



PAPER

# Linking language to sensory experience: Onomatopoeia in early language development

Yasamin Motamed<sup>1,2</sup> | Margherita Murgiano<sup>1</sup> | Pamela Perniss<sup>3</sup> |  
Elizabeth Wonnacott<sup>1,4</sup> | Chloë Marshall<sup>5</sup> | Susan Goldin-Meadow<sup>6</sup> |  
Gabriella Vigliocco<sup>1</sup>

<sup>1</sup>Division of Psychology and Language Sciences, University College London, London, UK

<sup>2</sup>Centre for Language Evolution, The University of Edinburgh, Edinburgh, UK

<sup>3</sup>University of Cologne, Köln, Germany

<sup>4</sup>Department of Education, University of Oxford, Oxford, UK

<sup>5</sup>Institute of Education, University College London, London, UK

<sup>6</sup>Department of Psychology, University of Chicago, Chicago, IL, USA

## Correspondence

Yasamin Motamed, Centre for Language Evolution, The University of Edinburgh, Dugald Stewart Building, 3 Charles Street, Edinburgh, EH8 9AD, UK.

Email: yasamin.motamed@ed.ac.uk

## Funding information

Economic and Social Research Council, Grant/Award Number: ES/P00024X/1; Foundation for the National Institutes of Health, Grant/Award Number: NICHD P01 HD40605; Royal Society, Grant/Award Number: WRM\R3\170016

## Abstract

A key question in developmental research concerns how children learn associations between words and meanings in their early language development. Given a vast array of possible referents, how does the child know what a word refers to? We contend that onomatopoeia (e.g. *knock*, *meow*), where a word's sound evokes the sound properties associated with its meaning, are particularly useful in children's early vocabulary development, offering a link between word and sensory experience not present in arbitrary forms. We suggest that, because onomatopoeia evoke imagery of the referent, children can draw from sensory experience to easily link onomatopoeic words to meaning, both when the referent is present as well as when it is absent. We use two sources of data: naturalistic observations of English-speaking caregiver-child interactions from 14 up to 54 months, to establish whether these words are present early in caregivers' speech to children, and experimental data to test whether English-speaking children can learn from onomatopoeia when it is present. Our results demonstrate that onomatopoeia: (a) are most prevalent in early child-directed language and in children's early productions, (b) are learnt more easily by children compared with non-iconic forms and (c) are used by caregivers in contexts where they can support communication and facilitate word learning.

## KEY WORDS

child-directed language, onomatopoeia, sound symbolism, word learning

## 1 | INTRODUCTION

How children learn words continues to be an unanswered question in the study of human development. A large body of work has provided insight into a number of key strategies that can support word learning, including children's sensitivity to temporal co-occurrence between labels and objects (Gleitman et al., 2005; Samuelson et al., 2011; Smith & Chen, 2008; Yu & Smith, 2011), to language-internal regularities (Mattys & Jusczyk, 2001; Pelucchi et al., 2009; Saffran et al., 1996) and powerful cues in

caregivers' communication such as their use of pointing (Goldin-Meadow, 2007; Iverson et al., 1994, 1999), eye gaze (Brand et al., 2007; Senju & Csibra, 2008) and prosodic speech modulations (Fernald & Simon, 1984; Herold et al., 2012; Nygaard et al., 2009; Vosoughi et al., 2010) to direct attention and convey meaning. However, these strategies, even taken together, cannot provide a full account of word learning. First and foremost, the majority of existing accounts assume that temporal overlap between referents and words is required for learning, but this is far from the norm. Language very often refers to objects and actions that are



not physically present, and indeed, *displacement* has long been considered a fundamental feature of human language (Hockett, 1960), one that occurs even in language directed to children (Tomasello & Kruger, 1992; Veneziano, 2001). Moreover, while statistical regularities in language input can provide reliable cues to word boundaries (Saffran et al., 1996), they cannot tell the child anything about what the word refers to, an issue known as the "symbol grounding problem" (Harnad, 1990; Searle, 1980).

Crucially, most existing proposals assume that the relationship between words and referents is purely arbitrary (de Saussure, 1916). However, this is not always the case. Onomatopoeia are non-arbitrary words whose form directly evokes the sound of real-world referents (e.g. *smack*, *drip*), and are well-represented across the world's languages. However, linguistic study has historically considered them as a marginal or trivial property of human languages, unworthy of scientific scrutiny (Newmeyer, 1992). For example, although onomatopoeia and sound-imitation featured centrally in historic theories of language origin, these theories have been heavily criticised (Muller, 1861; Thorndike, 1943), on the grounds that onomatopoeia differ across languages despite referring to the same real-world sounds (e.g. *cock-a-doodle-doo* in English and *chicchirichi* in Italian), and onomatopoeic forms being limited in the number of possible concepts they can express (those referring to sound).

Yet recent research suggests that onomatopoeia are not as marginal as previously thought. Onomatopoeia are among the most common words young children produce (Laing, 2014; Tardif et al., 2008), occur frequently in the input to children (Fernald & Morikawa, 1993) and may be further highlighted by the caregiver. For example, Laing et al. (2017) found that onomatopoeic words in child-directed speech were made more salient—e.g. vocalised at higher pitch and with longer duration—than their conventional counterparts. Onomatopoeic forms have also demonstrated a processing advantage (Laing, 2017), with 10- and 11-month-old infants in a picture mapping task fixating longer on pictures in response to onomatopoeic words compared to conventional word forms (e.g. 'baa' vs. 'sheep').

Here, we propose that onomatopoeia—by virtue of being imitative of real-world sounds—can support initial word acquisition in two key ways. Firstly, the imitative nature of these words allows infants to discover that the speech sounds they hear can be *meaningful*, referring to objects and actions in the world (Imai & Kita, 2014). Thus, onomatopoeia may focus infants' attention to speech as a source of meaningful information, and help tune the system to map these sounds to meanings. Once this is established, infants will be able to process and learn word meanings even when the words are not onomatopoeic. Secondly, onomatopoeia allow the child to begin building a lexical repertoire based in auditory sensory experience. By imitating sound properties of a specific referent (e.g. *meow* refers to a cat; *woof-woof* to a dog), onomatopoeia offer the first mappings between word and referent. Crucially, this is true both when referent and label co-occur, but also in cases where they do not (Perniss et al., 2018). When referents are present, they can single out the

## Research Highlights

- We combine observations of naturalistic caregiver-child interactions with a word learning experiment to investigate onomatopoeia in early word learning
- Onomatopoeia are more frequent in the vocabularies of younger children and more common in child-directed language at younger ages
- Onomatopoeia in child-directed language are more frequent when they can support learning: that is, when the referent is unfamiliar, or when label and referent are displaced
- Children learn onomatopoeia more easily than arbitrary words, even in contexts where label and referent are displaced

correct referent from a potentially vast array of possible referents. When referents are absent from the visual scene (as is the case for displacement), they can help to map the word to a memory trace of the referent.

We test three key predictions stemming from these proposals. Our first prediction is that, if onomatopoeia provide a mechanism that tunes infant's attention to speech as the source of meaningful information, caregivers should use more onomatopoeia early on, and their production of onomatopoeia should reduce substantially as the child's vocabulary expands beyond concepts with strong auditory properties that can be easily expressed using onomatopoeia. In parallel, children's production of onomatopoeia should follow a similar trajectory, reducing as their non-imitative vocabulary expands. There is already suggestive evidence in the literature compatible with this prediction (Jo & Ko, 2018; Laing, 2014; Perry et al., 2017). However, Laing (2014) shows evidence for this predicted trajectory for onomatopoeia in child productions only in a single child. Other studies offer evidence of the predicted trajectory for child-directed speech (Jo & Ko, 2018; Perry et al., 2017) but do not consider onomatopoeia separately from other iconic or sound-symbolic forms. Here, we look specifically at onomatopoeia, considering them as special, and particularly useful in early word learning. The imitative nature of onomatopoeia provides the child with a direct sound-to-sound mapping. In contrast, other forms of iconicity use less transparent mappings that may require an understanding of cross-modal mappings (i.e. that rounded vs. unrounded vowels can map onto size properties of the referent) that entail further abstraction.

While our first prediction concerns production of onomatopoeia by the caregiver and the child, our second prediction concerns the role of onomatopoeia in learning. We predict that children will learn words more easily when the mapping is onomatopoeic, compared to non-onomatopoeic words, because onomatopoeic words provide an imitative link between the sound of the word and its meaning. Our third prediction is that onomatopoeia should support learning



both when referents are present and visually accessible, as well as when they are absent (i.e. in displaced language use) because of their ability to imaginatively evoke properties of referents. Thus, we expect to see that (a) onomatopoeia are produced by caregivers more often in displaced contexts where reference cannot be supported by other cues such as pointing, directed eye gaze or shared actions; and (b) they have a greater role in learning in displaced contexts than in here-and-now contexts.

We test these predictions in two studies. In study 1, we analyse child-directed and child-produced onomatopoeia from a naturalistic longitudinal corpus, asking whether onomatopoeia production changes over development. We further investigate whether onomatopoeia is used more frequently by caregivers when most useful to the child—that is, when the caregiver produces a word unfamiliar to the child, and in displaced learning contexts, where the referent is not perceptually accessible. Showing that caregivers use onomatopoeia, however, is not the same as showing that children use onomatopoeia for learning. Therefore, in study 2, we investigate the link between onomatopoeia and word learning across both situated and displaced learning contexts, testing whether 2- to 3-year-olds learn onomatopoeic forms more easily than non-omatopoeic ones. Although previous research suggests a learning advantage for sound-symbolic forms, no study to date has focussed on onomatopoeia as a special case, which links form and meaning through a direct, unimodal association. Children in study 2 were taught novel labels for a set of events—those labels could be either onomatopoeic with respect to the event, or have no onomatopoeic relationship to the event. Furthermore, children learnt the novel labels in either a situated or a displaced learning context, before being tested in an immediate recall stage. With these two studies, we can provide evidence of how onomatopoeia are used in natural child-directed language and children's productions, and test specific predictions about onomatopoeia's facilitatory role in word learning in the experimental task.

## 2 | STUDY 1: ONOMATOPOEIA IN CHILD-DIRECTED LANGUAGE

In study 1, we analysed data from the Language Development Project (LDP) corpus, a longitudinal corpus of naturalistic interactions between children and caregivers, in which the language spoken is English (Goldin-Meadow et al., 2014). In the corpus, 64 typically developing children and their caregivers were videotaped for 90 min at 4-month intervals between 14 and 58 months, engaging in normal everyday activities at home. We randomly selected a sample of 40 participating families from the corpus to analyse here. Children in our sample were typically developing native English speakers from Chicago. Caregiver and child speech have previously been transcribed and tagged for part of speech with CHILDES part-of-speech tagging program (MacWhinney, 2000); we identified onomatopoeia based on these tags. Full

details of analyses here and throughout can be found at <https://osf.io/kvtv8a/>.

### 2.1 | Child-directed onomatopoeia and vocabulary development

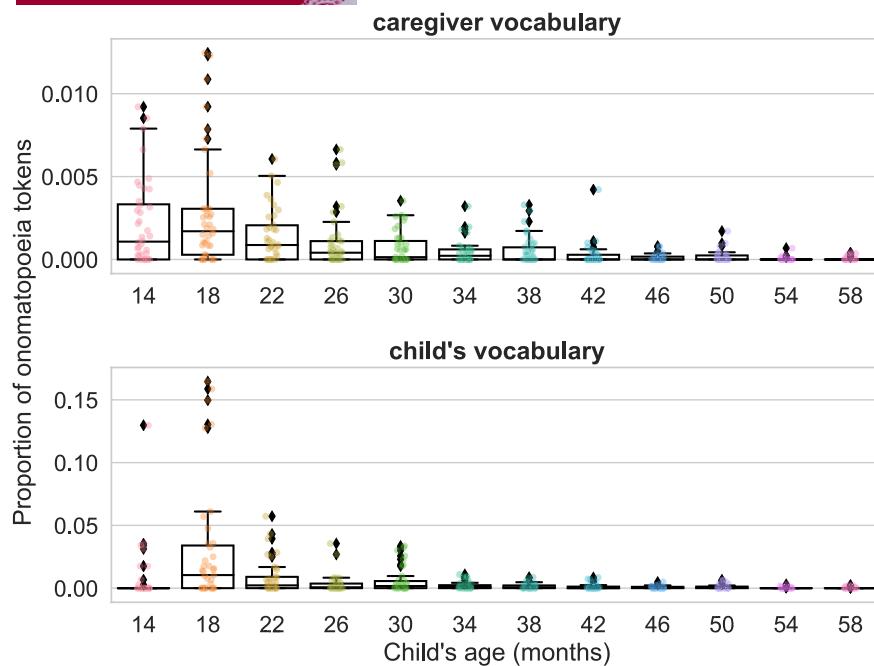
We predicted that onomatopoeia in child-directed language and children's productions should be most prevalent in early stages of lexical development. To assess this prediction, we analysed onomatopoeia use by caregivers and children in our sample, in relation to children's vocabulary development. For each caregiver and child in the sample, we obtained all transcribed lemmas in the corpus, and an index of socio-economic status (SES), calculated by transforming caregiver education and income variables to z-scores and then taking the mean value of z-scores for each family.

#### 2.1.1 | Results

We first investigated the relationship between onomatopoeia use and child's age (Figure 1), analysing the proportion of the total vocabulary (tokens) produced at each child age point in the corpus that contained onomatopoeia (as identified by the part-of-speech tags), for caregivers' and children's vocabularies. We removed data points more than 3 standard deviations from the mean value for a given age point. Both caregivers and children show considerable individual variation in their use of onomatopoeia, both within and across ages, as illustrated by the points in Figure 1 (note that some participants do not produce any onomatopoeia at a given time point).

To analyse this relationship, we performed a logistic mixed-effects model regression with R (R Core Team, 2013) and lme4 (Bates et al., 2015) for each speaker group (caregiver, child), using the proportion of onomatopoeia tokens as the outcome variable, weighted by the total number of tokens. Child's age was included as a centred predictor of interest, with the family SES index included as a centred control variable. In addition, we included a by-participant random intercept with a random slope term for age (centred).

Analysis of the caregiver results revealed a significant effect of child's age ( $\beta = -0.07$ ,  $SE = 0.008$ ,  $z = -10.34$ ,  $p < 0.001$ ). Caregiver use of onomatopoeic forms is higher for younger children, and decreases with the age of the child. The model did not demonstrate a significant relationship between SES and the proportion of onomatopoeia ( $\beta = 0.18$ ,  $SE = 0.11$ ,  $z = 1.57$ ,  $p = 0.12$ ). Analysis of children's speech revealed similar results: a significant negative association between onomatopoeia production and age ( $\beta = -0.10$ ,  $SE = 0.009$ ,  $t = -11.27$ ,  $p < 0.001$ ) and no evidence for a significant relationship between children's onomatopoeia production and SES ( $\beta = 0.09$ ,  $SE = 0.11$ ,  $t = 0.81$ ,  $p = 0.42$ ). Children's onomatopoeia usage also peaks early in vocabulary development, at approximately 18 months of age.



**FIGURE 1** Boxplots showing the proportion of onomatopoeia tokens for caregivers (top) and children (bottom) for each age point in the LDP corpus. Note that caregivers and children did not necessarily produce any onomatopoeia in a given session. Please note different y-axis scales for caregivers and children

## 2.2 | Contextual differences in child-directed onomatopoeia

We hypothesised that onomatopoeia support word learning by allowing children to access sensory properties of real-world referents, and that this characteristic of onomatopoeia is useful in displaced learning contexts, where referent and label do not co-occur. To test this hypothesis, we re-coded data from our 40 participants to note occurrences of onomatopoeia and the context—situated or displaced—in which they occurred.

### 2.2.1 | Lemma sampling and coding

We sampled data at 3 time points (18, 26 and 34 months), a period which sees extensive vocabulary growth. Given the time constraints of the project, we sampled sections from each participant's data at each session, which were then coded for onomatopoeia and context. Sampling of these sections was based on words from the caregivers' general vocabulary as transcribed from that session (i.e. not onomatopoeia), and were selected in the following way. First, we extracted all caregiver lemmas tagged as nouns and verbs and sampled 30 nouns and verbs for each participant and session, half of which in each group were known to the child and half of which were unknown, or new lemmas. We categorised each lemma according to the following procedure: categorise as known (a) any lemma found in the child's transcribed productions, in the current session or any previous sessions, (b) any lemma marked as known on the MacArthur Bates Communicative Development Inventory (Fenson et al., 1994)

for that session or any previous sessions, (c) any lemma with an average age of acquisition <4, according to Kuperman et al. (2012). Lemmas that did not match these criteria were marked as *unknown*. We resampled lemmas for the same participant and session until we had 15 known and 15 unknown lemmas for both nouns and verbs.

We then extracted a set of utterances in each session based on these selected lemmas. For each lemma, we took the first 5 utterances from that session containing tokens of the lemma, and took as our window of analysis 5 utterances before and after the lemma, giving a topic window of 11 utterances. Timestamps from the transcripts were used to locate utterances in the audio-visual data for coding. These topic windows comprised the data we annotated and analysed.

Data were coded by three coders, two of whom were blind to the hypotheses of the study. Utterances within each topic window were coded for onomatopoeia and context (situated/displaced) in relation to the target lemma. Onomatopoeia were coded as lexicalised (e.g. "oink") or as sound effects made with the mouth that replicated real-world sounds (e.g. making a snorting noise). Approximately 82% of onomatopoeia were coded as lexicalised. We assessed coding reliability for all unique lexicalised onomatopoeia in the sample ( $N = 118$ ) based on (a) entries in the Oxford English Dictionary (OED), (b) iconicity ratings from adult native English speakers from two studies (Lu, 2018; Perry et al., 2015), in which onomatopoeia were a subset of a larger set of words and (c) iconicity ratings from members of the research team (all native English speakers). Lemmas identified as onomatopoeic, imitative or echoic in the OED were accepted as iconic. For the remaining lemmas, we accepted our classification as iconic if it was rated as such by participants in at least one of the two rating studies (with ratings equal to or greater than the mid-point

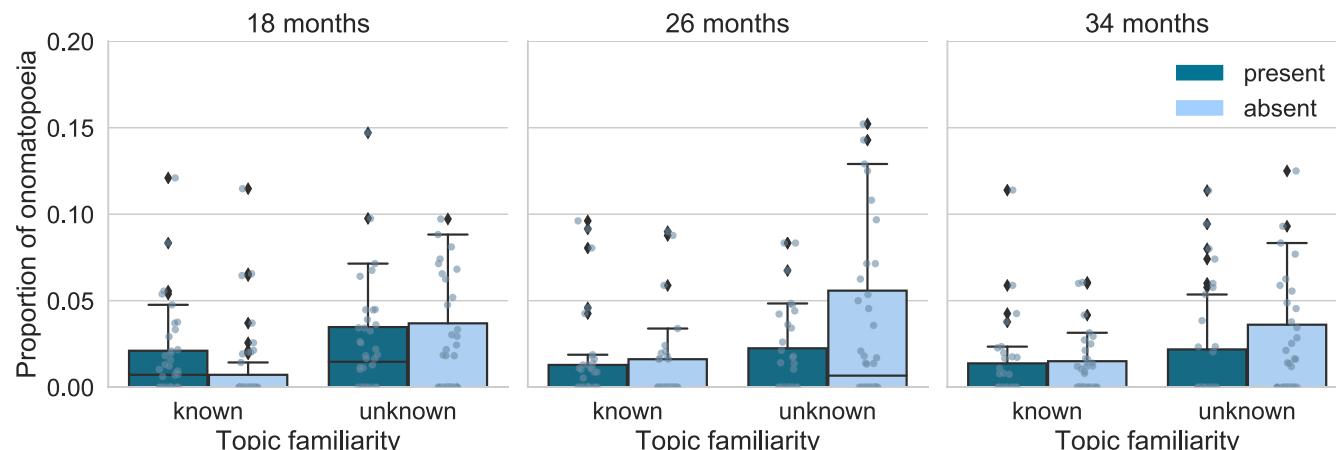
of the arbitrary to iconic scale). This procedure accounted for 91 of the forms coded as lexicalised onomatopoeia in our sample. For the remaining 27, two forms were rated below the mid-point of the scale in both studies and were excluded from analysis. A further 14 forms showed ratings below the scale mid-point in only one of the two studies, and the remaining 11 were not included in the stimuli for either study. We obtained iconicity ratings for the remaining 25 forms from research team members, which confirmed 3 additional forms as iconic. As such, we only included in analysis the 94 forms of lexicalised onomatopoeia for which we have evidence of iconicity.

Each utterance in the sample was coded for context. Broadly, we operationalise situated and displaced contexts in relation to the presence or absence of the referent from the visual scene. For nouns, this involved coding whether the referent was present or absent in the visual scene. For example, for the lemma *crayon*, we coded each topic-related utterance as present if the crayon was present in the audio-visual data, and absent if it was not.

For verbs, we coded the context as either present or absent according to how the event and utterance temporally co-occurred, following Tomasello and Kruger (1992). If the action occurred within the time frame of the topic-related utterance, the action was coded as *present* with respect to the label. For example, if the caregiver said, “are you bouncing it?” while the child plays with a ball, the action was annotated as present. If instead actions occurred before or after the topic-related utterance, they were defined as *absent*, as in the following examples: the caregiver said “You were bouncing your ball” after the child had finished playing, the caregiver said “Shall we bounce the ball” and then initiated play with the ball, or the caregiver said “Shall we bounce the ball” but then the final action was unrealised (e.g. the child decided to do something else).

## 2.2.2 | Results

Figure 2 shows the proportion of caregiver utterances that contain onomatopoeia, at each age point, for known and unknown target lemmas and for present and absent referents.



**FIGURE 2** Boxplots showing the proportion of onomatopoeia produced by caregivers. Results are shown at each age group, for present referents (dark blue) and absent referents (light blue) and for known and unknown topics

We conducted a logistic mixed-effects regression, analysing the effect of displacement (situated/displaced), topic familiarity (known/unknown based on the sampled lemmas) and age of the child on caregivers’ onomatopoeia production. We included displacement, familiarity and age of the child as centred fixed predictors, as well as the three-way interaction term and SES index as an additional centred fixed effect. We included a by-subject random intercept, with a random slope of displacement, familiarity and their interaction term. The model did not converge when age was included in the random slope structure. The model was fit using bound optimisation by quadratic approximation (bobyqa), which allowed convergence.

The model revealed a significant main effect of age ( $\beta = -0.02$ ,  $SE = 0.007$ ,  $z = -2.64$ ,  $p = 0.008$ ), consistent with the results found in Section 2.1, and a significant main effect of SES ( $\beta = 0.61$ ,  $SE = 0.15$ ,  $z = 4.10$ ,  $p < 0.001$ ). Overall, caregivers used fewer onomatopoeic forms as the child’s age increased, and higher SES was associated with proportionally higher use of onomatopoeia. Additionally, we found a main effect of topic familiarity ( $\beta = 0.79$ ,  $SE = 0.20$ ,  $z = 3.92$ ,  $p < 0.001$ ), with onomatopoeia occurring more frequently when the label was unknown to the child. Finally, although we did not find a main effect of displacement ( $\beta = 0.13$ ,  $SE = 0.18$ ,  $z = 0.74$ ,  $p = 0.46$ ), we found a significant interaction between displacement and child’s age ( $\beta = -0.03$ ,  $SE = 0.01$ ,  $z = -2.29$ ,  $p = 0.02$ ). Further breakdown of this interaction indicated onomatopoeia decreased less over age points for absent than for present referents; higher levels of onomatopoeia were therefore maintained in contexts where the topic referent was not visually present.

## 2.3 | Interim summary

In Study 1, we analysed caregiver and child productions of onomatopoeia in relation to children’s age, and in relation to the learning context: whether the topic lemma was known or unknown to the child, and whether the referent was visually accessible or not. In line with our first prediction, we found that caregivers produce a higher



proportion of onomatopoeia when interacting with younger children and younger children themselves produce more onomatopoeia than older children. Furthermore, we found that onomatopoeia were used more frequently in learning episodes, where the utterance topic is unknown to the child. Finally, although we did not find an overall effect of displacement, in line with our third prediction, we found an interaction between displacement and age, such that the reduction in onomatopoeic forms was less extreme for absent referents than present ones—i.e. use of onomatopoeia is maintained at a higher proportion across development in cases where the referent is not perceptually accessible. In this way, caregivers' language shows some sensitivity to the context in which learning takes place, suggesting a potential relationship between onomatopoeia in child-directed language and children's vocabulary development. In Study 2, we present an experimental study testing this prediction: do children learn onomatopoeia more easily than arbitrary words, and do onomatopoeia further facilitate word learning in displaced language contexts?

### 3 | STUDY 2: EXPERIMENTAL EVIDENCE FOR THE ROLE OF ONOMATOPOEIA IN WORD LEARNING

In Study 2, we ran an experimental investigation into whether onomatopoeia facilitate learning of novel words for events in 2- to 3-year-old English-speaking children, an age range where vocabulary

develops considerably, and verb learning in particular takes off. Participants were first trained on novel labels that were designed to be either onomatopoeic with respect to the meaning, or to have no relationship between word and meaning. Furthermore, labels and meanings were presented in two learning contexts, situated and displaced, to test whether onomatopoeia convey a particular advantage in displaced learning contexts.

#### 3.1 | Methods

##### 3.1.1 | Participants

A total of 44 typically developing children (aged 24–36 months) were recruited from nurseries around London, UK. Data from 7 children were excluded because the children were out of the age range on the day of testing (3 children), or did not complete the second of the two sessions (4 children), leaving a total of 37 children. All children in the study spoke English as their first language<sup>1</sup>.

##### 3.1.2 | Materials

Events for the experiment were devised as a set of actions associated with environmental sounds (e.g. splashing water, sneezing). Individual videos depicted each event separately (mean video duration 2,524 ms, range 2,090–3,020 ms), and videos were presented

Event name	Event description	Ono. label	Control label
Blow	Blowing air out of the mouth	Foofing	Yecking
Break	Breaking a toy egg against the table surface	Boshing	Ploffing
Chew	Chewing a sticky substance	Kurching	Wubwubbing
Clap	Clapping the hands together	Dapping	Gaffgaffing
Cough	Coughing	Yecking	Glipping
Cry	Crying	Wubwubbing	Jatting
Drink	Drinking water from a glass	Slooping	Boshing
Drop	Dropping a set of keys onto a table	Glinging	Prapping
Fall	Falling heavily backwards onto a chair	Ploffing	Dapping
Hit	Actor slapping herself on the arm	Jatting	Foofing
Knock	Knocking on the table	Konking	Glinging
Laugh	Laughing	Gaffgaffing	Slooping
Sneeze	Sneezing	Kooshing	Trunding
Splash	Hands splashing against water in a container	Glipping	Kurching
Squirt	Squirting water from the mouth	Prapping	Konking
Stomp	Stomping heavily across the floor	Trunding	Kooshing

TABLE 1 Event descriptions and corresponding onomatopoeic and control labels used in the experiment



with sound both in the studies described in this section and in the main experiment.

We developed an initial set of onomatopoeic labels for each event that conformed to English phonotactic constraints, and evoked the sounds in the event video. Label-video pairs were rated by between 10 and 12 raters for strength of onomatopoeic relationship on a 1–7 Likert scale (1: not iconic at all, 7: highly iconic) in two norming studies with English-speaking adults. In a first norming study, onomatopoeic labels were rated alongside either (a) randomly generated labels or (b) our devised onomatopoeic labels paired with events for which they are not iconic. Overall, the onomatopoeic labels were rated as more iconic than the non-iconic labels ( $t = 16.77, p < 0.001$ ; onomatopoeia:  $M = 3.51, SD = 0.99$ ; non-omatopoeic labels:  $M = 2.35, SD = 1.00$ ). We found no significant difference in ratings between the onomatopoeic words paired with events for which they are not iconic and the randomly generated labels ( $t = 0.73, p = 0.47$ ; re-used onomatopoeia:  $M = 2.43, SD = 1.18$ ; randomly generated labels:  $M = 2.27, SD = 0.47$ ). Therefore, we decided to re-use the onomatopoeic labels as control labels (paired with events for which they are not onomatopoeic), allowing form-internal consistency across both label sets (henceforth, we refer to these labels as non-omatopoeic controls). A second norming study confirmed that the onomatopoeic labels were rated higher than the control labels ( $t = 5.92, p < 0.001$ ; onomatopoeia:  $M = 4.84, SD = 1.16$ ; control:  $M = 2.37, SD = 1.60$ ), and no control event-label pair had a higher average rating than its corresponding onomatopoeic event-label pair. We selected a final set of 16 events (shown in Table 1), paired with unique onomatopoeic labels and unique control labels (i.e. no label was used more than once as a control label).

For our final 16 labels, we conducted an additional task to assess phonological similarity to existing English words for the same events, in order to exclude the possibility that performance in the main experiment could be driven by knowledge of existing words. Fourteen participants took part in an online task in which they were shown event videos and asked to note up to six ‘action’ words (examples were given in the instructions) to describe each event. We collated responses from this task to build a set of typical English words that describe our events. As some of the responses were unlikely to be known by the children in the study (e.g. ‘masticate’ for the chewing event), we removed words with an age of acquisition above 6 years of age (based on age of acquisition norms; Kuperman et al., 2012), leaving a total of 78 unique English words (see the OSF link provided below for full description of our inclusion criteria). For each of these words, we calculated the phonological distance between the corresponding onomatopoeic and control labels for the relevant event, using a feature edit distance weighted by phonological class from the PanPhon Python package (Mortensen et al., 2016). Using a related samples t-test, we analysed whether average phonological distance between the English word and the novel label was higher for our control labels compared with the onomatopoeic labels. Our analysis did not indicate a significant difference in phonological distance to existing English words across these two sets of labels ( $t = -0.46, p = 0.65$ ; onomatopoeia:  $M = 10.43, SD = 6.14$ , range = 1.19–30.75; control:  $M = 10.71, SD = 5.95$ , range = 1.19–31.94).

All stimuli as well as full descriptions of the iconicity norming and phonological distance analysis can be found at <https://osf.io/ktv8a/>.

### 3.1.3 | Design

For each child, the experiment consisted of two sessions, one using the situated, and one the displaced context. Within each session, each child took part in two blocks: onomatopoeic and non-omatopoeic, with the order of sessions and blocks counterbalanced across participants. Within each block, participants undertook two training trials immediately followed by two testing trials. Thus, each participant was taught and tested on eight event-label pairs in total.

For each child, we randomly sampled events from our full set of 16, with half randomly assigned to the situated and half to the displaced context. Within each context set, half were randomly paired with their onomatopoeic label, and half with their non-omatopoeic control label. Labels within a set of eight were unique to particular events: a participant would not learn the same label for more than one event. Furthermore, children were never tested on event pairs where the label could be the alternative label for the distractor event (i.e. the control, if the trial is onomatopoeic). We resampled the set of eight if this was the case.

### 3.1.4 | Procedure

The experiment was designed as a web application, using the p5 JavaScript library, and conducted using two Apple iPads. Children’s responses were recorded as touchscreen presses. The experiment was conducted in English, with novel verbs embedded in English carrier phrases.

Testing took place in a room separate from the main nursery class, with two experimenters, the participating child and nursery staff (or a caregiver) present in the room, out of the child’s view. One experimenter sat at a small table opposite the child. A second experimenter sat slightly behind the child, to note unusual responses (e.g. if the child pressed both screens, or responded too quickly). Sessions for both contexts were run on the same day, with a minimum of a half hour break in between, in which the child returned to their class. The length of the break depended on the nursery’s and the child’s schedule.

The experimenter sitting opposite the child conducted the task. Two iPads stood in stands on the table facing the child and away from the experimenter. The progression of the experiment was controlled with two Bluetooth keyboards, placed behind the screens. Experimenters took turns conducting the experiment, although the same experimenter conducted both sessions for any given child. All videos played with audio through the iPad speakers, set at the same level throughout. The experimenter conducting the experiment was given a cue card for that session that gave the labels to produce at each trial. The cue card did not indicate whether a trial was onomatopoeic or non-omatopoeic to minimise production differences



between onomatopoeic and control labels based on, for example, prosody (Laing et al., 2017). Experimenters aimed to produce labels across trials with prosody typical of child-directed speech.

Children first played with a set of wooden toys, with the experimenters. When they appeared comfortable, children were asked if they would like to play a game on the iPads, and, if the child consented, we proceeded to the experiment.

Participants were first introduced to the actor in the videos, "Holly" (a pseudonym), in a video showing her waving and saying hello. The child was told that Holly wanted to play a game with them to help them learn some new words. The child next completed two practice tests: first, in a picture selection task the child was shown an apple and a banana on screen and asked to find one of them. Both the target and position of each item were randomly assigned. The experimenter first asked, "Can you find the TARGET?" If further cues were required to get a response from the child, the experimenter asked, "Where is the TARGET?" The child made a response by pressing the screen, which changed to a block colour when touched (randomly selected from a set of 6). The second practice trial followed the same procedure, but used event videos not used in the main experiment, instead of pictures. The videos depicted Holly tearing a sheet of paper in one (target *tearing*), and chopping a carrot in the other (target *chopping*).

Next came the experimental blocks. Each block consisted of two training and two testing trials. In each training trial, participants were given three exposures to an event-label pair. The procedure differed depending on the condition (Figure 3): for *situated* trials, as the video played, the experimenter said "Look!" to direct the child's attention to the screen, followed by the carrier phrase, "Holly is TARGET-ing!". Between video plays, participants were shown a blank white screen. In the *displaced*

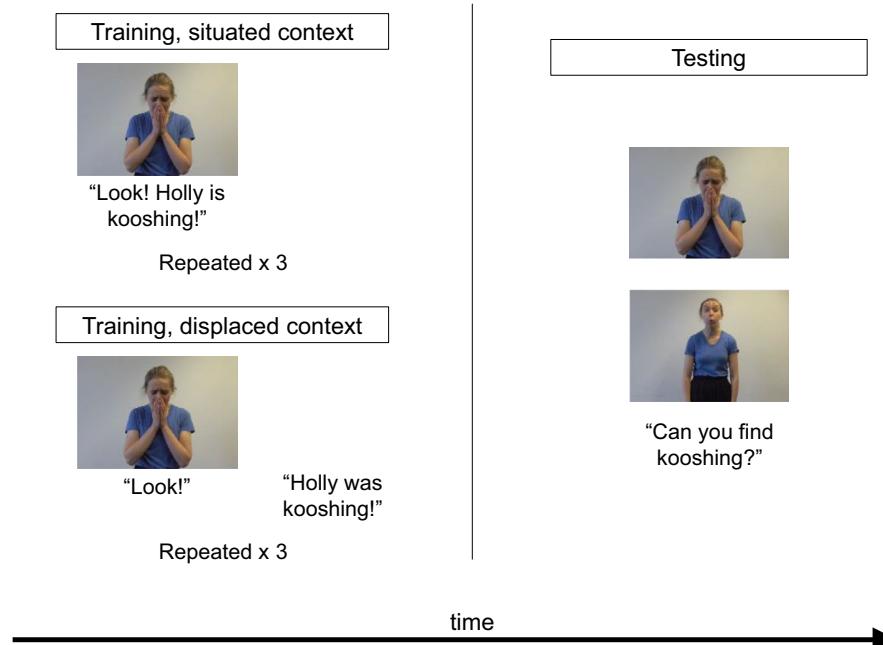
condition, the experimenter said "Look!" to draw the child's attention the video being played, but did not give the carrier phrase while the video played. Instead the experimenter said, "Holly was TARGET-ing!" after the video had stopped playing and the blank white screen showed. The screen on which each item was shown was randomly assigned, but balanced, with the two events shown on different screens.

Testing in each block immediately followed the training trials. For each trial, two videos were shown on screen, looping until participants made a response or the experimenter progressed the experiment. The second event video used in each block acted as the distractor item, such that participants saw a testing trial where the event was both target and distractor. The screen each video appeared on was randomly assigned for each trial. Once both videos had begun to play, the experimenter asked "Can you find TARGET-ing?" Further prompts were given if required, either "Where is TARGET-ing?" or "Which one is TARGET-ing?" The experimenter looked directly at the child throughout testing. The child made their response by pressing either screen, which changed colour upon being touched. After the first block, the experiment progressed immediately to the second.

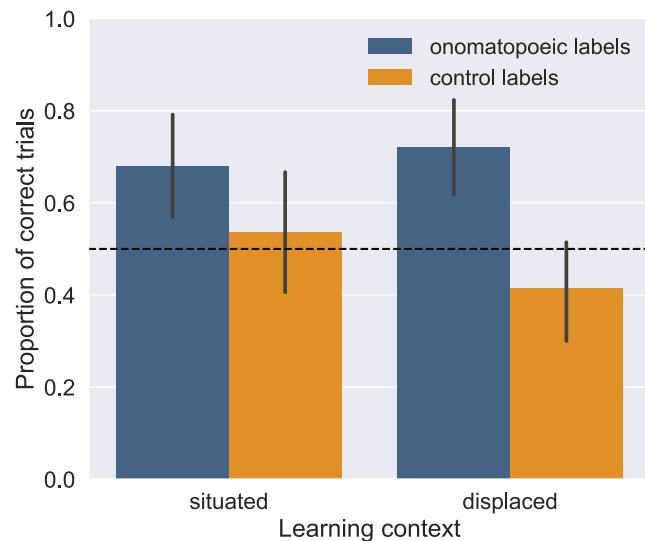
The full procedure described above took approximately 10 min to complete, and was repeated for both sessions.

### 3.2 | Results

Figure 4 shows children's performance in testing, by context and trial type. We subjected binary data (correct vs. incorrect response) to logistic mixed-effects models. Where children had touched both screens (usually due to over-excitement), we took their first response



**FIGURE 3** Outline of experimental procedure. Training is different in situated and displaced contexts, but testing is the same for both conditions. Testing occurs immediately following training for the two items in each block



**FIGURE 4** Proportion of correct responses across contexts (situated/displaced) and trials (onomatopoeic/control) in the experiment. The dashed black line represents chance performance (0.5). Error bars represent 95% confidence intervals

for analysis. Where children responded too quickly (before a label had been given), we removed these data points from analysis. Fixed effects were as follows: context (situated/displaced), iconicity (onomatopoeic/non-onomatopoeic control) and their interaction. We used a centred coding for fixed predictors so that the intercept corresponds to the grand-mean (in log-odds). We included a random intercept for participants and a by-participant random slope for iconicity (a full random-effects structure would not converge), and implemented optimisation by the Nelder-Mead method.

The model revealed that children learnt onomatopoeic labels better than control labels ( $M_{\text{onomat}} = 0.70$ ,  $M_{\text{control}} = 0.47$ ,  $\beta = 1.13$ ,  $SE = 0.3$ ,  $p < 0.001$ ). We found no evidence for better performance in situated over displaced trials ( $M_{\text{displaced}} = 0.57$ ,  $M_{\text{situated}} = 0.61$ ,  $\beta = 0.15$ ,  $SE = 0.27$ ,  $p = 0.58$ ), and no evidence for an interaction between iconicity and learning context ( $\beta = 0.74$ ,  $SE = 0.53$ ,  $p = 0.17$ ). Onomatopoeic labels were easier to learn than control labels, and this did not differ by learning context. We tested performance against chance in each trial type by fitting the same model with different intercepts for each trial type. This analysis showed evidence of above chance performance for onomatopoeic trials ( $\beta = 1.03$ ,  $SE = 0.28$ ,  $p < 0.001$ ) but not for control trials ( $\beta = -0.1$ ,  $SE = 0.18$ ,  $p = 0.58$ ).

In summary, the onomatopoeic labels were learnt significantly better than control labels in our experiment. Indeed, children struggled to learn the control labels above chance levels. However, we did not find evidence that children learnt either onomatopoeic or control labels differently in situated versus displaced learning contexts.

## 4 | DISCUSSION

We proposed that onomatopoeia play a special role in early vocabulary development, assisting young children in two ways: by helping

them to understand that speech sounds are meaningful, and by helping them to link specific forms to their referents. Based on this proposal, we made three key predictions. First, we predicted that onomatopoeia would be most frequently used early on in development, both by caregivers in child-directed communication and by children. This prediction is supported by the results from study 1, which showed that caregivers used proportionally more onomatopoeia when speaking to their children at younger ages than at older ages. Moreover, children also used proportionally more onomatopoeia at younger ages than at older ages. Importantly, these results extend previous research indicating that onomatopoeia and other iconic vocabulary are more common at the very early stages of children's language production (Kauschke & Hofmeister, 2002; Laing, 2014; Perry et al., 2017), are acquired earlier (Perry et al., 2015) and are common in child-directed speech (Jo & Ko, 2018; Perry et al., 2017).

Secondly, we predicted that the use of onomatopoeia occurs early because it supports word learning. This prediction is supported in two ways: the results from study 1 indicate that caregivers use more onomatopoeia when talking about referents unfamiliar to the child than when talking about referents already known to the child. Crucially, in study 2, we found that children learnt onomatopoeic words better than words that had no relation to their meaning. This experimental work is the first to look specifically at onomatopoeic forms (sound-to-sound relation), and, again, extends previous experimental research that found a learning advantage for sound-symbolic words over neutral forms, for both children (Imai et al., 2008; Kantartzis et al., 2011, 2019; Yoshida, 2012) and adults (Dingemanse et al., 2016; Lockwood et al., 2016). Furthermore, given that our onomatopoeic and control labels are formally identical, we suggest that, in this case, we can attribute our findings to the representational, iconic mappings offered by onomatopoeia, rather than the salience of phonological or prosodic features of onomatopoeia as suggested by Laing (2019). One important caveat for our study is that learning is indistinguishable from a general sensitivity to onomatopoeic forms (children could potentially map onomatopoeic labels to the correct referent in testing trials alone, without having learnt the words). However, previous work using sound-symbolic mappings finds similar results, pointing towards a learning account (Imai et al., 2008; Kantartzis et al., 2011, 2019). Additionally, Kantartzis et al. (2019) found that the learning advantage for sound symbolic forms could not be attributed solely to mapping at test alone; they demonstrated that children matched sound-symbolic items they had learnt more accurately than new sound-symbolic distractors that could simply be matched to the target event.

Finally, we predicted that onomatopoeia would be particularly helpful because they support learning both when the referent is present in the visual scene, as well as when the referent is visually (or otherwise) inaccessible, by providing an imagistic link between word and referent. The majority of research on children's vocabulary development focuses on situated contexts where label and referent co-occur. Although children are often faced with displaced language contexts (Veneziano, 2001), very few studies exist that investigate

what form child-directed communication takes in this case (Perniss et al., 2018; Tomasello & Kruger, 1992; Vigliocco et al., 2019). Our study investigated how context affects both the input the child receives and how they learn. In Study 1, we found that caregivers produced onomatopoeia both when referents were present and when they were absent. We also found an interaction between the age of the child and the presence of target referents, with onomatopoeia production maintained at higher levels over development when the referent was absent. In Study 2, contrary to our prediction, we did not find that children learnt labels differently in situated and displaced learning contexts. However, a null result is not evidence that no effect exists. Further Bayes factor analysis assessed the strength of the evidence in favour of our hypotheses over a corresponding null hypothesis (detailed at <https://osf.io/ktv8a/>) and suggested that evidence for the effect of context was inconclusive, such that we could neither firmly accept nor firmly reject an effect of displacement. Our findings from the experiment suggest that more work is needed to understand how children learn in displaced language contexts.

In summary, our findings suggest that onomatopoeia are present in the input children receive, used by caregivers in ways that might facilitate learning and used by children during learning episodes. However, the precise role that onomatopoeia play in learning episodes is somewhat complex to untangle. In a recent review of the iconicity and word learning literature, Nielsen and Dingemanse (2020) highlight the differences in evidence for the role that iconicity plays in local learning contexts—where iconic forms are easier to learn without necessarily conveying a more general learning advantage—and in general learning contexts, such that iconic forms, by virtue of being easier to learn, scaffold learning of non-iconic forms. They argue that, where evidence for a local learning enhancement attributed to iconicity is plentiful, little to no firm evidence exists of a general learning enhancement. We acknowledge here that study 2 focuses on local learning alone; we offer evidence that onomatopoeic labels are easier to learn, but not that the learning of onomatopoeic labels facilitates later acquisition of non-iconic labels. However, we assert that evidence from study 1 points to the role of onomatopoeia beyond the local learning context. Our assessment of utterance-level familiarity (known/unknown) is based on topic lemmas, not the onomatopoeic forms themselves—for example, the topic lemma might be 'bee' where the onomatopoeic form in the utterance might be 'buzz'. In this way, we demonstrate that caregivers use onomatopoeia more often when the child is less familiar with the *topic lemma*, not the onomatopoeic form. Although this is not evidence for general learning itself, it is suggestive of the broader role onomatopoeia play in language development, helping to scaffold interactions between caregivers and children when the child does not know the conventional lexical item. Caregivers can (and do) use onomatopoeia to support communication and provide conceptual and linguistic information about referents before the child has acquired the more adult-like conventional form.

Beyond these concerns about local versus general learning enhancement, we suggest that onomatopoeic forms are easier to learn

because the iconic relationship between form and meaning helps children to link linguistic forms with real-world referents. However, other explanations for the prevalence of onomatopoeia in children's early vocabularies and in child-directed language have been put forward. Laing (2019) reviews evidence that other characteristics of onomatopoeia facilitate comprehension, production and interaction with caregivers, concluding that, in addition to being iconic, onomatopoeia are made prosodically salient by caregivers and are phonologically simpler than non-iconic words. Furthermore, the high frequency of use by caregivers could, in itself, explain children's use of onomatopoeia. How, then, can we assess whether iconicity plays a role beyond these other characteristics? While for the naturalistic data in study 1 it is difficult to tease these factors apart, we suggest that our experimental design in study 2 offers evidence for an iconic learning advantage beyond these other characteristics. We re-use the same word forms across onomatopoeic and control trials, so the learning advantage cannot be attributed to relative phonological simplicity. Furthermore, we reduce the possible role that prosodic salience can play; the experimenter does not know which blocks are onomatopoeic and which are non-omatopoeic at the time of testing. They are also seated behind the iPad screens, and so do not see the event videos during play—as such, it is difficult to ascertain trial type from the co-occurrence of events with their iconic labels and therefore unlikely that prosodic salience drives learning enhancement in this case.

A further possibility for understanding the specific role that iconicity plays in the learning and use of onomatopoeic forms lies in our hypothesis about displaced learning contexts. While it is clearly the case that vocabulary learning takes advantage of co-occurrence between label and referent (Akhtar & Montague, 1999; Smith & Chen, 2008), learning can and does take place without co-occurrence (Tomasello & Barton, 1994; Tomasello et al., 1996), especially for words referring to events which unfold over time (Tomasello & Kruger, 1992). We hypothesised that iconicity would be specifically helpful in the absence of the referent, such that the link between linguistic form and real-world meaning can evoke properties of the referent and thus facilitate association. In contrast, we do not expect that prosodic salience or phonological simplicity would offer such an advantage in displaced learning contexts. While our results concerning displacement are somewhat mixed, we do find that children learn onomatopoeia more easily than control labels in displaced contexts, and that caregivers modify their use of onomatopoeia over development depending on the visual accessibility of the referent. The findings from the naturalistic data support previous work looking at caregiver-child interactions in a play-based context (Vigliocco et al., 2019), where iconic forms (both onomatopoeia and gestures) were more commonly used by caregivers in displaced contexts compared with situated contexts.

Therefore, we suggest that onomatopoeia are more learnable by young children by virtue of their relationship between form and meaning, in addition to other factors. Moreover, we assert that onomatopoeia play a special role beyond other iconic or sound symbolic



forms of communication in children's early language development. These findings have implications beyond acquisition; the ability of onomatopoeia to represent real-word properties of referents before conventional, arbitrary forms appear points to the possible role that onomatopoeia might play in the emergence of language (Perniss & Vigliocco, 2014). Onomatopoeia represent the simplest and most direct type of mapping between form and meaning in that they implement a unimodal (sound to sound), imitative mapping and refer to specific referents. For these reasons, they can be used by language learners who do not yet possess the cognitive resources or do not yet have the language experience that has been argued to be necessary to take advantage of other forms of iconicity (Irvine, 2016; Namy, 2008).

Iconicity more generally has been hypothesised to facilitate language learning in young children (Imai & Kita, 2014; Perniss & Vigliocco, 2014). Indeed, children show early sensitivity to sound symbolism unrelated to sound properties of the meaning—the bouba-kiki effect where bouba-like sounds map onto round shapes while kiki-like sounds map onto spiky ones has been found in children as young as 4 months (Ozturk et al., 2013). However, this effect may rely somewhat on the contrast between labels, rather than on the specific relationship between a word and its meaning (Dingemanse et al., 2016). Furthermore, most studies investigating iconicity in speech do not distinguish between onomatopoeia and other forms of sound symbolism, which do not provide the same direct access to meaning as onomatopoeia, and often represent cross-modal relationships (e.g. sound to shape, as with the bouba-kiki effect), requiring the child to recognise iconicity in a more abstract form that may not be as readily accessible as direct sound-to-sound mappings. This special role for onomatopoeia over other forms of lexical iconicity may find support from sign language acquisition, where iconic signs are the earliest acquired (Caselli & Pyers, 2017; Thompson et al., 2012; Vinson et al., 2008). Iconic signs (e.g. British Sign Language EAT, which involves putting the hand to the mouth as if lifting food to the mouth) are comparable to onomatopoeia, in that they provide a representation of meaning directly linked to the primary channel of articulation. Iconicity in other forms may support communication at a general level, but onomatopoeia have the potential to play a primary role in early spoken vocabulary learning, allowing linguistic infants (and possibly early linguistic humans) to build a lexical repertoire grounded in real-world sensory experience.

In sum, we bring together evidence from naturalistic caregiver-child interactions and experimental research on children's word learning to shed light on the role that onomatopoeia play in children's linguistic development. We found that children are able to learn onomatopoeic forms more easily than non-onomatopoeic ones, and that caregivers use onomatopoeia in ways that scaffold interactions with their children, and which may help learning. Furthermore, we have demonstrated that contextual factors such as displacement may affect the input that children receive. Given that most language acquisition research focuses solely on contexts where label and referent co-occur, we highlight the need to further understand how children learn in different contexts. Although

onomatopoeia has often been seen as a trivial and unimportant part of language, our findings suggest that it has the potential to play a substantial role in vocabulary learning. This role is most clearly evident in first language acquisition, but we also suggest that these findings could have wider implications for theories of how language evolved. In both cases, grounding linguistic form in sensory experience could facilitate form-meaning association for new learners, and could further explain why iconic forms persist on an evolutionary timescale and in everyday language use by adults. Far from being a trivial aspect of language, onomatopoeia may play a key role in how human languages emerge, evolve and are learnt by their users.

## ACKNOWLEDGEMENTS

This work was funded by an ESRC research grant (ES/P00024X/1) awarded to GV, PP, EW, CM and SGM, and by NICHD P01 HD40605 Program Project Grant awarded to SGM. GV was supported by a Royal Society Wolfson Research Merit Award (WRM\R3\170016). We thank Isabel Scharf and Alyssa Clemente for their contributions to data coding for study 1, and Michael Hochman and Kristi Schonwald who assisted us in accessing the LDP corpus. We thank Oisin Twomey-Brenner for his assistance with creating stimuli and collecting data for study 2. We also thank the reviewers whose valuable and constructive feedback helped improve this manuscript.

## CONFLICTS OF INTEREST

There are no conflicts of interest to declare.

## DATA AVAILABILITY STATEMENT

Data, stimuli and analysis files from the studies can be found at <https://osf.io/ktv8a/>. Data from study 2 are included here, but not data from study 1. The data in study 1 were sampled from the Language Development Project corpus, a longitudinal study. This study is ongoing and many of the children from the study are turning 18. As such, managers of the corpus are in the process of re-obtaining consent and therefore cannot yet make the data publicly available. We have included on the OSF page the analysis scripts for study 1.

## ORCID

Chloë Marshall <https://orcid.org/0000-0003-2405-1999>

## ENDNOTE

<sup>1</sup>Parents provided age and language information about the children. Ten children in the sample were listed as having a language in addition to English.

## REFERENCES

- Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, 19(57), 347–358. <https://doi.org/10.1177/014272379901905703>
- Bates, D., Bolker, B., Machler, M., & Walker, S. C. (2015). Fitting linear mixed-effects models using Lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>

- Brand, R. J., Shallcross, W. L., Sabatos, M. G., & Massie, K. P. (2007). Fine-grained analysis of motionese: Eye gaze, object exchanges, and action units in infant-versus adult-directed action. *Infancy*, 11(2), 203–214. <https://doi.org/10.1111/j.1532-7078.2007.tb00223.x>
- Caselli, N. K., & Pyers, J. E. (2017). The road to language learning is not entirely iconic: Iconicity, neighborhood density, and frequency facilitate acquisition of sign language. *Psychological Science*, 28(7), 979–987. <https://doi.org/10.1177/0956797617700498>
- de Saussure, F. (1916). *Cours de linguistique générale*. Payot.
- Dingemanse, M., Schuerman, W., Reinsch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language*, 92(2), e117–e133. <https://doi.org/10.1353/lan.2016.0034>
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., Pethick, S. J., ... Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5), 1–185. <https://doi.org/10.2307/1166093>
- Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to Infants. *Child Development*, 64(3), 637–656. <https://doi.org/10.2307/1131208>
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20(1), 104–113. <https://doi.org/10.1037/0012-1649.20.1.104>
- Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. (2005). Hard words. *Language Learning and Development*, 1(1), 23–64. <https://doi.org/10.1038/scientificamerican0491-138>
- Goldin-Meadow, S. (2007). Pointing sets the stage for learning language—and creating language. *Child Development*, 78(3), 741–745. <https://doi.org/10.1111/j.1467-8624.2007.01029.x>
- Goldin-Meadow, S., Levine, S. C., Hedges, L. V., Huttenlocher, J., Raudenbush, S. W., & Small, S. L. (2014). New evidence about language and cognitive development based on a longitudinal study: Hypotheses for intervention. *The American Psychologist*, 69(6), 588–599. <https://doi.org/10.1037/a0036886>
- Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1), 335–346. [https://doi.org/10.1016/0167-2789\(90\)90076-6](https://doi.org/10.1016/0167-2789(90)90076-6)
- Herold, D. S., Nygaard, L. C., & Namy, L. L. (2012). Say it like you mean it: Mothers' Use of Prosody to Convey Word Meaning. *Language and Speech*, 55(3), 423–436. <https://doi.org/10.1177/0023830911422212>
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203, 88–96.
- Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130298. <https://doi.org/10.1098/rstb.2013.0298>
- Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, 109(1), 54–65. <https://doi.org/10.1016/j.cognition.2008.07.015>
- Irvine, E. (2016). Method and evidence: Gesture and iconicity in the evolution of language. *Mind & Language*, 31(2), 221–247. <https://doi.org/10.1111/mila.12102>
- Iverson, J. M., Capirci, O., & Cristina Caselli, M. (1994). From communication to language in two modalities. *Cognitive Development*, 9(1), 23–43. [https://doi.org/10.1016/0885-2014\(94\)90018-3](https://doi.org/10.1016/0885-2014(94)90018-3)
- Jo, J., & Ko, E.-S. (2018). Korean mothers attune the frequency and acoustic saliency of sound symbolic words to the linguistic maturity of their children. *Frontiers in Psychology*, 9, 2225. <https://doi.org/10.3389/fpsyg.2018.02225>
- Kantartzis, K., Imai, M., Evans, D., & Kita, S. (2019). Sound symbolism facilitates long-term retention of the semantic representation of novel verbs in three-year-olds. *Languages*, 4(2), 21. <https://doi.org/10.3390/languages4020021>
- Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science*, 35(3), 575–586. <https://doi.org/10.1111/j.1551-6709.2010.01169.x>
- Kauschke, C., & Hofmeister, C. (2002). Early lexical development in German: A study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language*, 29(4), 735–757. <https://doi.org/10.1017/S0305000902005330>
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44(4), 978–990. <https://doi.org/10.3758/s1342-8-012-0210-4>
- Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. *First Language*, 34(5), 387–405. <https://doi.org/10.1177/0142723714550110>
- Laing, C. E. (2017). A perceptual advantage for onomatopoeia in early word learning: Evidence from eye-tracking. *Journal of Experimental Child Psychology*, 161, 32–45. <https://doi.org/10.1016/j.jecp.2017.03.017>
- Laing, C. E. (2019). A role for onomatopoeia in early language: Evidence from phonological development. *Language and Cognition*, 11(2), 173–187. <https://doi.org/10.1017/langcog.2018.23>
- Laing, C. E., Vihman, M., & Keren-Portnoy, T. (2017). How salient are onomatopoeia in the early input? A prosodic analysis of infant-directed speech. *Journal of Child Language*, 44(5), 1117–1139. <https://doi.org/10.1017/S0305000916000428>
- Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, 42(8), 1274–1281. <https://doi.org/10.1037/xlm0000235>
- Lu, B. (2018). *The neural correlates of processing onomatopoeia*. Masters, University College London, London, UK.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd edn). Lawrence Erlbaum.
- Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78(2), 91–121. [https://doi.org/10.1016/S0010-0277\(00\)00109-8](https://doi.org/10.1016/S0010-0277(00)00109-8)
- Mortensen, D. R., Littell, P., Bharadwaj, A., Goyal, K., Dyer, C., & Levin, L. S. (2016). PanPhon: A resource for mapping IPA segments to articulatory feature vectors. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (pp. 3475–3484). ACL.
- Muller, F. M. (1861). *Lectures on the science of language*, Vol. 1. Longmans Green.
- Namy, L. L. (2008). Recognition of iconicity doesn't come for free. *Developmental Science*, 11(6), 841–846. <https://doi.org/10.1111/j.1467-7687.2008.00732.x>
- Newmeyer, F. J. (1992). Iconicity and generative grammar. *Language*, 68(4), 756–796. <https://doi.org/10.2307/416852>
- Nielsen, A. K. S., & Dingemanse, M. (2020). Iconicity in word learning and beyond: A critical review. *Language and Speech*. <https://doi.org/10.1177/0023830920914339>
- Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning. *Cognition*, 112(1), 181–186. <https://doi.org/10.1016/j.cognition.2009.04.001>
- Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, 114(2), 173–186. <https://doi.org/10.1016/j.jecp.2012.05.004>
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80(3), 674–685. <https://doi.org/10.1111/j.1467-8624.2009.01290.x>
- Perniss, P., Lu, J. C., Morgan, G., & Vigliocco, G. (2018). Mapping language to the World: The Role of iconicity in the sign language input. *Developmental Science*, 21(2), e12551. <https://doi.org/10.1111/desc.12551>

- Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369, 20130300. <https://doi.org/10.1098/rstb.2013.0300>
- Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLoS One*, 10(9), 1–17. <https://doi.org/10.1371/journal.pone.0137147>
- Perry, L. K., Perlman, M., Winter, B., Massaro, D. W., & Lupyan, G. (2017). Iconicity in the speech of children and adults. *Developmental Science*, 21(3), 1–8. <https://doi.org/10.1111/desc.12572>
- R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4), 606–621. <https://doi.org/10.1006/jmla.1996.0032>
- Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. (2011). Grounding word learning in space. *PLoS One*, 6(12), e28095. <https://doi.org/10.1371/journal.pone.0028095>
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424. <https://doi.org/10.1017/S0140525X00005756>
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), 668–671. <https://doi.org/10.1016/j.cub.2008.03.059>
- Smith, L., & Chen, Y. U. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. <https://doi.org/10.1016/j.cognition.2007.06.010>
- Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology*, 44(4), 929–938. <https://doi.org/10.1037/0012-1649.44.4.929>
- Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British sign language. *Psychological Science*, 23(12), 1443–1448. <https://doi.org/10.1177/0956797612459763>
- Thorndike, E. L. (1943). The origin of language. *Science*, 98(2531), 1–6. <https://doi.org/10.1126/science.98.2531.1>
- Tomasello, M., & Barton, M. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, 30(5), 639–650. <https://doi.org/10.1037/0012-1649.30.5.639>
- Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, 19(2), 311–333. <https://doi.org/10.1017/S0305009000900011430>
- Tomasello, M., Strosberg, R., & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language*, 23(1), 157–176. <https://doi.org/10.1017/S030500090010138>
- Veneziano, E. (2001). Displacement and informativeness in child-directed talk. *First Language*, 21(63), 323–356. <https://doi.org/10.1177/014272370120106306>
- Vigliocco, G., Motamed, Y., Murgiano, M., Wonnacott, E., Marshall, C. R., Maillo, I. M., & Perniss, P. (2019). Onomatopoeias, gestures, actions and words in the input to children: How do caregivers use multimodal cues in their communication to children?. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 1171–1177). Austin, TX: Cognitive Science Society.
- Vinson, D. P., Cormier, K., Denmark, T., Schembri, A., & Vigliocco, G. (2008). The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity. *Behavior Research Methods*, 40(4), 1079–1087. <https://doi.org/10.3758/BRM.40.4.1079>
- Vosoughi, S., Roy, B. C., Frank, M. C., & Roy, D. (2010). Contributions of prosodic and distributional features of caregiver's speech in early word learning. In S. Ohlsson & R. Catrambone (Ed.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 1822–1827). Austin, TX: Cognitive Science Society.
- Yoshida, H. (2012). A cross-linguistic study of sound symbolism in children's verb learning. *Journal of Cognition and Development*, 13(2), 232–265. <https://doi.org/10.1080/15248372.2011.573515>
- Yu, C., & Smith, L. B. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science*, 14(2), 165–180. <https://doi.org/10.1111/j.1467-7687.2010.00958.x>

**How to cite this article:** Motamed Y, Murgiano M, Perniss P, et al. Linking language to sensory experience: Onomatopoeia in early language development. *Dev Sci*. 2020;00:e13066. <https://doi.org/10.1111/desc.13066>