping tests ( $t(35) = -.37$ ,  $p = .358$ ) (47% of the infants ( $n = 17$ ) looked longer at the target) and in the generalization tests ( $t(35) = -.03$ ,  $p = .486$ ) (52% of the infants ( $n = 19$ ) looked longer at the target).

For the With Mask Condition, a  $2 \times 2 \times 3$  repeated measures ANOVA was conducted and showed significant main effects of objects ( $F(1, 105) = 43.51$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .30$ ), test types ( $F(1, 105) = 6.62$ ,  $p = .011$ ,  $\eta^2_{\text{Partial}} = .06$ ) and words ( $F(2, 105) = 11.94$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .19$ ). The main effects of object and words were qualified by a significant interaction between objects and words, ( $F(2, 105) = 16.09$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .24$ ). No interactions were found for test types.

For the main effect of test types, a paired samples t-test revealed that fixation time on the objects was longer in the fast-mapping tests ( $M = 5.63$ ,  $SD = 2.44$ ) than in the generalization tests ( $M = 4.48$ ,  $SD = 2.27$ ),  $t(71) = 2.99$ ,  $p = .002$ ,  $d = .49$ ,  $r = .24$ , 95 % CI [.38, 1.91], again suggesting a general decline of attention in the test phases.

To further examine the interaction effect between objects and words, paired samples t-tests were carried out on fixation time for the objects in the fast-mapping and generalization tests (see Figure 6). Results revealed that in the With Mask Condition, after hearing the word “mi dou”, infants looked significantly longer at the target ( $M = 3.70$ ,  $SD = 2.25$ ) than at the distractor ( $M = 2.12$ ,  $SD = 1.67$ ) in the fast-mapping tests,  $t(35) = 3.01$ ,  $p = .002$ ,  $d = .80$ ,  $r = .37$ , 95 % CI [.52, 2.64]. This preference for the target was observed in 81% of the infants ( $n = 29$ ). Infants also looked significantly longer at the target ( $M = 2.95$ ,  $SD = 1.85$ ) than at the distractor ( $M = 1.38$ ,  $SD = 1.73$ ) in



**Figure 6.** Infants' mean looking time towards the target and distractor in the fast-mapping and generalization tests in the With Mask Condition (\* indicates significant differences,  $p < .05$ ).

the generalization tests,  $t(35) = 3.37, p < .001, d = .88, r = .40, 95\% \text{ CI } [.63, 2.51]$ . This was observed in 89% of the infants ( $n = 32$ ).

For the word “ding ge”, infants looked significantly longer at the target ( $M = 3.55, SD = 2.17$ ) than at the distractor ( $M = 1.83, SD = 1.53$ ) in the fast-mapping tests,  $t(35) = 3.65, p < .001, d = .92, r = .42, 95\% \text{ CI } [.76, 2.67]$ . This pattern was shown in 83% of the infants ( $n = 30$ ). They also looked significantly longer at the target ( $M = 3.45, SD = 2.01$ ) than at the distractor ( $M = 1.17, SD = 1.05$ ) in the generalization tests,  $t(35) = 6.04, p < .001, d = 1.42, r = .58, 95\% \text{ CI } [.38, 1.52]$ . This pattern was manifested in 89% of the infants ( $n = 32$ ).

Infants again did not discriminate between the target and distractor upon hearing the word “niu xing” in both the fast-mapping tests ( $t(35) = -1.00, p = .162$ ) (53% of the infants,  $n = 19$  looked longer at the target) and in the generalization tests ( $t(35) = -.93, p = .181$ ) (50% of the infants,  $n = 18$ ).

### Training trials analyses

Besides examining the test trials, the present study also analysed looking time in the training phase to uncover visual cues that infants relied on to achieve fast-mapping and generalization. Fixation patterns towards the speaker's eyes, mouth and the hand (pointing to the object) were compared between the two conditions. The objects were not included as AOIs because the primary interest of the training trials analyses was to identify visual cues initiated by the speaker that assisted infants to acquire the novel words. However, before analyses, the experimenter performed a prior examination of the eye-tracker recordings to ensure that all infants had at least one fixation point on the objects after the speaker named the object for the first time.

Proportion of looking time to the eyes, mouth and hand in relation to total looking time to the three visual cues were calculated (e.g., proportion of looking time to the eyes = looking time to the eyes / total looking time towards the eyes, mouth and hand) and compared within each condition. Paired-samples  $t$ -tests revealed that in the No Mask Condition, proportion of looking time did not differ between the eyes ( $M = .43, SD = .27$ ) and mouth ( $M = .42, SD = .29$ ),  $t(71) = .21, p = .419$ . However, infants looked significantly longer at the eyes than at the hand ( $M = .15, SD = .17$ ),  $t(71) = 6.54, p < .001, d = 1.24, r = .53, 95\% \text{ CI } [.19, .36]$ . They also looked longer at the mouth than at the hand,  $t(71) = 5.57, p < .001, d = 1.14, r = .49, 95\% \text{ CI } [.17, .35]$  (see Figure 7).

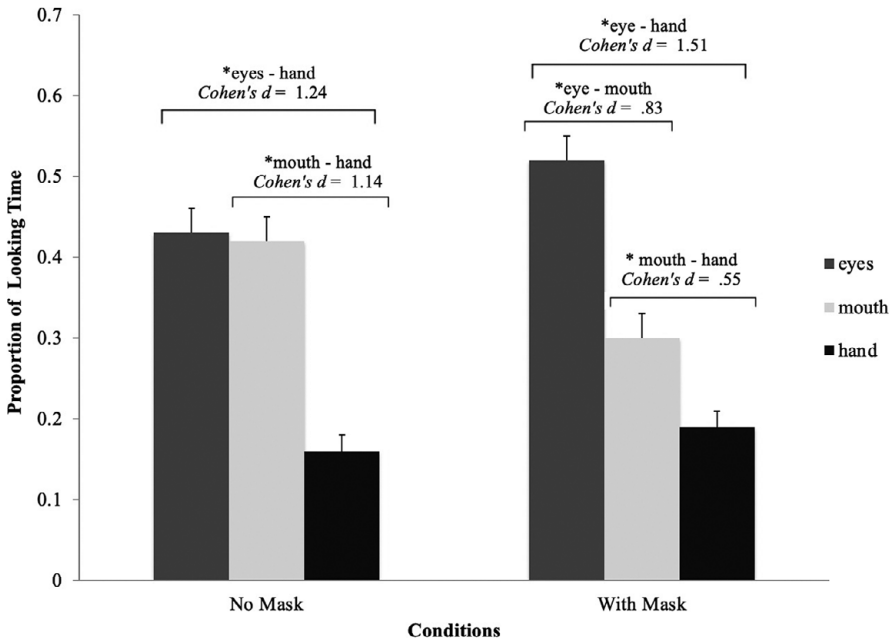


Figure 7. Infants' mean proportion of looking time towards the speaker's eyes, mouth and hand areas in each condition (\* indicates significant differences,  $p < .05$ ).

In the With Mask Condition, proportion of looking time to the eyes ( $M = .52$ ,  $SD = .27$ ) was larger than that of the mouth ( $M = .30$ ,  $SD = .26$ ) ( $t(71) = 3.70$ ,  $p < .001$ ,  $d = .83$ ,  $r = .38$ , 95 % CI [.10, .34] and the hand ( $M = .18$ ,  $SD = .17$ ) ( $t(71) = 7.82$ ,  $p < .001$ ,  $d = 1.51$ ,  $r = .60$ , 95 % CI [.24, .41]. Infants also looked longer at the mouth than they did at the hand,  $t(71) = 2.68$ ,  $p = .005$ ,  $d = .55$ ,  $r = .26$ , 95 % CI [.03, .19].

Additional analyses were then conducted to compare how proportion of looking time for each visual cue differed between the two conditions. A 3 (visual cues: eyes, mouth and hand)  $\times$  2 (conditions: no mask, with mask) mixed-design ANOVA (with visual cues as the within-subject factor) revealed a significant main effect of visual cues,  $F(2, 284) = 37.09$ ,  $p < .001$ ,  $\eta^2_{\text{Partial}} = .21$ . This was qualified by a significant interaction between visual cues and conditions, ( $F(2, 284) = 4.53$ ,  $p = .012$ ,  $\eta^2_{\text{Partial}} = .03$ ).

Follow-up independent samples t-tests suggested that infants allocated greater proportion of looking time towards the speakers' eyes in the With Mask Condition ( $M = .52$ ,  $SD = .27$ ) than in the No Mask Condition ( $M = .43$ ,  $SD = .27$ ),  $t(142) = -1.93$ ,  $p = .028$ ,  $d = .33$ ,  $r = .16$ , 95 % CI [-.18, .00]. In contrast, larger proportion of looking was allocated towards the mouth in the No Mask Condition ( $M = .42$ ,  $SD = .30$ ) than in the With Mask Condition ( $M = .30$ ,  $SD = .26$ ),  $t(142) = 2.56$ ,  $p = .012$ ,  $d = .43$ ,  $r = .21$ , 95 % CI [.03, .21]. There was no significant differences observed for the hand,  $t(142) = -1.10$ ,  $p = .272$ .

## Discussion

Previous studies produced mixed results regarding the impact of masks on basic word learning abilities such as word segmentation and familiar word recognition

(e.g., Frota et al., 2022; Sfakianaki et al., 2021; Singh, 2021). The present study expanded previous literature by exploring the potential impact of masks on fast-mapping and word generalization, which are also precursors of word learning (Houston-Price et al., 2005). The key research question was whether masks would impact infants' fixation on the objects in the fast-mapping and generalization tests. Results demonstrated that masks did not greatly impact fixation patterns: thus fast-mapping and word generalization occurred in both masked and not masked conditions. After being taught two novel words and required to identify the matching objects, 18-month-olds showed a consistent pattern of looking longer at the target object than they did at the distractor. Observation on an individual basis also indicates that the preference for targets upon hearing the words "mi dou" and "ding ge" was manifested in a majority of infants (i.e., more than 80%). Thus, infants in the present study could visually identify the referents of the two novel words after only three exposures of each, constituting an ability to fast-map new words (Houston-Price et al., 2005). In addition, when provided with the word "niu xing", infants did not discriminate between the target and distractor. This further indicates that longer fixation time on the targets for the word "mi dou" and "ding ge" reflected infants' understanding of the word-object associations rather than a case of random selections of referents. Furthermore, 18-month-olds also preferred to fixate more on the target rather than the distractor in the generalization tests (which was also observed in more than 80% of the infants). These findings indicate that after learning a novel word, infants expected different speakers to apply the same word to another object that belonged to the same category (i.e., an object that differed from the previous object only in colour).

Most importantly, the consistent preference for the target was observed in both the No Mask and With Mask Condition. However, the effect sizes of the differences in looking time between the target and distractor were generally larger in the No Mask Condition. The differences in the strength of effect sizes between the Mask and No Mask Condition could indicate that word learning in the With Mask Condition was comparatively less robust. That is, the With Mask Condition could contain less ostensive visual cues which in turn would make learning more difficult for infants. Nonetheless, the effect size values in both conditions were greater than .80 which could be qualified as large effect sizes (Bakker et al., 2019; Borenstein et al., 2009; Cohen, 1992; Whitehead et al., 2016). Singh (2021) found that opaque masks had no effect on familiar word recognition in 24-month-olds. The present findings add to previous literature by demonstrating that fast-mapping and word generalization in younger infants are not totally inhibited by masks.

Previous studies report that selective attention to different face areas, including the mouth, is important to language development (Brooks & Meltzoff, 2005; Lewkowicz, 2010; Weatherhead & White, 2017; Young et al., 2009). Results from the training trials analyses in the present study indicate that out of the three visual cues (i.e., eyes, mouth and hand), infants devoted the least attention to the hand in both conditions. For the other two cues, they allocated equal proportion of looking time towards the eyes and mouth in the No Mask Condition and looked longer at the eyes than the mouth in the With Mask Condition. These findings were further supported by the results that infants looked longer at the eyes in the With Mask than in the No Mask condition. Conversely, they looked longer at the mouth in the No Mask Condition compared to the With Mask Condition. Together, these results indicate that when a new word is presented without a mask, the speaker's mouth is as relevant as the eyes. However, when a speaker's mouth is occluded by a mask, infants tend to adjust their attention to look more at the eyes. The above findings are in concert with previous research (Cruz et al., 2020; Pejovic et al., 2021;

Sekiyama et al., 2021). For example, Frota et al. (2022) found that in response to masks, infants looked more at the eyes, whereas without the mask, they alternated between the eyes and mouth.

When a word is taught through a mask, infants' tendency to look more to the eyes could be explained by the finding that the eyes and mouth can both provide valuable cues that assist word learning and generalization. Gliga and Csibra (2009) found that 13-month-olds who saw a person looking and pointing at a position (called "ostensive cues") and providing a familiar object label (e.g., "a spoon") later expected a matching object to appear at the hidden location, indicating that they are sensitive to a range of visual cues to identify a referent of a word. Past studies also reveal that infants are prepared to interpret ostensively presented information as generalizable to other objects and other people (Csibra & Gergely, 2009). For example, only when provided with cues such as direct eye-contact and greetings, 9-month-olds modify their information encoding strategies to pay more attention to generalizable properties (i.e., the shape) of an object (Yoon et al., 2008). Therefore, 18-month-olds' fixation patterns in the training trials suggest that when taught by a speaker whose mouth is masked, infants can flexibly rely more on visual cues from eye gazes to derive valuable information such as the intention to name objects and shared meanings (Baldwin, 1995; Brooks & Meltzoff, 2008; Corkum & Moore, 1998; Langton et al., 2000; Lewkowicz & Hansen-Tift, 2012). This would in turn allow them to accomplish fast-mapping and word generalization, given that the speaker is making his or her communicative intentions clear (Csibra & Gergely, 2009; Farroni et al., 2002).

In addition, since the proportion of looking time to the hand was the smallest among the three visual cues and did not differ between the two conditions, the present study also suggests that the eyes and mouth might provide more valuable cues than the hand during fast-mapping and word generalization. However, previous research reveals that infants rely on pointing to learn new words (Grassmann & Tomasello, 2010; Paulus & Fikkert, 2014). One possible explanation of the present finding is that the action of pointing was always presented after the eye gazes and utterance of the novel word (which was a design based on previous word learning and generalization studies by Horváth et al., 2016; Liu & Sun, 2019). This could result in the hand receiving the least attention. Therefore, future studies should present the pointing hand in synchrony with the eye gazes and utterance of the word to further explore its importance to word learning.

One limitation of the study is that although the naming procedure was designed to be more natural than previous studies (infants viewed the speaker's naming actions instead of static images of a speaker's head and objects), it was still laboratory based thus lacked ecological validity. In particular, visual cues were prominently available to the infants in the experiment. That is, the speaker provided ostensive communicative cues such as unambiguous eye gazes and pointing to the object. However, such cues may not be always present in real life settings (Gampe et al., 2012). Thus, infants may rely more on the mouth to extract auditory cues, which could result in an adverse effect of masks on word learning. One future step would be to conduct a similar experiment without explicit cues.

The second limitation is that the experimental tasks could be too easy to reveal differences in learning across the No Mask and With Mask Condition. One related problem is that word learning entails both comprehension and production but the study only investigated the former. Young et al. (2009) demonstrated that attention to the mouth contributes to successful language learning since infants could form associations between mouth shapes and speech sounds. Tenenbaum et al. (2015) suggest that attention to the mouth and gaze following behaviours at 12-months correlated with productive vocabulary development at 18- and 24-months. Thus, future studies should investigate if

masks would influence infants' ability to pronounce novel words. Third, the generalization test phase was always presented after the fast-mapping phase. This was mainly because the generalization task posed more challenges for infants. If infants underwent the generalization tests first and found it difficult, they could potentially lose interest in the test trials and not focus on the fast-mapping tests at all. Results also indicate that although infants' fixation times on the objects reduced in the generalization tests, they still discriminated between target and distractor. In addition, since the old speaker who presented the training trials always appeared before the new speaker in the test trials, this could have alerted infants about how to respond for the new speaker. However, a past study that adopted the same design suggests that the order of speaker presentation is very unlikely to affect infants' looking responses in the two types of tests (Liu & Sun, 2019). Last, it is important for future studies to measure infants' vocabulary development and examine whether it is associated with infants' attention to different visual cues in the No Mask and With Mask Conditions.

In summary, the present study reveals that 18-month-olds can fast-map new words and generalize these words across people and objects of the same category when words were taught with or without a mask. The findings also shed light on the importance of visual cues in word learning by suggesting that when a word is taught by a speaker without a mask, infants rely equally on the eyes and mouth. However, when a speaker's mouth is masked, infants can flexibly rely more on the eyes to achieve fast-mapping and word generalization.

**Supplementary material.** The supplementary material for this article can be found at <http://doi.org/10.1017/S0305000923000697>.

**Competing interest.** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Atcherson, S. R., Mendel, L. L., Baltimore, W. J., Patro, C., Lee, S., Pousson, M., & Spann, M. J. (2017). The effect of conventional and transparent surgical masks on speech understanding in individuals with and without hearing loss. *Journal of the American academy of audiology*, *28*(01), 058–067.
- Bakker, A., Cai, J., English, L., Kaiser, G., Mesa, V., & Van Dooren, W. (2019). Beyond small, medium, or large: Points of consideration when interpreting effect sizes. *Educational studies in mathematics*, *102*, 1–8.
- Baldwin, D. A. (1995). Understanding the link between joint attention and language. *Joint attention: Its origins and role in development*, *131*, 158.
- Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1; 2. *Journal of child language*, *32*(1), 159–173.
- Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental science*, *19*(6), 901–917.
- Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. *The handbook of research synthesis and meta-analysis*, *2*, 221–235.
- Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental science*, *8*(6), 535–543.
- Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of child language*, *35*(1), 207–220.
- Buresh, J. S., & Woodward, A. L. (2007). Infants track action goals within and across agents. *Cognition*, *104*, 287–314.
- Butler, J., & Frota, S. (2018). Emerging word segmentation abilities in European Portuguese-learning infants: new evidence for the rhythmic unit and the edge factor. *Journal of child language*, *45*(6), 1294–1308.



- Cao, R., & Sarmah, P. (2007). "A Perception Study on the Third Tone in Mandarin Chinese," *UTA Working Papers in Linguistics* (2007), 2, 50–66.
- Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, 1(3), 98–101.
- Cohn, M., Pycha, A., & Zellou, G. (2021). Intelligibility of face-masked speech depends on speaking style: Comparing casual, clear, and emotional speech. *Cognition*, 210, 104570.
- Corey, R. M., Jones, U., & Singer, A. C. (2020). Acoustic effects of medical, cloth, and transparent face masks on speech signals. *The Journal of the acoustical society of America*, 148(4), 2371–2375.
- Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental psychology*, 34(1), 28.
- Crimon, C., Barbir, M., Hagihara, H., de Araujo, E., Nozawa, S., Shinya, Y., Abboub, N., & Tsuji, S. (2022). Mask wearing in Japanese and French nursery schools: The perceived impact of masks on communication. *Frontiers in psychology*, 13, 874264. <https://doi.org/10.3389/fpsyg.2022.874264>
- Cruz, M., Butler, J., Severino, C., Filipe, M., & Frota, S. (2020). Eyes or mouth? Exploring eye gaze patterns and their relation with early stress perception in European Portuguese. *J. Port. Linguist*, 19, 1–13.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in cognitive sciences*, 13(4), 148–153.
- de Boisferon, A. H., Tift, A. H., Minar, N. J., & Lewkowicz, D. J. (2018). The redeployment of attention to the mouth of a talking face during the second year of life. *Journal of experimental child psychology*, 172, 189–200.
- Desmond, J. E., & Glover, G. H. (2002). Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *Journal of neuroscience methods*, 118(2), 115–128.
- Du, X. (2010). *Pinyin and Chinese Children's Phonological Awareness* (Doctoral dissertation, University of Toronto).
- Eyer, J. A., Leonard, L. B., Bedore, L. M., McGregor, K. K., Anderson, B., & Viescas, R. (2002). Fast mapping of verbs by children with specific language impairment. *Clinical linguistics & phonetics*, 16(1), 59–77.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National academy of sciences*, 99(14), 9602–9605.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175–191.
- Fecher, N., & Watt, D. (2011, August). Speaking under Cover: The Effect of Face-concealing Garments on Spectral Properties of Fricatives. In *ICPhs* (pp. 663–666).
- Feng, S., Shen, C., Xia, N., Song, W., Fan, M., & Cowling, B. J. (2020). Rational use of face masks in the COVID-19 pandemic. *The Lancet respiratory medicine*, 8(5), 434–436.
- Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental psychology*, 43(1), 238.
- Frota, S., Pejovic, J., Cruz, M., Severino, C., & Vigário, M. (2022). Early word segmentation behind the mask. *Frontiers in psychology*, 13, 879123.
- Game, A., Liebal, K., & Tomasello, M. (2012). Eighteen-month-olds learn novel words through overhearing. *First language*, 32(3), 385–397.
- Gluga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological science*, 20(3), 347–353.
- Graham, S. A., & Poulin-Dubois, D. (1999). Infants' reliance on shape to generalize novel labels to animate and inanimate objects. *Journal of child language*, 26(2), 295–320.
- Graham, S. A., Stock, H., & Henderson, A. M. (2006). Nineteen-month-olds' understanding of the conventionality of object labels versus desires. *Infancy*, 9(3), 341–350.
- Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental science*, 13(1), 252–263.
- Greenhalgh, T., Schmid, M. B., Czypionka, T., Bassler, D., & Gruer, L. (2020). Face masks for the public during the covid-19 crisis. *Bmj*, 369.
- Henderson, A. M., Gerson, S., & Woodward, A. L. (2008). The birth of social intelligence. *Zero to three*, 28(5), 13.
- Henderson, A. M., & Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Developmental science*, 15(5), 641–652.
- Hessels, R. S., Andersson, R., Hooge, I. T., Nyström, M., & Kemner, C. (2015). Consequences of eye color, positioning, and head movement for eye-tracking data quality in infant research. *Infancy*, 20(6), 601–633.

- Horváth, K., Liu, S., & Plunkett, K. (2016). A daytime nap facilitates generalization of word meanings in young toddlers. *Sleep*, *39*(1), 203–207.
- Houston-Price, C., Plunkett, K. I. M., & Harris, P. (2005). ‘Word-learning wizardry’ at 1; 6. *Journal of Child Language*, *32*(1), 175–189.
- Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of child language*, *27*(1), 3–42.
- Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences*, *100*(18), 10568–10573.
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*(4577), 1138–1141.
- Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant behavior and development*, *7*(3), 361–381.
- Kwon, M., & Yang, W. (2023). Effects of face masks and acoustical environments on speech recognition by preschool children in an auralised classroom. *Applied acoustics*, *202*, 109149.
- Lalonde, K., Buss, E., Miller, M. K., & Leibold, L. J. (2022). Face masks impact auditory and audiovisual consonant recognition in children with and without hearing loss. *Frontiers in Psychology*, *13*, 874345.
- Lalonde, K., & Werner, L. A. (2019). Infants and adults use visual cues to improve detection and discrimination of speech in noise. *Journal of speech, language, and hearing research*, *62*(10), 3860–3875.
- Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in cognitive sciences*, *4*(2), 50–59.
- Lewkowicz, D. J. (2010). Infant perception of audio-visual speech synchrony. *Developmental psychology*, *46*(1), 66.
- Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child development*, *85*(2), 685–694.
- Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, *109*(5), 1431–1436.
- Liu, S., & Sun, R. (2019). Appreciating language conventions: thirteen-month-old Chinese infants understand that word generalization is shared practice. *Journal of child language*, *46*(4), 812–823.
- Magee, M., Lewis, C., Noffs, G., Reece, H., Chan, J. C. S., Zaga, C. J., Paynter, C., Birchall, O., Rojas Azocar, S., Ediriweera, A., Kenyon, K., Caverlé, M. W., Schultz, B. G., & Vogel, A. P. (2020). Effects of face masks on acoustic analysis and speech perception: Implications for peri-pandemic protocols. *The Journal of the Acoustical Society of America*, *148*(6), 3562. <https://doi.org/10.1121/10.0002873>
- Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental science*, *11*(1), 53–59.
- Mendel, L. L., Gardino, J. A., & Atcherson, S. R. (2008). Speech understanding using surgical masks: a problem in health care?. *Journal of the American academy of audiology*, *19*(09), 686–695.
- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility: Head movement improves auditory speech perception. *Psychological science*, *15*(2), 133–137.
- Nathani, S., Ertmer, D. J., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical linguistics & phonetics*, *20*(5), 351–369.
- Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469.
- Palmiero, A. J., Symons, D., Morgan, J. W., & Shaffer, R. E. (2016). Speech intelligibility assessment of protective facemasks and air-purifying respirators. *Journal of occupational and environmental hygiene*, *13*(12), 960–968.
- Paulus, M., & Fikkert, P. (2014). Conflicting social cues: Fourteen- and 24-month-old infants’ reliance on gaze and pointing cues in word learning. *Journal of cognition and development*, *15*(1), 43–59.
- Pejovic, J., Yee, E., & Molnar, M. (2021). Eyes can tell: Attention to the eyes and the mouth during audiovisual vowel processing in monolingual and bilingual infants. *PsyArXiv (Preprint)*. doi: 10.31234/osf.io/pytua
- Saeidi, R., Huhtakallio, I., & Alku, P. (2016, September). Analysis of Face Mask Effect on Speaker Recognition. In *Interspeech* (pp. 1800–1804).

- Saigusa, J. (2017). The effects of forensically relevant face coverings on the acoustic properties of fricatives. *Lifespans and styles*, 3(2), 40–52.
- Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child development*, 69(2), 309–320.
- Schwarz, J., Li, K. K., Sim, J. H., Zhang, Y., Buchanan-Worster, E., Post, B., Gibson, J. L., & McDougall, K. (2022). Semantic Cues Modulate Children's and Adults' Processing of Audio-Visual Face Mask Speech. *Frontiers in psychology*, 13, 879156. <https://doi.org/10.3389/fpsyg.2022.879156>
- Sekiyama, K., Hisanaga, S., & Mugitani, R. (2021). Selective attention to the mouth of a talker in Japanese-learning infants and toddlers: Its relationship with vocabulary and compensation for noise. *Cortex*, 140, 145–156.
- Sfakianaki, A., Kafentzis, G. P., Kiagiadaki, D., & Vlahavas, G. (2021, October). Effect of face mask and noise on word recognition by children and adults. In *Proceedings of 12th International Conference of Experimental Linguistics* (pp. 207–210). Athens: ExLing Society.
- Singh, L., Goh, H. H., & Wewalaarachchi, T. D. (2015). Spoken word recognition in early childhood: Comparative effects of vowel, consonant and lexical tone variation. *Cognition*, 142, 1–11.
- Singh, L., Tan, A., & Quinn, P. C. (2021). Infants recognize words spoken through opaque masks but not through clear masks. *Developmental science*, 24(6), e13117.
- Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of experimental child psychology*, 109(1), 132–140.
- Swingle, D. (2021). Infants' Learning of Speech Sounds and Word Forms. In *Oxford Handbook of the Mental Lexicon*, ed LR Gleitman, A Papafraou, JC Trueswell. Oxford: Oxford University Press
- Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, 108(3), 850–855.
- Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of child language*, 42(6), 1173–1190.
- Tobii, A. B. (2023). Tobii Pro Lab User Manual (Version 1.217). Tobii AB, Danderyd, Sweden.
- Tronick, E., & Snidman, N. (2021). Children's reaction to mothers wearing or not wearing a mask during face-to-face interactions. Available at SSRN 3899140.
- Weatherhead, D., & White, K. S. (2017). Read my lips: Visual speech influences word processing in infants. *Cognition*, 160, 103–109.
- Whitehead, A. L., Julious, S. A., Cooper, C. L., & Campbell, M. J. (2016). Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. *Statistical methods in medical research*, 25(3), 1057–1073.
- Wilkinson, K. M., & Mazzitelli, K. (2003). The effect of 'missing' information on children's retention of fast-mapped labels. *Journal of child language*, 30(1), 47–73.
- Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive psychology*, 49(2), 155–190.
- Yeung, W., Ng, K., Fong, J. N., Sng, J., Tai, B. C., & Chia, S. E. (2020). Assessment of proficiency of N95 mask donning among the general public in Singapore. *JAMA network open*, 3(5), e209670–e209670.
- Yoon, J. M., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences*, 105(36), 13690–13695.
- Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental science*, 12(5), 798–814.

Cite this article: Liu, S., Li, X., & Sun, R. (2024). The effect of masks on infants' ability to fast-map and generalize new words. *Journal of Child Language* 51, 637–655, <https://doi.org/10.1017/S0305000923000697>