A Role for Onomatopoeia in Early Language: Evidence from Phonological Development

Catherine Laing

Centre for Language and Communication Research,

Cardiff University, UK

laingc@cardiff.ac.uk

A role for onomatopoeia in phonological development

Abstract

Onomatopoeia appear in high quantities in many infants' earliest words, yet there is minimal research in this area. Instead, findings from the wider iconicity literature are generalised to include onomatopoeia, leading to the assumption that their iconic status makes them inherently learnable, thereby prompting their early production. In this review we bring together the literature on onomatopoeia specifically and iconicity more generally to consider infants' acquisition from three perspectives: perception, production, and interaction. We consider these findings in relation to Imai and Kita's (2014) 'sound symbolism bootstrapping hypothesis' to determine whether their framework can account for onomatopoeia alongside other iconic forms.

Keywords: *onomatopoeia, phonological development, iconicity, lexical development, language acquisition*

A role for onomatopoeia in phonological development

Compared with adult speech, infants' early vocabulary consists of a surprisingly high proportion of onomatopoeia: 20% of Menn and Vihman's (2011) appendix of 48 infants' first five words[i] are onomatopoeic, while Tardif and colleagues (2008) show even higher proportions in the first ten words of American English (29.5% of all words) and Cantonese (40.6%; Putonghua was lower, at 8.7%), according to parental vocabulary questionnaires of infants aged 8-16 months. These data suggest that there may be something particular to onomatopoeia that makes them suitable for early acquisition. This could be explained by Imai and Kita's (2014) 'sound symbolism bootstrapping hypothesis', which posits that the presence of iconicity in a word facilitates its early acquisition. However, this perspective derives largely from studies of sound symbolism that test infants' perception of non-words; their discussion of onomatopoeia is limited to the presence of these forms in infant-directed speech. If this framework sufficiently explains an advantage for iconic forms in early language, it should be applicable to onomatopoeia, as well as other iconic forms. This paper reviews the literature on onomatopoeia, including work from the present author (Laing, 2014, 2017; Laing, Vihman & Keren-Portnoy, 2017; Laing, forthcoming), to consider whether the sound symbolism bootstrapping hypothesis can explain infants' early acquisition of these forms, or whether there are other factors that prompt their presence in the early lexicon. Perspectives from perception, production, and interaction reveal a functional role for onomatopoeia in phonological development, which may derive from their inherent iconicity.

The literature on onomatopoeia – linguistic forms that imitate sounds from the environment, e.g. *choo choo* (train sound) – in language development is limited, drawn largely from theoretical viewpoints relating to the wider iconicity research. Imai and Kita (2014) suggest an early learning advantage for iconic words via 'referential insight', which "helps infants associate speech sounds and their referents and establish a lexical representation" (p.4). This hypothesis generalises over all iconic forms, including

onomatopoeia, and recent empirical evidence supports this (e.g. Perry, Perlman, & Lupyan, 2015; see Nielsen & Dingemanse, 2018). However, while the extant research presents a convincing argument towards the sound symbolism bootstrapping hypothesis in general, its perspective is limited; for a sufficient account we must consider the real-world learning and use of onomatopoeia if these forms are applicable to Imai and Kita's (2014) framework. In what follows, we consider the existing literature on onomatopoeia to determine whether the sound symbolism bootstrapping hypothesis (Imai & Kita, 2014) can sufficiently explain their predominance in infants' early words by considering these forms from three perspectives: perception, production, and interaction.

Onomatopoeia in perception

Infants may be sensitive to sound-meaning correspondences, or sound symbolism, by age four months (Ozturk, Krehm, & Vouloumanos, 2013), which could support word learning: English- and Japanese-speaking two-year-olds were better able to learn novel verbs with sound symbolic properties (e.g. *chokachoka* to represent 'fast walking with small steps'; Imai, Kita, Nagumo, & Okada, 2008, p. 57; Kantartzis, Imai & Kita, 2011). While findings testing non-words cannot be reliably generalised across real-word learning, more recent evidence shows that this is consistent in natural language. Lockwood, Dingemanse and Hagoort (2016) tested Dutch adults on their ability to learn Japanese ideophones[ii] against 'foils' with the opposite meaning. Their results showed convincingly that Dutch adults were better able to learn previously unknown Japanese words when they mapped iconically onto their referent.

Laing (2017) used eye-tracking to test infants' mapping of real onomatopoeic and non-onomatopoeic words to their referents. British 10-11-month-olds showed better knowledge of onomatopoeia over their object word counterparts, fixating significantly longer

to a picture of a dog after hearing *Where's the woof woof*? than after *Where's the doggie?* This supports the central tenets of the sound symbolism bootstrapping hypothesis (Imai & Kita, 2014), as responses suggested stronger form-meaning associations in the onomatopoeic condition. However, Laing (2017) also identified a role for the input: infants attended longest to words that were most commonly understood (according to the communicative development inventory; CDI, Fenson et al., 1994) , though fixations were greater after hearing onomatopoeia. Another eye-tracking study testing adults' perception of onomatopoeia showed no difference in reaction times between onomatopoeia and control words (Peeters, 2006). However, simultaneous ERP analyses revealed advantages for the onomatopoeia, suggesting facilitated lexical access for these forms. Hashimoto and colleagues (2006) compared participants' responses to onomatopoeia, nouns, and real-world animal sounds. Separate brain regions were activated when processing real-world animal sounds and (non-onomatopoeic) nouns, while onomatopoeia activated both regions. The authors posit that onomatopoeia "serve as a bridge between the processing of nouns and animal sounds" (2006, p.1768), reflecting the central tenet of the sound symbolism bootstrapping hypothesis (Imai & Kita, 2014). However, they also suggest that exaggerated prosodic features of onomatopoeia may generate stronger neurological activation (Hashimoto et al., 2006); if the sound effects that often accompany onomatopoeia support faster processing, this may signal a perceptual, rather than iconic, advantage for onomatopoeia in early word learning.

Evidence from infant-directed language/speech (IDL/IDS) shows that both signed and spoken iconic forms are produced with increased salience. Laing and colleagues (2017) found that mothers produced onomatopoeia with a higher pitch, wider pitch range, and longer duration than other words in naturalistic interactions with their eight-month-olds: onomatopoeia stood out even within the IDS register. Perniss, Lu, Morgan and Vigliocco

(2018) show similar findings in their analysis of deaf mothers' IDL. The mothers modified iconic signs more than non-iconic signs, thereby making iconic forms more salient. Frequency was also a factor in both studies: while there was no difference between overall number of onomatopoeia and conventional equivalents produced by the mothers in Laing and colleagues' (2017) study (e.g. *dog* in relation to *woof*), the preponderance of reduplication in onomatopoeia (*woof woof*, *quack quack*) meant that infants heard many more tokens of onomatopoeia overall. Similarly, Perniss and colleagues (2018) identified repetition as a key characteristic of sign modification.

Input may be essential in infants' acquisition of onomatopoeia. Laing (forthcoming) analysed home-recorded data of eight infants interacting with their caregivers to identify a significant correlation (ρ=.78, p=.01; Spearman's) between the number of onomatopoeia produced by caregivers and their infants. Ortega (2017) argues that form-meaning mappings cannot be considered as iconic if they occur without knowledge of the meaning in question. *Choo choo* is acquired early in English (Fenson et al., 1994), yet most infants have no experience of the real-world sound from which it derives. Indeed, as shown in (1), Naima (Providence corpus: Demuth, Culbertson & Alter, 2006; MacWhinney, 2000) produces onomatopoeia to represent a sea lion, a frog, an owl and an elephant:

(1) Naima, 1;01 (Demuth et al., 2006)

1	MOT:	what does an owl say?
2	CHI:	[hʌhuhu] {low pitch}
3	MOT:	an owl not a dog.
4	CHI:	[hʌhuhuhu] {high pitch}
5	MOT:	hoo hoo
6	CHI:	[u:u:]
7	MOT:	hoo hoo
8	MOT:	what does a sea lion say ?
9	CHI:	[ʊ ʊ]

```
10      MOT:    [æɪ æɪ æɪ æɪ]
11      MOT:    what about an elephant ?
12      CHI:    [uː uː]
13      MOT:    what about a frog?
14      CHI:    [ɪ ɪ] {high pitch}
15      MOT:    a frog?
16      CHI:    [i i] {low pitch}
```

These words were probably not learned through experience of the real-world sounds in question; instead, they were likely acquired from the input, produced by the caregiver with salient prosodic features that make them easily segmentable from the speech stream. This departs from Imai and Kita's (2014) framework to highlight input as a key factor in acquisition; onomatopoeia occur in early production if they are common in caregiver speech. This may be further supported by the presence of prosodic salience, leading to a processing advantage for these forms (Hashimoto et al., 2006). All else being equal, the iconic properties of onomatopoeia may support the establishment of form-meaning mappings, but so long as these forms are more prominent in the early input – through prosodic salience and increased frequency – it is not possible to isolate the role of iconicity in infants' early perceptual experience of onomatopoeia.

Onomatopoeia in production

Perry, Perlman and Lupyan (2015) compared the age of acquisition for words on the CDI in relation to their iconicity ratings (from -5: "words that sound like the opposite of what they mean" to +5: "words that sound like what they mean", p.12) to find that early-acquired words were rated as more iconic, even with onomatopoeia excluded from the data. Vinson and colleagues (2008) found parallel results in the sign acquisition literature: across 300 signs, those rated as most iconic were acquired earliest.

This may not be motivated solely by an iconic advantage. Infants' early words are phonologically simple, featuring canonical CV or CVCV forms much like the properties of

babble (Kim & Davis, 2015). Thompson, Vinson, Woll and Vigliocco (2012) showed that both phonological complexity and iconicity predicted age of acquisition for deaf children acquiring British Sign Language, with increasing phonological complexity as children got older. Massaro and Perlman (2017) analysed hearing children's acquisition to show that words with higher iconicity ratings were less phonologically complex, which they suggest might contribute to their early acquisition. The articulatory simplicity of onomatopoeia is noted by Kunnari (2002) in a study of Finnish infants' syllabification. Onomatopoeia were produced more accurately than non-onomatopoeic words, and Kunnari (*ibid*) concludes that onomatopoeia may be easier for infants to produce owing to their "easily mastered articulatory shape" (p.133). She suggests that their pragmatic status in language, as well as the prosodic features that render them salient in the input, might also be important in early acquisition. In another study of early phonological development, Stoel-Gammon and Cooper (1984) report data from three infants' first 50 words, showing striking variability in the acquisition of onomatopoeia. Two infants acquired three and zero onomatopoeia, while onomatopoeia accounted for 38% of the third infant's data. Reduplication was dominant across his lexicon, accounting for 66% of all words, including 14 of the 19 onomatopoeia.

The phonological structures common to onomatopoeia – reduplication (*woof woof*) and the open CV syllable (*moo*, *baa*) – might make them particularly suitable for early acquisition. These structures are typical of infants' early words (Kent, 1992; Vihman, 1978), and Laing (forthcoming) shows that onomatopoeia consistently match these structures. Laing (*ibid*) analyses the early lexicon of 16 infants acquiring a range of languages to show striking similarities in the phonological forms of early words. Consonant harmony (including reduplication) and the open CV syllable accounted for 36% (SD=.11) and 42% (SD=.16) of the data, respectively. Of the onomatopoeia, 57% (SD=.25) were produced with consonant harmony, and 31% (SD=.17) with the CV syllable. Laing (*ibid*) concludes that infants acquire

onomatopoeia in high proportion because their phonological structures are suited to early production. Moreover, onomatopoeia across languages typically fit one of these same two structures (e.g. reduplication in 'woof woof' across Dutch *woef woef*, Hebrew *how how*, and Greek *gav gav*; CV in 'moo' across Dutch *moe*, Hebrew *moo*, and Greek *moo*; Abbot, 2004); structures that are so-called 'universals' in early language production (Kent, 1992; Vihman, 1978). This is shown in cross-linguistic data from Laing (forthcoming) in Table 1.

Table 1: *Consonant harmony and CV structures in onomatopoeia across five infants' data. See Laing (forthcoming) for full set of examples.*

Infant	**Source**	**Consonant harmony**	**CV**
William	*US English* Demuth et al. (2006)	*beep beep* [biːpbiː] *quack quack* [gʊkwæ] *brm brm* [bɜbɜ]	*beep* [biʔ] *baa* [bɑˈɑː] *moo* [muː]
Nathan	*French* Demuth & Tremblay (2008)	*tchou tchou* 'choo choo' [tityː] *toot toot* 'beep beep' [kaka] *pin-pon* 'nee-naw' (siren) [papõ]	*boum* 'boom' [bɔː]
M	*Spanish-English bilingual* Deuchar & Quay (2001)	*quack/cuac* [kak] *bow-wow* [bəʊwəʊ] *woof woof* [wʊfwʊf]	*meow* [maʊ] *baa* [be]
Annalena	*German* Elsen (1991)	*tööt* 'toot' (train) [bɪp] *brum* [ʁm] *kikeriki* 'cock-a-doodle-doo' [kɪːkɪːç]	*meh* 'baa' [meː] *tööt* 'toot' [bɪ]
P	*Czech* Pačesová (1968)	*tudu* 'honk' (car) [tidiː] *kaka* 'quack' [kaːkaka] *ticktak* 'tick tock' (clock) [tsita]	*bebe* 'baa' [beː] *mnau* 'meow' [naː] *bac* 'bang' [baː]

So far, we have discussed onomatopoeia in relation to forms such as *quack quack*, that are conventionalised within a language and possess structural and segmental features that are typical of the language in question. However, it is widely reported (e.g. Elsen, 1991; Werner & Kaplan, 1963) that infants 'create' of onomatopoeia through imitation of real-world sounds. Elsen (1994) describes her daughter's spontaneous use of onomatopoeic sound-effects (as opposed to conventionalised onomatopoeia learnt from the input) to designate contrast between animals: [pɪpɪ] to represent small birds, [bagbagba] for ducks and geese, and [bɔa], produced with a hoarse voice to represent the sound of a crow (Elsen, 1994). She also reports the use of a 'snuffling' sound to represent hares or rabbits;

> "Although not produced orally, and therefore *not accepted as a word*, it was used appropriately only for hares and rabbits… The child was able to talk about these objects successfully, which meant a rather well-developed ability to map meanings on forms." (Elsen, 1994, p.309, italics added)

Onomatopoeic creations are reported in a number of diary studies, and may offer a phonetic advantage in early production. Stern and Stern's (1928, cited in Werner & Kaplan, 1963) daughter imitated the sound of various objects and events, including the production of "ö-ö-ö pronounced rhythmically, with effort, apparently signifying the strain of the horses involved in the pulling of [a] car" (pp.101-102; see Jakobson, 1968, pp.25-27, for further examples). Following a phonological analysis of one infant's early production of onomatopoeia, Laing (2014) proposes two kinds of onomatopoeia: conventionalised forms with a fixed structure ('standard'), and more expressive forms with a more flexible structure that depict the sound in question more closely, perhaps including prosodic modifications ('functional'). The latter may be particularly important in facilitating lexical expansion, as the rehearsal of individual segments, such as [z̪] to represent a bee's buzzing, or [bv:m] to represent a car's engine

(Elsen, 1991), allows the stabilization of the early phonological system through the flexible production of onomatopoeia that have no 'set' phonological form. As Laing (2014, p.403) explains, this "facilitate[s] production by providing an impressionistic template which incorporates lexical meaning while also allowing a wide margin of phonological error." Indeed, in Laing's (*ibid*) analysis we see the production of consonants such as /v/ and /ç/ in onomatopoeia (but not other forms) at 12 months; McLeod and Crowe (2018) report that these consonants aren't stabilised in the phonological inventory until after age 3. Onomatopoeia may allow infants access to segments that are otherwise not yet present in their phonological repertoire, thereby supporting expansion of the phonological system.

Similarly, Jakobson (1968) proposed that onomatopoeia may support the processes of rehearsing, memorising and retrieving phonological segments, leading infants to establish a stable and usable phonological system. He posited that this may facilitate early phonological development through "linguistic training" (1968, p.27): an infant may produce an /ɹ/-like form when imitating a car, thus rehearsing a segment that will eventually become the target phoneme /ɹ/ (stabilized after 5;0, according to McLeod and Crowe, 2018). Without the challenges provided by the surrounding segmental material of a whole word there is room for articulatory flexibility and expressiveness that is not possible with words such as *car* or *dog*. Since these forms are imitations of real-world sounds, and are therefore unattached to a conventionalized phonological form, they offer the freedom to produce segments that are not constrained by the surrounding articulatory demands necessary for target-like word production.

This flexibility in production is not limited to the phonology of onomatopoeia, nor is it exclusive to infant language. Dingemanse (2012) shows how ideophones differ from other words though flexibility in syllable-, morpheme- and phrase-level structures. They can be reduplicated (*woof* vs. *woof woof*), lengthened ([mu] vs. [mːuːː]), and can occur within a

phrase, phrase-finally, separated from the phrase by a pause, and as free-standing utterances. Dingemanse (*ibid*) analyses ideophones of all categories, but notes the "more peripheral syntactic realization" (p.657) of onomatopoeia in comparison with other ideophones. Dingemanse and Akita (2017) relate this to expressiveness: the more expressive a form is, the less grammatical integration is required. Ideophones (and thereby onomatopoeia) are a means to depict experiences through words, and as Kita[iii] (1997, p. 388) explains, they "[encode] the message in a different dimension from the rest of the sentence," – the 'affecto-imagistic' dimension. This dimension incorporates the use of prosodic effects and iconic gestures, harnessing expressiveness in production to convey a linguistic message.

The 'affecto-imagistic' properties of onomatopoeia may make them more memorable, and more producible on a phonetic and a phonological level. This raises the question of iconicity as a supportive mechanism in production; infants' recollection of concepts in relation to the sound that accompanies them provides access to meaningful production, which otherwise may not be articulatorily possible. Production does not feature specifically in Imai and Kita's (2014) framework, but here we find a convincing argument towards a bootstrapping role for onomatopoeia: they are more producible as a result of being iconic. With this in mind, perspectives from production could expand the sound symbolism bootstrapping hypothesis, since the expressiveness of these forms allows for more expressive – that is, more flexible – phonological, morphological, and syntactic realization.

Onomatopoeia in interaction

Sasamoto and Jackson (2016) show onomatopoeia to be a communicative phenomenon in adult language. They provide the interlocutor with evidence of meaning, illustrating sensory experiences by drawing on similarities between sounds produced in the vocal tract and sounds from the environment. This can be drawn upon in the absence of a varied lexicon, providing valuable linguistic material in early communication. This is

supported by two studies from Kauschke and colleagues who analyse changes in the composition of infant and caregiver speech over time (Kauschke & Hofmeister, 2002; Kauschke & Klann-Delius, 2007). Combined findings show parallels between the number of onomatopoeic words in the input and the output. Production of onomatopoeia is highest before 18 months, becoming less frequent as vocabulary grows. Kauschke and Klann-Delius (2007) propose that onomatopoeia capture infants' attention and allow them to take part in 'conversations' with the caregiver. They posit that onomatopoeia "[facilitate] a dialogue…when children's language comprehension and production are both limited" (p.198). We see evidence for this in Examples (2)-(4), taken from the Providence corpus (Demuth et al., 2006; MacWhinney, 2000):

(2) Lily, 1;4

1	MOT:	is that a bear?
2	MOT:	what does a bear say?
3	MOT:	arr!
4	CHI:	[æ]

(3) William, 1;4

1	MOT:	what's that?
2	CHI:	doggie.
		[dʌti]
3	MOT:	doggie, what's the doggie say?
4	CHI:	woof.
		[m̩]

(4) Naima, 1;5

1	MOT:	you're a chicken.
2	MOT:	buck buck buck buck buck buck baga:h.
3	CHI:	buck buck.
		[bæp bæp]

Examples (1)-(4) show how onomatopoeia support infant-caregiver dialogues in early development. The forms are phonologically simple but meaningful, and allow practice at turn-taking and the development of social, pragmatic and linguistic competence. In Examples (2) and (3) the mothers pose questions (Ex(2) line 2, Ex(3) lines 1 and 3), to which their infants respond appropriately with onomatopoeia (Ex(2) line 4, Ex(3) lines 2 and 4).

Returning to (1), we see an extended interaction, where a variety of onomatopoeia are used to allow seven conversational turns. Note that phonological form is largely unimportant: aside from /h/ in lines 2 and 4, Naima produces only vowels. Instead, prosodic and extra-linguistic features make these rudimentary forms meaningful. Naima draws on a range of para-linguistic features to produce onomatopoeia, which are conventionalised between mother and infant – this is clear when Naima produces the 'wrong' features in lines 2 and 14. Phonologically-speaking, the 'correct' form and the 'wrong' form are almost identical; only their prosody makes them distinct and, consequently, meaningful. We might even say that these para-linguistic features are contrastive, bringing their own distinct meanings to rudimentary syllables, and thereby allowing a conversation between mother and infant.

Onomatopoeia provide ample material for early turn-taking dialogues with adult-like sequencing within the constraints of the developing lexicon. The properties discussed here are specific to onomatopoeia, as these forms are particularly suitable for turn-taking routines: they are recursive (i.e. caregivers can prompt endless numbers of onomatopoeia), the nature of the routine can develop with an infant's phonological capacity, and caregivers' use of the non-onomatopoeic form may support infants' transition to further acquisition (i.e. from *woof* to *dog*; Laing, 2014). This suggests a supportive role for onomatopoeia in early development that cannot be generalised across iconic forms; they are not only phonetically and phonologically well-suited to early production, but they also allow for early turn-taking routines with the caregiver. Again, this departs from the sound symbolism bootstrapping

hypothesis (Imai & Kita, 2014), and highlights a specific role for onomatopoeia in early language development.

Iconicity and onomatopoeia

Evidence supports a learning advantage for onomatopoeia over non-onomatopoeic forms. But, broadly speaking, the advantages presented above are manifested in mechanisms known to facilitate language acquisition across the board: onomatopoeia are more salient in the input, they occur in isolation, they draw on prosodic, phonetic and phonological capacities that are available to young infants, and they facilitate turn-taking with the caregiver. Even so, an outright rejection of iconicity would be misleading. Onomatopoeia have a special status in language (Dingemanse & Akita, 2017), which derives from the iconic properties inherent in these forms. Indeed, it may be these very properties that make them well-suited for early acquisition across perception, production, and interaction.

Imai and Kita (2014) suggest that infants' sensitivity to iconic forms is driven by "a biologically endowed ability to map and integrate multi-modal input" (p.4). This is particularly relevant to onomatopoeia, which possess the simplest mapping between form and meaning (Dingemanse, 2012): the use of sound via the human voice to represent sound from the environment. Evidence from the multimodality literature supports this; Giard and Peronnet (1999) found that subjects responded faster and more accurately in a reaction-time task when bimodal (audio+visual), rather than unimodal stimuli (auditory or visual), were presented. Molholm, Ritter, Javitt and Fox (2004) found similar results in adults' responses to pictures of animals and animal vocalizations. Participants performed better when pictures were viewed alongside the animal sounds. Multisensory processing thus appears to support the integration of both linguistic and non-linguistic stimuli.

Masuda (2002) conducted an acoustic-phonetic analysis of onomatopoeia in relation to their natural sounds to show that onomatopoeia bear acoustic resemblance to their real-world referent. This may mean that the advantages shown for bimodal correspondences between real-world sound and meaning (Giard & Peronnet, 1999; Molholm et al., 2004) are particularly relevant to onomatopoeia. The phonological-acoustic correspondences that characterise these forms may support faster form-meaning mapping, which may have led to differing responses across Laing's (2017) results. This would support Imai and Kita's (2014) proposal, and motivate the early acquisition of onomatopoeia in particular. Of course, this would only be relevant if the infant were familiar with the corresponding real-world sound. Indeed, the infants in Laing's (2017) study responded consistently in some trials (cat, cow, dog, duck) but not others (sheep, car): results may have been affected by the infants' experience with the real-world sounds, and perhaps also the degree of acoustic match with the onomatopoeic form, which may prompt the acquisition of some onomatopoeia over others.

English has a relatively low number of ideophones compared with languages such as Japanese, which is rich in iconic forms known as 'mimetics' (see Kita, 1997). Mimetics relate to auditory (*kaki:n* 'intensive collision of metallic objects'), visual (*pika* 'a flash of light'), kinetic (*pyon* 'a swift jump'), emotional (*sowasowa* 'restlessness due to anxiety before an important event') or tactile experiences (*nurunuru* 'tactile sensation caused by a slimy object'), amongst others (Kita, 1997, pp.381-382). There are thousands of mimetic words in Japanese (Imai & Kita, 2014), and these are not reserved for child language or informal speech (Kita, 1997). Given that iconic forms are so common in Japanese, a consideration of Japanese acquisition may shed light on the extent to which input affects acquisition of these forms. Tsujimura (2005) shows that early-acquired mimetics represent either sound (onomatopoeia, e.g. *wanwan* 'woof woof') or manner (e.g. *poi* while throwing a ball). Forms

relating to visual and tactile experiences were acquired later. This fits with Dingemanse's (2012) 'implicational hierarchy of ideophones', which posits that "if a language has ideophones at all, it will have at least ideophones for sound (i.e. onomatopoeia)" (p.663). Sound-to-sound mappings are the simplest and perhaps most iconic of all ideophones. They may be acquired earlier because of the better-integrated multimodal mapping that sound-to-sound correspondences allow.

Furthermore, Fernald and Morikawa (1993) show a special status for onomatopoeia in Japanese IDS. Japanese mothers produced a higher number of onomatopoeia compared with American mothers, noted as "the most striking cultural difference" (p.645) in their sample. Across 30 Japanese mothers, onomatopoeia were used as labels for target objects (e.g. referring to a dog as a *woof woof*) in 52% of utterances; one mother used onomatopoeia in 78% of object labels. This was true for only one of 30 American mothers. Japanese mothers also used significantly more 'nonsense sounds' to accompany onomatopoeia, in imitation of real-world sounds. Mazuka, Kondo and Hayashi (2008) observed similar trends: in the IDS of 23 Japanese mothers, five of the seven most commonly-produced words were onomatopoeic. Many of these contained repetition, e.g. /waNwaN/ to represent 'dog' (*woof woof*), and reduplication was frequent throughout the data; 65% of infant-directed vocalizations contained reduplications or repetitions, while 39% constituted onomatopoeia or mimetics. When they compared these results with adult-directed speech, they found that the use of onomatopoeia and reduplication was particular to Japanese mothers' speech with their infants, not with other adults.

Frequency and salience of onomatopoeia in the input is no doubt an important factor in early acquisition. However, there may also be an indirect effect of iconicity in caregivers' production of onomatopoeia. Since they represent sounds from the environment, they are set up to be produced in a manner that separates them from the rest of the speech stream. *Meow*

is produced with a high pitch because the sound it represents is high-pitched; high pitch isn't a default feature of onomatopoeia in IDS. Laing et al. (2017) show that *ribbit*, for example, was produced at a lower pitch than *frog* because a frog's call is low-pitched. It is iconicity that accounts for these differences, not merely the affective speech-style typical of IDS. Imai and Kita (2014) discuss input only in relation to multimodality, but as well as the benefits of multimodality in onomatopoeia, infants may benefit from features that derive from iconicity, as caregivers draw upon the idiosyncratic status of onomatopoeia when modifying these words.

As well as being more salient in the input, onomatopoeia are phonologically simple and phonetically flexible, and thereby better-suited to early production than many non-onomatopoeic forms. Imai and Kita (2014) discuss the relevance of iconicity for word learning – that is, why these forms might be *understood* more easily – but they don't discuss how this relates to production. The advantages for onomatopoeia in production shown here highlight the need for an expansion of Imai and Kita's (2014) theory: their iconic status may be central to their producibility. As Laing (forthcoming) explains, it is to be expected that human production of non-human sounds does not require phonologically-complex structures. Onomatopoeia are simple in structure because their equivalent real-world sounds are also simple (to human ears, at least); they often feature repetition or pitch modification, which can be imitated easily in infancy. It is perhaps no accident that their rudimentary structures match the structures that are most accessible to the developing vocal tract.

Summary

There are advantages for onomatopoeia in early perception, production, and interaction. Prosodic modifications in caregivers' speech mean these forms stand out in the early input, and reduplication means they are often more frequent. They are phonologically simple, with structures thought to be 'universal' in early production (Kent, 1992; Vihman,

1978). They are also phonologically flexible, as infants draw upon prosodic and extra-linguistic features in the absence of reliable phonological ability. This makes them ideal for early interactions, and allows infants to take part in conversational routines despite limited linguistic capacity. Geographical and cultural limitations should be noted, however[iv]; Kaluli infants growing up in Papua New Guinea, for example, are not engaged in dyadic interactions with the caregiver (Ochs & Schieffelin, 2009), and so perspectives from populations with differing child-rearing practices would shed further light on the role of iconicity in early language. Indeed, expanding the consideration of onomatopoeia and sound symbolism to include a more diverse set of populations is a necessary next step in this field.

From the existing evidence we could conclude that infants are primed to acquire onomatopoeia; they offer learning advantages across a range of key learning mechanisms. In contrast with the sound symbolism bootstrapping hypothesis (Imai & Kita, 2014), this viewpoint does not require any advantage for iconicity. However, iconicity may be central to these properties: iconicity makes onomatopoeia more salient and phonologically simple. As explained by Dingemanse (2012, p.657), "they are made of the same material as ordinary words – the stuff of speech – but they use it in a different way." Perhaps that is what makes them important in early development; they stand separately from speech in ways that are relevant to early acquisition. In the broad picture of phonological development they enjoy no special status, but they are learned early because they are more learnable, and produced first because they are more producible.

References

Abbott, D. (2004). *Animal sounds.* [Online]. [Accessed 30th July 2018]. Available from: http://www.eleceng.adelaide.edu.au/personal/dabbott/animal.html.

Demuth, K., Culbertson, J. & Alter, J. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language and Speech*, **49**, 137-173.

Demuth, K. & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, **35**, 99-127.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Linguistics and Language Compass*, **6**, 654–672.

Dingemanse, M., & Akita, K. (2017). An inverse relation between expressiveness and grammatical integration: On the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics*, **53**, 501–532.

Elsen, H. (1991). *Erstspracherwerb: Der Erwerb des deutschen Lautsystems*. Wiesbaden: Deutscher Universitaets Verlag.

Elsen, H. (1994). Phonological constraints and overextensions. *First Language*, **14**, 305–315.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J. & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59**, i-185.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, **64**, 637–656.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of*

Cognitive Neuroscience, **11**, 473–490.

Hashimoto, T., Usui, N., Taira, M., Nose, I., Haji, T., & Kojima, S. (2006). The neural mechanism associated with the processing of onomatopoeic sounds. *NeuroImage*, **31**, 1762–1770.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **369**, 20130298.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, **109**, 54–65.

Jakobson, R. (1968). *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton.

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science*, **35**, 575–586.

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**, 735–757.

Kauschke, C. & Klann-Delius, G. 2007. Characteristics of maternal input in relation to vocabulary development in children learning German. In I. Gülzow & N. Gagarina (eds.) *Frequency Effects in Language Acquisition: Defining the Limits of Frequency as an Explanatory Concept* (pp. 181-204). Berlin, Germany: Walter de Gruyter GmbH.

Kent, R. D. (1992). The biology of phonological development. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Phonological development: Models, research, implications*

(pp. 65-90). Timonium, MD: York Press.

Kim, N., & Davis, B. L. (2015). A phonetic approach to consonant repetition in early words. *Infant Behavior and Development*, **40**, 193–203.

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics, *Linguistics*, **35**, 379–415.

Kunnari, S. (2002). Word length in syllables: Evidence from early word production in Finnish. *First Language*, **22**, 199–135.

Laing, C. E. (2014). A phonological analysis of onomatopoeia in early word production. *First Language*, **34**, 387–405.

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.

Laing, C. E. (forthcoming). Phonological motivation for the acquisition of onomatopoeia: An analysis of early words. *Language Learning and Development*.

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**, 1117–1139.

Lockwood, G., Dingemanse, M., & Hagoort, P. (2016). Sound-symbolism boosts novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **42**, 1274–1281.

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Massaro, D. W., & Perlman, M. (2017). Quantifying iconicity's contribution during language

acquisition: implications for vocabulary learning. *Frontiers in Communication*, **2**, 1–14.

Masuda, K. (2002). *A phonetic study of sound symbolism*. Unpublished PhD Thesis. University of Cambridge, UK.

Mazuka, R., Kondo, T. & Hayashi, A. (2008). Japanese mothers' use of specialised vocabulary in infant-directed speech: Infant-directed vocabulary in Japanese. In N. Masataka (Ed.) *The Origins of Language: Unravelling Evolutionary Forces*, Tokyo: Springer.

McLeod, S., & Crowe, K. (2018). Children's Consonant Acquisition in 27 Languages: A Cross-Linguistic Review. *American Journal of Speech-Language Pathology*, 1-26.

Menn, L., & Vihman, M. M. (2011). Features in child phonology: Inherent, emergent, or artefacts of analysis? In N. Clements & R. Ridouane (Eds.), *Where Do Phonological Features Come From?* (pp. 261–301). Amsterdam: John Benjamins.

Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: A high-density electrical mapping study. *Cerebral Cortex*, **14**, 452–465.

Nielsen, A. & Dingemanse, M. (2018). Iconicity in word learning and beyond: A critical review. https://doi.org/10.31234/osf.io/9hgsk

Ochs, E., & Schieffelin, B. (2009). Language acquisition and socialization: Three developmental stories and their implications. *Linguistic anthropology: A reader*, Oxford:Wiley Blackwell, pp. 263-301.

Ortega, G. (2017). Iconicity and sign lexical acquisition: A review. *Frontiers in Psychology*, **8**, 1–14.

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**, 173–186.

Pačesová, J. (1968). *The development of vocabulary in the child*. Brno, Czech Republic: Universita J. E. Purkyne.

Peeters, D. (2006). Processing consequences of onomatopoeic iconicity in spoken language comprehension. In *38th Annual Meeting of the Cognitive Science Society (CogSci 2016)* (pp. 1637–1642).

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**, e12551.

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**, e0137147.

Sasamoto, R., & Jackson, R. (2016). Onomatopoeia - showing-word or saying-word? relevance theory, lexis, and the communication of impressions. *Lingua*, **175**, 36–53.

Stoel-Gammon, C., & Cooper, J. A. (1984). Patterns of early lexical and phonological development. *Journal of Child Language*, **11**, 247–271.

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology*, **44**, 929–938.

Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: evidence from British Sign Language. *Psychologica Science*, **23**, 1443–1448.

Tsujimura, N. (2005). Mimetic verbs and innovative verbs in the acquisition of Japanese. In

Annual Meeting of the Berkley Linguistics Society (pp. 371–382).

Vihman, M. M. (1978). Consonant harmony: Its scope and function in child language. In J. H. Greenberg (ed.), *Universals of Human Language* (pp. 281-334). Stanford, CA: Stanford University Press.

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**, 1079–1087.

Werner, H. & Kaplan, B. (1963). *Symbol formation*. New York: John Wiley & Sons.

Table 1: *Consonant harmony and CV structures in onomatopoeia across five infants' data. See Laing (forthcoming) for full set of examples.*

Infant	**Data source**	**Consonant harmony**	**CV**
William	*US English* Demuth, Culbertson & Alter (2006)	*beep beep* [biːpbiː] *quack quack* [gʊkwæ] *brm brm* [bɜbɜ]	*beep* [biʔ] *baa* [bɑˈɑː] *moo* [muː]
Nathan	*French* Demuth & Tremblay (2008)	*tchou tchou* 'choo choo' [tityː] *toot toot* 'beep beep' [kaka] *pin-pon* 'nee-naw' (siren) [papõ]	*boum* 'boom' [bɔː]
M	*Spanish-English bilingual* Deuchar & Quay (2001)	*quack/cuac* [kak] *bow-wow* [bəʊwəʊ] *woof woof* [wʊfwʊf]	*meow* [maʊ] *baa* [be]
Annalena	*German* Elsen (1991)	*tööt* 'toot' (train) [bɪp] *brum* [ʁm] *kikeriki* 'cock-a-doodle-doo' [kɪːkɪːç]	*meh* 'baa' [meː] *tööt* 'toot' [bɪ]
P	*Czech* Pačesová (1968)	*tudu* 'honk' (car) [tidiː] *kaka* 'quack' [kaːkaka] *ticktak* 'tick tock' (clock) [tsita]	*bebe* 'baa' [beː] *mnau* 'meow' [naː] *bac* 'bang' [baː]

[i] Menn & Vihman (2011) judged a word to be 'acquired' if it was produced spontaneously by the infant during a recording session in the home between age 9-18 months.
[ii] Onomatopoeia (words that represent sounds; Laing, 2014), ideophones (words that represent all manner of sensory experiences; Dingemanse, 2012) and mimetics (words in Japanese that represent sensory experiences; Kita, 1997) are differentiated here, though these terminologies are interrelated.
[iii] Kita's account discusses mimetics specifically.
[iv] Thank you to the anonymous reviewer for this suggestion